



HAL
open science

On reconstructing Proto-Bantu grammar

Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois, Sara Pacchiarotti

► **To cite this version:**

Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois, Sara Pacchiarotti. On reconstructing Proto-Bantu grammar. Language Science Press, 2022, 978-3-96110-406-2. hal-04030985

HAL Id: hal-04030985

<https://hal.science/hal-04030985>

Submitted on 15 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On reconstructing Proto-Bantu grammar

Edited by

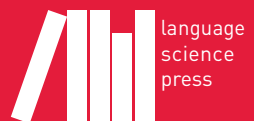
Koen Bostoen

Gilles-Maurice de Schryver

Rozenn Guérois

Sara Pacchiarotti

Niger-Congo Comparative Studies 4



Niger-Congo Comparative Studies

Chief Editor: Valentin Vydrin (INALCO – LLACAN, CNRS, Paris)

Editors: Larry Hyman (University of California, Berkeley), Konstantin Pozdniakov (INALCO – LLACAN, CNRS, Paris), Guillaume Segerer (LLACAN, CNRS, Paris), John Watters (SIL International, Dallas, Texas).

In this series:

1. Watters, John R. (ed.). East Benue-Congo: Nouns, pronouns, and verbs.
2. Pozdniakov, Konstantin. The numeral system of Proto-Niger-Congo: A step-by-step reconstruction.
3. Pozdniakov, Konstantin. Proto-Fula–Sereer: Lexicon, morphophonology, and noun classes.
4. Bostoen, Koen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.). On reconstructing Proto-Bantu grammar.

ISSN (print): 2626-3513

ISSN (electronic): 2627-0048

On reconstructing Proto-Bantu grammar

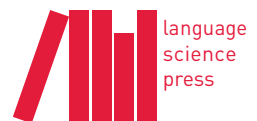
Edited by

Koen Bostoen

Gilles-Maurice de Schryver

Rozenn Guérois

Sara Pacchiarotti



Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.). 2022. *On reconstructing Proto-Bantu grammar* (Niger-Congo Comparative Studies 4). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/373>

© 2022, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: 978-3-96110-406-2 (Digital)

978-3-98554-064-8 (Hardcover)

ISSN (print): 2626-3513

ISSN (electronic): 2627-0048

DOI: 10.5281/zenodo.7560553

Source code available from www.github.com/langsci/373

Errata: paperhive.org/documents/remote?type=langsci&id=373

Cover and concept of design: Ulrike Harbort

Typesetting: Gilles-Maurice de Schryver, Yanru Lu, Felix Kopecky, Sebastian Nordhoff

Illustration: Jennifer Gale, Sebastian Nordhoff

Proofreading: Alexandra Fosså, Amy Amoakuh, Craevschi Alexandru, Elliott Pearl, Janina Rado, Jeroen van de Weijer, Maria Zielenbach, Jean Nitzke, Patricia Cabredo-Hofherr, Prisca Jerono, Rebecca Madlener, Sandra Auderset, Teodora Mihoc, Thera Crane, Tom Bossuyt, Alena Witzlack-Makarevich

Fonts: Libertinus, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}T_EX

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

<http://langsci-press.org>

Storage and cataloguing done by FU Berlin

Freie Universität  Berlin

Contents

Acknowledgments	iii
An introduction to Reconstructing Proto-Bantu Grammar Koen Bostoen	v
I Proto-Bantu phonology	
1 Double reflexes in north-western Bantu and their implications for the Proto-Bantu consonant system Gérard Philippson	3
2 Sorting out Proto-Bantu * <i>j</i> Jeffrey Wills	59
II Proto-Bantu verbal morphology	
3 Tense in Proto-Bantu Derek Nurse & John R. Watters	105
4 Reconstructing the development of the Bantu final vowels Jeff Good	173
5 The relevance of Bantoid for the reconstruction of Proto-Bantu verbal extensions Roger M. Blench	235
6 Causative and passive high tone in Bantu: Spurious or proto? Larry M. Hyman	281
7 Reconstructable main clause functions of Proto-Bantu applicative * <i>-id</i> Sara Pacchiarotti	309

Contents

8	Reconstructing suffixal phrasemes in Bantu verbal derivation Koen Bostoen & Rozenn Guérois	343
III Proto-Bantu clausal morphosyntax and information structure		
9	Predicate structure and argument indexing in early Bantu Tom Güldemann	387
10	On reconstructing the Proto-Bantu object marking system Benji Wald	423
11	Agreement on Proto-Bantu relative verb forms Mark Van de Velde	465
12	On subject inversion in Proto-Bantu relative clauses Fatima Hamlaoui	495
13	Predicate partition for predicate-centred focus and Meeussen's Proto-Bantu "advance verb construction" Tom Güldemann & Ines Fiedler	537
14	Proto-Bantu existential locational construction(s) Maud Devos & Rasmus Bernander	581
15	Diachronic typology and the reconstruction of non-selective interrogative pronominals in Proto-Bantu Dmitry Idiatov	667
	Index	738

Acknowledgments

Along with all contributors to the present volume, who bore with us and patiently went through several rounds of revisions, we wish to thank all other colleagues who contributed to the *International Conference on Reconstructing Proto-Bantu Grammar*, which was organised in November 2018 by the Service of Culture & Society of the Royal Museum for Central Africa (RMCA) in Tervuren, and the UGent Centre for Bantu Studies (BantUGent) at Ghent University (UGent). We also acknowledge the funding bodies which financed that conference, i.e. the Research Foundation – Flanders (FWO), the UGent Faculty of Arts and Philosophy, and the RMCA. In the process of editing this book, the editors were financially supported by the Special Research Fund of Ghent University (KB, GMdS, RG), the French National Centre for Scientific Research (RG), the FWO (SP), as well as Consolidator’s Grant n°724275 of the European Research Council (KB).

An introduction to Reconstructing Proto-Bantu Grammar

Koen Bostoen

Ghent University

This book is about reconstructing the grammar of Proto-Bantu, the ancestral language at the origin of the African linguistic family commonly known as Bantu. It is about how to retrieve the phonology, the morphology and the syntax the earliest Bantu speakers used to communicate with each other. In §1, I explain how this book came about. In §2, I offer a short presentation of its contents. In §3, I reflect critically on a number of methodological issues. Finally, in §4, I attempt to assess to what extent the new research presented in this volume requires a revision of Meussen (1967).

1 Raison d'être for *Reconstructing Proto-Bantu Grammar*

Why would Proto-Bantu (PB) matter? Why would one put so much intellectual effort into recomposing a dead language, and especially its grammar, which unlike vocabulary tells us more about its internal functioning than about the outer world? What is the broader relevance of this academic endeavour?

First, Bantu is Africa's principal linguistic family, not only by language count, but also in terms of speakers' numbers and geographical extent (Bostoen & Van de Velde 2019: 3). This is the main reason why Niger-Congo, of which Bantu is a low-level branch, is today the world's biggest phylum as far as number of languages is concerned (Eberhard et al. 2022). Delving into the history of Bantu languages and their speakers is therefore inquiring into significant episodes of Africa's past. The history of Bantu as a distinct language family is assumed to have begun some 5,000 to 4,000 years ago when Bantu speakers started to migrate southwards from their putative homeland in the current-day borderland of southern Cameroon and Nigeria (Vansina 1995: 52; Blench 2006: 126; Bostoen



2018, but see Idiatov & Van de Velde 2021: 98 who propose a more northerly location). The historical origins of Bantu languages and their ancestral speakers are not well known among the wider public, neither inside nor outside Africa, not even among populations currently living in the homeland area (John R. Watters, p.c.). Roughly four to five millennia ago is the approximate time by which PB, their most recent common ancestral language, would have started to diverge into different daughter languages. ‘Proto’ here means that this ancestral language is a reconstruction from present-day Bantu languages, not known from actual historical records.

As writing is a relatively late human invention, i.e. only some 5,000 to 6,000 years old (Pae 2018: 1), written attestations of language from that very period are actually extremely rare worldwide. Cuneiform, a logographic and syllabic script which developed in Mesopotamia out of earlier economy-related sign systems and whose oldest attestations date back to around 3,300 BCE, is commonly seen as the first graphic representation of language (Goody 1986: 47–49). Closer to the Bantu homeland, and on the African continent, is hieroglyphic writing, of which the earliest inscriptions are also dated ca. 3,300 BCE (Kahl 2001: 102), with the first instance of a complete sentence in Old Egyptian from 2,690 BCE (Allen 2013: 2). Thus, the world’s two oldest writing systems, viz. cuneiform and hieroglyphs, hardly predate the assumed advent of Bantu itself. Other early writing systems are considerably younger. For example, Proto-Sinaitic, an intermediary form between Egyptian hieroglyphs and early Semitic alphabets from which later alphabetic scripts (e.g. Greek, Latin, Arabic) evolved, is thought to have been invented over 3,500 years ago (LeBlanc 2017). Similarly, Oracle Bone, the earliest known ancient Chinese script and the ancestor of modern Chinese, is estimated at about 3,300 years of age (Han et al. 2020: 228). In Mesoamerica, embryonic forms of writing only appeared around 700–500 BCE (Kettunen & Helmke 2019: 12).

In other words, Bantu is the rule rather than the exception among the world’s languages in not having written records of its ancestral language, and definitely so for the period around 4 to 5 millennia ago. Apart from the Swahili world where writing in Arabic characters mediated through Islam might be older but without any surviving documentation (Mugane 2015: 175–181), literacy only entered the Bantu-speaking world as part of the so-called Columbian Exchange, i.e. “the exchange of diseases, ideas, food crops, technologies, populations, and cultures between the New World and the Old World after Christopher Columbus’ voyage to the Americas in 1492” (Nunn & Qian 2010: 163). The oldest surviving Bantu text dates from the 17th century CE, i.e. a Kongo translation of a catechism by the Portuguese Jesuit Matheus Cardoso (1584–1625) from 1624 (cf. Cardoso

1624; Bontinck & Ndembe Nsasi 1978). Documentation and description of most Bantu languages – if any – did not start before the late 19th century. In order to retrace the history of the Bantu languages and their speakers, we therefore must go upstream, that is from the recent past back to the source.

Second, even if vocabulary may give more direct access to the history of human culture and society, through the so-called ‘words-and-things method’ (cf. Dimmendaal 2011: 334–336), historical grammar studies also offer insights into how the intricacies of the human mind evolved through time. Bantu Grammatical Reconstructions are particularly relevant in that regard if one reckons how the complexities of Bantu languages at different levels have advanced the development of linguistic theories over the past decades. For example, the intricate tonal systems of Bantu languages such as Ganda JE15 and Tonga M64, along with that of Igbo (Benue-Congo), encouraged Goldsmith (1976) to establish his theory of Autosegmental Phonology, which matured and went in new directions thanks to more theoretically-informed tone studies on a range of different Bantu languages (cf. Clements & Goldsmith 1984; Goldsmith 1987; Hyman & Kisseberth 1998; Kisseberth & Odden 2003; Marlo & Odden 2019). Likewise, tone spreading in the southern Bantu language Shona S10 was one of the case studies in Prince & Smolensky (1993) launching Optimality Theory, which led to many more studies in Bantu phonology (e.g. Downing 1995; Leitch 1996; Myers 1997; Kadenge 2014; Kadenge & Simango 2014) and extended to other domains of Bantu languages such as morphology (e.g. Lusekelo 2012) and syntax (e.g. Harford & Demuth 1999; de Vos & Mitchley 2012). The impact that morphosyntactic data from (mostly Eastern) Bantu languages has had on formal syntactic approaches such as Relational Grammar (e.g. Gary & Keenan 1977; Perlmutter & Postal 1983; Rosen 1984), Lexical Functional Grammar (e.g. Bresnan & Kanerva 1989; Alsina & Mchombo 1990; Bresnan & Moshi 1990; Alsina & Mchombo 1993; Bresnan & Moshi 1993), Government and Binding (e.g. Marantz 1981; 1982; Baker 1985; Baker 1988; 1990; 1992) and subsequent developments such as Minimalism (e.g. Pylkkänen 2000b,a; McGinnis 2001; 2008; Pylkkänen 2008) is immense. To give just one example, it was on the basis of data from Chaga E60 and Chewa N31b applicative constructions that Bresnan & Moshi (1990) developed the by now well-known distinction between symmetrical and asymmetrical object-type languages. Bantu languages such as Swahili G42d, Bemba M42, Rangi F33 and Swati S43 were also instrumental in the creation and expansion of the Dynamic Syntax formalism (e.g. Marten 2003; Gibson 2012; Marten 2013; Gibson & Marten 2016; Chatzikyriakidis & Gibson 2017), first developed in the early 2000s (Kempson et al. 2001; Marten 2002).

The significance of diachronic Bantu studies, and African historical-comparative linguistics more generally, for the birth and growth of linguistic typology

is by now universally acknowledged. Joseph Greenberg, with his work on universals (Greenberg 1966; Greenberg et al. 1978), is generally seen as the founding father of language typology (cf. Hyman 2018: 3). Not only did Greenberg propose a genealogy of African languages (Greenberg 1963), but he also contributed to the reconstruction of Proto-Afro-Asiatic (Greenberg 1958), as well as PB (Greenberg 1948) and its homeland (Greenberg 1972). He also carried out comparative Bantu research (e.g. Greenberg 1951). Ever since, the fields of (historical-)comparative Bantu grammar and language typology have been in an inspiring, mutually feeding relation (e.g. Givón 1971a; 1974; Poulos 1984; 1985; Güldemann 1996; 1999a,b; Odden 1999; Güldemann 2003b; Ngo-Ngijol Banoum 2004; Fleisch 2006; Van de Velde 2006; Maslova 2007; Van de Velde 2009; Devos et al. 2010; Devos & van der Auwera 2013; Aunio 2015; Guérois 2017; Dom et al. 2018; Pacchiarotti 2020). This is also clearly reflected in the current volume on PB grammar, in which several authors propose reconstructions that are strongly informed by typology (cf. *infra*). Given the importance of variation in Bantu grammar for linguistic theory and typology, reconstructing the foundations out of which it developed definitely deserves some scholarly scrutiny.

The importance of Bantu for Africa's past and present (for academic and popular audiences, both inside and outside Africa) and for the field of linguistics are the two main reasons why we thought it timely, half a century after A.E. Meeussen's seminal work *Bantu Grammatical Reconstructions* (1967), to devote a new book to the reconstruction of PB grammar. The present multi-authored volume is the result of this joint effort. Given the way in which Bantu linguistics developed over the past 50 years and the variety of approaches and theoretical frameworks it entails, our book could not be a systematic update of Meeussen (1967). An update of PB grammar cannot simply be resumed where it was left more than five decades ago; for one thing because Meeussen (1967) provides neither factual data nor explicit argumentation for his grammatical reconstructions. Moreover, no unanimity exists on the assumptions, principles and methods underlying Bantu Grammatical Reconstructions, a situation begging for critical reflection. In addition, the huge mass of newly available data has different implications for different aspects of PB grammar. For these reasons, our book is about reconstructing different ancestral grammatical features of Bantu languages rather than an actual comprehensive reconstruction of PB grammar.

2 Historical background to *Reconstructing Proto-Bantu Grammar*

In 2017, *Bantu Grammatical Reconstructions* (1967) by the Belgian linguist A.E. Meeussen celebrated its 50th anniversary. His treatise was the first systematic attempt at a reconstruction of all categories of PB grammar, even though several others before him had succeeded in identifying numerous grammatical cognates between Bantu languages, starting with Bleek (1869) and Meinhof (1899), based on a very small set of languages from different parts of the domain. In order to commemorate the golden jubilee of this important milestone in the history of Bantu linguistics, the *International Conference on Reconstructing Proto-Bantu Grammar* took place in Ghent and Tervuren (Belgium), on November 19–23, 2018. This commemorative event, proposed by Larry M. Hyman (University of California at Berkeley) and Jenneke van der Wal (Leiden University), was co-organised by the RMCA Service of Culture & Society (i.e. the linguists at the Royal Museum for Central Africa in Tervuren), which used to host the research program in comparative Bantu studies known as *Lolemi* (meaning ‘tongue; language’, a reflex of PB **dv-dimi*) led by Achiel Emiel Meeussen (cf. Polak-Bynon 1964; Doneux 1965; Meeussen 1965), and BantUGent (i.e. the UGent Centre for Bantu Studies), founded in 2016 to promote a transdisciplinary approach to the past and present of Bantu languages. This RMCA-UGent collaboration was firmly rooted in a shared history and existing partnerships within the field of Bantu linguistics.

The conference’s organising committee consisted of Gilles-Maurice de Schryver (BantUGent), Maud Devos (RMCA & BantUGent), Sebastian Dom (then BantUGent, now University of Gothenburg), Rozenn Guérois (then BantUGent, now LLACAN, Paris), Hilde Gunnink (BantUGent), Jacky Maniacky (RMCA), Sara Pacchiarotti (BantUGent) and Koen Bostoen (BantUGent). The scientific committee comprised Maud Devos (RMCA & BantUGent), Larry M. Hyman (University of California at Berkeley), Jacky Maniacky (RMCA), Derek Nurse (independent scholar; emeritus), Gérard Philippson (DDL, Lyon; emeritus), Thilo C. Schadeberg (Leiden University; emeritus), Jenneke van der Wal (Leiden University), Mark Van de Velde (LLACAN, Paris) and Koen Bostoen (BantUGent).

Instead of simply being a commemoration, the conference intended to gather today’s junior and senior scholars with the most relevant expertise in comparative Bantu studies in order to reflect together on how to realise a state-of-the-art update of Meeussen (1967). Given the large amount of Bantu language data that have become available since 1967, the vastness of the Bantu language family, and the wide array of grammatical topics to be addressed, such an update can

nowadays no longer be a one-person project. It is inevitably a collaborative effort building on the expertise of numerous scholars, a necessity which Meeussen (1973: 18) himself recognised: “future research in comparative Bantu should consist mainly in team work, in which all available evidence, examined critically, is taken into account”.

The conference attempted to advance first and foremost the reconstruction of grammatical features of PB. Even if contributors were not used to adopting a historical-linguistic approach in their comparative Bantu research, they were asked to do so for their contribution to the conference. They were invited to revisit the comparative evidence on which they had been working for many years with the specific aim of identifying shared retentions with a current-day distribution across the family’s subgroups significant enough to qualify for reconstruction back to PB. In this endeavour, following Meeussen (1967) himself, participants were requested to establish, whenever possible, specific associations of form and function/meaning that are likely to go back to PB.

The conference hosted more than 50 participants from four different continents (Africa, North America, Asia and Europe) representing a fine mix of junior and senior scholars in Bantu linguistics. The academic parts of the final program are reproduced below.

Monday November 19, 2018 (UGent)

Opening

09.15 Opening address by the organising committee (Koen Bostoen)

Chair Koen Bostoen

09.30 Thilo C. Schadeberg (Leiden University)

Reconstructing Proto-Bantu Grammar Half a Century after Meeussen (1967)

10.15 Rebecca Grollemund (University of Missouri)

Reconstructing Proto-Bantu in the Light of the Latest Insights into Bantu Phylogeny

Proto-Bantu Phonology

Chair Rozenn Guérois

11.30 Nancy C. Kula (University of Essex)

Proto-Bantu Segmental Phonology

12.15 Gérard Philippson (DDL, Lyon)

‘Double Reflexes’ Revisited: Implications for the Proto-Bantu Consonant System

An introduction to Reconstructing Proto-Bantu Grammar

- 14.00 Lotta Aunio (University of Helsinki) & Jacky Maniacky (RMCA, Tervuren)
Proto-Bantu Nominal Tone
- 14.45 Michael R. Marlo (University of Missouri)
Proto-Bantu Verbal Tone
- 15.30 Larry M. Hyman (University of California at Berkeley)
Causative and Passive H Tone: Spurious or Proto?
- 16.45 Round table discussion

Tuesday November 20, 2018 (UGent)

Proto-Bantu Verbal Form

Chair Gilles-Maurice de Schryver

- 09.30 Jeff Good (University at Buffalo) & Tom Güldemann (Humboldt University of Berlin)
Proto-Bantu Verbal Form

Proto-Bantu Verbal Derivation

Chair Gilles-Maurice de Schryver

- 10.15 Roger M. Blench (Kay Williamson Educational Foundation)
Proto-Bantu Verbal Extensions from a Bantoid Perspective
- 11.30 Sara Pacchiarotti (Ghent University)
On the Reconstructable Main Clause Functions of Proto-Bantu
Applicative Suffix *-ɪd
- 12.15 Rozenn Guérois (Ghent University)
Proto-Bantu Passive Constructions
- 14.00 Sebastian Dom (Ghent University) & Leonid Kulikov (Ghent University)
Proto-Bantu Middle Voice: From Meeussen to Schadeberg and Beyond
- 14.45 Koen Bostoen (Ghent University)
Non-Compositional Complex Verbal Derivation Suffixes and the
Semantic Reconstruction of *-an in Proto-Bantu
- 16.00 Round table discussion

Wednesday November 21, 2018 (Royal Museum for Central Africa in Tervuren)

Welcome

- 10.15 Welcome address at the RMCA (Jacky Maniacky)
- Chair Sebastian Dom
- 10.30 Maud Devos (RMCA, Tervuren)

Recent Research on the Biography of Achiel Emiel Meeussen in
Relation to Bantu Grammatical Reconstructions 1967

Proto-Bantu Tense, Aspect and Polarity

Chair Sebastian Dom

- 11.30 Derek Nurse (Independent Scholar)
Proto-Bantu Tense and Aspect
- 12.15 John R. Watters (SIL International)
Proto-Bantu Tense from a Benue-Congo Perspective
- 14.00 Thera M. Crane (University of Helsinki) & Bastian Persohn
(University of Hamburg)
Proto-Bantu Lexical Aspect
- 15.00 Round table discussion
- 16.15 Guided pre-view and visit of the renovated Royal Museum for
Central Africa

Thursday November 22, 2018 (UGent)

Proto-Bantu Verbal Morphosyntax

Chair Sara Pacchiarotti

- 09.00 Mark Van de Velde (LLACAN, Paris)
Proto-Bantu Relative Clauses
- 09.45 Hannah C. Gibson (University of Essex)
Proto-Bantu Auxiliary Constructions
- 11.00 Rasmus Bernander (University of Helsinki) & Maud Devos (RMCA,
Tervuren)
Proto-Bantu Existentials

Proto-Bantu Clausal Syntax and Information Structure

Chair Hilde Gunnink

- 11.45 Benji Wald (University of California at Los Angeles)
Some Problems in the Information Structure of Proto-Bantu (& its
Descendants)
- 13.30 Fatima Hamlaoui (University of Toronto)
Proto-Bantu Word Order
- 14.15 Yukiko Morimoto (Humboldt University of Berlin) & Nobuko
Yoneda (Osaka University)
Proto-Bantu Subject and Topic
- 15.00 Jenneke van der Wal (Leiden University)
Proto-Bantu Focus Constructions

16.15 Round table discussion

Friday November 23, 2018

Proto-Bantu Clausal Syntax and Information Structure (Continued)

Chair Jacky Maniacky

09.00 Laura J. Downing (Gothenburg University)
Prosodic Phrasing in Proto-Bantu

09.45 Tom Güldemann (Humboldt University of Berlin)
Meeussen's (1967) 'Advance Verb Construction' – What to
Reconstruct?

10.30 Dmitry Idiatov (LLACAN, Paris)
Proto-Bantu Question Words

Proto-Bantu Nominal Morphosyntax

Chair Maud Devos

11.45 Josephat M. Rugemalira (University of Dar es Salaam)
Proto-Bantu Noun Phrase Structure

13.30 Jean Paul Ngoboka (University of Rwanda)
Proto-Bantu Locatives

14.15 Jean-Georges Kamba Muzenga (Lubumbashi University)
Proto-Bantu Substitutives and Possessives

Closure

15.30 Gilles-Maurice de Schryver (Ghent University)
Bibliometrics in Bantu Lexical and Grammatical Reconstructions:
A.E. Meeussen and Beyond

16.15 Round table discussion + round-up

17.30 Closing words (Koen Bostoen)

Recordings of the talks and the round table discussions are available at the BantUGent website: <https://www.bantugent.ugent.be/events/orpbgconference/>.

After the conference, all presenters were invited to submit texts on the topics they developed for the conference. We received seventeen manuscripts (including one from a participant in the audience). Following double-blind peer review, fifteen chapters were eventually accepted. These were then assigned to one of the thematic sections in the current book (see Table of Contents) even though, unsurprisingly, most chapters treat issues that could belong to more than one thematic section.

3 Methodological issues in *Reconstructing Proto-Bantu Grammar*

In this section I go through some matters of method regarding the reconstruction of PB grammar and discuss how they are variably dealt with in the different contributions to this volume. I treat the Comparative Method (§3.1), genealogical classification (§3.2), and grammaticalisation theory and typology (§3.3).

3.1 The Comparative Method and Bantu grammatical reconstruction

Recovering the estimated 5,000-year-old ancestor of the Bantu language family needs to be based on more or less synchronic data that are mostly younger than 150 years, whether it concerns phonology, the lexicon or grammar. To do so, historical linguists rely first and foremost on the Comparative Method (CM). The reconstruction of proto-languages is one of the primary objectives of the CM. Without historical language sources, the CM is a necessary, effective and bottom-up approach for recreating past languages from cognate morphemes attested in its present-day descendants (cf. Nurse 1997: 361; Weiss 2014: 127). Reconstruction through the CM attempts to “reduce synchronic variation to earlier invariance and in doing so, to recover prehistoric linguistic changes” (Hock 1991: 581). As discussed in Bostoen (2019: 208–209), the CM has been particularly successful for the reconstruction of PB for at least three reasons: (1) the CM is a method for confirming or rejecting genetic affinity rather than for generating hypotheses about it, and such a hypothesis has existed for Bantu ever since Bleek (1862); (2) thanks to their close genealogical affinity, identifying cognate lexemes and grammemes between Bantu languages is relatively straightforward; (3) the efficacy of the CM depends on the quantity of synchronic data available, which is quite favourable in the case of Bantu, especially from a broader African perspective. As a consequence, since its first application by Meinhof (1899) to pave the way for his *Ur-Bantu*, the CM has greatly contributed to the reconstruction of PB phonology, the PB lexicon and PB grammar.

Bantu fulfils the three minimal conditions which Baldi (1990: 1–3) deems necessary for the CM to be used as fruitfully as in Finno-Ugric and Indo-European, where its main empirical foundations were laid during the 19th century: (1) a significant percentage of cognates in core vocabulary to establish genealogical relatedness; (2) the recurrence of systematic correspondences between related languages; (3) regular sound change. As soon as two languages comply with these conditions, the CM can be put to work to reconstruct their ancestral language, but many more languages can of course be added to the reconstruction

equation. The emphasis on regularity and systematicity betrays the legacy of the 19th century Neogrammarians, for whom sound change had no inexplicable exceptions. It also indicates a predilection of the CM for diachronic phonology, where change tends to be more regular and systematic than in other domains of language. Even though its full regularity and systematicity are doubtful in many cases (for a discussion of irregularity in diachronic sound change in Bantu, see Janssens 1993, Pacchiarotti & Bostoen 2022, Philippson 2022 [this volume]), phonological change still has what Baldi (1990: 5) calls a ‘ripple effect’ on other domains of language. It transforms morphs and can therefore eventually lead to the restructuring of grammatical categories and processes. So once the CM succeeds in undoing the sound shifts undergone by the languages of a given family, one can not only reconstruct the phonology of the proto-language, but one can also retrieve the proto-forms of those cognate morphemes, both lexical and grammatical, which were originally used to establish regular sound correspondences by ‘triangulating backwards’ from each of the comparative series (cf. Nurse 2008: 228). For example, the cognate series listed in Table 1, along with other ones, not only led Meinhof et al. (1932) to reconstruct the voiceless bilabial stop **p* to PB and establish the regular sound correspondences between its reflexes in five distant Bantu languages, i.e. Duala A24, Swahili G42d, Kongo H16, Herero R30, and Northern Sotho S32. It also allowed them to reconstruct the form and meaning of three verb stems and two grammemes, i.e. the locative prefix of class 16 and an interrogative particle (see also Idiatov 2022 [this volume]).

The reliance of the CM on cognate series of lexical and grammatical morphemes to establish regular sound correspondences explains why from the early days of historical-comparative Bantu linguistics phonological, lexical and grammatical reconstruction happened concurrently. In his pioneering study of Bantu phonology, *Grundriß einer Lautlehre der Bantusprachen* (‘Outline of a phonology of the Bantu languages’), Meinhof (1899) not only reconstructed a Proto-Bantu sound system, but also identified numerous lexical and grammatical cognate series for which he proposed corresponding reconstructions. His *Grundriß* was soon followed by his *Grundzüge einer vergleichenden Grammatik der Bantusprachen* (‘Basics of a comparative grammar of the Bantu languages’) (Meinhof 1906), the forerunner of *Bantu Grammatical Reconstructions* by Meeussen (1967). The enterprise of reconstructing Proto-Bantu grammar has thus always been firmly rooted in the CM. This explains Meeussen’s insistence on correspondences as a key notion for the reconstruction method.

In a short methodological assessment of Malcolm Guthrie’s *Comparative Bantu* (Guthrie 1967; 1970; 1971), Meeussen (1973) mentions the concept of correspondence no less than 59 times. Exactly for that reason, he is very critical of Guthrie’s

Table 1: Cognate series identified by Meinhof et al. (1932)

Duala ^a A24	Swahili G42d	Kongo ^b H16	Herero R31	N. Sotho S32	PB ^c
<i>w-an-a</i> 'bring'	<i>pa</i> 'give'	<i>va</i> 'give'	<i>pa</i> 'give'	<i>fa</i> 'give'	* <i>pá</i> 'give'
/	<i>pa-a</i> 'scrape'	<i>val-a</i> 'scrape'	<i>par-a</i> 'scrape'	<i>fal-a</i> 'scrape'	* <i>pád</i> 'scrape'
<i>wos-o</i> 'twist tog.'	<i>pot-e-a^d</i> 'go astray'	<i>vot-a</i> 'twist'	<i>pot-a</i> 'be mixed'	<i>for-a</i> 'plait'	* <i>pót</i> 'twist'
<i>wa^e</i> 'here'	<i>pa-</i> loc. class 16	<i>va-</i> loc. class 16	<i>pa-</i> loc. class 16	<i>fa-</i> loc. class 16	* <i>pa-</i> loc. class 16
<i>we</i> 'where?'	<i>-pi</i> 'which?'	<i>-vi</i> 'inter. suffix'	<i>pi</i> 'where?'	<i>-fe</i> 'which?'	* <i>-pi^f</i> 'inter. suffix'

^aFor reasons of uniformity, the examples of Meinhof et al. (1932) are not rendered in the original spelling here, but in IPA spelling.

^bKongo was not part of the original sample of Bantu languages which Meinhof used for his Grundriß (Meinhof 1899). Central Kongo data was added to the revised English version (Meinhof et al. 1932). Neither Makonde P23, which was already used in the first edition, nor Zulu S42, which was added to the revised edition, are included in Table 1 for reasons of space constraints.

^cFor reasons of uniformity, the original toneless Ur-Bantu reconstructions are not given here, but rather the PB reconstructions as found in Bastin et al. (2002), except for the last two which do not occur in the latter.

^dMeinhof et al. (1932: 220) list this lexicalised applicative verb stem along with the deverbative noun *upote/phote* 'bowstring'. As Benji Wald (p.c.) pointed out, Swahili also has the underived base verb stem *pota* 'twist (strings by rolling them between the fingers or on the knee)' (Sacleux 1939: 759).

^eMeinhof et al. (1932) do not provide this reflex; it is found in the Duala dictionary by Helmlinger (1972: 505).

^fIdiatov (2009) reconstructs this interrogative particle as **pà-í* 'where?' [CL16- 'what?'].

so-called “two-stage method of Comparative Bantu study” (Guthrie 1962), which consists of: (1) the construction of ‘Common Bantu’ (CB) by establishing comparative series of synchronic correspondences (comparable to what tend to be called ‘cognate sets’ in historical linguistics), which in Guthrie’s view should be absolutely free from irregularities or exceptions and are symbolised by ‘starred forms’; (2) the true reconstruction of Proto-Bantu as a hypothesis on Bantu prehistory. Meeussen (1973: 16) considers this “explicit distinction between two successive stages in comparative work” as dispensable and at odds with the basic principles of the CM. In his view, the CM provides sufficient inherent guarantees for circularity not to creep in.¹

For example, regarding the final vowel correspondences in several parallel starred forms in CB, such as **na*, **ne*, **ni*, and **nayi* ‘four’ and **da*, **de*, and **dai* ‘long’, Meeussen (1973: 6) judges:

There is a group of synonymous sets of forms in *CB* which differ only in the final vowel. [...] Each of these forms is given as a separate correspondence [...] the attempts at unifying these divergences in prehistory are different from case to case. [...] In a two-stage comparative method it is extremely difficult to obtain more than the observations and conclusions just reported. In an adequately developed one-stage method one is led to try and make full use of all kinds of data in order to reduce as much as possible the variations found between similar correspondences. In the present case it proves possible to view not only each of the clusters [...] as a simplex lexical correspondence, but also the set of these clusters, apart from their consonants, as one complex phonetic correspondence.

Meeussen strongly stresses here that in order to reconstruct the original ancestral language the CM should strive to reduce as much as possible synchronic variation by maximally establishing correspondences – even complex and indirect ones – between present-day languages (see also Hock 1991: 581, cited above). This same methodological emphasis on cross-linguistic form-meaning correspondences made us require contributors to this volume to be as explicit as possible on the specific associations of form(s) and function(s)/meaning(s) they propose to reconstruct to Proto-Bantu and to present sufficient and convincing evidence from present-day languages to substantiate these reconstructions. We furthermore asked them to be explicit with regard to their arguments when considering a given form-function pairing as either a shared retention (i.e. reconstructable to Proto-Bantu) or a shared innovation (i.e. not reconstructable to Proto-Bantu).

¹Circularity is what Guthrie (1962: 1) terms ‘feed-back’, i.e. the introduction of some of the results of an investigation into the conduct of the investigation itself.

3.2 Genealogical classification and Bantu grammatical reconstruction

Our request to authors in this volume to position their (PB) reconstructions in the phylogenetic tree of the Bantu family by Grollemund et al. (2015) and to consider Bantu-external evidence is also prompted by methodological recommendations spelled out in Meeussen (1973). Criticising Guthrie's distinction between PB-X (i.e. the earliest PB stage), PB-A (i.e. the Western 'dialect' of PB) and PB-B (i.e. the subsequent Eastern 'dialect') (see also Dalby 1975; 1976), Meeussen (1973: 17–18) observes that:

all considerations about PB-A and PB-B must remain extremely vague and general, whereas PB-X is purely speculative since it refers to an utterly unattainable stage. Pending the construction of an acceptable genealogical tree for Bantu, we can have reconstructions for one period of Bantu only (the "threshold"). [...] But there is an extremely powerful means of ascertaining the value of a reconstruction by showing that it is required by other, more distantly related languages, in the first place Benue-Congo languages in the case of Bantu.

Not only did pioneers in Bantu reconstruction miss a compass in terms of internal classification, but they also had to work without a widely accepted hypothesis on the Bantu homeland. It is therefore not surprising that Meeussen (1967; 1969) gave less prominence to data from north-western Bantu than we do today with the insights into Bantu classification accumulated over the past five decades.

Although there is still no comprehensive Bantu genealogy based on the CM (Nurse & Philippson 2003; Schadeberg 2003; Philippson & Grollemund 2019), consecutive quantitative approaches using basic vocabulary – mainly lexicostatistical and phylogenetic – considerably enhanced our understanding of the external and internal classification of Bantu since Meeussen (1973). We asked authors to refer to the lexicon-based phylogeny in Grollemund et al. (2015), not because we consider it to be the definitive statement on the internal divergence of the Bantu family, but rather because it is the latest and most comprehensive phylogenetic classification which basically confirms – some deviations notwithstanding – the main results of earlier quantitative approaches such as Bastin et al. (1999), the last and most complete lexicostatistical study for Bantu (see also Bastin & Piron 1999).² Grollemund et al. (2015) sub-classify the Bantu family into five major clades, i.e. North-Western, Central-Western, West-Western, South-Western

²The recent publication of a new phylogeographic analysis of the Bantu language expansion by Koile et al. (2022) shows that Grollemund et al. (2015) is indeed not a definitive internal classification of the Bantu family. As our book was sent off for production in July 2022, it was

and Eastern, which is a substantial simplification of the actual divergence their tree displays (see Bostoen Forthcoming for a detailed assessment of this tree). What is important to retain here is that Grollemund et al. (2015) confirm earlier studies in showing that the north-western part of the Bantu domain, more specifically Cameroon and northern Gabon, is linguistically the most diverse. Their so-called *North-Western clade* actually lumps five discrete monophyletic groups (Pacchiarotti & Bostoen 2020: 156–157). Moreover, Grollemund et al. (2015) corroborate previous studies in demonstrating that after the initial diversification in the north-west, only four major clades occupy the rest of the Bantu domain. Three of them cover the western half, i.e. (1) *Central-Western* aka *North Zaire* or *Congo*, (2) *West-Western* aka *West-Coastal*, and (3) *South-Western*, while all Bantu languages spoken in eastern and south-eastern Africa belong to a single *Eastern* branch (Vansina 1995; Bastin et al. 1999; Bastin & Piron 1999; Bostoen et al. 2015; de Schryver et al. 2015). What is more, *South-Western* and *Eastern* are as a matter of fact not discrete clades in Grollemund et al. (2015), but form one single superclade (cf. Pacchiarotti & Bostoen 2020: 156–157). In other words, the linguistic diversity in the north-west is extremely high compared to the remainder of the Bantu domain. Consequently, a feature occurring in North-Western and Eastern Bantu, for instance, has more relevance for Proto-Bantu reconstruction than one only attested in West-Western and South-Western Bantu or even in South-Western and Eastern Bantu, except of course if it also occurs elsewhere in Benue-Congo or Niger-Congo outside of Bantu. If one admits that Eastern Bantu is indeed a lower-level offshoot in the Bantu family tree, a feature attested in North-Western Bantu and one or more of the other Western clades but not in Eastern Bantu could also be considered for reconstruction into Proto-Bantu, which we situate at the level of either node 1 (excluding Grassfields Bantu) or node 0 (including Grassfields Bantu) in the tree of Grollemund et al. (2015).

The crucial importance of evidence from both North-Western Bantu and Benue-Congo, or even Niger-Congo, outside of Bantu for the reconstruction of PB is an insight that is broadly shared by scholars who contributed to this volume. While several chapters consider evidence from outside Bantu, both **Blench** and **Nurse & Watters** really place Bantoid or Wide Bantu, as opposed to Narrow Bantu, i.e. Bantu as defined by the referential classification of Guthrie (1948; 1971),

too late to take into account this new research published online on August 1, 2022. In any event, for the purposes of this book, their maximum clade credibility tree has no significant implications since its typology is broadly in line with Grollemund et al. (2015). The most important difference regarding the family's internal divergence is that the Western-Western or West-Coastal branch and part of the Central-Western Bantu languages share a most recent common ancestral node which they do not share in the phylogeny of Grollemund et al. (2015).

in the forefront. Comparative data from Benue-Congo and Kwa languages, and even from Niger-Congo languages far beyond, also play a prominent role, along with data from mainly north-western Bantu languages, in the revision of verbal argument cross-reference in PB by **Güldemann**. Likewise, **Philippon**'s chapter focuses specifically on North-Western Bantu. Several other chapters reanalyse earlier PB reconstructions by giving more historical weight to north-western Bantu data than Meeussen (1967; 1973) ever did. For example, more data from the North-Western Bantu branches play an important role in **Wills**'s revision of PB *j in several Bantu lexical reconstructions. Likewise, **Nurse & Watters** and **Bostoen & Guérois** question the PB status of the anterior final suffix *-ide (Bastin 1983) and the passive suffix *-ibv (Stappers 1967) respectively, because they miss reflexes in present-day North-Western Bantu. They rather consider these suffixes to be innovations that emerged at a later node of the Bantu family tree after the ancestral North-Western Bantu branches had split off. A thorough review of comparative north-western Bantu data also leads **Good** to conclude that the system of final inflectional vowels reconstructed by Meeussen (1967: 110) is to be seen as an innovation that rather happened at node 2 in the tree of Grollemund et al. (2015). For the same reasons, **Hamlaoui** disputes the hypothesis of both Meeussen (1967: 120–121) and Nsuka-Nkutsi (1982) according to which lexical subjects would have followed the verb in PB object relative clauses. Being largely absent from the North-Western Bantu branches, Hamlaoui considers VS order in relative clauses to be an innovation that possibly only arose at node 2 or 3 in the tree of Grollemund et al. (2015). North-western Bantu data are also crucial in **Devos & Bernander**'s reconsideration of non-inverted existential locational constructions as a possible archaism. The reconstruction of such existentials to PB would imply that the main clause type reconstructed to PB by Meeussen (1967: 120) as 'anastasis', better known today as 'subject inversion' (cf. Marten & van der Wal 2014), is also a later innovation. Absence from North-Western Bantu is also for **Güldemann & Fiedler** a conclusive argument to consider 'preverbal preposed verb focus doubling', one of the constructions possibly corresponding to the so-called 'advance verb construction' which Meeussen (1967: 121) reconstructs to PB, as a post-PB innovation, unlike 'in-situ verb focus doubling' and 'initial preposed verb focus doubling' which can be ascribed to PB. **Wald** too reviews ample data from north-western Bantu languages in his chapter on PB object marking. Although he agrees with Polak (1986) in observing that north-western Bantu languages generally do not admit more than one object prefix per verb form, he disagrees with her conclusion that multiple object marking is an innovation posterior to PB. In doing so, **Wald** goes against the possible misconception that if a given feature is not in North-Western Bantu, it cannot have been in PB. It is not

because North-Western Bantu consists of older clades than the rest of Bantu that its features (or lack thereof) must be older and presumably closer to PB. Wald interprets the diversity of object indexing systems in north-western Bantu languages as the outcome of progressively ordered stages of change away from the state of affairs in PB, which is more conservatively preserved in more recently formed clades. As such he rather sides with Meeussen (1967: 112) in reconstructing a PB object marking system that allowed for sequences of object prefixes in one and the same verb form, even though both authors seemingly have different views on the functional motivation for prefix ordering (cf. *infra*).

All in all, the general picture that emerges from our volume is that when checked against increasing insights into Bantu internal classification, several PB grammatical reconstructions proposed by Meeussen (1967) turn out to be not as old as previously thought. Rather than go back to the most recent common ancestor of all (Narrow) Bantu languages, i.e. the “threshold” which Meeussen (1973: 18) had in mind, they seem to go back no further than the one that emerged after the ancestors of several North-Western Bantu branches had split off. Methodologically, it shows the importance of genealogical classification for a judicious appraisal of the relative time depth of reconstructions. In terms of chronology, it calls for a general reassessment of the actual time depth of Proto-Bantu grammar as reconstructed by Meeussen (1967), which goes beyond the scope of this book.

The insight that Proto-Bantu as traditionally conceived is in all likelihood considerably younger than commonly assumed, even within Narrow Bantu, is also highly relevant for future reconstruction work within Bantoid and more widely Benue-Congo or even Niger-Congo. As Watters (2018: 16) points out:

It is tempting, whether conscious or subconscious, to take a Bantu-centric view and begin conceiving Proto-Bantoid as being equivalent to Proto-Bantu, and even perhaps extending the temptation and conceiving Proto-EBC [Proto-East Benue-Congo] as being equivalent to Proto-Bantu. Bantu has received the attention of a multitude of linguists for more than a century and Proto-Bantu has been reconstructed in ways to which no other Bantoid subgroup can compare. [...] It can be easy to [...] forget that Proto-Bantu and its own subgroups and individual languages have their own history of retentions, innovations, and borrowings. So, in reconstructing Bantoid and EBC, caution has to be taken. [...] Care is needed not to attribute everything found in Proto-Bantu to Proto-Bantoid, and in Proto-Bantoid to Proto-EBC.

Such care and caution are even more warranted if one reckons that several typical Bantu features that have commonly been seen as retentions from PB turn

out to be later innovations. Hence, Bantoid or EBC did not necessarily lose what Bantu retained. Bantu also developed morphology and syntax that its ancestors never had.

3.3 Grammaticalisation and typology in Bantu grammatical reconstruction

Meeussen's strong reliance on the CM and his emphasis on regular correspondences explains why his *Bantu Grammatical Reconstructions* focuses on phonology and morphology rather than on syntax, to which he nonetheless dedicates some pages. It also accounts for the fact that his reconstructions are prominently biased towards form to the detriment of meaning and function. The CM does not have a distinct approach to phonological vs. morphological reconstruction (Hoenigswald 1991; Koch 1996; 2014). Morphological and syntactic reconstruction are known to be more challenging than their phonological counterpart (Hock 1991; Koch 1996). Morphological and syntactic changes also happen independently of phonological change, and not necessarily in a systematic way reflected in regular correspondences. Hence, the undoing of such changes with the aim of reconstruction is considerably more difficult, not only because non-phonological changes are much less regular but also because we have much less insight into their natural direction (Hock 1991: 610). Due to analogy, regular sound changes might be blocked or undone in morphemes. This is especially so in inflectional paradigms, where grammatical morphemes are easily affected by reanalysis of their external boundaries and therefore become more readily eroded than lexical morphemes (e.g. Traugott & Heine 1991). Gildea (2000) also sees the absence of regular laws of grammatical change as one of the main reasons why it is so difficult for comparative linguists to identify cognates among grammatical constructions and morphosyntactic patterns to the extent that some would even consider grammar unreconstructable.

Grammaticalisation theory fortunately came to the rescue of morphosyntactic reconstruction by identifying recurrent patterns of grammatical evolution across languages, most prominently “the almost universal directionality from independent, concrete lexical item to bound, abstract grammatical morpheme” (Gildea 2000: vii). This theory allows for establishing possible cognates between lexemes and grammemes and distinguishing between likely sources and later innovations. Initially such patterns were mainly observed in historical language documents (i.e. based on attested change through time) and by means of internal reconstruction (i.e. based on language-internal synchronic variation reflecting successive diachronic developments).

When going by language-internal evidence, synchronically irregular or anomalous forms are crucial for morphological reconstruction, since regular forms can always result from analogical levelling, i.e. the principle of ‘archaic heterogeneity’ (cf. Hetzron 1976). Likewise, it is important to compare archaic patterns surviving in peripheral areas of grammar and/or idiomatic expressions. To do so, comparative evidence from closely or more distantly related languages might be essential to identify archaisms and argue for the plausibility of a specific levelling or reanalysis scenario or for a given pathway of grammaticalisation (cf. Bybee et al. 1994; Heine & Kuteva 2002). That is why, in the absence of historical data, “one must become a typologist to motivate the evolutionary scenario” (Gildea 2000: viii). Thanks to Bernd Heine and his team (cf. Heine & Reh 1984; Heine et al. 1993), African languages greatly contributed to the efflorescence of the typological literature on grammaticalisation.

Unsurprisingly, both grammaticalisation and typology also play an important role in this volume, not only in the chapters of **Güldemann**, one of Heine’s most prolific disciples, but also in many other chapters. For instance, in **Pacchiarotti**’s chapter on the main clause functions of the PB applicative **-rd*, whose formal reconstruction she considers to be established, paths of change from allative to benefactive, which are numerous attested in the grammaticalisation literature, constitute a main argument in favour of reconstructing the suffix with an original Spatial Goal or Location-oriented function. Obviously, grammaticalisation also plays an important role in the reconstruction by **Nurse & Watters** of how tense emerged and evolved in ancestral Bantoid and Bantu. The pre-stem domain in Bantu is known to be particularly productive in attracting lexical verbs for the expression of grammatical categories of tense, aspect and mood/modality, first as free auxiliaries and subsequently as bound prefixes (Güldemann 2003a; Nurse 2008; Nurse & Devos 2019). Alongside grammaticalisation, typology is given a lot of argumentative power in several chapters, especially in the third thematic part on clausal morphosyntax and information structure. Authors tend to deal there with abstract patterns, such as agreement and word order, rather than with specific morphological constructions. **Devos & Bernander** and **Idiatov** are exceptions in that they do target specific form-meaning associations in the domains of existential locationals and non-selective interrogative pronominals respectively. They come up with what **Idiatov** calls “typologically informed reconstructions”. In other words, the CM and typology go hand in hand. **Idiatov** provides a general methodological discussion of the issue of variation in functional elements and the possible ways of dealing with it in reconstruction as well as an overview of the diachronic typology of non-selective interrogative pronominals. He does not reconstruct specific morphosyntactic constructions to any given node in the

Bantu family tree, but rather identifies recurrent formal types of non-selective interrogatives as starting points for further reconstruction. **Devos & Bernander** do come up with specific existential locational constructions to which they attribute variable time depths according to their present-day distribution across major Bantu clades. **Idiatov's** formal types, on the contrary, could easily emerge as convergent innovations due to repeated cycles of the accretion and reduction of the same inherited substance. The attestation of similar interactions between accretion and reduction but with different morphemes in other language families of the world leads **Idiatov** to the conviction that several interrogatives from present-day Bantu languages are nothing but seeming cognates, which seriously hampers proper reconstruction. A bottom-up approach starting out from low-level Bantu branches might shed new light on **Idiatov's** diachronic typology.

Cyclicity in the reanalysis of morpheme sequences also plays a major role in **Van de Velde's** historical interpretation of how agreement evolved in Bantu relative verb forms. He contests the direct and indirect relative clause constructions which Meeussen (1967: 120–121) reconstructed for PB, not so much because these would be unattested in present-day Bantu languages or insufficiently spread across subgroups, but because no logically possible scenario of morphosyntactic change within Bantu relative clause constructions can derive present-day variation in Bantu from these reconstructions. Despite their widespread distribution across the Bantu family and their relative uncommonness in the world's languages, **Van de Velde** refutes, contra Meeussen (1967: 120–121) and Nsuka-Nkutsi (1982), the assumption that relative verbs agreeing with the antecedent are shared retentions inherited from PB. Just like **Idiatov's** formal types of non-selective interrogatives are possibly the outcome of convergent evolutions, **Van de Velde** considers these widespread relative constructions as parallel innovations of the “Bantu Relative Agreement cycle”. However, relative verbs agreeing with their subject which he proposes as the alternative PB starting point is strictly speaking not a reconstruction, but a default situation, both typologically and within Bantu and Bantoid. It could have occurred at any stage in the evolution of Bantu, Benue-Congo and Niger-Congo. In my view, it is impossible to say whether attestations in present-day Bantu languages of what **Van de Velde** identifies as the PB source constructions are shared retentions or the outcomes of convergent evolution. It might prove interesting to test his typologically informed top-down proposal for PB via a bottom-up approach focusing on low-level Bantu subgroups.

Such bottom-up testing could also be applied to **Güldemann's** hypotheses on predicate structure and argument indexing in early Bantu, which result from what he describes himself as “primarily an arguably viable exercise in diachronic (and partly areal) typology”. The so-called ‘Macro-Sudan Belt’ in northern Sub-Saharan Africa, a linguistic macro-area stretching between Senegal and Ethiopia

and including the Bantu homeland (cf. Clements & Rialland 2008; Güldemann 2008; 2018; Idiatov & Van de Velde 2021), plays a key role in his areal-typological considerations. In his contribution to our volume, **Güldemann** further buttresses his earlier claim that the PB verb template was not highly agglutinative, as reconstructed by Meeussen (1967: 108–111) and defended by Hyman (2004; 2011), but rather a split predicate structure with free pronouns or person-inflected portman-teau morphemes simultaneously encoding tense, aspect, modality, and polarity. This is the typological profile which is most prominent today in North-Western Bantu, including the Bantu homeland, and in Niger-Congo outside of Bantu. Strongly relying on grammaticalisation theory and areal typology, Güldemann (2011) argues that the direction of change from Proto-Bantu to most of present-day Bantu beyond the north-west was from analyticity towards agglutination by way of phonological fusion. Relying on what he considers to be relic features in North-Western Bantu and Niger-Congo beyond Bantu, Hyman (2011) advocates the opposite direction of change from agglutination towards analyticity by way of erosion and loss of bound morphology. The two poles of this debate adopt a top-down approach relying on very similar and selective samples of distantly-related Niger-Congo languages to argue for “today’s morphology is yesterday’s syntax” (**Güldemann**), aka “grammaticalisation” or “morphologisation through desyntactisation” (cf. Givón 1971b), vs. “today’s syntax is yesterday’s morphology” (Hyman 2011), aka “degrammaticalisation” (cf. Norde 2009). Unlike in Güldemann (2011), **Güldemann** does go beyond typology and grammaticalisation in his contribution to this volume by performing a comparative study of concrete morphemes, i.e. subject and object indexes involved in verbal cross-referencing. He shows that the prefixes reconstructed by Meeussen (1967) deviate considerably from the (free) pronoun forms, which prevail in North-Western Bantu. The latter would correspond to those which can be assumed for earlier Benue-Kwa and Niger-Congo (cf. Güldemann 2017) and can therefore be considered as archaisms in his view. As a consequence, Meeussen’s reconstructions of bound participant cross-reference are to be seen as later innovations. Their emergence is to be situated after the branching off of North-Western Bantu clades (cf. *supra*) and be seen as intimately linked with the development of a more agglutinative verb template. This hypothesis merits to be tested through a contemporary and cross-linguistically informed bottom-up application of the CM for morphosyntactical reconstruction, as in **Pacchiarotti**’s ongoing post-doctoral research project focusing on a specific Bantu clade, i.e. West-Coastal Bantu aka West-Western Bantu.³

³See <https://research.flw.ugent.be/en/projects/directionality-morphosyntactic-change-west-coastal-bantu-historical-test-case-linguistic>.

4 Reconsidering *Bantu Grammatical Reconstructions*

As discussed above, a systematic revision of the PB grammar reconstructed by Meeussen (1967) is not feasible at this stage and goes beyond the scope of the current volume. Nonetheless, by way of closing the introduction to this book, I run through its chapters and discuss succinctly how each of them revises (or not) Meeussen's *Bantu Grammatical Reconstructions*.

Philippon brings up a long-standing question in Bantu historical linguistics, i.e. the so-called *double reflexes*. It is the phenomenon, particularly common in North-Western Bantu, whereby one and the same proto-consonant has two or more reflexes in a given language which cannot be accounted for by phonological conditioning and/or lexical borrowing. Such unexplainable exceptions to the Neogrammarian principle of regular sound change raise the question whether an additional series of consonants subsequently lost through phonemic merger should be reconstructed in PB, or whether a specific conditioning which caused phonemic split became opaque. To shed new light on this question, **Philippon** systematically reviews comparative evidence from North-Western Bantu, whose internal classification he summarises in his own view. He concludes that double reflexes of voiced PB oral stops can to a large extent be accounted for by a tonal conditioning that was lost, but that the situation regarding voiceless PB consonants is much blurrier. This is definitely the case for a recurrent set of stems whose reconstructed **t* systematically escapes the lenition that is regular in other stems. He relies on the lexical diffusion model of sound change to explain these irregular retentions. All things considered, he concludes that for the time being his survey does not warrant a revision of the PB consonant system proposed by Meeussen (1967).

Wills does contest one specific segment in Meeussen's PB consonantal phoneme inventory, i.e. **j*, for which Guthrie distinguished between **j* and **y*. **Wills** systematically reviews the comparative lexical evidence across Bantu, with special attention to North-Western Bantu. Based on this broad survey, he argues that most stem-initial segments in present-day Bantu languages, such as in /y/, /z/ or /j/, are the outcome of later developments universally common at morpheme boundaries. They should not be seen as regular reflexes of PB **j*, as Meeussen (1967; 1969) and his disciples (cf. Coupez et al. 1998; Bastin et al. 2002) proposed. As a consequence, many *Bantu Lexical Reconstructions* with initial **j* should be reconstructed with a stem-initial vowel instead and both **ny* and **nj* should be

reconstructed as distinct phonemes. However, **Idiatov**, in the appendix to his chapter, argues why several PB roots reconstructed with **j* did have an initial consonant, even if the initial **j* seen in *Bantu Lexical Reconstructions* confounds several PB consonants, including minimally **s*, **z*, **ʃ*, **y*, and **g*.

Following the two chapters on PB phonology, **Nurse & Watters** open the section on PB verbal morphology. Their chapter and the following by **Good** focus on verbal inflection. **Nurse & Watters** consider, predominantly though not exclusively, tense and aspect morphology in the pre-stem domain, while **Good** (2022 [this volume]) deals with verb endings involved in the expression of tense, aspect, mood, and polarity. As discussed above, **Nurse & Watters** review extensive new data from Bantoid, which **Watters** accumulated and in the light of which **Nurse's** earlier historical-comparative research on tense and aspect in Bantu is reassessed (cf. **Nurse** 2003; **Nurse & Philippson** 2006; **Nurse** 2008). Their main new idea is that tense as a grammatically encoded category emerged in Benue-Congo (or more narrowly in Bantoid) not long before the rise of PB itself. It was innovated in the most recent common ancestor of Narrow Bantu and those Bantoid languages spoken along and to the east of the Cameroon Volcanic Line. Early Benue-Congo (or more strictly Bantoid) ancestral languages must have been aspect-prominent, i.e. without grammatically contrastive tense categories, as is still the case for many Niger-Congo languages today. In other words, **Nurse & Watters** confirm **Meeussen's** reconstruction of both tense and aspect morphology to PB, but posit that tense-related grammemes were a relatively recent development at that stage. When it comes to specific tense/aspect constructions, i.e. verbal conjugations involving prefixes and/or suffixes, the revisions of the PB *tense formulae* proposed by **Meeussen** (1967: 112–113) are basically the same as those already proposed in **Nurse** (2008), as nicely summarised in **Nurse & Watters's** Table 10 in their conclusions, except for two suffixes involved in several of those *tense/aspect forms*. As discussed above, **Nurse & Watters** consider verb-final **-ide* as a later innovation and reconstruct instead **-i* as the verb ending involved in two PB conjugations, i.e. present and past retrospective (perfect). Similarly, they propose **-ag* instead of **-ang* (*-nga-* in **Meeussen** 1967), as the pre-final suffix in two PB conjugations, i.e. present and past imperfective. Direct reflexes of **-ag* are also attested in Bantoid, while direct reflexes of **-ang* do not occur outside of Narrow Bantu (see also **Sebasoni** 1967).

Without stating it explicitly, **Good** actually contests **Nurse & Watters's** reconstruction of the verb ending **-i* to PB, because he considers the entire PB reper-

toire of inflectional verb endings proposed by Meeussen (1967: 110) as an innovation that only emerged after the first North-Western Bantu branches had split off. His extensive review of final vowel patterns in fifteen North-Western Bantu languages of Guthrie's zones A and B leads to the observation that the northernmost languages of the survey area, all belonging to the first North-Western Bantu branches, i.e. those splitting off before ancestral node 2 in the tree of Grollemund et al. (2015), generally miss the reconstructed inventory of final vowels. Relics only surface in the southern part of the survey region, i.e. in languages belonging to later North-Western Bantu branches as well as West-Western Bantu. Good (2022 [this volume]) prudently interprets this situation as suggesting that Meeussen's relevant reconstructions may be better associated with a later stage corresponding roughly to node 2 in the tree of Grollemund et al. (2015). He also reconstructs a plausible historical path for the development of the canonical Bantu final vowel system that involves the gradual integration of postverbal elements coding tense/aspect/mood/polarity (TAMP) categories into the verb form, their subsequent reduction and reanalysis to vocalic suffixes, and the analogical extension of these to all verb forms. He admits, nonetheless, that its time depth remains unclear. The existence of inflectional final vowels in several Bantoid languages surveyed in the chapter of **Nurse & Watters** might suggest that, contra **Good**, their emergence actually did pre-date PB, or that they are parallel innovations. If they would be older than PB, their absence in the North-Western Bantu languages in **Good**'s sample would have to be the outcome of loss instead of reflecting the original system, as **Wald** argues, for example, with regard to multiple object marking (cf. supra).

Blench is the first of four chapters dealing with verbal derivation morphology. Through a survey in a set of languages belonging to different Bantoid branches, he assesses the relevance of their repertoires of verbal extensions (i.e. derivational suffixes) for the reconstruction of PB verbal extensions. Rather than being a true historical-linguistic exercise in reconstruction, his chapter is a comparative overview of relevant morphology in the most well-known Bantoid subgroups in close proximity of the putative Bantu homeland, i.e. Dakoid, Mambiloid, Tivoid, Beoid, Grassfields, and Mbe-Ekoid. It does not directly lead to revisions of the PB derivational verb suffixes reconstructed by Meeussen (1967: 92). **Blench** (2022 [this volume]) observes that apart from the long causative suffix **-ic*, clear traces of the reconstructed PB system can only be found in Grassfields and may also be reconstructed to their most recent common ancestor. However, formal resemblances between extensions attested in other Bantoid languages and extensions in some languages of Guthrie's zone A, which do not appear to be cognate with

any of the established PB reconstructions, lead **Blench** to the conclusion that the PB inventory of verbal derivation suffixes might need to be enlarged with suffixes that were never reconstructed before. This hypothesis needs to be tested via a thorough application of the CM, especially to exclude that superficial resemblances between certain extensions in zone A Narrow Bantu languages and those in nearby Bantoid are not false cognates or later contact-induced innovations.

Hyman revises a specific feature of the PB verbal derivation system, i.e. the high tone which Meeussen (1967: 92) tentatively sets up for the causative **-i* and passive **-v* suffixes. The possible high tone of these two suffixes is historically relevant, because along with their exceptional vowel shape it is one of the two formal features that makes them stand out compared to all other verb derivational suffixes reconstructed with a low tone and a VC form. Moreover, both suffixes tend to be stacked after all other derivational suffixes, i.e. just before the final vowel (Hyman 2003; Good 2005). These three odd features have been interpreted as indications that they could be old Niger-Congo voice suffixes, which were integrated later on in the verbal derivational system (see Hyman 2007: 161). **Hyman** demonstrates, however, that the high tone on short causative and passive suffixes is attested almost exclusively in some Eastern Bantu languages of the Great Lakes region, where Meeussen was very active as a descriptive linguist. **Hyman** also elaborates different morphological and phonological scenarios in which the high tone on these suffixes could have developed. He concludes that causative and passive high tone does not go back to PB confirming Meeussen's own hesitations on its reconstructability.

With her diachronic approach to the semantics and syntax of PB applicative **-id*, **Pacchiarotti** fills a void in Meeussen (1967), not only with regard to this specific suffix, but also more generally with regard to the semantic and syntactic reconstruction of PB grammemes. As discussed above, Meeussen's efforts focused on the reconstruction of form to the detriment of meaning and function. Relying on her earlier comparative research gathering data from all major Bantu branches (cf. Pacchiarotti 2020), **Pacchiarotti** reconstructs the main clause functions of **-id*. This is quite a challenge given the semantic underspecification and the high degree of polyfunctionality of the applicative suffix in present-day Bantu languages. The suffix further stands out with respect to other Bantu verbal derivational suffixes in that it performs dedicated discourse functions. She argues that the traditional view of PB **-id* as a purely valence-increasing syntactic device should be abandoned. She identifies three interrelated functional retentions that

are sufficiently shared among current-day reflexes of **-id* to be reconstructed to PB: (1) syntactically, introducing a non-Actor semantic role which can otherwise not be conveyed in the main clause; originally, this was likely a Spatial Goal or a Location-related role; (2) semantically, adding notions such as completeness, iterativity or thoroughness to the verb root's meaning; and (3) pragmatically, signalling narrow focus on a Location-related noun phrase.

Bostoen & Guérois introduce the concept of 'suffixal phrasemes' in the field of Bantu verbal derivation and assess whether any non-compositional suffix sequences can be reconstructed to PB. They argue that the coinage of such suffixal phrasemes is first and foremost a morphological strategy on which Bantu languages have repeatedly relied to innovate verbal derivation morphology, though using suffixes inherited from PB. Across Bantu, semantically non-compositional aggregations of suffixes are common in verb derivational categories as diverse as the pluractional, neuter, intensive, reciprocal, passive and causative. The rise of suffixal phrasemes started within the paradigm of causative morphology. **Bostoen & Guérois** show that PB did not only inherit from older Benue-Congo ancestors causative **-i* and **-ic*, as reconstructed by Meeussen (1967: 92), but also innovated **-idi*, a non-compositional reanalysis of PB applicative **-id* and causative **-i*. After North-Western Bantu split off, **-iki* (itself probably resulting from the phraseologisation of neuter **-ik* and causative **-i*) was added to the causative repertoire. As for the passive, they agree with Meeussen (1967: 92) in only reconstructing **-v* and not the suffixal phraseme **-ibv* as proposed by Stappers (1967), which only emerged when the main North-Western subgroups had branched off. They argue that the middle suffix **-Vb*, the first component of **-ibv*, does in all likelihood go back to the most recent common ancestor of all Bantu languages and should be added to the inventory of extensions reconstructed by Meeussen (1967: 92).

Güldemann argues that the morphologically compact predicate with bound argument cross-reference on the agglutinative verb form reconstructed by Meeussen (1967: 108–111) for PB, is a later innovation. According to his historical-linguistic analysis, PB rather had a split predicate structure with free pronouns or person-inflected portmanteau morphemes also encoding tense, aspect, modality, and polarity, as is still the case in many present-day North-Western Bantu languages and in Niger-Congo languages beyond Bantu. In support of this line of argumentation, he reviews comparative evidence for the morphosyntax of verbal argument cross-reference and the basic segmental shape of its exponents across

Bantu, especially the form of speech-act participant cross-reference morphemes. From the bound 1SG/PL and 2SG/PL subject and object prefixes (eight in total) proposed by Meeussen (1967: 97), only the bound 1SG prefix **n-* (for both subject and object syntactic functions, possibly with a front vowel following the nasal) can be maintained (see his Table 10). **Güldemann** considers it as a potential retention from earlier Benue-Kwa that co-existed with a 1SG free pronoun and therefore had functional restrictions to specific contexts. All other prefixes reconstructed to PB by Meeussen (1967: 97), i.e. **ʊ-* (2SG subject), **kʊ-* (2SG object), **tʊ-* (1PL subject/object), and **mʊ-* (2PL subject/object), only emerged at later stages according to **Güldemann**. In his PB reconstruction, predicate arguments were chiefly marked through independent pronouns inherited from ancestral Benue-Kwa, i.e. **mi* (1SG), **(B)U* (2SG), **tU* (1PL) and **nU* (2PL). **Güldemann** prefers to remain agnostic on the specific consonant and/or vowel qualities of the last three pronouns and indicates this with capital letters.

In the same vein as **Pacchiarotti** does for PB applicative **-id*, **Wald** focuses on the function rather than the form of the PB object marking system. As discussed above, he agrees with Meeussen (1967: 110) in reconstructing a PB verb form that allowed for the prefixation of more than one object index. In doing so, he does not only disagree with Polak (1986), who considers multiple object marking (MOM) as a later innovation of the PB single object marking (SOM) system, but probably also with **Güldemann** (2022 [this volume]) above who reconstructs **SBJ OBJ STEM*, **[SBJ=TAMP] OBJ STEM*, and **[SBJ=TAMP] [OBJ=STEM]* as the three major PB morphosyntactic patterns of predicates involving object marking. Although **Güldemann** is not really explicit on the number of bound object markers, he seems to reconstruct both no object marking (NOM) and SOM to PB. **Wald** suggests that “a major problem of **Güldemann**’s dependence on typology is the timing of the $V-O_{\text{PRO}} > O_{\text{PRO}}-V$ change relative to PB”, i.e. when free postverbal object pronouns shifted into pronominal object prefixes. For **Wald**, situating this change after PB is problematic because there is a relic area of full object marking systems among the North-Western Bantu languages that first split off according to Grollemund et al. (2015). He resolves this question by projecting **Güldemann**’s reconstruction back to a stage earlier than PB, which itself would then already have had a MOM system. In so doing, **Wald** further notes that retrofitting **Güldemann**’s proposal to pre-Bantu is compatible with a MOM system at the PB stage, because it allows for multiple object pronouns in a single predicate simultaneously morphologising into object prefixes. While Meeussen remains silent on how the ordering of object prefixes was semantically conditioned in the PB MOM system, **Wald** does come up with a functional motivation. Based on his extensive

comparative review of pragmatic and syntactic factors determining variation in object marking systems across Bantu, he reconstructs for PB a MOM system with contextual topicality as the decisive principle for the selection and ordering of object prefixes. The leftmost prefix, i.e. the subject prefix before any object prefix, marks the referent with the highest topicality, i.e. the one which is the oldest, most given or deducible, according to the discourse context. Thereafter each object prefix continues in next leftmost order according to the higher contextual topicality of its referent relative to the referent of any object prefix to its right. This proposal differs from that of Meeussen (1967: 110), who proposes, without any further argumentation, a PB object prefix ordering which corresponds to the mirror-image of the order of postverbal object noun phrases.

Van de Velde challenges the PB reconstruction by Meeussen (1967: 113–114) of both direct and indirect relative clause constructions that agree with the head noun by means of an agreement morpheme belonging to the paradigm of so-called *pronominal prefixes (PPs)*. Although relative verb forms agreeing with the relativised noun phrase are common in present-day Bantu languages, **Van de Velde** does not consider them to be shared retentions. Rather, he posits them as the outcome of convergent evolution through the so-called *Bantu Relative Agreement (BRA) cycle*, whereby erstwhile independent relativisers occurring between the relativised noun phrase and the relative clause gradually get integrated into the relative verb form. In this way, unbound morphemes of diverse origins, such as demonstratives, personal pronouns, and connective relators, turned into bound relative agreement prefixes by means of parallel, independent innovations. In indirect relative constructions, the agreement prefixes may precede the subject prefix agreeing with the subject of the relative clause and occupy the verb form's so-called *pre-initial* slot (cf. Meeussen 1967: 108). According to **Van de Velde**, they should not be reconstructed to PB either. Although the BRA cycle in itself does not exclude the existence of bound relative agreement on the verb in PB and some of the PP in present-day relative verb forms could be shared retentions, **Van de Velde** rejects this possibility, because “[t]he only logically possible starting point from which the currently attested typological variation in Bantu relative clause constructions could have evolved is one in which relative verbs agreed with their subject”.

Hamlaoui also focuses on PB relative clauses, specifically the position of subject noun phrases in indirect relative clauses. She tests the hypothesis that a *free subject* (i.e. lexically overt subject), if any, follows the verb in PB indirect relative clauses, as claimed by Meeussen (1967: 220) and confirmed by Nsuka-Nkutsi

(1982). To do so, she enlarges Nsuka-Nkutsi's original sample to 167 languages, viz. 151 Narrow Bantu and 16 other Niger-Congo languages, and observes that VS is still the most frequent word order. Nonetheless, SV-only word order prevails in Bantu zone A as well as Niger-Congo beyond Narrow Bantu. What is more, SV-only is attested in a significant portion of Eastern Bantu. The hypothesis that SV-only would be an innovation linked with the assumed shift from more synthetic to more analytic, as argued by Nurse (2007) and Hyman (2017), and the concomitant loss of argument cross-reference on the verb does not hold for the highly agglutinative Eastern Bantu languages with SV order. Given its present-day distribution within and outside Narrow Bantu, SV-only could be posited as a shared retention from PB. If so, like several other reconstructions in Meeussen (1967), VS order in indirect relative clauses would be a later innovation that only emerged at the level of nodes 2 or 3 in the tree of Grollemund et al. (2015).

Güldemann & Fiedler closely examine the so-called *advance verb construction* which Meeussen (1967: 121) reconstructs to PB as “[a] peculiar kind of sentence, with twice the same verb, the first occurrence being an infinitive”, but without much functional elaboration, i.e. “[t]he meaning varies between stress of « reality », stress of « degree », and even « concession »”. **Güldemann & Fiedler** present a detailed comparative review of the structure and function of this and related constructions and come up with a diachronic interpretation of the synchronic variation they manifest across Bantu. In the end, they ascribe two verb doubling constructions to PB, i.e. one whose non-finite verb occurs in-situ and one where it is preposed to clause-initial position before the subject/agent noun phrase. Both constructions had the function of signalling focus on the state-of-affairs expressed by the verb. Structurally speaking, **Güldemann & Fiedler** consider verb doubling constructions whose non-finite verb occurs immediately before the finite verb, which are recurrent outside of North-Western Bantu, as later innovations. Functionally speaking, they interpret the expansion from state-of-affairs focus to general predicate-centred focus (i.e. including polarity, truth value and TAM), and further to temporal predicate meanings (first to focus-sensitive progressive aspect and then to proximal future tense), as posterior to PB.

Devos & Bernander present the results of their comparative study of existential constructions in a convenience sample of 180 Bantu languages with a special focus on existential locationals. The two most widespread constructions are one with a locative copula and (formal) locative inversion, i.e. *[LOC.NP

#) LOC.SM-*di* # NP (# LOC.NP)] (# = word boundary), and another one with a locative subject marker and a comitative copula, i.e. *[(LOC.NP #) LOC.SM-*di* (#) *na* (#) NP (# LOC.NP)]. Despite their wide distribution across Bantu, **Devos & Bernander** doubt their reconstructability to PB, because of their scarcity in the North-Western and Central-Western Bantu branches. As discussed above, this might imply that the common Bantu main clause type known as ‘subject inversion’ and reconstructed to PB by Meeussen (1967: 120) as *anastasis* might also be a later innovation. North-Western and Central-Western Bantu languages tend to have non-inverted existential locationals, which are nevertheless uncommon elsewhere in Bantu and in the world’s languages. The rare non-inverted constructions outside of North-Western and Central-Western Bantu could be seen as instances of archaic heterogeneity, which would support their interpretation as a shared retention and thus their reconstruction to PB. **Devos & Bernander** are uncertain, however, whether this is the most plausible scenario, because North-Western and Central-Western Bantu do have “inverted constructions which are not easily interpreted as independent innovations but rather seem to involve traces of a former full-fledged concord system with locative agreement”. Inverted constructions could therefore be an archaism from PB after all. In that case, the emergence of the cross-linguistically uncommon non-inverted existential locationals needs to be accounted for. **Devos & Bernander** think that such an innovation could have been triggered by the reduction of the agreement system and the loss of locative agreement, which is widespread in the north-western Bantu periphery and possibly an effect of contact with non-Bantu languages.

Idiatov, lastly, deals with non-selective interrogative pronominals in PB and thus partially reviews the “fragmentary system of interrogative nouns with stem *-í* : 7 *ki-í* ‘what’, 16 *pa-í* (17 *ku-í*, 18 *mu-í*) ‘where’; but 1a *n(d)ái* ‘who’” (Meeussen 1967: 103), which Meeussen reconstructs, with some hesitance on whether the last interrogative is really part of it, because “an element *n(d)á-* [...] is not attested otherwise” (Meeussen 1967: 103). **Idiatov** shows that there is no such thing as an element *n(d)á-*, but that such sequences may have popped up independently through Bantu language history due to the accretion of inherited morphology. In the same vein, he concludes that no ‘who?’ stem can be reconstructed for PB. The form *n(d)ái* “results from univerbation and nominalisation, either by conversion or by means of an overt nominaliser such as the augment, of a clause-level interrogative cleft construction”. Reconstructable PB non-selective interrogatives originate in complex constructions that were created earlier on at some ancestral Southern Bantoid stage, i.e. **à ndé yé-yà* (~ **à ndé yé-là*) [3SG COP NMLS₁-which?]

‘it is which one?’ and *à ndé yé-yà-yé (~ *à ndé yé-là-yé) [3SG COP NMLS₁-which?-NMLS₂] ‘it is which one exactly?’. The last one led to *n(d)ái*-like ‘who?’ interrogatives but also to question words meaning ‘what?’ or both ‘who?’ and ‘what?’. For PB ‘what?’, **Idiatov** reconstructs something like *yíí or *yíí, probably going back to the same pre-PB structure *yé-yà-yé (~ *yé-là-yé) [NMLS₁-which?-NMLS₂]. Given the complex constructional origin of non-selective interrogatives, **Idiatov** also touches upon several other issues of Bantu historical morphosyntax, such as deictics (both spatial and discourse ones), the so-called augment and more generally referential status marking, nominalisation, noun classes, subject indexation, copulas, cleft constructions, relative clause constructions, constituent order, and root.

Acknowledgements

I wish to thank Tom Güldemann, Larry M. Hyman, Dmitry Idiatov, Sara Pacchiarotti, Mark Van de Velde, Benji Wald, John R. Watters and Jeffrey Wills for their feedback on a previous version of this introduction.

Abbreviations

BRA	Bantu Relative Agreement	PB	Proto-Bantu
C	consonant	PL	plural
CM	Comparative Method	PP	pronominal prefix
COP	copula	SBJ	subject
LOC	locative	SG	singular
MOM	multiple object marking	SM	subject marker
NMLS	nominaliser	SOM	single object marking
NOM	no object marking	TAM(P)	tense/aspect/mood/(polarity)
NP	noun phrase	V	verb; vowel
OBJ	object	#	word boundary
O _{pro}	object pronoun; pronominal object		

References

- Allen, James P. 2013. *The Ancient Egyptian language: An historical study*. Cambridge: Cambridge University Press.
- Alsina, Alex & Sam A. Mchombo. 1990. The syntax of applicatives in Chicheŵa: Problems for a theta theoretic asymmetry. *Natural Language & Linguistic Theory* 8(4). 493–506.

- Alsina, Alex & Sam A. Mchombo. 1993. Object asymmetries and the Chicheŵa applicative construction. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar 1* (CSLI Lecture Notes 38), 17–45. Stanford, CA: Center for the Study of Language & Information.
- Aunio, Lotta. 2015. A typological perspective on Bantu nominal tone: The case of Ikoma-Nata-Isenye in Western Tanzania. *Southern African Linguistics and Applied Language Studies* 33(3). 359–371.
- Baker, Mark C. 1988. Theta theory and the syntax of applicatives in Chichewa. *Natural Language and Linguistic Theory* 6(3). 353–389.
- Baker, Mark C. 1990. Elements of a typology of applicatives in Bantu. In John P. Hutchison & Victor Manfredi (eds.), *Current approaches to African linguistics 7* (Publications in African Languages and Linguistics 11), 111–124. Dordrecht: Foris Publications.
- Baker, Mark C. 1992. Thematic conditions on syntactic structures: Evidence from locative applicatives. In Iggy M. Roca (ed.), *Thematic structure: Its role in grammar* (Linguistic Models 16), 23–46. Berlin: Walter de Gruyter.
- Baker, Mark Cleland. 1985. *Incorporation: A theory of grammatical function changing*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Baldi, Philip. 1990. Introduction: The comparative method. In Philip Baldi (ed.), *Linguistic change and reconstruction methodology 1* (Trends in Linguistics. Studies and Monographs 45), 1–13. Berlin: Mouton de Gruyter.
- Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en bantou* (Annales – Série in-8° – Sciences humaines 114). Tervuren: Royal Museum for Central Africa.
- Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and divergence in the Bantu languages: Perspectives from a lexicostatistic study* (Annals – Human Sciences 162). Tervuren: Royal Museum for Central Africa. https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/lexicostatistic-study-bantu-languages.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bastin, Yvonne & Pascale Piron. 1999. Classifications lexicostatistiques : bantou, bantou et bantoïde (de l'intérêt des « groupes flottants »). In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 149–163. Stanford, CA: Center for the Study of Language & Information.

- Bleek, Wilhelm H. I. 1862. *A comparative grammar of South African languages, Part 1: Phonology*. London: Trübner & Co.
- Bleek, Wilhelm H. I. 1869. *A comparative grammar of South African languages, Part 2: The concord*. London: Trübner & Co.
- Blench, Roger M. 2006. *Archaeology, language, and the African past*. Lanham, MD: AltaMira Press.
- Blench, Roger M. 2022. The relevance of Bantoid for the reconstruction of Proto-Bantu verbal extensions. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 235–280. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575823.
- Bontinck, François & D. Ndembe Nsasi. 1978. *Le catéchisme kikongo de 1624 : réédition critique* (Classe des Sciences Morales et Politiques 44-5). Brussels: Académie royale des Sciences d’Outre-Mer.
- Bostoen, Koen. 2018. The Bantu expansion. In *Oxford research encyclopedia of African history*. Published online April 26, 2018. DOI: 10.1093/acrefore/9780190277734.013.191.
- Bostoen, Koen. 2019. Reconstructing Proto-Bantu. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 308–334. Milton Park, Abingdon: Routledge.
- Bostoen, Koen. Forthcoming. Language and (pre)history in Bantu-speaking Africa. In Lutz Marten, Nancy C. Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages* (Oxford Guides to the World’s Languages 2). Oxford: Oxford University Press.
- Bostoen, Koen, Bernard Clist, Charles Doumenge, Rebecca Grollemund, Jean-Marie Hombert, Joseph Koni Muluwa & Jean Maley. 2015. Middle to late Holocene paleoclimatic change and the early Bantu expansion in the rain forests of Western Central Africa. *Current Anthropology* 56(3). 354–384.
- Bostoen, Koen & Mark Van de Velde. 2019. Introduction. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 1–13. Milton Park, Abingdon: Routledge.
- Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.
- Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2). 147–185.
- Bresnan, Joan & Lioba Moshi. 1993. Object asymmetries in comparative Bantu syntax. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar 1* (CSLI

- Lecture Notes 38), 47–92. Stanford, CA: Center for the Study of Language & Information.
- Bybee, Joan L., Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago, IL: University of Chicago Press.
- Cardoso, Mattheus. 1624. *Dovtrina Christaã. Composta pelo P. Marcos Iorge da Companhia de IESV Doutor em Theologia. Acrescentada pelo Padre Ignacio Martinz da mesma Companhia Doutor Theologo. De nouo traduzida na lingoa do Reyno de Congo, por ordem do P. Mattheus Cardoso Theologo, da Companhia de IESV, natural da cidade de Lisboa. Ao muito poderoso, & catholico Rey de Congo dom Pedro Affonso Segundo deste nome* [Christian Doctrine. Composed by Father Marcos Jorge of the Society of Jesus, PhD. Augmented by Father Ignacio Martinz of the same Society, PhD. Translated into the language of the Kongo Kingdom under the supervision of Mattheus Cardoso, Theologian of the Society of Jesus, born in the city of Lisbon. To the very powerful and catholic King of the Kongo Dom Pedro Affonso II of this name]. Lisbon: Geraldo da Vinha.
- Chatzikyriakidis, Stergios & Hannah C. Gibson. 2017. The Bantu-Romance-Greek connection revisited: Processing constraints in auxiliary and clitic placement from a cross-linguistic perspective. *Glossa: A journal of general linguistics* 2(1). 1–39.
- Clements, George N. & John Goldsmith (eds.). 1984. *Autosegmental studies in Bantu tone*. Dordrecht: Foris Publications.
- Clements, George N. & Annie Rialland. 2008. Africa as a phonological area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 36–85. Cambridge: Cambridge University Press.
- Coupez, André, Yvonne Bastin & Evariste Mumba (eds.). 1998. *Bantu lexical reconstructions 2 / Reconstructions lexicales bantoues 2*. Tervuren: Royal Museum for Central Africa.
- Dalby, David. 1975. The prehistorical implications of Guthrie's 'Comparative Bantu'. Part I: Problems of internal relationship. *Journal of African History* 16(4). 481–501.
- Dalby, David. 1976. The prehistorical implications of Guthrie's 'Comparative Bantu'. Part II: Interpretation of cultural vocabulary. *Journal of African History* 17(1). 1–27.
- de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87–162 + supplementary material online.

- de Vos, Mark & Hazel Mitchley. 2012. Subject marking and preverbal coordination in Sesotho: A perspective from Optimality Theory. *Southern African Linguistics and Applied Language Studies* 30(2). 155–170.
- Devos, Maud, Michael Kasombo Tshibanda & Johan van der Auwera. 2010. Jespersen cycles in Kanincin: Double, triple and maybe even quadruple negation. *Africana Linguistica* 16. 155–181.
- Devos, Maud & Johan van der Auwera. 2013. Jespersen cycles in Bantu: Double and triple negation. *Journal of African Languages and Linguistics* 34(2). 205–274.
- Dimmendaal, Gerrit J. 2011. *Historical linguistics and the comparative study of African languages*. Amsterdam: John Benjamins.
- Dom, Sebastian, Leonid Kulikov & Koen Bostoen. 2018. Valency-decreasing derivations and quasi-middles in Bantu: A typological perspective. *Southern African Linguistics and Applied Language Studies* 36(3). 165–173.
- Doneux, Jean Léonce. 1965. Bibliographie du programme Lolemi. *Africana Linguistica* 2. 201–221.
- Downing, Laura J. 1995. The metrical domain of register raising in Jita: An Optimality approach. In Francis X. Katamba (ed.), *Bantu phonology and morphology* (LINCOS Studies in African Linguistics 6), 28–39. Munich: LINCOS Europa.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2022. *Ethnologue: Languages of the world*. 25th edn. Dallas, TX: SIL International.
- Fleisch, Axel. 2006. Agent phrases in Bantu passives. In F.K. Erhard Voeltz (ed.), *Studies in African linguistic typology* (Typological Studies in Language 64), 93–111. Amsterdam: John Benjamins.
- Gary, Judith O. & Edward L. Keenan. 1977. On collapsing grammatical relations in Universal Grammar. In Peter Cole & Jerrold M. Sadock (eds.), *Syntax and semantics 8: Grammatical relations*, 83–120. New York, NY: Academic Press.
- Gibson, Hannah C. 2012. *Auxiliary placement in Rangi: A Dynamic Syntax perspective*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Gibson, Hannah C. & Lutz Marten. 2016. Variation and grammaticalisation in Bantu complex verbal constructions: The dynamics of information growth in Swahili, Rangi and siSwati. In Léa Nash & Pollet Samvelian (eds.), *Approaches to complex predicates* (Syntax and Semantics 41), 70–109. Leiden: Brill.
- Gildea, Spike. 2000. Preface. In Spike Gildea (ed.), *Reconstructing grammar: Comparative linguistics and grammaticalization* (Typological Studies in Language 43), vii–xiv. Amsterdam: John Benjamins.
- Givón, Talmy. 1971a. Dependent modals, performatives, factivity, Bantu subjunctives and what not. *Studies in African Linguistics* 2(1). 61–81.

- Givón, Talmy. 1971b. Historical syntax and synchronic morphology: An archaeologist's field trip. *Chicago Linguistic Society* 7. 394–415.
- Givón, Talmy. 1974. Syntactic change in Lake-Bantu: A rejoinder. *Studies in African Linguistics* 5(1). 117–139.
- Goldsmith, John. 1987. Stem tone patterns of the Lacustrine Bantu languages. In David Odden (ed.), *Current approaches to African linguistics*, vol. 4 (Publications in African Languages and Linguistics 7), 167–177. Dordrecht: Foris Publications.
- Goldsmith, John A. 1976. *Autosegmental phonology*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Good, Jeff. 2005. Reconstructing morpheme order in Bantu: The case of causativization and applicativization. *Diachronica* 22(1). 3–57.
- Good, Jeff. 2022. Reconstructing the development of the Bantu final vowels. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pachiariotti (eds.), *On reconstructing Proto-Bantu grammar*, 173–234. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575821.
- Goody, Jack. 1986. *The logic of writing and the organization of society* (Studies in Literacy, the Family, Culture and the State). Cambridge: Cambridge University Press.
- Greenberg, Joseph H. 1948. The tonal system of Proto-Bantu. *Word* 4(3). 196–208.
- Greenberg, Joseph H. 1951. Vowel and nasal harmony in Bantu languages. *Zaire : revue congolaise / Zaïre: Congolees tijdschrift* 5. 813–820.
- Greenberg, Joseph H. 1958. The labial consonants of Proto-Afro-Asiatic. *Word* 14(2–3). 295–302.
- Greenberg, Joseph H. 1963. *The languages of Africa* (Indiana University Research Center in Anthropology, Folklore and Linguistics 25). Bloomington, IN; The Hague: Indiana University; Mouton.
- Greenberg, Joseph H. 1966. *Language universals: With special reference to feature hierarchies*. The Hague: Mouton.
- Greenberg, Joseph H. 1972. Linguistic evidence regarding Bantu origins. *The Journal of African History* 13(2). 189–216.
- Greenberg, Joseph H., Charles A. Ferguson & Edith A. Moravcsik (eds.). 1978. *Universals of human language, Vol 1. Method & theory, Vol. 2. Phonology, Vol. 3. Word structure, Vol. 4. Syntax*. Stanford, CA: Stanford University Press.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.

- Guérois, Rozenn. 2017. Conditional constructions in Cwabo. *Studies in African Linguistics* 46(1–2). 193–212.
- Güldemann, Tom. 1996. *Verbalmorphologie und Nebenprädikationen im Bantu: Eine Studie zur funktional motivierten Genese eines konjugationalen Subsystems* (Bochum-Essener Beiträge zur Sprachwandelforschung 27). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Güldemann, Tom. 1999a. The genesis of verbal negation in Bantu and its dependency on functional features of clause types. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 545–587. Stanford, CA: Center for the Study of Language & Information.
- Güldemann, Tom. 1999b. Toward a grammaticalization and typological account of the ka-possessive in Zulu. *Journal of African Languages and Linguistics* 20(2). 157–184.
- Güldemann, Tom. 2003a. Grammaticalization. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 182–194. London: Routledge.
- Güldemann, Tom. 2003b. Present progressive vis-à-vis predication focus in Bantu: A verbal category between semantics and pragmatics. *Studies in Language* 27(2). 323–360.
- Güldemann, Tom. 2008. The Macro-Sudan Belt: Towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151–185. Cambridge: Cambridge University Press.
- Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas: With special reference to Africa* (Tokyo University of Foreign Studies, Studies in Linguistics 2), 109–141. Amsterdam: John Benjamins.
- Güldemann, Tom. 2017. A shared pronominal canon in the Macro-Sudan Belt: Typological, areal and genealogical perspectives. In Raija Kramer & Roland Kießling (eds.), *Mechthildian approaches to Afrikanistik: Advances in language based research in Africa. Festschrift für Mechthild Reh*, 101–146. Cologne: Rüdiger Köppe.
- Güldemann, Tom. 2018. Areal linguistics beyond contact, and linguistic areas in Afrabia. In Tom Güldemann (ed.), *The languages and linguistics of Africa* (The World of Linguistics 11), 448–545. Berlin: Mouton de Gruyter.

- Güldemann, Tom. 2022. Predicate structure and argument indexing in early Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 387–421. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575831.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1962. A two-stage method of comparative Bantu study. *African Language Studies* 3. 1–24.
- Guthrie, Malcolm. 1967. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 1: The comparative linguistics of the Bantu languages*. Farnborough: Gregg Press.
- Guthrie, Malcolm. 1970. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part II, Volumes 3 & 4: A catalogue of Common Bantu with commentary*. Farnborough: Gregg International.
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Han, Xu, Yuzhuo Bai, Keyue Qiu, Zhiyuan Liu & Maosong Sun. 2020. Isobs: An information system for oracle bone script. In Qun Liu & David Schlangen (eds.), *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 227–233. Stroudsburg, PA: Association for Computational Linguistics.
- Harford, Carolyn & Katherine Demuth. 1999. Prosody outranks syntax: An Optimality approach to subject inversion in Bantu relatives. *Linguistic Analysis* 29(1–2). 47–68.
- Heine, Bernd, Tom Güldemann, Christa Kilian-Hatz, Donald A. Lessau, Heinz Roberg, Mathias Schladt & Thomas Stolz. 1993. *Conceptual shift: A lexicon of grammaticalization processes in African languages* (Afrikanistische Arbeitspapiere 34-35). Cologne: Institut für Afrikanistik.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.
- Helmlinger, Paul. 1972. *Dictionnaire duala – français, suivi d’un lexique français – duala* (Langues et littératures de l’Afrique noire 9). Paris: Librairie C. Klincksieck (pour L’Université fédérale du Cameroun).
- Hetzron, Robert. 1976. Two principles of genetic reconstruction. *Lingua* 38(2). 89–108.

- Hock, Hans H. 1991. *Principles of historical linguistics*. Second edition, revised and updated (Mouton Textbook). Berlin: Mouton de Gruyter.
- Hoeningswald, Henry M. 1991. Morphemic change, typology, and uniformitarianism: A study in reconstruction. In Winfred P. Lehmann & Helen-Jo Jakusz Hewitt (eds.), *Language typology 1988: Typological models in the service of reconstruction* (Current Issues in Linguistic Theory 81), 17–26. Amsterdam: John Benjamins.
- Hyman, Larry M. 2003. Suffix ordering in Bantu: A morphocentric approach. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2002*, 245–281. Dordrecht: Kluwer Academic Publishers.
- Hyman, Larry M. 2004. How to become a ‘Kwa’ verb. *Journal of West African Languages* 30(2). 69–88.
- Hyman, Larry M. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville, MA: Cascadilla Press.
- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.
- Hyman, Larry M. 2017. Multiple argument marking in Bantoid: From syntheticity to analyticity. In Walter Bisang & Andrej Malchukov (eds.), *Unity and diversity in grammaticalization scenarios* (Studies in Diversity Linguistics 16), 67–95. Berlin: Language Science Press. DOI: 10.5281/zenodo.823230.
- Hyman, Larry M. 2018. What is phonological typology? In Larry M. Hyman & Frans Plank (eds.), *Phonological typology* (Phonology and Phonetics 23), 1–20. Berlin: Walter de Gruyter.
- Hyman, Larry M. & Charles W. Kisseberth (eds.). 1998. *Theoretical aspects of Bantu tone*. Stanford, CA: Center for the Study of Language & Information.
- Idiatov, Dmitry. 2009. A Bantu path towards lack of differentiation between ‘who?’ and ‘what?’ *Africana Linguistica* 15. 59–76.
- Idiatov, Dmitry. 2022. Diachronic typology and the reconstruction of non-selective interrogative pronominals in Proto-Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 667–737. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575843.
- Idiatov, Dmitry & Mark Van de Velde. 2021. The lexical distribution of labial-velar stops is a window into the linguistic prehistory of Northern Sub-Saharan Africa. *Language* 97(1). 72–107.
- Janssens, Baudouin. 1993. *Doubles réflexes consonantiques : quatre études sur le bantu de zone A (bubi, nen, bafia, ewondo)*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).

- Kadenge, Maxwell. 2014. Comparing nasal-obstruent clusters derived from /mu-/ reduction in Shangwe and Zezuru: An Optimality Theory analysis. *Language Matters, Studies in the Languages of Africa* 45(1). 40–62.
- Kadenge, Maxwell & Silvester R. Simango. 2014. Comparing vowel hiatus resolution in ciNsenga and chiShona: An Optimality Theory analysis. *Stellenbosch Papers in Linguistics Plus* 44. 105–127.
- Kahl, Jochem. 2001. Hieroglyphic writing during the fourth millennium BC: An analysis of systems. *Archéo-Nil* 11. 101–134.
- Kempson, Ruth M., Wilfried Meyer-Viol & Dov M. Gabbay. 2001. *Dynamic Syntax: The flow of language understanding*. Oxford: Wiley-Blackwell.
- Kettunen, Harri & Christophe Helmke. 2019. *Introduction to Maya hieroglyphs*. Sixteenth revised edition. Wayeb. <https://www.wayeb.org/wp-content/uploads/2019/04/wh2019english.pdf>.
- Kisseberth, Charles W. & David Odden. 2003. Tone. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 59–70. London: Routledge.
- Koch, Harold. 1996. Reconstruction in morphology. In Mark Durie & Malcolm Ross (eds.), *The comparative method reviewed: Regularity and irregularity in language change*, 218–263. Oxford: Oxford University Press.
- Koch, Harold. 2014. Morphological reconstruction. In Claire Bowerman & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 286–307. London: Routledge.
- Koile, Ezequiel, Simon J. Greenhill, Damián E. Blasi, Remco Bouckaert & Russell D. Gray. 2022. Phylogeographic analysis of the Bantu language expansion supports a rainforest route. *Proceedings of the National Academy of Sciences* 119(32). e2112853119.
- LeBlanc, Paul D. 2017. *Deciphering the Proto-Sinaitic script: Making sense of the Wadi el-Hol and Serabit el-Khadim early alphabetic inscriptions*. Ottawa: Subclass Press.
- Leitch, Myles Francis. 1996. *Vowel harmonies in the Congo Basin: An Optimality Theory analysis of variation in the Bantu zone C*. Vancouver: University of British Columbia. (Doctoral dissertation).
- Lusekelo, Amani. 2012. *Inflectional and derivational morphology in Optimality Theory. Multiple object-nouns and co-occurrence of verbal extensions in Kinyakyusa*. Gaborone: University of Botswana. (Doctoral dissertation).
- Marantz, Alec Paul. 1981. *On the nature of grammatical relations*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

- Marantz, Alec Paul. 1982. Affixation and the syntax of applied verb constructions. In Daniel P. Flickinger, Marlys Macken & Nancy Wiegand (eds.), *Proceedings of the First West Coast Conference on Formal Linguistics*, 330–340. Stanford, CA: Center for the Study of Language & Information.
- Marlo, Michael R. & David Odden. 2019. Tone. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 150–171. Milton Park, Abingdon: Routledge.
- Marten, Lutz. 2002. *At the syntax-pragmatics interface: Verbal underspecification and concept formation in Dynamic Syntax*. Oxford: Oxford University Press.
- Marten, Lutz. 2003. The dynamics of Bantu applied verbs: An analysis at the syntax-pragmatics interface. In Kézié K. Lévikaza (ed.), *Actes du 3e congrès mondial de linguistique africaine. Lomé 2000*, 207–221. Cologne: Rüdiger Köppe.
- Marten, Lutz. 2013. Formal syntax and African languages: Dynamic Syntax case studies of Swahili, Bemba, and Hadiyya. *SOAS Working Papers in Linguistics* 16. 195–213.
- Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318–368.
- Maslova, Elena S. 2007. Reciprocal and polyadic (remarkable reciprocals in Bantu). In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions* (Typological Studies in Language 71), 335–352. Amsterdam: John Benjamins.
- McGinnis, Martha. 2001. Variation in the phase structure of applicatives. In Pierre Pica (ed.), *Linguistic variation yearbook 2001* (Linguistic Variation Yearbook 1), 105–146. Amsterdam: John Benjamins.
- McGinnis, Martha. 2008. Applicatives. *Language and Linguistics Compass* 2(6). 1225–1245.
- Meeussen, Achiel Emiel. 1965. Vijf jaar Lolemi-activiteit. *Africa-Tervuren* 15. 1–3.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeussen, Achiel Emiel. 1969. *Bantu lexical reconstructions*. Unpublished ms., 28 pp. Tervuren: Royal Museum for Central Africa.
- Meeussen, Achiel Emiel. 1973. Comparative Bantu: Test cases for method. *African Language Studies* 14. 6–18.
- Meinhof, Carl. 1899. *Grundriß einer Lautlehre der Bantusprachen, nebst Anleitung zur Aufnahme von Bantusprachen. Anhang: Verzeichnis von Bantuwortstämmen* (Abhandlungen für die Kunde des Morgenlandes 11(2)). Leipzig: F.A. Brockhaus.
- Meinhof, Carl. 1906. *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. Berlin: Dietrich Reimer.

- Meinhof, Carl, Alice Werner & Nicolaas J. Van Warmelo. 1932. *Introduction to the phonology of the Bantu languages*. Being the English version of *Grundriß einer Lautlehre der Bantusprachen*. Berlin: Dietrich Reimer (Ernst Vohsen).
- Mugane, John M. 2015. *The story of Swahili*. Athens, OH: Ohio University Press.
- Myers, Scott. 1997. OCP effects in Optimality Theory. *Natural Language & Linguistic Theory* 15(4). 847–892.
- Ngo-Ngijol Banoum, Bertrade B. 2004. Bantu gender revisited through an analysis of Basaá categories: A typological perspective. In Akinbiyi Akinlabi & Olusye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 445–458. Cologne: Rüdiger Köppe.
- Norde, Muriel. 2009. *Degrammaticalization* (Oxford Linguistics). Oxford: Oxford University Press.
- Nsuka-Nkutsi, François. 1982. *Les structures fondamentales du relatif dans les langues bantoues* (Annales – Série in-8° – Sciences humaines 107). Tervuren: Royal Museum for Central Africa.
- Nunn, Nathan & Nancy Qian. 2010. The Columbian exchange: A history of disease, food, and ideas. *Journal of Economic Perspectives* 24(2). 163–188.
- Nurse, Derek. 1997. The contributions of linguistics to the study of history in Africa. *Journal of African History* 38(3). 359–391.
- Nurse, Derek. 2003. Aspect and tense in Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 90–102. London: Routledge.
- Nurse, Derek. 2007. Did the Proto-Bantu verb have a synthetic or an analytic structure? In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 239–256. London: Department of Linguistics, School of Oriental & African Studies.
- Nurse, Derek. 2008. *Tense and aspect in Bantu* (Oxford Linguistics). New York, NY: Oxford University Press.
- Nurse, Derek & Maud Devos. 2019. Aspect, tense and mood. In Mark Van de Velde, Koen Bostoën, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 204–236. Milton Park, Abingdon: Routledge.
- Nurse, Derek & Gérard Philippson. 2003. Towards a historical classification of the Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 164–181. London: Routledge.
- Nurse, Derek & Gérard Philippson. 2006. Common tense-aspect markers in Bantu. *Journal of African Languages and Linguistics* 27(2). 155–196.
- Odden, David. 1999. Typological issues in tone and stress in Bantu. In Shigeki Kaji (ed.), *Proceedings of the Symposium “Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Typology, and Related Topics”, December 10–12, 1998, Takino-*

- gawa City Hall, Tokyo*, 187–215. Tokyo: Institute for the Study of Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.
- Pacchiarotti, Sara. 2020. *Bantu applicative constructions* (Stanford Monographs in African Languages). Stanford, CA: Center for the Study of Language & Information.
- Pacchiarotti, Sara & Koen Bostoen. 2020. The Proto-West-Coastal Bantu velar merger. *Africana Linguistica* 26. 139–195.
- Pacchiarotti, Sara & Koen Bostoen. 2022. Erratic velars in West-Coastal Bantu: Explaining irregular sound change in Central Africa. *Journal of Historical Linguistics* 12(3). 381–445.
- Pae, Hye K. 2018. Written languages, East-Asian scripts, and cross-linguistic influences: An introduction. In Hye K. Pae (ed.), *Writing systems, reading processes, and cross-linguistic influences. Reflections from the Chinese, Japanese and Korean languages* (Bilingual Processing and Acquisition 7), 1–22. Amsterdam: John Benjamins.
- Perlmutter, David M. & Paul M. Postal. 1983. Toward a universal characterization of passivization. In David M. Perlmutter (ed.), *Studies in relational grammar 1*, 3–29. Chicago, IL: University of Chicago Press.
- Philippon, Gérard. 2022. Double reflexes in north-western Bantu and their implications for the Proto-Bantu consonant system. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 3–58. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575815.
- Philippon, Gérard & Rebecca Grollemund. 2019. Classifying Bantu languages. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippon (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 335–354. Milton Park, Abingdon: Routledge.
- Polak, Louise. 1986. Les infixes (“prefixes objets”) du bantou et leur reconstruction. *Africana Linguistica* 10. 365–421.
- Polak-Bynon, Louise. 1964. Lolemi : un programme de dépouillement de grammaires bantoues. *Africa-Tervuren* 10. 63–66.
- Poulos, George. 1984. A typological approach to Bantu linguistic studies. *African Studies* 43(2). 241–248.
- Poulos, George. 1985. Typological trends in South-Eastern Bantu. *South African Journal of African Languages* 5(1). 17–23.
- Prince, Alan S. & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in Generative Grammar*. Tech. rep. Piscataway, NJ: Rutgers Center for Cognitive Science, Rutgers, the State University of New Jersey.

- Pylkkänen, Liina. 2000a. What applicative heads apply to. *University of Pennsylvania Working Papers in Linguistics* 7(1). 197–210.
- Pylkkänen, Liina. 2008. *Introducing arguments* (Linguistic Inquiry Monographs 49). Cambridge, MA: The MIT Press.
- Pylkkänen, Liina. 2000b. Deriving adversity. In Roger Billerey & Brook D. Lillehaugen (eds.), *Proceedings of the 19th West Coast Conference on Formal Linguistics*, 399–410. Somerville, MA: Cascadilla Press.
- Rosen, Carol G. 1984. The interface between semantic roles and initial grammatical relations. In David M. Perlmutter & Carol G. Rosen (eds.), *Studies in relational grammar 2*, 38–77. Chicago, IL: University of Chicago Press.
- Sacleux, Charles. 1939. *Dictionnaire swahili – français* (Travaux et mémoires de l’Institut d’ethnologie 36–37). Paris: Institut d’ethnologie.
- Schadeberg, Thilo C. 2003. Historical linguistics. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 143–163. London: Routledge.
- Sebasoni, Servilien. 1967. La préfinale du verbe bantou. *Africana Linguistica* 3. 122–135.
- Stappers, Leo. 1967. Het passief suffix -u- in de Bantoe-talen. *Africana Linguistica* 3. 137–145.
- Traugott, Elizabeth C. & Bernd Heine (eds.). 1991. *Approaches to grammaticalization* (Typological Studies in Language 19). Amsterdam: John Benjamins.
- Van de Velde, Mark. 2006. Multifunctional agreement patterns in Bantu and the possibility of genderless nouns. *Linguistic Typology* 10(2). 183–221.
- Van de Velde, Mark. 2009. Agreement as a grammatical criterion for proper name status in Kirundi. *Onoma* 44. 219–241.
- Vansina, Jan. 1995. New linguistic evidence and ‘the Bantu expansion’. *Journal of African History* 36(2). 173–195.
- Watters, John R. 2018. East Benue-Congo. In John R. Watters (ed.), *East Benue-Congo: Nouns, pronouns, and verbs* (Niger-Congo Comparative Studies 1), 1–25. Berlin: Language Science Press. DOI: 10.5281/zenodo.1314337.
- Weiss, Michael. 2014. The comparative method. In Claire Bowerman & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 127–146. London: Routledge.

Part I

Proto-Bantu phonology

Chapter 1

Double reflexes in north-western Bantu and their implications for the Proto-Bantu consonant system

Gérard Philippon

DDL - Dynamique du Langage

A number of languages in the north-westernmost area of the Bantu domain have been claimed to present two different reflexes of originally unitary Proto-Bantu (PB) phonemes. A solution to this surprising situation has been sought in the presence of some assumed phonological conditioning, whereas other authors have proposed to reconstruct new proto-phonemes. The present chapter establishes that for voiced PB phonemes, a tonal conditioning can indeed be found; but for voiceless PB phonemes, the situation is more confused, and specifically there emerges a small but consistent sub-group of reconstructed stems which escape the general “weakening” of the proto-phoneme *t, without any obvious conditioning. The hypothesis is that according to a wave model, those items were not touched by the weakening innovation at the time of its spread.

1 Introduction

The Comparative Method in historical linguistics aims at establishing series of regular sound correspondences among related languages with the ultimate goal of reconstructing the sound system of the ancestor language. It has succeeded in numerous cases, mainly, to be sure, among closely related languages. However, irregularities in correspondences often occur in a somewhat haphazard manner from which no general conclusions can be drawn. In other cases, a considerable part of the lexicon is affected by such irregularities and comparative linguists have tended to approach the question in two different ways, either: (a) by considering that the change considered has not (yet) affected all the eligible lexical



items at time t , which is often formulated as some variant of a “wave” model – for a recent summary, see François (2014); or (b) by positing a new series of proto-phonemes, or some prosodic conditioning, as in the case of Verner’s Law (see, for instance, Halle 1997) to account for the different correspondences.

Many Bantu languages of the north-western part of the domain have been shown to exhibit divergent correspondences for some of the putative Proto-Bantu (PB) phonemes without any apparent conditioning, a problem known in historical Bantu linguistics as that of “double reflexes” (cf. Van Leynseele & Stewart 1980; Bancel 1988; Janssens 1991; Teil-Dautrey 1991a; Botne 1992a; Janssens 1993). The aim of this chapter is to examine this situation in detail, to assess whether it has implications for the reconstruction of PB and provide a tentative solution for this apparent challenge to the Comparative Method. To maintain the size of the chapter within reasonable limits, we focus here on the C_1 position. Consonants in C_2 are left for a later study.

1.1 Classification of the north-western Bantu languages

Without attempting a complete review of all the proposals aiming at delimitating “Bantu” from “non-Bantu” languages (cf. Watters & Leroy 1989a,b; Grollemund 2012; Philippson & Grollemund 2019), I will refer to the most complete phylogenetic classification to date, i.e. Grollemund et al. (2015), while commenting on it.

But first it is necessary to set out the list of the languages which will be examined here. It is usual among Bantuists to identify languages by an alphanumeric code, first devised by Malcolm Guthrie (1948; 1953; 1971) and expanded by Jouni Maho (2003; 2009) – see also Hammarström (2019). All the languages covered by this chapter have a referential code beginning with A, except for Seki B21. In this sense, my sample is complementary to the one of Pacchiarotti & Bostoen (2022), who deal with the same problem, but focus on the irregular reflexes of PB velar stops in West-Coastal Bantu, i.e. Guthrie’s groups B40–80, H10, H30–40 (except Mbala H41), and Samba L12a.

In theory, languages sharing the same letter and first digit (e.g. A41, A42, A43 etc.) should be closely related, but this is not always true, as will be seen below. In order to avoid relying too much on the alphanumeric codes as if they had a genealogical value, I will propose names, sometimes geographical, for the various groups, in the same way that Ehret (1998) on the one hand, and Nurse & Philippson (1980b,a) did for Eastern Bantu languages. I borrow some of these names from others, for example “Sawabantu” from Ebobissé (2015), even if I give it a broader compass. Below I summarise my current view on the classification

1 *Double reflexes in north-western Bantu*

of the north-western Bantu languages dealt with in this chapter – see also Appendix A for an overview of the languages included in this study and the sources consulted for each language.

1) Mbam

Western: Nen A44; Nyokon A45; Maande A46; Tuotomb A461

Sanaga: Tuki > Ki, etc. A601

Yambasa > Yangben A62A; Mmala A62B; Elip A62C; Baca A621; Gunu A622

Mbule A623

Yambeta A462 is difficult to affiliate but seems a little closer to Sanaga

2) Bafia

Fa', Zakaan, Maja, Balom A51; Dimbong A52; Kpa, Pe A53; Bea (Ngayaba) A54

3) Bubi

Northern A31a; South-West A31b; South-East A31c

4) Sawabantu

Oroko: Lundu A11; Ngolo A111; Bima A112; Batanga A113; Lokoko A114; Londo ba Diko A115; Lue A12; Mbonge A121; Kundu A122; Ekombe A123

Central: Mboko A21; Kpe A22; Bubia A221; Su A23; Kole A231; Duala A24; Bodiman A241; Oli A25; Pongo A26; Mongo A261; Limba A27

Southern: Noho A32a; Bapuku A32b; Batanga A32C; Yasa A33a; Kombe A33b

Benga: Benga A34

5) Manenguba¹

North-east: Mbuu, Mboo A15A

¹For all these languages minus A13c, see classification and data in Hedinger (1987). Nkongho A151 is only known from a wordlist supplied by Hedinger (1987). It is definitely not part of Manenguba and might in fact not be Narrow Bantu. But no valid conclusion can be reached on the basis of such meagre data. Note that it is part of Manenguba in Grollemund et al. (2015), but to my mind, this is due to somewhat dubious cognate identifications.

North-west: Myenge A15B

Central: A15C

Eastern: Mkaa, Mwahed, Mwaneka, Belon, etc.

Western: Akoose, Elung, Mbo, Nnenong, etc.

Balong/Bafo: Balong A13; Bafo (Lefo) A141; (Bonkeng A14 ?)

6) Basaa

Lombi A41; Abo (Bankon) A42; Basaa A43a; Bakoko (North Kogo) A43b; South Kogo A43c ²

7) Beti

Eton A71; Ewondo A72a; Bulu A74a; Fang A75 (Northern: Ntumu, etc. A75A; Southern: Okak A75B; Atsi A75D; Mvai A75F) + Njowi A63 ³

8) Nyong-Dja

Northern: Makaa A83; Kol A832; Njem A84; Bajwee A841; Koonzime A842; Bekwel A85b; Mpiemo, etc. A86c + Polri A92a (and Pomo A92b ?)

Southern: Gyeli A801; Shiwa A803; Kwasio A81

9) Kwakum

Kwakum A91; Kako A93; Seki B21

The Mbam languages are placed by lexicon-based quantitative classifications (both lexicostatistical and phylogenetic) as standing outside the rest of Narrow Bantu, which is partially confirmed by their diachronic phonology.

As for the rest of zone A languages, Bubi (group 3) and Sawabantu (group 4) differ in their phonological structure from the others (in that they have mostly CVCV stems) and they also have agglutinative verb structure. However, I see these as retentions which do not suggest any very close proximity between the two. Likewise, I consider the lexicostatistical closeness of Bubi to the Mbam languages as an artefact of lexicostatistics (Philippson 2018). Basaa (group 6) and Beti (group 7) have much in common and might well form a genuine clade, as supported by lexicon-based quantitative studies. Nyong-Dja (group 8) and Kwakum

²It is likely that the very poorly known Hijuk A501 also belongs to this group.

³Guthrie's (1953) categorisation of Njowi in the A60 group is due to a confusion of the ethnic name Mengisa which covers two linguistic entities, i.e. Leti (undescribed), which probably is most closely related to the other members of the Sanaga group A60, and Njowi A63, which is a close relative of Eton A71.

A91 share a specific (and rare) innovation, i.e. devoicing of voiced pre-nasalised stops such as **mb* > (*m*)*p*^(h). It is equally attested in Northern Bubi A31a and sporadically throughout Bantu. However, this innovation is shared by neither Kako A93 nor Seki B21 and therefore is probably due in Kwakum to contact with a Nyong-Dja language. Whether the Nyong-Dja and Kwakum groups as a whole form a sub-clade, as lexicon-based quantitative studies indicate, is not clear at present.

In the purely lexically-based classification of Grollemund et al. (2015), all the Grassfields languages are separated from Narrow Bantu, i.e. those which Guthrie considered as Bantu. However, Jarawan languages also belong to this “non-Grassfields” Bantu group. Not having had the opportunity to look in detail at Jarawan, I leave it out of consideration here. Within Narrow Bantu, Grollemund et al. (2015) have a first branching, including (alongside Jarawan) Mbam (group 1) and Bafia (group 2) as well as Bubi spoken on the island of Bioko.

After having looked carefully at the data (Philippson 2018), not only do I not see any close proximity between Mbam and Bubi, in particular the absence of common lexical innovations, this can also be seen in Grollemund (2012), which is a more detailed survey of the north-western Bantu languages. The lexicon of Bubi is highly idiosyncratic and certainly innovated. This is a case where the lexicon cannot be taken as a valid clue to genetic affiliation. An examination of the phonology and inflectional morphology of Bubi shows it to be much closer to the bulk of the north-western Bantu languages than to Mbam (Philippson 2018).

In Grollemund et al. (2015), groups 4 to 7 belong to the same branch, but are separate from groups 8 and 9, which cluster into another clade alongside several languages belonging to Guthrie’s B20 group. I have not had time yet to look at the latter languages in detail and will leave them out of the discussion, but there is no doubt that they do appear to exhibit similarities to groups 8 and 9 above. Nevertheless, I consider groups 8 and 9 to belong to a common clade with the other groups cited (apart from Mbam). As it is impossible to deal fully with this hypothesis in the context of a chapter devoted to double reflexes, I regretfully have to defer my arguments to another publication.

One more remark: In most discussions of Bantu diachronic phonology, much attention is generally paid to Bantu Frication (BF) (cf. Hyman 2003b; Hyman & Merrill 2016), also known as Bantu Spirantisation (cf. Schadeberg 1994; Bostoen 2008), i.e. the process by which stops are affected by a following [+high] vowel. It so happens that among the languages mentioned above, BF only concerns the southernmost languages, i.e. southern varieties of Fang and southern Nyong-Dja languages. It will thus be referred to only occasionally. Note that it should not be

confused, as it sometimes is, with a process of palatalisation before front vowels, the latter being quite active in our region of study.

1.2 The problem of double reflexes

Meeussen (1967: 83) lists the PB phonemes presented in Table 1, here in a reorganised way.

Table 1: Chart of PB phonemes in Meeussen (1967: 83)

*p	*t	*c	*k
*b	*d	*j	*g
*m	*n	*ɲ	
*mp	*nt	*ɲc	*ɲk
*mb	*nd	*ɲj	*ɲg
*mm	*nn	*ɲɲ	

Although left unmentioned in Meeussen (1967), the problem of “different consonant shifts” – termed “*dualité de reflexes*” in Van Leynseele & Stewart (1980), the first paper to systematically address the subject – in certain north-western Bantu languages was discussed by Guthrie (1967), who attributed the duality of certain consonantal reflexes to the quantity of the following vowel. Witness the following statement: “(T)here are cases in this area [zone A and the adjacent parts of zone B] where the shift in a starred consonant with *VV is different from that with *V” (Guthrie 1967: 58). He then immediately admitted: “The occurrence of this special sound shift with *VV necessitates the use of a double vowel in the starred form of some C.S., even though the vowel distinction *VV/*V is missing in all the entries [...]” (Guthrie 1967: 58). The latter explanation is tantamount to acknowledging that he used vowel length simply as a diacritic to identify the different reflexes.⁴

The zone A languages explicitly mentioned by Guthrie as exhibiting the phenomenon of double reflexes are the following: Lundu A11, Duala A24, Benga

⁴A fairly large number of stems with *NC₂ are entered as *CVVNC(V) in *Comparative Bantu* (Guthrie 1967; 1970a,b; 1971). Although nowhere stated explicitly, it would appear that Guthrie based himself primarily on the B50 languages which do offer length distinctions in such contexts, e.g. *kààng ‘tie up, seize’ > Tsaangi B53 *kaàng*, or *kóóndá ‘pigeon sp.’ > Nzebi B52 *lakoond(a)* vs. *bángá ‘jaw’ > Duma B51 *mubángá*, or *góóngò ‘hoe’ > Nzebi B52 *ləngong*. Bantu Lexical Reconstructions (Bastin et al. 2002) considers all *CVVNC(V) reconstructions as spurious and rejects them, on the principle that no length contrast is possible in Bantu languages before pre-nasalised stops. The data just cited show that this is not necessarily so.

1 *Double reflexes in north-western Bantu*

A34, Basaa (Mbene) A43a, Nen A44, Yambasa A62, Bulu A74a, and Mvumbo A81. Furthermore, his charts in Guthrie (1971: 32–33) mention three other languages: Maande A46, Fa’ A51 and Kwakum A91. However, only one alternant pair is given for each of those, whereas for the other languages cited, all the reconstructed voiceless stops exhibit two series of reflexes as can be seen from Table 2, drafted from Guthrie’s correspondence lists.

Table 2: Double reflexes in some zone A languages (cf. Guthrie 1967; 1971)

	*p/_V	*p/_VV	*t/_V	*t/_VV	*k/_V	*k/_VV	*b	*d	*g
Lundu A11	∅	p	l	t	∅	k	b	∅	k
Duala A24	w	p	l	t	∅	k	b	∅	k
Basaa A43a (C ₁)	h	p	∅	t	(h)	k	b	(l)	k
Bulu A74a (C ₁)	v, h	f	l	t	∅	k	b	y	k
Yambasa = Gunu A622	h	f	d	t	g	k	b	n	∅

Two points should be noted at this stage. First, in Table 2 only voiceless stops exhibit double reflexes, although *d in Basaa is mentioned as alternating between /l/ and ∅ (this is what Guthrie’s brackets mean). However, the account is not complete, even on Guthrie’s own terms. For Yambasa A622,⁵ a closer examination of the data shows that the three voiced proto-stops also exhibit double reflexes namely: *b > b / f, *d > l / n and *g > k / ∅. We return below to the situation in Mbam (to which Yambasa belongs) and show that for voiced proto-stops some conditioning factor can be detected, which also holds for double reflexes of *d in Basaa.

Second, as far as the labial and dorsal voiceless stops are concerned, the difference between the two sets can definitely be seen as one of “strength”, as discussed below. The reflex in front of *V is mostly ∅, or a glottal or a glide (in two cases a voiced fricative or stop), whereas the reflex in front of *VV is a voiceless stop or at most a voiceless fricative. The case of the coronal stop is rather different, however. It is not so obvious that the lateral should be considered as a “weak”

⁵Judging from Guthrie (1953), the Yambasa data come from his own field notes. ‘Yambasa’ is of course a cover term, but judging from the material appearing in *Comparative Bantu* (Guthrie 1967; 1970a,b; 1971), his source is probably Gunu A622.

form of a voiceless stop and furthermore there appears to be a possibility of overlapping with the reflexes of **d*, a subject which we discuss at greater length later on.

The main support for Guthrie’s hypothesis on the existence of double reflexes came from John Stewart. In several articles (Stewart 1973; 1975; Stewart & Van Leynseele 1979; Van Leynseele & Stewart 1980; Stewart 1983; 1989) he attempted to demonstrate that Proto-Bantu had two series of stops (voiced as well as voiceless), which he termed “lenis” and “fortis” respectively. The proposal emerged from his work on the reconstruction of Potou-Tano (aka Potou-Akanic or Greenberg’s Akan), a branch of the Kwa languages spoken in Ghana and Ivory Coast, comprised of the Lagoon languages Cama (Ebrié) and Mbatto on the one hand and the Akanic languages (Anyi-Baule, Ahanta, Fante, etc.) on the other. As the Potou languages retain a contrast between stops that Stewart analysed as “fortis” and “lenis” respectively (in Cama for both voiced and voiceless series, in Mbatto reduced to voiced stops only), he reconstructed those sounds for the group’s ancestor language, although the Akanic branch shows no evidence for them.

In his contribution to the *International Colloquium on the Bantu Expansion* held in April 1977 (published as Van Leynseele & Stewart 1980), Stewart seems to have put forward for the first time the hypothesis that the Bantu double reflexes correspond to the *fortis/lenis* contrast, which he had reconstructed for Volta-Congo, i.e. the most recent common ancestor of Kwa and Benue-Congo. Note that this included more double reflexes than Guthrie admitted. As we saw in Table 2, Guthrie did not posit double reflexes for reconstructed voiced stops, apart from the marginal case of Basaa. On the other hand, for Nen, the main focus of their contribution, Van Leynseele & Stewart (1980: 428) had the following Table 3 in which “lenis” stops are preceded by an apostrophe.

Table 3: Double reflexes in Nen (Van Leynseele & Stewart 1980: 428)

*'p	*p	*'t	*t	*'k	*k	*'b	*b	*'d	*d
h	f	l	t	∅	k	f	b	n	l

Present in their analysis was the notion that *fortis* and *lenis* consonants generally tended to harmonise at C₁ and C₂ positions (Van Leynseele & Stewart 1980). Stewart (1989) attempted to synthesise his position with Guthrie’s long vowel contrast, so that a long vowel tended to produce long (i.e. “fortis”) stops both preceding and following it.

The definition of “fortis” and “lenis” has been the object of considerable debate in phonology (perhaps particularly diachronic phonology) of which an enlight-

ening and very complete summary is to be found in Honeybone (2008). Without entering the discussion, it might be said that many authors would entertain the following approximative hierarchy from “strongest” or most “fortis” to “weakest” or most “lenis”: voiceless stops > voiced stops > (voiced or voiceless) fricatives > approximants > zero. Using labials as examples, the hierarchy would be $p > b > \beta > w > \emptyset$, or alternatively: $p > \phi > h > \emptyset$. This is of course a simplification and a more complete chart can be found in Hock (1991: 83).

Although airstream mechanisms are not often mentioned in such hierarchies, Stewart (1993) considers, on the basis of realisations in Cama, that the most probable phonetic definition of his two series was the following (for labials): voiceless “fortis” = [p^h] (aspirated voiceless plosive); voiceless “lenis” = [p] (voiceless plosive); voiced “fortis” = [b] (voiced plosive); voiced “lenis” = [ɓ] (voiced implosive). Hence, the hierarchy would be as follows: $p^h > p > b > \beta > w > \emptyset$.

At the same time, Stewart (1993) proposed that plain voiceless and aspirated series merged in PB to plain voiceless⁶, e.g. *p, whereas the implosive and plain voiced stops merged to implosives, e.g. *ɓ, thus in effect disposing of double reflexes in Bantu! This was due to the detailed criticism of Guthrie’s position in Janssens (1991), according to whom the distribution of double reflexes was in fact not conditioned by vowel length but mostly by the diachronic presence of a nasal prefix. Stewart (1993) still maintained the question of “consonantal harmony” with the source being now attributed to C₁, since this is where nasal prefixes could have produced “fortis” stops. A PB voiced C₂ consonant would then devoice if C₁ was “fortis” (i.e. diachronically pre-nasalised). Janssens (1991) was much more hesitant on this point.

In reaction to Stewart’s earlier proposal, several authors had expressed either their (partial) approbation (Nsuka-Nkutsi 1980; Hedinger 1987; Bancel 1988) or more decisively their opposition (Blanchon 1991; Janssens 1991; 1993). Summing up the latter’s arguments, one can posit three main objections to Stewart’s hypothesis:

- a. no consonant harmony at C₁ and C₂ can be statistically established (Janssens 1991; 1993);
- b. many languages exhibit variants for the same lexeme with both “fortis” and “lenis” reflexes (Blanchon 1991; Janssens 1991; 1993);

⁶Note that in later publications (e.g. Stewart 2002), while still retaining two series, he again regarded all the voiceless stops (in C₁ position) as implosives. Due to the importance of Stewart’s conceptions, we will discuss them at length in §3.

- c. there is no correspondence between languages: for the same root some languages have “fortis” for “lenis” and vice-versa (Janssens 1993).

After a careful examination of the evidence I reach the following assessment of the objections presented above:

- a. Objection a. is now uncontroversial and Guthrie’s attempt to link the putative harmony to vowel length is spurious, in spite of Stewart’s (1989) attempt to salvage it;
- b. Objection b. is certainly supported by a reasonable number of examples. However, the fact that each language can make use of and reorganise a phonemic opposition for its own needs (reanalysis, morphological leveling, etc.) does not necessarily invalidate the diachronic origin of this opposition;
- c. As for objection c., I argue in this chapter against Janssens (1993) that there is in fact a rather high degree of correspondence between the forms exhibited by different languages – amounting in some cases to striking identity.

2 Double reflexes in zone A: Synchronic variation

Let us now turn to the distribution of double reflexes in zone A languages. Here the Mbam languages stand out against the rest. For most languages of our area, PB voiced stops are not concerned by any duality of reflexes, apart from **d* in Kpa A53 and a couple of other languages. In Mbam, however, all PB stops exhibit some duality of reflexes, with the partial exception of PB **t* which is only affected in Nen – see also Appendix B for a list of reflexes of **t* and **p* in Nen vs. Maande. Nonetheless, for the PB voiced stops, this duality of reflexes can be shown to be largely conditioned by the tone of V_1 (since we are only concerned here with C_1 reflexes). The first mention of this tonal conditioning in Nen is to be found in Botne (1992b),⁷ which is an important contribution, but rather overshoots its target. It claims that a similar tonal conditioning also explains the double reflexes in voiceless stops, which is not supported by the evidence at my disposal. For the voiceless stops, different possible types of conditioning are examined here and I conclude that the evidence robustly confirms the validity of double reflexes only for the voiceless coronal stop **t*. I attempt to show that these cannot be traced to

⁷Teil-Dautrey (1991a) had already observed it in Basaa.

an opposition in PB, but developed during the course of the phonological evolution of certain sub-groups.⁸

2.1 Reflexes of voiced stops

As seen above, Stewart recognised that in Nen, the only Mbam language he dealt with, even voiced stops had an opposition between ‘fortis’ and ‘lenis’ and thus yielded different reflexes:⁹ **b > f* vs. **b > b*¹⁰ and **d > l* vs. **d > n*; **g* and **g* would have merged very early in PB and thus left no duality of reflexes (cf. Van Leynseele & Stewart 1980). He was not struck by the fact that the different reflexes were also largely correlated with a difference in the tone of the following vowel. In fact, he paid very little attention to tone as can be seen in several correspondences he proposed. If he had, he might have been put on the track by his own example set (11) in Stewart (1989), where he clearly set out that Akan /ɲ/, or /y/ in non-nasal contexts, corresponds to PB **d* followed by L, whereas Akan /d/ corresponds to PB **d* followed by H.

It is indeed the case that the duality of reflexes for voiced stops is in good part conditioned by the tone of the following vowel, as well established by Botne (1992b) for Nen. Since **g* is not involved, the situation must be evaluated for **b* and **d*. A very important difference must be noticed at the outset. **d* is affected throughout the Mbam languages; furthermore, the same situation obtains in Basaa and a couple of other north-western languages. On the other hand, **b* undergoes this tonally-conditioned split in part of the Mbam group only, and this fairly independently of internal sub-divisions. Tuki and Gunu do not seem affected, albeit Gunu is otherwise a fairly close congener of the Yambasa sub-group consisting of Yangben, Mmala, Elip and Baca. The other outlier of Yambasa, i.e. Mbule, would also seem not to be affected, but this is a very little-known language and the available data are meagre.

Since the situation appears to be due to some tonal conditioning, it does not concern double reflexes which by definition should not be conditioned. I will thus

⁸I will adopt for the synchronic data a broad phonetic transcription, following the IPA with the exception of <ɣ> instead of IPA [j], as is the usual practice for most Africanists.

⁹I will treat pre-nasalised stops as unit phonemes and deal with them only sporadically, since they do not exhibit any duality of reflexes. They are mostly “fortis” and only rarely subject to weakening. I am of course aware of the extensive discussion in general phonology about the phonemic status of such pre-nasalised sounds and will decline to enter it here. A good review is Downing (2005), among others. My decision to treat them all – including voiceless pre-nasalised stops – as units is a purely practical one since it makes the statement of correspondences much simpler.

¹⁰Stewart took pains to explain that he did not consider Nen reflexes as exhibiting a “fortis”/“lenis” distinction, but that it only applied to the proto-phonemes

not delve too deeply into this fascinating and puzzling situation here. I will only chart the reflexes of the various proto-phonemes in the languages concerned and exemplify the case of **d* in Basaa and Kpa, with a complement on Kwakum.

Note that in Nen and Maande the bilabial stop can be realised indifferently voiceless or voiced. There appears to be no social or regional conditioning since even individual realisations are in free variation. I always transcribe < *p* >. In Yambeta, I transcribe as pronounced, because the realisation is conditioned by context: [*p*] initially and finally, [*b*] intervocalically. See Table 4.

Table 4: Tonally-conditioned split of **b* and **d* in zone A languages

	* <i>b</i> + L	* <i>b</i> + H	* <i>d</i> + L	* <i>d</i> + H
Nen A44	<i>f</i>	<i>p</i>	<i>l</i>	<i>n</i>
Maande A46	<i>f</i>	<i>p</i>	<i>l</i>	<i>n</i>
Yambeta A462	∅	<i>p</i> ~ <i>b</i>	<i>l</i>	<i>n</i>
Tuki A601	<i>b</i>	<i>b</i>	<i>r</i>	<i>n</i>
Gunu A622	<i>b</i>	<i>b</i>	<i>l</i> ? ^a	<i>n</i>
Yangben A62A	∅	<i>p</i>	<i>l</i>	<i>n</i>
Mmala A62B	∅	<i>b</i>	<i>l</i>	<i>n</i>
Elip A62C	<i>h</i>	<i>b</i>	<i>l</i>	<i>n</i>
Baca A621	<i>f</i> ~ <i>h</i> ^b	<i>p</i>	<i>l</i>	<i>n</i>
Mbule A623	<i>p</i>	<i>p</i>	<i>l</i> ?	<i>n</i>
Basaa A43a	<i>b</i>	<i>b</i>	∅	<i>l</i>
Fa' A51	<i>b</i>	<i>b</i>	<i>l</i>	<i>d</i>
Kpa A53	<i>b</i>	<i>b</i>	<i>r</i>	<i>l</i> ~ <i>d</i> ^c
Kwakum A91	<i>b</i>	<i>b</i>	<i>l</i>	<i>d</i>

^aFew valid examples.

^bVery few examples: 1 case of /*h*/, 2 of /*f*/.

^c*d*/__*i* / *u*, *l* elsewhere.

As detailed example, Basaa and Kpa reflexes of **d* before H and L are listed in (1) and (2) respectively.¹¹

¹¹I give code numbers for both Guthrie's Comparative Series (C.S., cf. Guthrie 1970a,b) and Ter-
vuren's Bantu Lexical Reconstructions, version 3 (BLR, cf. Bastin et al. 2002).

- (1) Basaa A43a and Kpa A53 reflexes of **d* before H
- a. **dáád* ‘sleep’ (C.S. 455, BLR 795) > *lâl* (A43a), *lál* (A53)
 - b. **dém*¹² ‘be crippled’ (C.S. 531, BLR 914) > *lém* (A43a), *kì-lém* ‘lameness’ (A53)
 - c. **dók* ‘vomit’ (C.S. 695, BLR 1179) > *ló* (A43a), *lóó* (A53)
 - d. **dámb* ‘cook’ (C.S. 486, BLR 842) > *lám* (A43a), *lám* (A53)
 - e. **démà* ‘bat’ (C.S. 532, BLR 916) > *n-léém* (A43a), *kì-lém* (A53)
 - f. **dó* ‘sleep’ (C.S. 633, BLR 1080) > *hì-ló* (A43a), *fì-ló* (A53)
 - g. **dóbò* ‘fish-hook’ (C.S. 640, BLR 1093) > *n-lóp* (A43a), *fì.lóp* (A53)
- (2) Basaa A43a and Kpa A53 reflexes of **d* before L
- a. **dà(i)p* ‘be long’ (C.S. 504, BLR 784/873) > *àp* (A43a), *ràp* (A53)
 - b. **did* ‘cry’ (C.S. 561, BLR 959) > *èè* (A43a), *rèn* (A53)
 - c. **dòg* ‘bewitch’ (C.S. 644, BLR 1100) > *òk* ‘curse’ (A43a), *rò?* ‘poison’ (A53)
 - d. **ditò* ‘heavy’ (C.S. 631, BLR 1076) > *yèr / gwèr* ‘weight’ (A43a), *rì?* (A53)
 - e. **dògù* ‘wine, beer’ (C.S. 649, BLR 1108) > *màòk* (A43a), *màrò?* ‘palm wine’ (A53)
 - f. **dèdù* ‘beard’¹³ (C.S. 519, BLR 897) > *liy=èé* (A43a), *fìrēē* (A53)
 - g. °*dimbà* ‘witchcraft’¹⁴ > *lièmb* (A43a), *mλ.rèm* (A53)¹⁵

In (3) are the only two exceptions I found to the tonal conditioning illustrated in (1) and (2), interestingly attested in both languages¹⁶ with /l/ where Ø/r would be expected.

¹²Upon request by the editors, I adopt here the PB vowel notation system of BLR, i.e. /i, ɪ, e, a, o, ɔ, u/, for reasons of uniformity across the volume. Personally, I consider the /i, ɪ, e, a, ɔ, ɔ, u/ transcription preferable, because closer to the phonetic reality of many present-day Bantu languages.

¹³Although both Guthrie and BLR give a LL tone pattern for this stem, the Basaa and Kpa data indicate LH.

¹⁴Although not reconstructed by Guthrie nor BLR, this is a very widespread stem in north-western Bantu, found even in Mankon (Eastern Grassfields).

¹⁵If the meanings can be shown to fit, this is possibly another case: **dègid* ‘be slack’ (C.S. 523, BLR 902) > *y=ègèp* ‘be dejected’ (A43a), *rλ?* ‘soften’ (A53).

¹⁶The same exceptions are found in Mbam alongside several others, with /n/ instead of /l/.

- (3) Exception to tonal conditioning of **d* reflexes in Basaa A43a and Kpa A53
- a. **dòd* ‘be bitter’ (C.S. 684, BLR 1162) > *lòl* (A43a), *lòl-ɛn* (A53)
 - b. **dòŋgò* ‘kinship’ (C.S. 665, BLR 1135) > *lòŋ* ‘country’ (A43a), *kì-lòŋ* ‘village’ (A53)

Teil-Dautrey (1991b) also mentions **dà* ‘intestine’ (C.S. 442, BLR 773) > *n-là* (A43a) and **dàdà* ‘grandchild’ (ps. 145,¹⁷ BLR 798) > *n-làlà* (A43a). She suggests that the cl. 1 and 3 prefixes might explain the retention of /l/; neither stem is found in the available Kpa data. Teil-Dautrey (1991b) also has the exception *lèl* ‘rock baby’ (A43a) < **dèd* (C.S. 510-1, BLR 882), which is not attested in my Kpa database. The first two exceptions at least are also found in Mbam: Maande *nòò-nà* and *ò-nàna*.

The tonal conditioning in Basaa was first mentioned by Teil-Dautrey (1991a,b), but she did not refer to the Kpa correspondences. She points out that whereas the influence of the [+voice] feature in consonants on the emergence of L tone is well-known, we seem to be faced here with the reverse influence, i.e. the tone of the vowel determines the segmental realisation.

Note that in Basaa, the last stage (Ø) must be fairly recent, since an empty onset subsists as can be seen with the cl. 5/6 prefixes, for instance: *li-èmb* ‘witchcraft’, *mà-òk* ‘palm wine’. Conversely, the deletion of **k* must be ancient, since the result is always identical with vowel-initial stems,¹⁸ e.g. **kómi* ‘ten’ (C.S. 1208, BLR 2027): *ɔ̄-òm* / *m-òm* (A43a), and not ***li-óm* / *mà-óm*. In spite of the fact that the tonal conditioning does not qualify the results as double reflexes, the question of the reflexes of **d* will have to be considered further on, alongside those of **t* due to the partial overlap between them.

I must add one tantalising fact, which cannot be pursued further with the available data. Kwakum offers a handful of cases which might be related to what has just been discussed. In this language, the regular reflex of **d* is /d/ in front of [–high] vowels, as shown in (4).¹⁹

- (4) Kwakum A91 reflexes of **d* before [–high] vowels
- a. **dómè* ‘male, husband’ (C.S. 697, BLR 1182-3) > *n-dóm* / *à-dóm*
 - b. **dímè* ‘tongue’ (C.S. 571, BLR 971) > *dém*
 - c. **dó* ‘sleep’ (C.S. 633, BLR 1080) > *dó*

¹⁷The abbreviation “ps.” in Guthrie stands for “partial series”, not well-supported and more tentative.

¹⁸Meaning those written with initial **y* by Guthrie. Cf. Wills (2022 [this volume]) for discussion.

¹⁹It is /ɖ/ in front of [+high, –back] vowels; the only example of **d* in front of a [+high, +back] vowel is given in (5), i.e. **dùt* ‘pull’.

However, I found five examples where the reflex is /l/ and they are all followed by L tone, as shown in (5).

- (5) Kwakum A91 reflexes of *d before L
- a. *dògù ‘wine, beer’ (C.S. 649, BLR 1108) > ò-lòkù
 - b. *dùt ‘pull’ (C.S. 749, BLR 1267) > lùt-ò
 - c. *dìd ‘cry’ (C.S. 561, BLR 959) > lèn-ò
 - d. °dimbà ‘witchcraft’²⁰ > i-lèmbò
 - e. *dà(i)p ‘be long’ (C.S. 504, BLR 784/873) > làw-ààwè

The number of L stems beginning with *d is rather limited, but I found one exception and it is identical to one of those cited above for Basaa and Kpa (and Mbam), i.e. *dòd ‘be bitter’ (C.S. 684, BLR 1162) > dòl-ààwè.

I have no explanation to offer for this apparently shared evolution, but contact seems out of the question, since Kwakum is spoken far to the east. In spite of the very deficient information, it would seem that the closely related Seki B21 shares this characteristic with Kwakum A91. The reflex of *d is /d/ before H-tone vowels, but we also find /l/ in front of L-tone vowels, as illustrated in (6). The matter should be further investigated.

- (6) Seki B21 reflexes of *d before L
- a. *dìd ‘cry’ (C.S. 561, BLR 959) > lèl-ò
 - b. *dìb(ad) ‘forget’ (C.S. 556a, BLR 953) > lèb-idye (cf. Kwakum lèè-faa ?)
 - c. *dòg ‘bewitch’ (C.S. 644, BLR 1100) > lòkò (not attested in Kwakum)

Apart of course from the Mbam languages mentioned above in the case of *b, *b and *g do not exhibit this tonal conditioning. For one, as established by Teil-Dautrey (2004), *g at C₁ is practically always followed by a L-toned syllable. For instance, in Guthrie’s Common Bantu list with more than 170 stems with C₁ *g, only 30 appear with a H-tone first syllable. Of those, six are likely to be vowel-initial stems where the *g appears as an artefact of Guthrie’s method (cf. Wills (2022 [this volume])); seven are “osculant” (cf. Bostoen 2001; Ricquier & Bostoen 2008; Bostoen & Bastin 2016) with an initial *k as alternative (and one with *b). This would leave us with a bare dozen, hardly 10% of the total with *g + H.

For *b, Teil-Dautrey (2004: 153–155) finds that for verbal roots there are twice as many reconstructions where the voiced bilabial C₁ is followed by a H than by

²⁰See footnote 14.

a L tone. She then attributes this imbalance to the fact that **b* was probably an implosive [ɓ]²¹ whose affinity for H tone is well-known. Indeed, many languages in the north-west have a [ɓ] realisation for **b*, even if it appears in complementary distribution with [b] in some languages, for instance Duala where **b* > [ɓ], except for **b/_i, u* > [b]. There are thus no traces of unconditioned double reflexes here.

2.2 Reflexes of voiceless stops

Turning now to the reflexes of voiceless stops, we see here a rather different situation. As the best case for double reflexes can be made for **t*, we examine it first.

Apart from Nen A44,²² all Mbam languages as well as Bubi A31 and the Kwakum group, i.e. Kwakum A91, Seki B21 and Kako A93,²³ regularly have /t/ as the reflex of **t* in C₁ position. In part of the Yambasa A62 group, the reflex is voiced /d/. Since those languages either have no voice contrast for the stops, or else only voiced stops in reflexes of inherited vocabulary, I hold the voicing to be secondary. Selected examples of reflexes of **t* are given in (7). As (7d) illustrates, Northern Bubi and Kwakum manifest a tendency for palatalisation in front of the close front vowel **i*.

- (7) Reflexes of **t* in the Mbam, Bubi, Kwakum and Yambasa groups
- a. **tímà* ‘heart’ (C.S. 1738, BLR 2895) > *ò-tím* (A62A), *ò-dím* (A62C), *bò-tébá* (A31a), *mò-témá* (A31c), *témò* (A93, B21) (with Ø-prefix of cl. 3 NP)
 - b. **tóm* ‘send’ (C.S. 1831, BLR 3055) > *tòm* (A62A), *dóm* (A62C), *tòbá* (A31a), *tóm-à* (A31b), *tòm* (A91), *tom-u* (A93) [tones uncertain]
 - c. **támbò* ‘trap’ (C.S. 1661, BLR 2766) > *i-dám* (A462), *i-támbú* (A601), *bò-tápô* ‘fish-trap sp.’ (A31a), *i-táá’mbó* (A91)
 - d. **tíg* ‘leave’ (C.S. 1746, BLR 2910): *fi?-à* (A31a), *fi’k-ó* (A91), cf. *tsík-ò* (B21)

²¹Grimm (2019) queries the existence of genuine implosives in some of the north-western Bantu languages and considers the sounds as pre-glottalised explosives instead. While her reasoning is quite sound and she provides good instrumental evidence to support her point, Greenberg’s (1970) conclusion, i.e. that there is no contrast between implosives and pre-glottalised voiced consonants in any language described, still stands. I will just stick to the traditional definition of those sounds as implosives here.

²²See Appendix B for a list of reflexes of **t* and **p* in Nen vs. Maande.

²³Recall that I put Polri A92a, and tentatively its close relative Pomo A92b, in the Nyong-Dja group with the A80 languages.

The other languages show two distinct reflexes for C_1 **t*, either a strong /t/ or a weak lateral/zero, partly with clear conditioning. In front of the highest vowels **i* and **u*, the normal reflex is strong /t/ in Bafia, Sawabantu, Manenguba, Basaa, Beti and Nyong-Dja, as shown in (8), with partial exceptions in Beti (group 7, cf. §1.1) and Nyong-Dja (group 8, cf. §1.1), which are discussed below. A couple of exceptions should be noted, which appear with the weak reflexes before **i* and **u*. We also discuss them later on.

- (8) Strong reflexes of **t* in north-western languages other than in (7)
- a. **túúb* ‘pierce’ (C.S. 1860, BLR 3100) > *túβá* (A11), *túbà* (A24), *tú* (A13), *túp* (A15C, A53, A72a), *tóp* (A43a), *túw* (A63), *túbò* (A832)
 - b. **túúdi* ‘shoulder’ (C.S. 1862, BLR 3103, 3987) > *è-túri* (A11), *è-túli* (A22), *è-tú* (A141), *è-tút* (A43c), *è-túù* (A75A)
 - c. °*tìd* ‘write’²⁴ > *tìl-à* (A32C), *tìl* (A44a, A13), *tèl* (A15C), *tìlè* (A842)
 - d. **tínd(ik)*/**tíínd(ik)* ‘push’ (C.S. 1758, BLR 2933-4) > *tíndè* (A11, A43a), *tíndiy* (A33a), *tíi* (A15B), *tínd* (A63), *tín-lò* (A92a)

The southern Nyong-Dja languages, i.e. Gyeli A801, Shiwa A803 and Kwasio A81, are affected by BF, which produces affricates in front of high vowels. Since this affects all stops it is better left for a special treatment. The same applies to the southern Beti varieties, e.g. Atsi A75D. On the other hand, all other Beti languages (except A63?) have **t*/___**i* > *tʃ* ~ *ts*, e.g. **tíitó* ‘animal, meat’ (C.S. 1767, BLR 2952) > *tsít* (A75A), *tsít* (A72a), *tít* (A63). Similarly, **d*/___**i* > *ɟ* ~ *ɟʃ* in the same languages, but there is no affrication before **u*. So, the process is probably not to be seen as an instance of BF but rather of palatalisation, followed by a fronting to [+ant], a rather frequent phenomenon universally.

Other than in front of **i* and **u*, the normal reflex of C_1 **t* is not strong /t/, but a variety of weaker reflexes, including, the weakest of all, i.e. Ø, as shown in (9). The most widespread reflex is a lateral; two Sawabantu languages, i.e. Kpe A22 and Bubia A221, have **t* > **l* > Ø.

- (9) Weak reflexes of **t* in north-western languages other than in (7)
- a. **tóm* ‘send’ (C.S. 1831, BLR 3055) > *lóm-à* (A11, A24, A25), *óm-à* (A22), *lóm* (A13, A15C), *lòm* (A75A, A72a), *lúm-è* (A801)²⁵, *lòm-ɔ* (A92a)²⁶

²⁴ °*tìd* ‘write’ is not reconstructed but widespread in the area.

²⁵ Although synchronically 7V languages, the southern Nyong-Dja have mostly merged **ɔ* with **u* and **ɪ* with **i*, which parallels the development of BF.

²⁶ Polri A92a tones are as given by Wéga Simeu (2016).

- b. **tónd* ‘be full’ (C.S. 1840, BLR 3067) > *lón-d-à* (A11, A24, A25), *ón-d-à* (A22), *lón* (A13, A15C), *lód-àl* (A85b), *lón-d-ólò* (A86c), *lúnd-à* (A801)
- c. **téndé* ‘palm tree’ (C.S. 1712, BLR 2849) > *léndé* (A24, A25), *là-lénd* (A63), *à-lón* (A75A), *è-lén* (A842), *lè-léndé* (A801)
- d. **táng* ‘read, count’ (C.S. 1672, BLR 2786) > *láng-à* (A11, A32C, A24), *láy-à* (A25), *láj* (A13, A15C, A63), *lá-à* (A85b, A842), *làng-lò* (A92a)
- e. **tímà* ‘heart’ (C.S. 1738, BLR 2895) > *mò-lémà* (A11, A33a), *ɣm-émà* (*ɣm* < *mw-*) (A22), *n-lémà* (A25), *n-lém* (A13, A15C, A75F), *líma* (A803)
- f. **tóng* ‘crow (rooster), sing, whistle, etc.’ (C.S. 1793, BLR 2994) > *lóng-ò* (A32C, A24, A25), *lój* (A15C, A71), *lwanj* (A81)²⁷

However, the Basaa group (A41-3) and Kpa A53 have the weak reflex of **t* as Ø and /r/ respectively, which is identical to the reflex of **d* before a L tone vowel (cf. §2.1). Note, though, that Fa’ A51, which is very closely related to Kpa, has /l/ as reflex of **t*. Not all Basaa stems are attested in Kpa. See the examples in (10).

(10) Weak reflexes of **t* in the Basaa group (A41-3) and Kpa A53

- a. **tóm* ‘send’ (C.S. 1831, BLR 3055) > *óm* (A41, A43a), *róm* (A53)
- b. **tímà* ‘heart’ (C.S. 1738, BLR 2895) > *ɲ-ém* (A43a) *n.dém*²⁸ / *mÀ-rém* (A53)
- c. **támbò* ‘trap’ (C.S. 1661, BLR 2766) > *ò-ám* (A43b), *fí-rám* (A53)
- d. **tóng* ‘crow (rooster), sing, whistle, etc.’ (C.S. 1793, BLR 2994) > *ój* (A43a)
- e. **táng* ‘read, count’ (C.S. 1672, BLR 2786) > *áj* (A43a)

Table 5 summarises the partially overlapping correspondences for **d* and **t*. Note that I have adopted a conservative position in considering that the C₁ reflex of **d* in Manenguba and Beti is a palatal or palato-alveolar. In other languages, where similar sounds appear they can be shown to be (originally!) epenthetic onset-fillers. It is probable that we might have the same situation in Manenguba and Beti, but a detailed examination of the problem would require a chapter of its own.

Nevertheless, most striking is a third group, where the reflex of **t* is strong /t/ without any apparent conditioning. These items are not extremely numerous, as

²⁷No tones are available for Kwasio A81.

²⁸The reflex of “lenis” **t* after syllabic nasal is /d/. The normal /r/ reflex is visible in the plural.

Table 5: Weak reflexes of C₁ **d* and **t* in some north-western languages

	* <i>d</i>	* <i>t</i>
Nen A44	<i>l</i>	<i>l</i>
Maande A46	<i>l</i>	<i>t</i>
Elip A62C	<i>l</i>	<i>d</i>
Lundu A11	∅	<i>r</i>
Duala A24	∅	<i>l</i>
Kpe A22	∅	∅
Oli A25	∅	<i>l</i>
Bubi A31	<i>l</i>	<i>t</i>
Balong A13	<i>y</i>	<i>l</i>
Akoose A15C	<i>f</i>	<i>l</i>
Mkaa A15C	<i>ɕ</i>	<i>l</i>
Basaa A43a	<i>l</i> ~ ∅	∅
Fa' A51	<i>d</i>	<i>l</i>
Kpa A53	<i>l</i> ~ <i>r</i>	<i>r</i>
Ewondo A72a	<i>y</i>	<i>l</i>
Gyeli A801	∅	<i>l</i>
Kwakum A91	<i>d</i> (~ <i>l</i>)	<i>t</i>

shown in (11), but they are quite consistent between groups. In (11), I also mention the few deviations.²⁹

(11) Unexpected strong reflexes of **t* without any conditioning

- a. **táánò/ò* 'five' (C.S. 1662, BLR 2768 & 2769) > *tâ* (A11), *tá* (A122), *tánù* (A24), *táà* (A22), *táàn* (A15C, A53), *tánò* (A32C, A33a), *tân* (A43a, A75A), *tán* (A71), *tên* (A84), *tánè* (A801); only Nen has /l/, i.e. *lánò* (compare **tátò* 'three' for which all the languages mentioned have the weak reflex)
- b. **tòòg* 'boil up, bubble up' (C.S. 1777, BLR 2966-7) > *tòk-ó* (A13), *tò* (A24, A25), *tó* (A22), *tòk-ò* (A32C, A33a), *tòk* (A15C, A75F), *twà?* (A842), *tòg-i* (A86c)

²⁹Nen, the only Mbam language to have double reflexes of **t*, sometimes does not coincide with the other languages. This plus the fact that the phenomenon is absent from its close relative Maande raises the question of the origin of the split in Nen, which might be a recent innovation.

- c. **tóóbá* ‘six’ (C.S. 1815, BLR 3034) > *mù-tóbá* (A24), *mò-tóbá* (A22), *ñ-tóbá* (A25, A32C), *ñ-tóóp* (A15C), *tóbó* (A842), *n-tùú* (A801)
- d. **tédam* ‘stand’ (C.S. 1692½, BLR 2816) > *tíním* (A44), *té(mè)* (A22, A24, A25, A32C, A33a), *tyéè-m* (Akoose A15C), *téé-bé* (Mkaa A15C), *tél-è* (A43c), *tél-êp* (A42a), *tél-ê* (A75A), *tél-i* and intr. *té-bê* (A71), *tál-i* (A803), but note *rél-ì* (A53) with the weak reflex
- e. **tóná* ‘spot, speckle’ (C.S. 1785, BLR 2976) > *dì-tónó* (A11), *tón/mà-tón* (A24), *à-tón* (A15C, A72a), *lì-tón* (A43a), *à-twán* ‘pimple’ (A75A), *è-tón* (A85b), *à-tōnī* (A86c)

Other items exhibiting the same correspondence are less well represented, not because of contradictory data, but because they happen not to be present in all groups, as shown in (12).

(12) Unexpected strong reflexes of **t* without any conditioning

- a. **tónd*/*tóond* ‘desire’ (C.S. 1788, BLR 2980) > *tóndò* (A24, A25, A32C, A33a), *tóndá* ‘worship’ (A43a) [Sawabantu and Basaa]
- b. **támbí*/*táámbí* ‘sole of foot, shoe’ (C.S. 1659, BLR 2761) > *è-támbí* (A12, A24), *ì-támbí* (A22), *támbí* (A25), *à-támbé* (A15C), *támb* (A43a) [Sawabantu, Manenguba and Basaa]
- c. **tòdú*³⁰ ‘navel’ (C.S. 1776, BLR 2965) > *mù-tòdì* (A24), *ñ-tòlì* (A25), *ì-tódù* (A32C), *twôl / mò-* (A832), *twáli ~ twéli* (A803) [Sawabantu and Nyong-Dja]
- d. **tèk* ‘become soft’ (ps. 434, BLR 2827) > *tà?* (A75A), *tèk’* (A63) *tyè?* (A842), *tàk* (A83) *tia’* (A81), but note *rà?* (A53) with the weak reflex [Beti and Nyong-Dja]
- e. **tàndá* ‘invertebrate: spider; spider’s web’ (BLR 9730)³¹ > *è-tàndà* (A122), *è-tàndá* (A24, A22, Mkaa A15C), *tàndá* (A25), *è-tàndó* ‘insect sp.’ (A43C), *tàndí* ‘grasshopper’ (A43a), *ì-tàndág* (A63), *ñ-tàntà / bì-n-tàntà* ‘grasshopper’ (A803), but *è-làndàni* (A32C) and *è-làndì* (A33a) with the weak reflex [Sawabantu, Basaa and Nyong-Dja]

³⁰The tones reconstructed by Guthrie are suspect: they are only supported by the Ngiri C30 languages and Mongo C61. The Abo A42 form given by Guthrie is not cognate but is a reflex of **kóbú* (C.S. 1098, BLR 1865). The tone patterns of zone A languages point to LL or HL. Note also the V₂ differences.

³¹**tàndá* is not reconstructed by Guthrie.

- f. **tààtá* ‘father’ (C.S. 1686, BLR 2806) > *tátà* (A11), *tàtá* (A122), *tàà* (Mkaa A15C), *tàtá* (A43b), *tàtâ* (A43a), *tààtá* (A53), *tàdá* ~ *tàrá* (A71), *tá* (A801)³²

As shown in (13), I found only one clear example where the reflexes differ sharply among groups, i.e. **tóí* ‘ear’.

- (13) Mixed strong/weak reflexes of **t* in **tóí* ‘ear’ (C.S. 1813 etc., BLR 3030)
- a. Strong reflex in Sawabantu and Manenguba: *tóì/mà-tóì* (A24) *li-tóò* (A22), *ì-tóì* (A32C), *ì-tô* (A13), *è-túù* (A15C)
 - b. Weak reflex elsewhere: *mù-lwá* (A44), *óó* (A43a), *ì-réè* (A53), *à-ló* (A75A, A92a), *lè-ló* (A801)

It should be noted that Kpa stands out among our languages, as illustrated in (14). It is the only language that has the strong reflex in a number of other stems. It also does not share many of the previous instances, but this might be due to gaps in the lexical documentation.

- (14) Strong reflexes of **t* in Kpa A53 vs. weak reflexes elsewhere
- a. **téndé* ‘palm tree’ (C.S. 1712, BLR 2849) > *rì-tén* (A53) vs. *li-én* (A43c), see also (9)
 - b. **tóng*/**tóng* ‘build’ (C.S. 1848, BLR 3081) > *tón* (A53) vs. *ì-lóngà* (A32C), *ón* (A43c), *lôn* (A75A), *lòò* (A85b)
 - c. **tòé* ‘head’ (C.S. 1800, BLR 3007, 3023) > *n-tó* (A53) vs. *mò-ló* (A44), *mò-ló* (A32C), *n-ló* (A15C), *n-ló* (A75A), *ò-ó* (A43a), *lò* (A86c)
 - d. **tí* ‘tree’ (C.S. 1729, BLR 2881) > *kì-té* (A53) vs. *pò-lí-á* (A44), *ò-ré* (A122), *è-é* / *bì-é* (A43c), *è-lé* (A75A), *lé* (A85b)

The reverse situation remains to be considered: weak reflexes in front of [+high] vowels, as exemplified in (15). Here also, the items tend to be the same across groups, although examples are less numerous and often affect some of the groups only, the lexical items in question being unattested in the others.

- (15) Weak reflexes of **t* in front of [+high] vowels
- a. **tíg* ‘leave’ (C.S. 1746, BLR 2910) > *y=ék* (A43a),³³ *lík* (A71, A63, A85b), *lí?* (A842)

³²I have added this cognate set, even if I am wary of correspondences in nursery words.

³³Note that the second-degree vowel in the stem is due to a regular ablaut process (see Hyman 2003a).

- b. **túd* ‘hammer, forge’ (C.S. 1861, BLR 3101) > *lún* (A44), *lúl-è* (A24), *lúù* (A15C), *óó* ‘make’ (A43a), *lwî* (A72a), *lûl* (A842)
- c. **túkò* ‘night’ (C.S. 1864, BLR 3105) > *pù-lwá* (A44), *è-lúù* (A33a), *ú / ma-ú* (A43a), *ì-rú* (A53), *a-lú* (A63, A71, A75A), *lè-lú* (A84)
- d. **túm* ‘stab’ (C.S. 1865 & 1866, BLR 3108) > *lúm* ‘hit with missile’ (A44), *lúmà* ‘sew’ (A21) (< C.S. 1865, BLR 3107 **túm* ‘sew, plait’), *lúm* (A72a), *lúm* (A75F)
- e. °*tí* ‘clear forest’³⁴ > *é*, reflexive *í-βâ* (A43a), *lí* (A71, A75D), *lî* (A72a), *lyâ* (A803), *líyâ* (A801)

The two examples in (16) are the only ones with mixed reflexes among the group. One of them is **tá* ‘war’, whose frequent reanalysis as °*yitá* has sometimes led to the consonant being placed in C₂ position, which is not treated in this chapter. The other one is **tím* ‘dig’, which is absent from Sawabantu and most of Manenguba.

(16) Mixed reflexes of **t* in **tá* ‘war’ and **tím* ‘dig’

- a. **tá* ‘war’ (C.S. 1630, BLR 2704, 9206)
 - Weak in Nen and Sawabantu: *pì-lá* (A44), *bì-lá* (A24), *bì-lá* (A25); as C₂: *b-ól* (A15A), *gw-ěr* (A43a), *w-ēl* (A53)
 - Strong in Beti: *bì-tá* (A72a), *wì-tá* (A63)
 - (Not attested in Nyong-Dja)
- b. **tím* ‘dig’ (C.S. 1752, BLR 2918)
 - Strong in Nen, Basaa and Kpa: *tímà* (A44), *tém* (A43a), *tím* (A53)
 - Weak in Akoose and Nyong-Dja: *lím* (A15C), *lím-à* (A832), *lúm-à* (A83), *à-līm-à* (A86c)
 - (Not attested in Sawabantu and most of Manenguba)

The other two PB voiceless stops offer also some duality of reflexes, but in either a more clearly conditioned or else more haphazard way. We consider **p* first, which has mostly a weak reflex in our area, including Mbam.³⁵ As shown in (17), the degree of weakening is quite varied, ranging from /f/ to Ø, rather independently of genetic groups, which would tend to indicate that the weakening is somewhat recent.

³⁴ °*tí* ‘clear forest’ is a regional stem reconstructed by neither Guthrie nor BLR.

³⁵ It is difficult to find Mbam cognates due to the fact that the lexicon of Mbam is rather different from the other languages.

(17) Supposedly weak reflexes of *p

- a. *pínd ‘(be) black’ (C.S. 1555, BLR 2577) > *índ-à* (A11, A22), *wínd-à* (A24), *víndà* (A33a), *fín* (A141, A53), *hín* (A15C), *hénd* (A43a), *vín* (A71), *vínd* (A63), *wínd-áá* (A86c), *yínd-ò* (A93)
- b. *píná ‘pus’ (C.S. 1553, BLR 2574) / *pídá³⁶ (C.S. 1547, BLR 2565) > *lò-wíná* (A24), *mà-víná* (A32C, A33a), *ò-hín* (A15B), *dì-hên* (A43a), *à-vín* (A72a, A75A), *wínó* (A86c), *è-Ƶ=ínó* (A832),³⁷ *dì-vínó* (B21), *ro-ia* (A11),³⁸ *è-hilá* (A31a), *è-sílá ~ è-hílá* (A31b), *è-víl* (A63, A71), *filó* (A91)
- c. *pémb ‘blow nose’ (C.S. 1471, BLR 2440) > *bí-fím* (A44), *wémb-è* (A24), *émb-é* (A22), *è-vémb-è* (A33a), *hém* ‘blow’ (?) (A43a), *à-w^yémb-ò* (A86c), *f^yémb-làà* (A91)

(18a–18d) have as the strong reflex /p/³⁹ instead of one of the weak reflexes of *p in (17). (18c–18d) are found in part of the area only. (18e) has an even more restricted distribution and manifests a mix of strong and weak reflexes of *p.

(18) Supposedly strong reflexes of *p

- a. *pàpá ‘wing’ (C.S. 1447, BLR 2410) > *ki-pàpós* (A62A), *dì-ƆàƆé* (A11, A22), *è-ƆàƆá* (A122), *pàpá* (A25), *lò-pàpá* (A31), *à-pàp* (A15C), *lì-pàƆáƶ* (A43a),⁴⁰ *ε-păp* (A71), *ā-fāp* (A75A), *pàbá* (A803), *pàpò* (A93)
- b. *pínd ‘plait’ (C.S. 1524, BLR 2523) > *pèndà* (A24), *pèn* (A15C), *pèn(d)* (A42), *pèn* (A71), *p^yèn* (A84), *píndà* (A803)
- c. *pùùpà ‘wind’ (ps. 420, BLR 2691) > *m-pùpé* (A24), *m-pùpê* (A22), *è-pù?* (A15A), *pùp* (A85b), *è-pùbò* and *pùb-lò* ‘to blow’ (A842), *pƆùƆ-èlè* ‘blow’ (A801), *ki-pùp-ùl* (A91) [not in Basaa, Kpa or Beti]
- d. °pùmá ‘fruit’⁴¹ > *è-Ɔùmá* (A11, A12, A22), *è-pùmá* (A24, A32C), *è-pùm* (A15C), *pùmá* ‘orange’ (A43a),⁴² but *è-(h)mámá* (perhaps [-mámá]?) in A31a (which would appear to be weak: *p > h) [not in Kpa, Beti, Nyong-Dja or Kwakum]

³⁶Meeussen (1976) quite rightly corrected Guthrie’s HL tone pattern (cf. Guthrie 1971: 153) to HH.

³⁷The /Ƶ/ is not a reflex of *p but an (originally) epenthetic onset-filler. I cannot develop this important point here, but I demarcate those onset-fillers by the equal (=) sign (see also Wills (2022 [this volume])).

³⁸Tones are not given in the source.

³⁹/Ɔ/ in those Sawabantu languages where the weak reflex is Ø. Most Beti languages have /f/ in these items, but this is a recent development, since the northernmost lects Eton A71 and Njowi A63 do have /p/ and the weakening even applies to /p/ < *mp.

⁴⁰[β] is the reflex of *p in C₂ position intervocalically.

⁴¹Not reconstructed by Guthrie nor BLR but obviously related to *bùmá (C.S. 228, BLR 374).

⁴²This Basaa term for ‘orange’ is quite possibly a loan.

- e. **pèèp* ‘blow (as wind)’ (C.S. 1489, BLR 2469) and **pèèpè* ‘wind’ (C.S. 1491, BLR 2476)
- Strong: *pìp* ‘fan’ (A44), *pìpà* ‘winnow’ (A601), *dì-φèφè* ‘wind’ (A122), *pèp* (A13), *è-pàp* ‘wind’ (A15A), *fəp* (A72a)
 - Weak: *fép* (A53), which however is properly a reflex of **pép* ‘blow (as wind); fly; winnow’ (C.S. 1487, BLR 2463), cf. *fəbà* ‘blow with mouth’ (A15A) and *vàbà* ‘breathe’ (A75D), see also *pìpà* ‘winnow’ (A601).

Regarding the items in (18), there is some partial conditioning, in so far as three of these five items have **p* at both C_1 and C_2 . Furthermore, they have a meaning linked with air movement for which some ideophonic origin can at least be suspected (cf. the consonant clusters *bl*, *fl*...at the beginning of the English translations). There are about half a dozen more comparative series found in various parts of the Bantu domain with the same **pVp* structure and referring to the same semantic field. In other words, the forms in (18) are at least partly motivated semantically and can therefore not be taken as pure instances of double reflexes. Finally, alternative (i.e. “osculant”) reconstructions exists for several items, as shown in (19).

- (19) Alternative reconstructions for items in (18) involving **b*
- a. **bàbá* ‘wing’ (C.S. 6, BLR 11)⁴³ > *è-bǎp* (A85b), *lè-mpʷàb* (A832)⁴⁴
 - b. **bìnd* ‘plait’ (C.S. 126, BLR 206) > *mò-bèndà* (n.) (A22), *m-bèndà* (n.) (A25)⁴⁵
 - c. **bùmá* ‘fruit’ (C.S. 228, BLR 374) > *è-bùmá* (A72a, A75A), *bǔm* (A85a), *bvàmá* (A803) [i.e. in Beti and Nyong-Dja]

In neighbouring and sometimes closely related languages, the three items in (19) exhibit either a strong reflex of **p* or a regular reflex of **b*. We seem to see here some overlap of [+voiced] and [–voiced] stops. I conclude that nothing more can be asserted at this point and that there is no convincing evidence for double reflexes of C_1 **p*.

The case of **k* turns out to be fairly straightforward. In spite of Guthrie’s claim (see Table 2 earlier in this chapter), there seems to be no valid evidence for a

⁴³This root is attested in the northern Nyong-Dja languages, while the Southern ones have reflexes of **p* as seen in (18).

⁴⁴/b/ and /mp/ are the regular reflexes of **mb* in Bekwel A85b and Kol A832 respectively (Cheucle 2014).

⁴⁵These few reflexes in Sawabantu can only reflect **b*.

contrast between weak and strong reflexes of *k. Except in Mbam, where the situation is more diverse (cf. infra), the general reflex of *k is Ø as the examples in (20) demonstrate.⁴⁶ As (20f) shows, Basaa has *h* as an onset-filler in a couple of stems (otherwise /h/ < *p).

(20) Reflexes of *k

- a. *kákà ‘pangolin’ (C.S. 991, BLR 1684) > *i-ǭ=á* (A12, A22), *fè-á* (A13), *wù-y=á?* (A141); other languages have this item in cl. 9 where C₁ **ɲk* > *k*
- b. *kádà ‘charcoal’ (C.S. 980, BLR 1662) > *m-ǎà* (A24), *m-â-g* (A71),⁴⁷ *d-áá* (A85b), *lè-gy=â* (A801), *i-ǭ=áá’lò* (A91), *dy-álà-kò* (B21)
- c. *kómì ‘ten’ (C.S. 1208, BLR 2027) > *d-óm* (A24), *rì-y=ómè* (A22), *ǭ-óm* (A33a), *dy-ôm* (A141, A15C), *ǭ-òm* (A43a), *à-w=ôm* (A72a, A75A), *rè-w=úmò* (A801), *dy-óómù* (B21)
- d. *kútà ‘oil’ (C.S. 1278, BLR 2138) > *m-ùlà* (A24, A25), *bù-útá* (A31a), *mù-útá* (A31b, A31c), *m-ól* (A15C), *m-òó* (A43a), *mà-w=ù(l)* (A72a), *m-ùl* (A85b), *ɲ-g=ú’tó* (A91),⁴⁸ *m-útò* (B21), *m-útò* (A93)
- e. *káng ‘fry, roast’ (C.S. 1009, BLR 1718) > *ángá* (A122), *ángà* (A24), *áyà* (A25), *y=án* (A13, A15C, A63, A71),⁴⁹ *w=án* (A43a), *án* (A43c), *y=án* (A53), *ɲ=án* (A832), *gy=âl-ε* (A801), *ǭ=áá* (A91)
- f. *kómb ‘scrape’ (C.S. 1134, BLR 1916) > *ómbò* (A11, A24), *w=óm* (A15C), *h=óm* (A43a)

The case of the Mbam languages is more diverse and puzzling. The Western Mbam languages and Tuki have the Ø reflex. However, Yambeta, Yambasa and Gunu, alone among all zone A languages, have kept the strong reflex /k/, in some languages as voiced /g/ either contextually or across the board. This parallels the fact that those languages (but also Tuki) have Ø as the normal reflex of *g,⁵⁰ whereas the latter has shifted to /k/ in Western Mbam as in the rest of zone A.

⁴⁶Recall that the equal (=) sign separates onset-fillers (cf. footnote 37).

⁴⁷Guthrie (1970a: 259) proposes a deviant source *kágà, but consideration of the Seki B21 (and Western Mbam) forms rather suggests *kádà to which -aga is suffixed, i.e. *kádàgà ‘charcoal’ BLR 2335.

⁴⁸In Kwakum, /g/ is the normal onset-filler before stem-initial back vowels.

⁴⁹Many Fang dialects have a strong form *kán* (Medjo Mvé 1997), at least as a variant. This is not the case in the most northern lects (Njowi, Eton, Ewondo or Bulu), perhaps due to contact. Galley (1964) has *y=án* and *k’yén*, both ‘make roast’.

⁵⁰Ki has thus merged the C₁ reflexes of *k and *g to Ø, an evolution it shares with some of the Sawabantu languages, e.g. Duala.

The number of putative strong reflexes of *k is extremely small, much smaller than for *t or even *p. To give an example, Table 6 shows the proportions I found for Duala excluding pre-nasalised stops.

Table 6: Proportions of ‘lenis’ and ‘fortis’ reflexes in Duala A24

Putative reflexes	*p	*t	*k
“lenis”	17 (65%)	18 (53%)	17 (81%)
“fortis”	9 (35%)	16 (47%)	4 (19%)
Total	26	34	21

Furthermore, a quick look at the Duala data indicates that for *k the cognacy with reconstructed items is doubtful at best. For three of the four items the tones do not fit: Duala *kòl* ‘be large’ (cf. **kód* ‘grow up’, C.S. 1190, BLR 1197); *kés* ‘cut’ (cf. **kèc* ‘cut’, C.S. 1028, BLR 1752); *kwàt* ‘scrape’ (cf. **kóát* ‘seize’ [!], C.S. 1172, BLR 1974). In fact, only one stem is attested widely enough to give rise to some interrogation, i.e. **kòògú*, **kòògó* ‘sugar-cane’ (C.S. 1201, BLR 2017-8). Interestingly enough, the only normal (weak) reflex is found in idiosyncratic Bubi, i.e. *b-o?ó* (A31a), *m-o?ó* (A31b, A31c), where *k > Ø and *g > ? are perfectly regular. Other languages exhibit a strong reflex, but generally also some other unexpected peculiarity (H tone on the NP, change of final vowel etc.). The number of irregularities leads one to suspect numerous borrowings for this culture item. In view of the lack of convincing examples, it may thus be safely concluded that genuine double reflexes for *k are non-existent.

3 Double reflexes in zone A: Diachronic evolution

Having surveyed the putative double reflexes for reconstructed voiceless stops, we have concluded that they do not affect the voiceless velar at all, and can be shown to be partly motivated for the voiceless labial. There remains the coronal. As for the voiced stops, we have convincingly established that double reflexes for *d are in fact conditioned by the tone of the first stem vowel. This seems also to be the case for *b, even if it is restricted to a few languages of the Mbam group and might be of no considerable antiquity. As for *g, we concur with Stewart (1989; 1993) that no trace of a dual development can be evidenced in the north-western languages we have examined. On the other hand, there does seem to be two reflexes of *t in some languages, namely the unconditioned reflex /l/ as well

as /t/ most often found in front of reconstructed [+high] vowels and in a small number of stems with no determinable conditioning factor.

We follow Stewart (1993) in admitting that PB had a voiced coronal phoneme *d (or perhaps better *l?) with two conditioned allophones, i.e. *d/___V[+high], and *l elsewhere (see also Hyman 2019: 142). Note that once the two allophones were established, they tended to evolve into genuine contrastive phonemes, among other things due to loans. For instance, a quick glance at the small Noho A32a vocabulary of Adams (1907) shows nine verb stems with /d/, i.e. [d], in front of [–high] vowels, where only /l/ is expected, versus 14 with regular /l/, contrasting e.g. *dàngwa* ‘travel’ with *langwa* ‘say’ (tones not noted).

There is thus some partial overlap between the weak reflexes of *t and *d in front of [–high] vowels, i.e. /l/, whereas in front of [+high] vowels we normally encounter their strong reflexes, i.e. /t/ and /d ~ d/, respectively.

In order to get a more precise idea of this overlap, I find it convenient to now summarise all the C₁ reflexes in one table,⁵¹ not only for *t and *d, but also for their pre-nasalised congeners, considering peculiarities of context when necessary. This is done in Table 7. The capital letter T stands for the unconditioned strong reflex. When there are no double reflexes, “n.a.” is put into that column. Contrary to Table 5, I have decided not to include what I consider onset-filling glides in Manenguba and Beti, for the sake of clarity.

As can be seen, the few languages having retained /l/ as the general reflex of *d are those which do not have the /l/ reflex for *t, suggesting a relationship between the two processes with the exception of Nen and Fa’.

Before turning to the detailed examination of the possible diachronic paths leading to the present situation it might prove worthwhile to briefly consider the situation in the closest relatives of Narrow Bantu, i.e. the Grassfields Bantu languages. To be sure, we do not have at our disposal reconstructed diachronic databases of the calibre of BLR (Bastin et al. 2002) or Guthrie’s *Comparative Bantu* (Guthrie 1967; 1970a,b; 1971). However, there is a very valuable collection of Proto-Eastern Grassfields (PEG) roots (Elias et al. 1984), which can be compared to the unpublished *Index of Proto-Grassfields Bantu Roots* (PG) by Larry M. Hyman.⁵² Both lists reconstruct a proto-phoneme *t and also *d and *l. Glancing cursorily through available data, it is clear that there is no sign of double reflexes for *t, the unconditioned reflex being uniformly /t/, at least in Eastern Grassfields. As for

⁵¹The reflexes in the prefixes are generally the same, but not always, as elsewhere in the Bantu domain, e.g. Saghala E741 *b > Ø, but the reflex of PB *bà- (cl. 2) is βa- (Gérard Philippon, unpublished fieldwork notes, 1981–1984).

⁵²My thanks to Larry M. Hyman for graciously letting me have access to a digital version of his Index.

Table 7: Reflexes of *t and *d in various environments

	*t	*t/___*i, *u	T	*nt	*d	*d/___*i, *u	*nd
Nen A44	l	t/l	t	nd	n ~ l	n	ɾ ^a
Maande A46	t	t	n.a.	nd	n ~ l	n	?
Yambeta A462	t	t	n.a.	?	n ~ l	n	?
Elip A62C	d	d	n.a.	nd	n ~ l	n	?
Lundu A11	r	t	t	t	∅	d [r] ^b	Nd
Kundu A122	r	t	t	t	∅	d [r]	Nd
Duala A24	l	t	t	t	∅	d	Nd
Kpe A22	∅	t	t	t	∅	l	Nd
Oli A25	l	t	t	t	∅	d	Nd
	l	t	t	t	∅	d	Nd
Batanga A32C							
Yasa A33a	l	t	t	t	∅	d	Nd
N. Bubi A31a	t	f	n.a.	f(?)	l/∅	r	T
S.W. Bubi A31b	t	f	n.a.	(?)	l/∅	r	nd
S.E. Bubi A31c	t	f	n.a.	(?)	l/∅	r	d
Balong A13	l	t	t	t	∅	d	nd
Bafo A141	l	t	t	t	∅	d	nd
Mbuu A15A	l	t	t	t	∅	d	nd
Myenge A15B	l	t	t	t	∅	d	nd
Akoose A15C	l	t	t	t	∅	d	nd
Mkaa A15C	l	t	t	t	∅	d	nd
Basaa A43a	∅	t	t	t	l ~ ∅	l ~ ∅	nd
Bakoko A43b	∅	t	t	t	l	d	nd
Fa' A51	l	t	t	d	d ~ l	d	d
Kpa A53	r	t	t ^d	d	l ~ r	d	d

^aNo clear example in C₁.^bFor a discussion of the partly individual variants of this sound, see Friesen (2002: 24ff).^cC₁ *d is sometimes realised as ∅ in Bubi varieties with no consistency, e.g. *dám̄b 'cook' (C.S. 486, BLR 842) with initial /l/ everywhere except two A31c varieties which have a ∅ reflex in C₁, i.e. áb̄á. On the other hand *dóm̄è 'husband' (C.S. 697, BLR 1183) has ∅ everywhere, except in the A31b variety of Batete which has /l/, i.e. m̄ò-lóm̄ε. However, other roots, such as *dób 'fish with line' (C.S. 638, BLR 1088), have /l/ everywhere in C₁. Note that the tendency for *d > ∅ is much stronger in C₂. It is clear that the areal shift *l > ∅ does not entirely bypass Bubi.^dAs mentioned above, there are more strong reflexes of *t in Kpa A53 than in other languages.

1 Double reflexes in north-western Bantu

	*t	*t/_*i, *u	T	*nt	*d	*d/_*i, *u	*nd
Eton A71	l	t	t	t	∅	d/_*u, dz/_*i	nd
Njowi A63	l	t	t	t	∅	d	nd
Ewondo A72a	l	t	t	t	∅	d	nd
Ntumu A75A	l	t/_*u, ts/_i	t	t	∅	d/_*u, dz/_*i	nd
Atsi A75D	l	ts	t	t	∅	dz	nd
Mvai A75F	l	t/_*u, ts/_i	t	t	∅	d/_*u, dz/_*i	nd
Bekwel A85b	l	t	t	(?)	∅	d	ɗ
Kol A832	l	t	t	t(?)	∅	d	nt
Koonzime A842	l	t	t	(?)	∅	d	nt
Njem A84	l	t	t	t(?)	∅	d	nt
Makaa A83	l	tʃ	t ^a	(?)	∅	ʃ	ntʃ
Mpiemo A86c	l	t(?)	t	t(?)	∅	d	nt
Polri A92a	l	t(?)	t	t(?)	∅	d	ndɕ(?)
Kwasio A81	l	ts/_*i, (?)/_ *u	t	t(?)	∅	dz/_*i, (?)/_ *u	nd ^h
Gyeli A801	l	ts/_*i, tʃ/_ *u	t	t(?)	∅	ɕ/_*i, (?)/_ *u	nd
Kwakum A91	t	tʃ/_*i(?)	n.a.	t(?)	d ~ l	ɕ/_*i, d ~ l/_ *u(?)	nd(?)
Seki B21	t	ts	n.a.	(?)	d ~ l	d	(?) ^b

^aI have found only a single convincing instance of a strong reflex in Makaa A83.

^bIn C₂ *nd > nd, but there is virtually no example of C₁ *nd in the Seki B21 sources, contrary to *mb and *ŋg. This rarity of *nd seems to be an areal phenomenon.

**d* vs. **l*, Elias et al. (1984: 48) state that “[t]he distinction between initial **d* and **l* is not always clear”, but consider that the two must be distinguished on the basis of different reflexes in the Northern Eastern Grassfields languages (Limbum, Adere, etc.). Although, in this case also, much more research is needed and any conclusion must for the time being remain impressionistic, one can notice some apparent tonal conditioning, as exhibited in (21) by the reflexes in Northern Eastern Grassfields languages, compared with relevant data from Mankon, a member of the Ngemba branch of Eastern Grassfields.

(21) Reflexes of Proto-Eastern Grassfields (PEG) **d* and **l*

- a. PEG **d* > /r/
 - **dàl* ‘bridge’ > Lus *rà*, Mankon *ì-là*
 - **dil* ‘beard’ > Lus *rà*, Mankon *nì-lù-à*
 - **dùk* ‘palm wine’ > Nkot *rùk*, Mankon *mì-lù?-ù*
 - **dùn* ‘be old’ > Nkot *ràn*, Mankon *lvùn*
- b. PEG **l* > /l/
 - **lé*m ‘blood’ > Lus *lé*⁵³, Mankon *à-lém-à*
 - **lón* ‘beg’ > Nkot *lón*, Mankon *lón*
 - **lák* ‘village’ > Lus *lò?*, Mankon *à-lá?-á*

This digression to Grassfields Bantu is very superficial and further examination might shed new light on the question. However, a comparison of the PG and PEG lists indicates that, if we limit ourselves to items coinciding both in form and meaning in the two lists, out of 11 stems reconstructed with **d*, nine are followed by L tone, while out of 11 stems with **l*, eight are followed by H. There is thus at least a suspicion for the tonal conditioning of a reflex split, as we saw in Narrow Bantu, and since Grassfields **t* > *t* in all cases, we shall conclude that Grassfields data cannot help us in our search for the partial merging of PB **t* and **d*.

We must then come back to Stewart’s proposals since they are the only ones trying to flesh out the diachronic developments of the PB phonemes. Stewart clearly saw that to explain the /l/ reflexes of PB **t* in north-western Bantu languages some merging of **d* (or **l*) and **t* must have occurred. For Nen, Van Leynseele & Stewart (1980) propose the stages in (22), starting from PB with couples of ‘fortis’ and ‘lenis’ consonants **t* / *’*t*; **d* / *’*d*.

- (22) Nen A44 reflexes of **t* / *’*t*; **d* / *’*d* (Van Leynseele & Stewart 1980)
- 1) *’*d* nasalises to *n*

⁵³The symbol ° signals a non-downgliding L tone.

- 2) highly marked $*'t$ shifts to $*'d$ (drag chain?)
- 3) $*d$ also shifts to $*'d$, thus merging the reflexes of $*'t$ and $*d$
- 4) $*'d > l$

The shifts summarised in (22) lead to the following situation in Nen: $*t > t$, $*'t > l$, $*d > l$, $*'d > n$. This solution works but at the cost of positing a ‘Duke of York’ type of change (Pullum 1976; Yates & Zukoff 2018), where the diachrony gets rid of one phoneme ($*'d$) to reintroduce it in the next move.

Having determined that PB did not have an implosive as ‘lenis’ counterpart to $*d$, but a lateral instead, Stewart (1989) changed his approach. This did not really improve on the previous solution, since now it was $*l$ that nasalised to $/n/$, only to be reintroduced from ‘lenis’ $*f$ through a stage $*d$, thus $*f > *d > l$. The final reflexes for Nen were then $*t > t$, $*f > l$, $*d > l$ (also through a $*d$ stage) and $*l > n$. The two proposals are summarised in Table 8.

Table 8: Stewart’s successive conceptions of PB coronals (Van Leynseele & Stewart 1980; Stewart 1989)

1980	PB	$*t$	$*'t$	$*d$	$*'d$
	Nen	t	$*'d > l$	$*'d > l$	n
1989	PB	$*t$	$*'t$	$*d$	$*l$
	Nen	t	$*'d > l$	$*'d > l$	n

Stewart (1993) abandoned his view of double reflexes in Bantu (cf. §1.2), but he still proposed a diachronic path for $*t > l$ in ‘North-Western Bantu’.⁵⁴ Surprisingly, he posited a development $*t > *ǃ > l$, while admitting that “[... in present-day North-Western Bantu languages ...] $*ǃ$ appears never to have the direct reflex $ǃ$. $ǃ$ is however a plausible source for the various reflexes that do occur; the most common reflex is l [...]” (Stewart 1993: 19). Contrary to Stewart (1993), I consider this development rather implausible. To the best of my knowledge, the West Kele B22a and Ngom B22b varieties of the Gabonese language Kele B22, which is geographically remote, are the only ones in our general area to have $/ǃ/$. Moreover, their $/ǃ/$ is a reflex of $*d$ and not of $*t$ (Guthrie 1967: 34).

So, we should try to define more precisely the phonetic content of the putative proto-phonemes. In other words, what sounds do the comparative symbols $*t$ and $*d$ stand for, since this should allow us to discern how reflexes of $*t$ and $*d$ came partly to overlap?

⁵⁴Nowhere does Stewart (1993) define the coverage of this ‘North-Western Bantu’ group, but an examination of the proposed reflexes shows that it could not include Nen.

As far as the voiceless coronal is concerned, there is room for little hesitation. Its strong reflexes, whether conditioned by [+high] vowels, pre-nasalised or unconditioned, are always [+coronal] [–voice] [–continuant], so a voiceless coronal stop /t/. The few exceptions are due to affrication processes triggered by [–back] [+high] vowels, as in Bubi or Kwakum for example, or to the beginning of BF, as seen in the Fang varieties and the southern Nyong-Dja languages, but even then the result is a voiceless coronal affricate. Furthermore, an examination of the whole Bantu field clearly shows that by far the most widespread reflex is also /t/, leaving little doubt that this was the identity of the proto-sound. What we would like to know, but have very little evidence to go by for, is the precise place of articulation [±anterior] and the precise laryngeal setting, as this would help to understand the weakening trajectory. As for the place of articulation, the only thing which can be said is that in those Bantu languages where a [±anterior] contrast exists, the reflex of *t is always [–anterior]. For instance, in Amu G42a or Makhuwa P31, /t/ is a reflex of *t and /t̥/ of *c. In Mashati E623B, /t/ is also a reflex of *t, while /t̥/ is from extraneous sources. Even in languages where no contrast exists, such as Unguja Swahili G42d, the realisation of /t/ is audibly [–anterior] with most speakers.

The situation for *d is much more difficult. First of all, there are extremely few languages where its unconditioned reflex is [d]. Guthrie (1967: 62) even claims there are none,⁵⁵ but this is proven wrong by two languages in our area, namely Kwakum and its close relative Seki, which both have /d/, but also /l/ with tonal conditioning (cf. §2.1). In most Bantu languages outside our area, the unconditioned reflex is /l/ often weakening to Ø. The strong reflex, i.e. /d ~ d̥/, is found in the same environments as for *t, i.e. in front of [+high] vowels (with the same peculiarities of affrication as mentioned above), and also in pre-nasalised position, where *nd is maintained as /nd/, apart from Kpa and Bubi varieties which denasalise.

The choice for the proto-sound is obviously between *l, which was the solution of Meinhof (1899) who posits no voiced stops at all, and *d chosen by Guthrie (1967: 62) with some hesitation, admitting that it might have gone to *l very early in Bantu language history. Meeussen (1967: 83) is rather non-committal about it: “[...] one might just as well use the symbol [...] /l/ instead of /d/”. Nevertheless, he reasoned by analogy that since the contrast in reflexes was mostly [+voice] vs. [–voice] in the labial and dorsal series (such as *p / b ~ β* and *k / g ~ γ*), even if spirantised, the coronal series must have exhibited originally the same sort of contrast, i.e. *t / d ~ l*. The fact that Meeussen (1967) also accepts the lateral grapheme,

⁵⁵Guthrie (1967: 62) probably did not check his own notes, as Guthrie (1971: 33–34) does state the correct Kwakum and Seki correspondences.

shows that he himself was hesitant on this point. Many close and less close relatives of Narrow Bantu exhibit /l/ in corresponding items and Elias et al. (1984) reconstruct *l for their PEG, while Stewart (1989) posits *l alongside *d for his PB. I assume here that the PB phoneme was indeed *l with a [d ~ d̥] allophone. In the case of pre-nasalisation, *N + l > nd* is expected by spreading of the [–continuant] feature of the first part onto the second. Notice that /l/ is somewhat paradoxical: it is articulatorily both [+continuant] since the airflow can escape laterally, but also [–continuant] since some part of the tongue makes a contact with a passive articulator (typically the hard palate). Generally, the [–continuant] part of the sound’s identity plays no phonological role, but in contact with [+high] vowels where aperture is minimal, it can be considered to become exclusive, hence the realisation [d ~ d̥]. In Kwakum and Seki, the /d/ reflex must be considered a case of strengthening and we have seen that the weak reflex /l/ is attested in front of L tone. However, this problem does not impinge on the question of *t, since those two languages exhibit no double reflexes for it.

We shall thus turn to the well-attested double reflexes of *t. In (23), I present again the maximum list of items with unconditioned strong reflexes of *t which I could establish (see also (11)–(12) in §2.2).

(23) Reconstructed roots manifesting a strong reflex of *t

a. Fairly well distributed

**táánò/ò* ‘five’ (C.S. 1662, BLR 2768 & 2769)

**tòòg* ‘boil up, bubble up’ (C.S. 1777, BLR 2966-7)

**tóóbá* ‘six’ (C.S. 1815, BLR 3034)

**tédám* ‘stand’ (C.S. 1692½, BLR 2816)

**tóná* ‘spot, speckle’ (C.S. 1785, BLR 2976)

b. More restricted distribution (due to lexical variation)

**tónd* ‘desire’ (C.S. 1788, BLR 2980)

**támbí* ‘sole of foot, shoe’ (C.S. 1659, BLR 2761)

**tòdú* ‘navel’ (C.S. 1776, BLR 2965)

**tèk* ‘become soft’ (ps. 434, BLR 2827)

c. Somewhat doubtful item

**tàndá* ‘invertebrate: spider; spider’s web’ (BLR 9730)

d. Nursery word

**tààtá* ‘father’ (C.S. 1686, BLR 2806)

Although these items are not very numerous, they are nevertheless striking in their regularity. Only one item shows systematic non-correspondence between languages and it is also often irregular as far as its vowel is concerned, to the extent that Guthrie reconstructs no less than three C.S.s for it: **tói* ‘ear’ (C.S. 1801, 1809, 1813, BLR 3030) (variants **tói(ò)* and **tói(é)*). Furthermore, many Benue-Congo languages attest a final *ŋ* for this item.

Since we have posited that the phonetic content of PB **t* must have been /t/, these items then show *retention* of the original sound, just as it was retained in front of [+high] vowels in contradistinction to most other items where it shifted to /l/. To what extent can this shift to /l/ be considered as weakening?

The weakening trajectories we have been considering above would posit (in a logical, step by step fashion) the following stages: **t* > **ɾ* (first weakening), then either **ɾ* > *h* (second weakening) or **ɾ* > *r* (strengthening of marked sound).⁵⁶ These stages are well-attested in some north-eastern and south-eastern Bantu languages, e.g. Rimi F32 **t* > *ɾ*, Pokomo E71 **t* > *h*, and Gweno E65, Ngazija G44a, Cuwabo P34, etc. **t* > *r*. However, they are unattested in our area with the lone exception of Kpa, which incidentally has fewer cases of weakening than the others. For the other languages, the reflex is always /l/, further weakened to Ø in Kpe (for Basaa see below). One could possibly consider that /l/ is a further weakening of /r/.⁵⁷ Nevertheless, this appears unlikely to me. Most Bantu languages, apart from those mentioned above, do not have a distinctive contrast between a lateral and a rhotic and realise their liquid phoneme (PB **d* or in Stewart’s PB **l*), either as one or the other, in some cases in clearly defined contexts. Sometimes, as is the case of Oroko A101, the liquid is realised as an alveolar tap, giving the auditory impression of a sound intermediate between [l] and [r] – cf. Nida (1964: 20), cited in Friesen (2002: 25), with reference to Oroko orthography. In the other languages, however, the lateral character of the liquid is strongly asserted by all the sources and the unlikely path **t* > *r* > *l* is not supported by an intermediate stage.⁵⁸ A particularly suggestive case is provided by Fa’, a language closely related to Kpa (see Table 7). Whereas the latter opposes /r/ (reflex of both **t* and **d* before L) and /l/ (reflex of **d* before H), Fa’ has no /r/. Instead, it has /l/, wherever

⁵⁶Maddieson (1984) has just three languages with voiceless /ɾ/ versus 130 with voiced /r/.

⁵⁷I owe this suggestion to Jean-Marie Hombert (p.c.). Support for this might be seen in the fact that in Bubi /r/ appears instead of /l/ in front of [+high] vowels, which environment conditions strong reflexes (mostly /d ~ d/) in the other languages. So /l/ is weaker than /r/.

⁵⁸I know of only one Bantu language where **t* > *r* > *l* is attested, namely Lozi K21, but this evolution is clearly due to contact. Being a language of S30 origin with an initial r/l contrast, Lozi lost this opposition by accommodation to the articulatory habits of the majority of speakers after having been transplanted to linguistic surroundings with no such contrast (cf. Gowlett 1989).

Kpa has /r/, and /d/ for Kpa /l/, as well as for Kpa /d/, the positional allophone of /l/ before [+high] vowels. I would thus suggest that Fa' presents the original situation (similar to the one offered by all the other languages) and that Kpa for unknown reasons strengthened /l/ to /r/, and later weakened /d/ to /l/ in some contexts. The conclusion that, in our area at least, /r/ cannot constitute a weakening stage in an assumed trajectory $*t > r > l$ is quite convincing.

What appears is rather that somehow the reflexes of $*t$ have shifted to occupy the place of PB $*l$, the latter having weakened to zero. Now this complete weakening is not unknown in the rest of Bantu. Although absent from many parts of the domain, it is quite frequent in the north-eastern quadrant, especially its north-easternmost part, i.e. Sabaki and Kilimanjaro Bantu mostly, with a few isolated cases like Rimi F32, Kamba E55 or Shambaa G23. Often, but not systematically, those languages where $*l$ has weakened have also weakened $*t$, e.g. Mashami E621B $*l > \emptyset / *t > \beta$, Lower Pokomo E71B $*l > \gamma / *t > h$, Dawida E74a $*l > \emptyset / *t > d$, etc. Counter-examples are Kamba E55, Shambaa G23, Unguja Swahili G42d and a few others, which have $*l > \emptyset / *t > t$, where contact can be suspected to be the cause of $*l > \emptyset$, since their closest relatives do not exhibit the change (except in the case of Unguja Swahili). In none of those languages is $*t > l$ attested.

I conclude that in our area the initial change must have been $*l > \emptyset$, except in strong environments. It is only then that the change $*t > l$ could occur. If it had occurred before, this new /l/ would also have gone to \emptyset . Indeed, a number of Central Sawabantu languages followed this course as seen in Table 7, but since their closest relatives have retained /l/, the development must be recent. Note that this shift did not remove /t/ from the phoneme inventory since it subsisted in the very same strengthening environments just mentioned. Instead, it reintroduced a liquid phoneme, so that the languages in question still presented a full roster of coronal stops and laterals: /l/, /t/, /d ~ d/ (before [+high] vowels, see Table 7), /nd/.

Nevertheless, during the course of this change, a reduced number of items (the ones mentioned above) were bypassed by it. As seen earlier, there are more in Kpa, which perhaps significantly is the northernmost of the languages treated. Since they are exactly the same in all the languages and designate mostly non-cultural items, it is very unlikely that their presence is due to borrowing, except for $*t\acute{a}mb\acute{i} / *t\acute{a}mb\acute{i}$ (Guthrie 1970b: 90, C.S. 1659; BLR 2761), whose original meaning is 'sole of foot', but spread in our area with the meaning 'shoe', possibly from Duala. Their exemption from $*t > l$ must be a characteristic of the putative ancestor language of the languages concerned, that is the common ancestor of the Sawabantu, Manenguba, Beti and Nyong-Dja languages, a north-western clade

characterised by the old change $*k > \emptyset$.⁵⁹ But there is no evidence that would lead us to put this situation back to PB, especially since there is no sign of a similar split in Grassfields languages.

For Guthrie, as already mentioned, the items characterised by retention of the strong reflex had a long stem vowel. As for our list in (23) above, independent evidence for a long vowel is only robust for $*táánò/\grave{v}$ ‘five’ and $*tónđ/*tóónđ$ ‘desire’ (Guthrie 1970b: 118, C.S. 1788; BLR 2980). There is some independent evidence of a short vowel for $*tédam$ ‘stand’, as Tsaangi B53 and Kongo H16 have a short vowel, $*tèk$ ‘become soft’, as Luba-Kasai L31a and Bemba M42 have a short vowel, and $*tóná \sim *tónì$ ‘spot, speckle’ with a short vowel in Kongo and Bemba.⁶⁰ No decision about vowel length can be made for lack of independent evidence in the case of $*tòòg$ ‘boil up, bubble up’, which has an “osculant” short-vowel form $*tòk$ (Guthrie 1970b: 117, C.S. 1778; BLR 2967), $*tósóbá$ ‘six’, and $*tòdú$ ‘navel’.

Recall from §1.2 that Janssens posited pre-nasalisation as the source for strong reflexes. This might conceivably apply to nouns, which would have originally belonged to classes 9/10 or 11/10 and thus acquired a nasal prefix ($*nt > t$ in all languages concerned, apart from Nen), which they would have retained even when placed in other noun classes. Indeed, this fact can be easily seen, when C_1 belongs to the voiced stop series, as the nasal is normally retained, as in Du-ala $mù-η-gàngà$ ‘medicine-man’ ($< *NP_1-NP_9-gàngà$) or Basaa $lì-η-géñéé$ ‘bell’ ($< *NP_5-NP_9-gèñgédéé$), or again Seki $di-m-bílo$ ‘oil palm’ ($< *NP_5-NP_9-bidà$).

This might explain items which appear with /t/ even when not in cl. 9/10 or 11/10, provided there is some evidence they might have originally belonged there. Unfortunately, such items in our list ($*tóná \sim *tónì$ ‘spot, speckle’; $*tòdú$ ‘navel’) never appear in cl. 9 anywhere. In fact, they are solidly attested in cl. 5. Now, it is known from Eastern Bantu that cl. 5 also can have a strengthening effect on stem-initial consonants, but this does not appear to be the case in our area, at least I have not observed any traces of this conditioning. Bachmann (1989) claims that it does apply, but I find his few examples unconvincing.

For other items, an anonymous reviewer remarks quite correctly that $*táánò/\grave{v}$ ‘five’ has an “osculant” form with $C_1 *c$, i.e. $*cáànò/\grave{v}$ (Guthrie 1970a: 82, C.S. 275-6; BLR 446, 448). This is true, but as far as I can see, it is restricted to Eastern

⁵⁹The clade presumably includes, apart from the languages treated here, several of the B20 languages and a large part of the forest languages grouped by Guthrie under zone C. I must postpone this discussion to a later publication.

⁶⁰Identical stems with seemingly related meanings like ‘drip’, ‘drop’ or ‘rain’ appear in a number of Eastern languages with a long vowel, e.g. Shi JD53 $rōop$ ‘drip’, $óómúrōopi$ ‘drop (n.)’, Gusii JE42 $tóóni$ ‘drip’, Nilamba F31 $tʰóóni/matóóni$ ‘drop’, etc. So, there might be some doubt about the length.

Bantu and does not affect our area. Furthermore, it was originally (and is still synchronically in some languages) conditioned by a numeral prefix **i-* and the same conditioning applies to **tátw* ‘three’, which presents an /l/ reflex in all our languages.

There is thus no conclusive evidence for the “anomalous” items in (23) having had some phonological characteristic that would make them impervious to the **t > l* shift. We are thus forced to conclude that we are faced with a change progressing through the lexicon, but failing to reach certain words, in a ‘wave’ pattern. In view of the overall evidence, this is not an ongoing situation but the frozen result of a process long spent, since the same few items are affected. The various phonemes (/l/, /t/, ...) are well established and serve as basis for the introduction of new lexical items, through borrowings, internal derivation, etc., as a detailed examination of the various lexicons would show.

4 Conclusion

In spite of considerable achievements in the domain of comparative Bantu phonology, few diachronic processes were reconstructed in detail. Guthrie’s *Comparative Bantu* (1967; 1970a; 1970b; 1971) mostly aimed at establishing Common Bantu forms, i.e. series of synchronic correspondences between individual languages. It is true that with his two-stage method (Guthrie 1962), he attempted to deduce from these correspondences what he considered as Proto-Bantu reconstructions. However, due to his uncertain methodology, his “Proto-Bantu” turned out to be not much more than a glorified “Common Bantu”. Thorough criticisms of Guthrie’s method can be found in Meeussen (1973) and Möhlig (1976). Scholars from the Tervuren school did some very valuable work on specific points (e.g. Grégoire & Doneux 1977; Bastin 1983), but as far as I know never published a general survey of consonant systems. The only real attempt in this direction was made by Stewart, as discussed repeatedly in this chapter. However, since his ultimate goal was setting up a Proto-Bantu-Potou-Tano, which could eventually constitute a basis for a Proto-Niger-Congo as expressed in Stewart (2002), there were constraints on his Proto-Bantu reconstructions due to the necessity of establishing cognates with Cama, Mbatto and Akan. He therefore reconstructed the “fortis”/“lenis” opposition, which is not well supported within Bantu, as he eventually admitted himself (Stewart 1993).

I have nevertheless followed Stewart’s lead up to a point. Although he did not contribute to solve the puzzle of double reflexes, his positing of **l/*d* as a PB phoneme and his suggestions as to the voicing and ‘lateralisation’ of PB **t* seem

to me on the right track. Although the present chapter has not really established the origin of the duality of reflexes for PB **t* in some north-western Bantu languages, it has at least confirmed its existence. As for the other voiceless stops reconstructed for PB, double reflexes are not really an issue for **k* and those observed for **p* can be demonstrated to be partly conditioned. When it comes to the voiced PB stops, **g* does not manifest real double reflexes, as is the case for its voiceless counterpart **k*. Double reflexes of **d* can be shown to be conditioned by the tone of the first stem vowel, which also holds for **b*, but to a lesser extent and possibly due to more recent development. For the time being, my survey of putative double reflexes in north-western Bantu languages does not warrant the revision of the PB consonant system. This being said, any conclusion on the PB consonant system in general would be premature at this stage, because it would have to be based on all Bantu languages, and not just the sample I considered here. Minimally, it should also take in the Grassfields Bantu languages from outside Narrow Bantu. In my view, slow, careful, bottom-up reconstruction is of paramount importance here. Whether the occurrence of double reflexes in north-western Bantu languages supports the “phonetically abrupt and lexically gradual” model of sound change as proposed by Wang (1969) is also a point that should be argued further, perhaps by extending and refining the database.

With reference to these two last points, i.e. slow, careful, bottom-up reconstruction and Wang’s lexical diffusion model of sound change, it is worth mentioning a recent article by Pacchiarotti & Bostoen (2022) on the multiple reflexes of the PB **g* and **k* in C₂ position within West-Coastal Bantu, a major discrete branch within the Bantu family (cf. de Schryver et al. 2015; Grollemund et al. 2015; Pacchiarotti et al. 2019; Philippson & Grollemund 2019), situated south of the study area of this chapter.

Lastly, Pacchiarotti & Bostoen (2022) plead for the recognition by comparativists of irregularities in correspondences alongside the regular application of the Comparative Method. That such irregularities are well-attested in Bantu languages is easy to confirm. As an illustration, a rapid survey of the data presented by Guthrie (1970a) shows that out of some 285 comparative series with **b* in C₁, about 135, so almost half, present at least one ‘skewed’ entry, i.e. one judged by Guthrie to exhibit an irregularity in correspondence (indicated in his data by being placed inside square brackets). However, these appear to be fairly haphazard and individual – pending some more detailed study which definitely needs to be undertaken – and thus different from the rather systematic “strong” vs. “weak” reflexes of PB **t* treated in this chapter. Furthermore, the C₂ position is notoriously “weak” in north-western Bantu languages, unlike elsewhere in the Bantu area where C₁ and C₂ positions are normally not marked by different reflexes.

This fact might be attributed for north-western languages to some prosodic factor that demarcates the first stem syllable (see, for example, Paulian 1975), whereas elsewhere in Bantu the penult constitutes the most salient position (cf. Philippson 1991; Hyman 2013). Micro-variation in C_2 reflexes would thus appear to be less significant than those in C_1 (the ‘prosodically salient’ position). This does not mean that a detailed study of C_2 reflexes, such as presented by Pacchiarotti & Bostoën (2022) for West-Coastal Bantu is unnecessary, quite the opposite. Such a study is underway for the languages covered in the present chapter and its results will tell whether and to what extent the scenario outlined above needs to be modified.

Abbreviations

BF	Bantu Frication
BLR	Bantu Lexical Reconstructions (Bastin et al. 2002)
C.S.	Comparative Series (Guthrie 1970a,b)
C	consonant
C_1	stem-initial consonant
C_2	second consonant
CL.	class
H	high tone
INTR.	intransitive
L	low tone
N	nasal
NP	noun prefix
PEG	Proto-Eastern Grassfields
PG	Proto-Grassfields
ps.	partial series (Guthrie 1970a,b)
V	vowel

Appendix A Languages covered and sources used

Oroko	A101	Friesen (2002)
Lundu	A11	Kuperus (1985)
Kundu	A122	Bufe (1910–1911); Atta (1993)
Balong	A13	Bufe (1910–1911); Kouoh Mboundja (2004)
Bafo	A141	Hedinger (1987)
Manenguba	A15	Hedinger (1987)
Akoose	A15C	Dorsch (1911–1912a,b; 1912–1913); Hedinger & Hedinger (1977); Hedinger (1985; 1987)
Mkaa	A15C	Ewane Etame & Hedinger (2017)
Nkongho	A151	Hedinger (1987)
Mboko	A21	Ebobissé (2015)
Kpe	A22	Ardener (1956); Hawkinson (1986); Monikang (1989); Kagaya (1992); Ebobissé (2015)
Bubia	A221	Chia (1993); Ebobissé (2015)
Su	A23	Meinhof (1889–1890); Ebobissé (2015)
Kole	A231	Ebobissé (2015)
Duala	A24	Dinkelacker (1914); Ittmann (1939); Paulian (1971); Helmlinger (1972)
Bodiman	A241	Ebobissé (2015)
Oli	A25	Hagège (1967); Ebobissé (2015)
Pongo	A26	Ebobissé (2015)
Mongo	A261	Ebobissé (2015)
Limba	A27	Ebobissé (2015)
Bubi	A31	Baumann (1887–1888); Rurangwa (1989); Bolekia Boleká (1991; 2008; 2009)
Noho	A32a	Adams (1907); Ebobissé (2015)
Bapuku	A32b	Ebobissé (2015)
Batanga	A32C	Ebobissé (2015)
Yasa	A33a	Bôt (1992); Blench (2010b); Ebobissé (2015)
Kombe	A33b	Fernandez (1951); Elimelech (1976)
Benga	A34	Nassau (1892); Guthrie (1967; 1970a,b; 1971)
Lombi	A41	Lamberty (2002)
Abo	A42	Atindogbé (1996); Lamberty (2002)
Basaa	A43a	Janssens (1986); Mous & Breedveld (1986); Teil-Dautrey (1991a,b); Hyman (2003a); Njock (2005)

1 *Double reflexes in north-western Bantu*

Bakoko	A43b	Edika (1990); Kenmogne (2000); Mathaus & Anderson (2010)
Nen	A44	Dugast (1967; 1971); Mous (2003); Boyd (2015); Boyd (2019b)
Nyokon	A45	Richardson (1957); Mous & Breedveld (1986); Lovestrand (2011)
Maande	A46	Mous & Breedveld (1986); Nomaande Language Committee et al. (2003); Boyd (2015)
Tuotomb	A461	Mous & Breedveld (1986)
Yambeta	A462	Mous & Breedveld (1986); Mongo & Bolioki (2012); Boyd (2015); Boyd (2020)
Fa'	A51	Guarisma & Paulian (1986); Perrin (1986); Isaac (2014)
Dimbong	A52	Guarisma & Paulian (1986)
Kpa	A53	Guarisma (2000)
Bea	A54	Guarisma & Paulian (1986)
Tuki	A601	Mous & Breedveld (1986); Hyman & Biloa (1992); Boyd (2015); Boyd (2016e)
Yangben	A62A	Paulian (1986); Boyd (2015); Boyd (2016f); Bébiné (2018)
Mmala	A62B	Mous & Breedveld (1986); Paulian (1986); Boyd (2009); Boyd (2015); Boyd (2016d)
Elip	A62C	Paulian (1986); Prittie (2002); Boyd (2015); Boyd (2016b)
Baca	A621	Paulian (1986); Boyd (2015); Boyd (2016a)
Gunu	A622	Paulian (1986); Patman & Robinson (1989); Yukawa (1992); Hyman (2001); Boyd (2015); Boyd (2019a)
Mbule	A623	Boyd (2015); Boyd (2016c)
Njowi	A63	Blench (2010a)
Eton	A71	Van de Velde (2006)
Ewondo	A72a	Tsala (1956); Abéga (1971); Angenot (1971); Essono (2000)
Bulu	A74a	Bates & Johnson (1926); Alexandre (1966)
Fang	A75	Galley (1964); Kelly (1974); Andeme Allogo (1991); Hombert (1991); Medjo Mvé (1997); Mékina (2012)
Gyeli	A801	Grimm (2015)
Shiwa	A803	Puech (1989b); Ollomo Ella (2013); Cheucle (2014)
Kwasio	A81	Guthrie (1967; 1970a,b; 1971); Cheucle (2014)
So	A82	Beavon & Beavon (2018)
Makaa	A83	Heath & Heath (1982); Heath (2003); Cheucle (2014)

Kol	A832	Henson (2007); Cheucle (2014)
Njem	A84	Beavon (2003; 2005); Cheucle (2014)
Koonzime	A842	Beavon (1983; 2020); Cheucle (2014)
Bekwel	A85b	Puech (1989a); Bouka (1995); Cheucle (2014)
Mpiemo	A86c	Thornell & Nagano-Madsen (2004); Cheucle (2014); Festen & Murrell (2020)
Kwakum	A91	Belliard (2005); Njantcho Kouagang (2018)
Polri	A92a	Wéga Simeu (2016)
Kako	A93	Ernst (1996a,b); Medjo Mvé (2008)
Seki	B21	Jacquot (1983); Mickala Manfoumbi (2005); Puech (S.d.)

Appendix B Reflexes of **t* and **p* in Nen A44 vs. Maande A46

- Nen /l/
 - **tá* ‘saliva’ (C.S. 1629, BLR 2703) > *mà-lá*, cf. Maande *maa-tá*
 - **táánò/ò* ‘five’ (C.S. 1662, BLR 2768 & 2769) > *lánv*
 - **támb* ‘set trap’ (ps. 429, BLR 2759) > *lámb*
 - **tátv* ‘three’ (C.S. 1689, BLR 2811) > *lálv*, cf. Maande *tátv*
 - **tém* ‘cut down (tree)’ (C.S. 1703, BLR 2832) > *lím-á* ‘clear field’, cf. Maande *tám-a*
 - **tí* ‘tree’ (C.S. 1729, BLR 2881) > *pò-lí-á*, cf. Maande *pòv-tí*
 - **tímà* ‘heart’ (C.S. 1738, BLR 2895) > *mò-límá*, cf. Maande *v-témá*
 - **tó* ‘ashes’ (C.S. 1769, BLR 2954) > *mò-lv*, cf. Maande *mòv-tá*
 - **tóng* ‘crow (rooster)’ (C.S. 1793, BLR 2994) > *lón*
 - **tó* ‘head’ (C.S. 1800, BLR 3007) > *mò-lv*, cf. Maande *aa-tv*
 - **tóm* ‘send’ (C.S. 1831, BLR 3055) > *lóm*, cf. Maande *tóm-a*
- Nen /l/ before [+high vowels]
 - **túd* ‘forge’ (C.S. 1861, BLR 3101) > *lún*, cf. Maande *tún-ə*
 - **túkò* ‘night’ (C.S. 1864, BLR 3105) > *pù-lw-á*, cf. Maande *pu-tú*
 - **túm* ‘stab’ (C.S. 1866, BLR 3108) > *lúm* ‘hit with missile’, cf. Maande *túm-ə* ‘stick into’

- Nen /t/
 - **tákò* ‘buttock’ (C.S. 1650, BLR 2741) > *i-tá*
 - **tédam* ‘stand’ (C.S. 1692½, BLR 2816) > *tínim*
 - **tíáb* ‘gather firewood’ (C.S. 1735, BLR 2889) > *tʸáp-á*, cf. Maande *tyáp-a*
 - **tśád* ‘carry on head’ (C.S. 1806, BLR 3017) > *twán*
 - **tśóg* ‘draw water’ (C.S. 1826, BLR 3048) > *tśk*, cf. Maande *tśk-a*
- Nen /t/ before [+high vowels]
 - **tím* ‘dig’ (C.S. 1752, BLR 2918) > *tím-à*, cf. Maande id.
 - **tú* ‘spit’ (C.S. 1857, BLR 3096) > *tú*, cf. Maande id.
 - **túútú* ‘bump’ (C.S. 1882, BLR 3137) > *i-tútú*, cf. Maande *ji-tútú*
- Mbam reflexes of **p*
 - **pá(an)* ‘give’ (C.S. 1404(a), BLR 2344) > Nen *hán*, Tuki, Mmala *fá*, Yangben *fà*, Elip *hʷá*, Baca, Gunu *fã*
 - **pèèm* ‘breathe’ (C.S. 1468, BLR 2436) > Nen *him* ‘breathe noisily’ ~ *fim* ‘blow’, Maande *bí-fáma* ‘blow nose’, Yambeta *fimit* ‘blow’, Mmala *bí-fémà* ‘blow nose’, Baca *fúmà*
 - **pép(id)* ‘blow (as wind)’ (C.S. 1487, BLR 2463) > Nen *fifà*
 - **pépɔk* ‘be light in weight’ (C.S. 1494, BLR 2480) > Nen *háh-án*
 - **pìng* ‘exchange’ (C.S. 1530, BLR 2539) > Nen *hìŋ* ‘replace’
 - **pàc* ‘split’ (C.S. 1405, BLR 2346) > Yambeta *pàsa* ‘carve’
 - **pèèp(id)* ‘fan’ (C.S. 1489, BLR 2469) > Nen *pìp*, Tuki *pìp-á* [tones?] ‘fan’ ~ *pìp-à* ‘winnow’

References

- Abéga, Prosper. 1971. *Grammaire ewondo*. Yaoundé: Université Fédérale du Cameroun.
- Adams, Gustaf A. 1907. Die Sprache der Banôho. *Mittheilungen des Seminars für Orientalische Sprachen an der Königlichen Friedrich-Wilhelms-Universität zu Berlin* 10(3). 34–83.

- Alexandre, Pierre. 1966. *Système verbal et prédicatif du bulu* (Langues et littératures de l'Afrique noire 1). Paris: Librairie C. Klincksieck.
- Andeme Allogo, Marie-France. 1991. *Morpho-syntaxe du ntumu : dialecte fang, langue bantu du Gabon*. Paris: Institut National des Langues et Civilisations Orientales. (Doctoral dissertation).
- Angenot, Jean-Pierre. 1971. *Aspects de la phonétique et de la morphologie de l'ewondo*. Leiden: Leiden University. (Doctoral dissertation).
- Ardener, Edwin. 1956. *Coastal Bantu of the Cameroons (the Kpe-Mboko, Duala-Limba and Tanga-Yasa groups of the British and French Trusteeship Territories of the Cameroons)* (Ethnographic Survey of Africa: Western Africa 11). London: International African Institute.
- Atindogbé, Gratien Gualbert. 1996. *Bankon (A40) : éléments de phonologie, morphologie et tonologie* (Grammatical Analyses of African Languages 7). Cologne: Rüdiger Köppe.
- Atta, Samuel Ebongkome. 1993. *The phonology of Lukundu (Bakundu)*. Yaoundé: Université de Yaoundé I. (MA thesis).
- Bachmann, Armin R. 1989. Zum 'Fortis/Lenis-Kontrast' in den nordwestlichen Bantu-Sprachen. *Afrikanistische Arbeitspapiere* 19. 23–31.
- Bancel, Pierre. 1988. Double reflexes in Bantu A.70 languages. *Pholia* 3. 7–16.
- Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en bantou* (Annales – Série in-8° – Sciences humaines 114). Tervuren: Royal Museum for Central Africa.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bates, George L. & Silas F. Johnson. 1926. *Handbook of Bulu, with Bulu – English vocabulary*. Revised edition, with the vocabulary revised and much enlarged by the author and Silas F. Johnson, M.D. Elat, Cameroon: Halsey Memorial Press.
- Baumann, Oskar. 1887–1888. Beiträge zur Kenntniss der Bube-Sprache auf Fernando-Póo. *Zeitschrift für afrikanische Sprachen* 1. 138–155.
- Beavon, Keith H. 1983. A phonology of Konzime. *Africana Linguistica* 9. 109–136.
- Beavon, Keith H. 2003. *Njyem – French – English lexicon*. Cameroon: SIL Cameroon. <https://www.sil.org/resources/archives/47267>.
- Beavon, Keith H. 2005. *A phonology of Njyem*. Unpublished ms. Yaoundé: SIL Cameroon.
- Beavon, Keith H. 2020. *Dictionnaire nzime*. SIL International. <https://www.webonary.org/nzime/>.

- Beavon, Keith H. & Mary Beavon. 2018. *Dictionnaire swo*. SIL International. <https://www.webonary.org/swo/>.
- Bébiné, Adriel J. 2018. *Description phonologique et morphosyntaxique du nuasúe (A62A)*. Yaoundé: University of Yaoundé I. (Doctoral dissertation).
- Belliard, François. 2005. *Instruments, chants et performances musicales chez les Kwakum de l'arrondissement de Doume (est-Cameroun): étude ethnolinguistique de la conception musicale d'une population de langue bantu A91*. Paris: Université Paris 7. (Doctoral dissertation).
- Blanchon, Jean A. 1991. Le pounou (B43), le mpongwe (B11a) et l'hypothèse fortis/lenis. *Pholia* 6. 49–83.
- Blench, Roger M. 2010a. *Mangisa wordlist*. Unpublished ms. Cambridge: Kay Williamson Educational Foundation. <https://www.rogerblench.info/Language/Niger-Congo/Bantu/A70/Mangisa/Mangisa%20wordlist.pdf>.
- Blench, Roger M. 2010b. *Yasa – French wordlist*. Unpublished ms. Cambridge. <https://www.rogerblench.info/Language/Niger-Congo/Bantu/A30/Yasa/Yasa-French%20wordlist.pdf>.
- Bolekia Boleká, Justo. 1991. *Curso de lengua bubí* (Colección Ensayos 8). Malabo: Centro Cultural Hispano-Guineano.
- Bolekia Boleká, Justo. 2008. *Lingüística bantú a través del bubí* (Acta Salmanticensia Estudios filológicos 321). Salamanca: Ediciones Universidad de Salamanca.
- Bolekia Boleká, Justo. 2009. *Diccionario español – bubí / Ě ribúkku ra balláa bēbōbē – lēepanná*. Madrid: Akal.
- Bostoen, Koen. 2001. Osculance in Bantu reconstructions: A case study of the pair °-kádang-/°-káng- ('fry', 'roast') and its historical implications. *Studies in African Linguistics* 30(2). 121–146.
- Bostoen, Koen. 2008. Bantu spirantization: Morphologization, lexicalization and historical classification. *Diachronica* 25(3). 299–356.
- Bostoen, Koen & Yvonne Bastin. 2016. Bantu lexical reconstruction. In *Oxford handbooks online*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199935345.013.36.
- Bôt, Dieudonné Martin Luther. 1992. *Phonologie générative du yasa*. Yaoundé: University of Yaoundé. (Doctoral dissertation).
- Botne, Robert. 1992a. Double reflexes in eastern and southern Bantu. *Afrikanistische Arbeitspapiere* 29. 131–148.
- Botne, Robert. 1992b. Phonemic split in nen (A44) – A case of tonal conditioning of glottalic Proto-Bantu consonants. *Afrika und Übersee* 75(2). 207–225.
- Bouka, Léonce-Yembi. 1995. *Structures phonologiques et structures prosodiques : le modèle bekwel*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).

- Boyd, Virginia. 2009. *Précis d'orthographe de langue numála*. 2nd revised edition. Yaoundé: SIL Cameroon.
- Boyd, Virginia. 2016a. *Dictionnaire baca – français*. Dallas, TX: SIL International. <https://www.webonary.org/baca/>.
- Boyd, Virginia. 2016b. *Dictionnaire elip – français*. Dallas, TX: SIL International. <https://www.webonary.org/elip/>.
- Boyd, Virginia. 2016c. *Dictionnaire mbule – français*. Dallas, TX: SIL International. <https://www.webonary.org/mbule/>.
- Boyd, Virginia. 2016d. *Dictionnaire mmala – français*. Dallas, TX: SIL International. <https://www.webonary.org/mmala/>.
- Boyd, Virginia. 2016e. *Dictionnaire tuki – français*. Dallas, TX: SIL International. <https://www.webonary.org/tuki/>.
- Boyd, Virginia. 2016f. *Dictionnaire yangben – français*. Dallas, TX: SIL International. <https://www.webonary.org/yangben/>.
- Boyd, Virginia. 2019a. *Dictionnaire gunu – français*. Dallas, TX: SIL International. <https://www.webonary.org/gunu/>.
- Boyd, Virginia. 2019b. *Dictionnaire tunen – anglais*. Dallas, TX: SIL International. <https://www.webonary.org/nen/>.
- Boyd, Virginia. 2020. *Dictionnaire yambeta – français*. Dallas, TX: SIL International. <https://www.webonary.work/yambeta/>.
- Boyd, Virginia Lee. 2015. *The phonological systems of the Mbam languages of Cameroon with a focus on vowels and vowel harmony*. Leiden: Leiden University. (Doctoral dissertation).
- Bufe, E. 1910–1911. Die Dualasprache in ihrem Verhältnis zu den Dialekten des Nordgebiets der Station Bombe. *Zeitschrift für Kolonialsprachen* 1(1). 25–36.
- Cheucle, Marion. 2014. *Étude comparative des langues makaa-njem (bantu A80) : phonologie, morphologie, lexique – Vers une reconstruction du proto-A80*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Chia, Agnes F. S. 1993. *Phonology of Bubia*. Yaoundé: University of Yaoundé. (MA thesis).
- de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87–162 + supplementary material online.
- Dinkelacker, Ernst. 1914. *Wörterbuch der Duala – Sprache* (Abhandlungen des Hamburgischen Kolonialinstituts. Reihe B. Völkerkunde, Kulturgeschichte und Sprachen 10). Hamburg: L. Friederichsen & Co. (L. & R. Friederichsen).
- Dorsch, Heinrich. 1911–1912a. Vocabularium der Nkosi-Sprache (Kamerun). Teil 1. *Zeitschrift für Kolonialsprachen* 2(3). 161–193.

- Dorsch, Heinrich. 1911–1912b. Vocabularium der Nkosi-Sprache (Kamerun). Teil 2. *Zeitschrift für Kolonialsprachen* 2(4). 324–330.
- Dorsch, Heinrich. 1912–1913. Vocabularium der Nkosi-Sprache (Kamerun). Teil 3. *Zeitschrift für Kolonialsprachen* 3(1). 34–62.
- Downing, Laura J. 2005. On the ambiguous segmental status of nasals in homorganic NC sequences. In Marc van Oostendorp & Jeroen van de Weijer (eds.), *The internal organization of phonological segments*, 183–216. Berlin: Mouton de Gruyter.
- Dugast, Idelette. 1967. *Lexique de la langue tûnen (parler des Banen du Sud-Ouest du Cameroun)* (Langues et littératures de l’Afrique noire 2). Paris: Librairie C. Klincksieck.
- Dugast, Idelette. 1971. *Grammaire du tûnen* (Langues et littératures de l’Afrique noire 8). Paris: Librairie C. Klincksieck.
- Ebobissé, Carl. 2015. *Sawabantu: Eine vergleichende Untersuchung der Küstensprachen Kameruns (Bantu A.20 und A.30)* (Schriften zur Afrikanistik / Research in African Studies 24). Frankfurt am Main: Peter Lang.
- Edika, E. Solange F. 1990. *Esquisse phonologique du bakoko (parler de Dibombari)*. Yaoundé: University of Yaoundé. (MA thesis).
- Ehret, Christopher. 1998. *An African classical age: Eastern and Southern Africa in world history, 1000 BC to 400 AD*. Charlottesville, VA: University Press of Virginia.
- Elias, Philip, Jacqueline Leroy & Jan Voorhoeve. 1984. Mbam-Nkam or Eastern Grassfields. *Afrika und Übersee: Sprachen, Kulturen* 67(1). 31–107.
- Elimelech, Baruch. 1976. Noun tonology in Kombe. In Larry M. Hyman (ed.), *Studies in Bantu tonology* (Southern California Occasional Papers in Linguistics 3), 113–130. Los Angeles, CA: Department of Linguistics, University of Southern California.
- Ernst, Urs. 1996a. *Alphabet et orthographe du Kakɔ (Kakɔ-est)*. Yaoundé: SIL Cameroon.
- Ernst, Urs. 1996b. *Alphabet et orthographe du mbonjɔ (kakɔ-ouest)*. Yaoundé: SIL Cameroon.
- Essono, Jean-J. Marie. 2000. *L’ewondo : langue bantu du Cameroun : phonologie – morphologie – syntaxe*. Yaoundé: Presses de l’Université catholique d’Afrique centrale.
- Ewane Etame, Jean & Robert Hedinger. 2017. *Lexique mkaa’ – français et index français – mkaa’ de Jean Ewane Etame*. 2nd revised edition. Yaoundé: SIL Cameroon.
- Fernandez, Galilea L. 1951. *Diccionario español – kômbé*. Madrid: Instituto de Estudios Africanos.

- Festen, Bradley & Paul Murrell. 2020. *Dictionnaire mpyemo – français*. SIL International. <https://www.webonary.org/mpyemo/>.
- François, Alexandre. 2014. Trees, waves and linkages: Models of language diversification. In Claire Bowerman & Bethwyn Evans (eds.), *The routledge handbook of historical linguistics*, 161–189. London: Routledge.
- Friesen, Dan T. 2002. *Oroko orthography development: Linguistic and sociolinguistic factors*. Grandforks, ND: University of North Dakota. (MA thesis).
- Galley, Samuel. 1964. *Dictionnaire fang–français et français–fang, suivi d’une grammaire fang*. Neuchâtel: Éditions Henri Messeliller.
- Gowlett, Derek F. 1989. The parentage and development of Lozi. *Journal of African Languages and Linguistics* 11(2). 127–149.
- Greenberg, Joseph H. 1970. Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics* 36(2). 123–145.
- Grégoire, Claire & Jean Léonce Doneux. 1977. Une séquence vocalique à la finale de quelques thèmes en protobantou. *Africana Linguistica* 7. 159–199.
- Grimm, Nadine. 2015. *A grammar of Gyeli*. Berlin: Humboldt University of Berlin. (Doctoral dissertation).
- Grimm, Nadine. 2019. Implosives in Bantu A80? The case of Gyeli. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and description in African linguistics: Selected papers from the 47th Annual Conference on African Linguistics* (Contemporary African Linguistics 4), 135–153. Berlin: Language Science Press. DOI: 10.5281/zenodo.3367134.
- Grollemund, Rebecca. 2012. *Nouvelles approches en classification : application aux langues bantu du nord-ouest*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guarisma, Gladys. 2000. *Complexité morphologique – simplicité syntaxique. Le cas du bafia, langue bantoue périphérique (A 50) du Cameroun* (Société d’études linguistiques et anthropologiques de France 379). Paris: Peeters.
- Guarisma, Gladys & Christiane Paulian. 1986. Dialectométrie lexicale de quelques parlers bantou de la zone A. In Gladys Guarisma & Wilhelm J. G. Möhlig (eds.), *La méthode dialectométrique appliquée aux langues africaines*, 93–176. Berlin: Dietrich Reimer.

- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1953. *The Bantu languages of Western Equatorial Africa* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1962. A two-stage method of comparative Bantu study. *African Language Studies* 3. 1–24.
- Guthrie, Malcolm. 1967. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 1: The comparative linguistics of the Bantu languages*. Farnborough: Gregg Press.
- Guthrie, Malcolm. 1970a. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part II, Volume 3: A catalogue of Common Bantu with commentary*. Farnborough: Gregg International.
- Guthrie, Malcolm. 1970b. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part II, Volume 4: A catalogue of Common Bantu with commentary (continued)*. Farnborough: Gregg International.
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Hagège, Claude. 1967. Description phonologique du parler wori (sud-ouest du Cameroun). *Journal of West African Languages* 4(2). 15–34.
- Halle, Morris. 1997. On stress and accent in Indo-European. *Language* 73(2). 275–313.
- Hammarström, Harald. 2019. An inventory of Bantu languages. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 17–78. Milton Park, Abingdon: Routledge.
- Hawkinson, Ann Katherine. 1986. *Bakweri verb morphology*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).
- Heath, Daniel & Teresa Heath. 1982. *A phonology of the Məkaá language*. Yaoundé: SIL Cameroon.
- Heath, Teresa. 2003. Makaa (A83). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 335–348. London: Routledge.
- Hedinger, Robert. 1985. The verb in Akɔɔse. *Studies in African Linguistics* 16(1). 1–55.

- Hedinger, Robert. 1987. *The Manenguba languages (Bantu A.15, Mbo cluster) of Cameroon*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Hedinger, Robert & Sylvia Hedinger. 1977. *Phonology of Akó̄sē (Bakossi)*. Yaoundé: Summer Institute of Linguistics.
- Helmlinger, Paul. 1972. *Dictionnaire duala – français, suivi d'un lexique français – duala* (Langues et littératures de l'Afrique noire 9). Paris: Librairie C. Klincksieck (pour L'Université fédérale du Cameroun).
- Henson, Bonnie J. 2007. *The phonology and morphosyntax of Kol*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).
- Hock, Hans H. 1991. *Principles of historical linguistics*. Second edition, revised and updated (Mouton Textbook). Berlin: Mouton de Gruyter.
- Hombert, Jean-Marie. 1991. Quelques critères de classification des parlers fang. *Pholia* 6. 145–153.
- Honeybone, Patrick. 2008. Lenition, weakening and consonantal strength: Tracing concepts through the history of phonology. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition* (Studies in Generative Grammar 99), 9–92. Berlin: Mouton de Gruyter.
- Hyman, Larry M. 2001. Vowel harmony in Gunu. *Studies in African Linguistics* 30(2). 147–170.
- Hyman, Larry M. 2003a. Basaá (A43). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 257–282. London: Routledge.
- Hyman, Larry M. 2003b. Sound change, misanalysis, and analogy in the Bantu causative. *Journal of African Languages and Linguistics* 24(1). 55–90.
- Hyman, Larry M. 2013. Penultimate lengthening in Bantu: Analysis and spread. In Balthasar Bickel, Lenore A. Grenoble, David A. Peterson & Alan Timberlake (eds.), *Language typology and historical contingency* (Typological Studies in Language 104), 309–330. Amsterdam: John Benjamins.
- Hyman, Larry M. 2019. Segmental phonology. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 128–149. Milton Park, Abingdon: Routledge.
- Hyman, Larry M. & Edmond Biloa. 1992. Transparent low tone in Tuki. In Laura A. Buszard-Welcher, Jonathan Evans, David Peterson, Lionel Wee & William Weigel (eds.), *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society: Special Session on the Typology of Tone Languages*, 104–127. Berkeley, CA: Berkeley Linguistics Society, University of California.

- Hyman, Larry M. & John Merrill. 2016. Morphology, irregularity, and Bantu frication: The case of Lulamogi. In Jean Léo Léonard (ed.), *Actualités des néogrammariens* (Mémoires de la Société de Linguistique de Paris, Nouvelle Série 23), 139–157. Paris: Peeters.
- Isaac, Kendall M. 2014. *Noun classes in Lefa* (ALCAM 581). Yaoundé: SIL Cameroon.
- Ittmann, Johannes. 1939. *Grammatik des Duala (Kamerun) unter Mitarbeit von Carl Meinhof* (Beihefte zur Zeitschrift für Eingeborenen-Sprachen 20). Berlin: Dietrich Reimer.
- Jacquot, André. 1983. *Les classes nominales dans les langues bantoues des groupes B.10, B.20 et B.30 (Gabon-Congo)* (Travaux et documents de l'ORSTOM 157). Paris: Éditions de l'Office de la recherche scientifique et technique outre-mer.
- Janssens, Baudouin. 1986. Eléments de phonologie et de morphologie historique du basaa (bantou A43a). *Africana Linguistica* 10. 147–211.
- Janssens, Baudouin. 1991. Doubles réflexes apparents en ewondo ou les chassés-croisés de la dérivation. *Pholia* 6. 155–180.
- Janssens, Baudouin. 1993. *Doubles réflexes consonantiques : quatre études sur le bantu de zone A (bubi, nen, bafia, ewondo)*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- Kagaya, Ryohei. 1992. *A classified vocabulary of the Bakueri language* (Bantu Vocabulary Series 8). Tokyo: Institute for the Study of Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.
- Kelly, John. 1974. Close vowels in Fang. *Bulletin of the School of Oriental and African Studies* 37(1). 119–123.
- Kenmogne, Michel. 2000. *The lexical phonology of Bakoko*. Buea: University of Buea. (Doctoral dissertation).
- Kouoh Mboundja, Christian Josué. 2004. *Bàlòŋ (bantou A13) : description phonologique et morphologique*. Lausanne: University of Lausanne. (Doctoral dissertation).
- Kuperus, Juliana. 1985. *The Londo word: Its phonological and morphological structure* (Annals – Human Sciences 119). Tervuren: Royal Museum for Central Africa.
- Lamberty, Melinda. 2002. *A rapid appraisal survey of the Abo and Barombi speech communities. South West and Littoral Provinces, Cameroon*. Survey report Journal of Language Survey Reports 2002-075. Dallas, TX: SIL International.
- Lovstrand, Joseph. 2011. *Notes on Nyokon phonology (Bantu A.45, Cameroon)*. Yaoundé: SIL Cameroon.
- Maddieson, Ian. 1984. *Patterns of sounds* (Cambridge Studies in Speech Science and Communication). Cambridge: Cambridge University Press.

- Maho, Jouni F. 2003. A classification of the Bantu languages: An update of Guthrie's referential system. In Derek Nurse & Gérard Philippon (eds.), *The Bantu languages* (Language Family Series 4), 639–651. London: Routledge.
- Maho, Jouni F. 2009. *NUGL online: The online version of the new updated Guthrie list, a referential classification of the Bantu languages*. https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf.
- Mathaus, Njeck & Stephen C. Anderson. 2010. *Bakoko orthography guide*. Yaoundé: Cameroon Association for Bible Translation & Literacy.
- Medjo Mvé, Pither. 1997. *Essai sur la phonologie panchronique des parlers fang du Gabon et ses implications historiques*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Medjo Mvé, Pither. 2008. Les kakɔ et le kakɔ de la région de Bitam, données préliminaires sur une communauté et une langue du Gabon peu connues. In Daniel F. Idiata (ed.), *Éléments de description des langues du Gabon*, vol. 2 (Collections sciences humaines et sociales), 69–104. Libreville: Les Éditions du Cenarest.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeussen, Achiel Emiel. 1973. Comparative Bantu: Test cases for method. *African Language Studies* 14. 6–18.
- Meeussen, Achiel Emiel. 1976. Notes on tone in Bantu nominal stems. *African Languages / Langues Africaines* 2. 60–70.
- Meinhof, Carl. 1899. *Grundriß einer Lautlehre der Bantusprachen, nebst Anleitung zur Aufnahme von Bantusprachen. Anhang: Verzeichnis von Bantuwortstämmen* (Abhandlungen für die Kunde des Morgenlandes 11(2)). Leipzig: F.A. Brockhaus.
- Meinhof, Carl. 1889–1890. Das Verbum in der Isubu-Sprache: Versuch einer grammatischen Darstellung. *Zeitschrift für afrikanische Sprachen* 3. 206–234.
- Mékina, Émilienne-Nadège. 2012. *Description du fang-nzaman, langue bantoue du Gabon : phonologie et classes nominales*. Nancy: University of Lorraine. (Doctoral dissertation).
- Mickala Manfoumbi, Roger. 2005. Vocabulaire seki. Unpublished ms.
- Möhlrig, Wilhelm J. G. 1976. Guthrie's Beitrag zur Bantuistik aus heutiger Sicht. *Anthropos: Internationale Zeitschrift für Völker- und Sprachenkunde* 71(5–6). 673–715.
- Mongo, Raoul & Léonard-Albert Bolioki. 2012. *Kálata' ε káábáá kó kaakan tufuu tó yanpedá 1 / Syllabaire en langue yambeta 1*. 2nd revised edition. Bafia: YALICO.
- Monikang, Evelyn Neh. 1989. *Phonology of Mokpwe*. Yaoundé: University of Yaoundé. (MA thesis).

- Mous, Maarten. 2003. Nen (A44). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 283–306. London: Routledge.
- Mous, Maarten & Anneke Breedveld. 1986. A dialectometrical study of some Bantu languages (A. 40 – A. 60) of Cameroon. In Gladys Guarisma & Wilhelm J. G. Möhlig (eds.), *La méthode dialectométrique appliquée aux langues africaines*, 177–241. Berlin: Dietrich Reimer.
- Nassau, Robert H. 1892. *Mackey's grammar of the Benga-Bantu language*. New York, NY: American Tract Society.
- Nida, Eugene A. 1964. What is phonemics? In William A. Smalley (ed.), *Orthography studies: Articles on new writing systems*, 18–21. London: United Bible Societies.
- Njantcho Kouagang, Elisabeth. 2018. *A grammar of Kwakum*. Paris: Institut National des Langues et Civilisations Orientales. (Doctoral dissertation).
- Njock, Pierre E. 2005. *Dictionnaire basaa – français – anglais – allemand*. SIL International. <https://www.webonary.org/basaa/>.
- Nomaande Language Committee, Terri R. Scruggs, Carolyn Taylor & Patricia Wilkendorf. 2003. *Nomaande – French lexicon*. Yaoundé: SIL Cameroon.
- Nsuka-Nkutsi, François. 1980. Quelques réflexes du proto-bantou en punu. In François Nsuka-Nkutsi (ed.), *Éléments de description du punu* (Linguistique et sémiologie. Travaux du centre de recherches linguistiques et sémiologiques de Lyon), 129–178. Lyon: Université Lumière – Lyon 2.
- Nurse, Derek & Gérard Philippson. 1980a. Historical implications of the language map of East Africa. In Luc Bouquiaux (ed.), *L'expansion bantoue : actes du colloque international du CNRS, Viviers (France) – 4–16 avril 1977*, vol. 3 (Société d'études linguistiques et anthropologiques de France, Numéro spécial 9), 685–714. Paris: SELAF.
- Nurse, Derek & Gérard Philippson. 1980b. The Bantu languages of East Africa: A lexicostatistical survey. In Edgar C. Polomé & Charles P. Hill (eds.), *Language in Tanzania*, 26–67. London: Oxford University Press (for the International African Institute).
- Ollomo Ella, Régis. 2013. *Description linguistique du shiwa, langue bantu du Gabon : phonologie, morphologie, syntaxe, lexique*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Pacchiarotti, Sara & Koen Bostoen. 2022. Erratic velars in West-Coastal Bantu: Explaining irregular sound change in Central Africa. *Journal of Historical Linguistics* 12(3). 381–445.

- Pacchiarotti, Sara, Natalia Chousou-Polydouri & Koen Bostoen. 2019. Untangling the West-Coastal Bantu mess: Identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica* 25. 155–229.
- Patman, Keith E. & Clinton D. W. Robinson. 1989. *Règles d'orthographe pour la langue ngunu*. Yaoundé: SIL Cameroon.
- Paulian, Christiane. 1971. Esquisse phonologique du duala (République Fédérale du Cameroun). In André Jacquot, Christiane Paulian, Yves Monino & Paulette Roulon-Doko (eds.), *Etudes bantoues, Tome I* (Société d'études linguistiques et anthropologiques de France 25), 53–88. Paris: SELAF.
- Paulian, Christiane. 1975. *Le kukuya, langue teke du Congo : phonologie – classes nominales* (Société d'études linguistiques et anthropologiques de France 49–50). Paris: SELAF.
- Paulian, Christiane. 1986. Les parlers yambasa du Cameroun (Bantou A. 62), dialectométrie lexicale. In Gladys Guarisma & Wilhelm J. G. Möhlig (eds.), *La méthode dialectométrique appliquée aux langues africaines*, 243–279. Berlin: Dietrich Reimer.
- Perrin, Mona. 1986. *Report on a brief study of Balom*. Yaoundé: SIL Cameroon.
- Philippson, Gérard. 2018. Where exactly does Bobe (A31) fit in the Bantu genealogy? Paper presented at the 48th Colloquium on African Languages and Linguistics, Leiden University, 30 August – 1 September 2018.
- Philippson, Gérard. 1991. *Tons et accent dans les langues bantu d'Afrique orientale : étude comparative typologique et diachronique*. Paris: Université Paris V 'René Descartes'. (Doctoral dissertation).
- Philippson, Gérard & Rebecca Grollemund. 2019. Classifying Bantu languages. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 335–354. Milton Park, Abingdon: Routledge.
- Prittie, Rebecca. 2002. *Grammar sketch of Nulibie*. Yaoundé: SIL Cameroon.
- Puech, Gilbert. 1989a. L'opposition implosives/mi-voisées en bekwel. Paper presented at the 19th Colloquium on African Languages and Linguistics, Leiden University.
- Puech, Gilbert. 1989b. Les constituants suprasyllabiques en jíwá (bantou A-80). *Pholia* 4. 217–227.
- Puech, Gilbert. S.d. Seki vocabulary. Unpublished ms.
- Pullum, Geoffrey K. 1976. The Duke of York gambit. *Journal of Linguistics* 12(1). 83–102.
- Richardson, Irvine. 1957. *Linguistic survey of the Northern Bantu borderland*. Vol. 2. London: Oxford University Press (for the International African Institute).

- Ricquier, Birgit & Koen Bostoen. 2008. Resolving phonological variability in Bantu lexical reconstructions: The case of ‘to bake in ashes’. *Africana Linguistica* 14. 109–149.
- Rurangwa, Innocent M. 1989. Enquête linguistique sur le lubi, langue bantou insulaire de Guinée Equatoriale : phonologie et système des classes. In Théophile Obenga (ed.), *Les peuples bantou : migrations, expansion et identité culturelle*, 77–100. Libreville; Paris: Ciciba; L’Harmattan.
- Schadeberg, Thilo C. 1994. Spirantization and the 7-to-5 vowel merger in Bantu. *Belgian Journal of Linguistics* 9. 73–84.
- Stewart, John M. 1973. The lenis stops of the Potou Lagoon languages and their significance for pre-Bantu reconstruction. In Mary E. Kropp Dakubu (ed.), *Papers in Ghanaian linguistics* (Transactions of the Linguistic Circle of Accra 2), 1–49. Legon: Institute of African Studies, University of Ghana.
- Stewart, John M. 1975. Lenis stops and the origins of Volta-Comoe consonant mutation. In Robert K. Herbert (ed.), *Proceedings of the Sixth Conference on African Linguistics, held at the Ohio State University, Columbus, April 12–13, 1975* (Working Papers in Linguistics 20), 16–31. Columbus, OH: Department of Linguistics, Ohio State University.
- Stewart, John M. 1983. The high unadvanced vowels of proto-Tano-Congo. *Journal of West African Languages* 13(1). 19–36.
- Stewart, John M. 1989. Fortis/lenis and vowel length in Proto-Bantu. *Journal of African Languages and Linguistics* 11(1). 45–88.
- Stewart, John M. 1993. The second Tano consonant shift and its likeness to Grimm’s law. *Journal of West African Languages* 23(1). 3–39.
- Stewart, John M. 2002. The potential of Proto-Potou-Akanic-Bantu as a pilot Proto-Niger-Congo, and the reconstructions updated. *Journal of African Languages and Linguistics* 23(2). 197–224.
- Stewart, John M. & Helene Van Leynseele. 1979. Underlying cross-height vowel harmony in Nen (Bantu A.44). *Journal of African Languages and Linguistics* 1(1). 31–54.
- Teil-Dautrey, Gisèle. 1991a. Conditionnement tonal de certains “doubles réflexes” en basaa (A 43a). *Pholia* 6. 181–190.
- Teil-Dautrey, Gisèle. 1991b. *Evolution phonétique d’une langue bantoue du nord-ouest : le bàsàà A.43a*. Lyon: Université Lumière – Lyon 2. (MA thesis).
- Teil-Dautrey, Gisèle. 2004. *Lexiques proto-bantous : étude des cooccurrences segmentales et supra-segmentales*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Thornell, Christina & Yasuko Nagano-Madsen. 2004. Preliminaries to the phonetic structure of the Bantu language Mpiemo. *Africa & Asia* 4. 163–180.

- Tsala, Théodore. 1956. *Dictionnaire ewondo – français*. Lyon: Emmanuel Vitte.
- Van de Velde, Mark. 2006. *A description of Eton: Phonology, morphology, basic syntax and lexicon*. Leuven: Katholieke Universiteit Leuven. (Doctoral dissertation).
- Van Leynseele, Helene & John M. Stewart. 1980. Harmonie consonantique en pré-nen. In Luc Bouquiaux (ed.), *L'expansion bantoue : actes du colloque international du CNRS, Viviers (France) – 4–16 avril 1977*, vol. 2 (Société d'études linguistiques et anthropologiques de France, Numéro special 9), 387–400. Paris: SELAF.
- Wang, William S-Y. 1969. Competing changes as a cause of residue. *Language* 45(1). 9–25.
- Watters, John R. & Jacqueline Leroy. 1989a. Bantoid overview. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 401–429. Lanham, MD: University Press of America.
- Watters, John R. & Jacqueline Leroy. 1989b. Southern Bantoid. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 430–449. Lanham, MD: University Press of America.
- Wéga Simeu, Abraham. 2016. *Grammaire descriptive du pôlri : éléments de phonologie, morphologie et syntaxe* (Grammatical Analyses of African Languages 51). Cologne: Rüdiger Köppe.
- Wills, Jeffrey. 2022. Sorting out Proto-Bantu *j. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 59–101. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575817.
- Yates, Anthony D. & Sam Zukoff. 2018. The phonology of Anatolian reduplication. *Synchrony and diachrony. Indo-European Linguistics* 6(1). 201–270.
- Yukawa, Yasutoshi. 1992. A tonological study of Yambasa verbs. In Ryohei Kagaya (ed.), *Studies in Cameroonian and Zairean languages* (Bantu Linguistics 3), 1–46. Tokyo: Institute for the Study of Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.

Chapter 2

Sorting out Proto-Bantu *j

Jeffrey Wills

Ukrainian Catholic University

The most problematic of the consonants that Meeussen reconstructed for Proto-Bantu (PB) phonology is *j, for which Guthrie used both *j and *y. Earlier generations had also sometimes omitted either in favour of vowel-initial roots. Recent progress in establishing a solid family tree of the Bantu languages allows the evidence to be re-evaluated based on phylogenetic significance, especially with the help of more data from the North-Western Bantu branches. It has long been recognised that Meeussen's *j has various outcomes throughout the Bantu area based on phonological or morphological environments. The primary method of this chapter is to sort out the evidence for PB *j into different phonological and morphological environments, and then consider possible scenarios for reconstruction of those categories. In most roots with initial *j, there is no support for a PB stop and an initial vowel or glide should be reconstructed. That includes common verbs like *(y)àd 'spread' and *(y)óm 'be dry', and nouns like *ícò 'eye' or *óbà 'sun'. Most modern reflexes in /z/ or /j/ are the result of developments at morpheme boundaries after the PB stage. Both *ny and *nj/nz are reconstructed as distinct phonemes.

1 Introduction

In his *Bantu Grammatical Reconstructions*, Meeussen (1967: 83) put forth the following Proto-Bantu (PB) reconstructions for simple consonants (with a parallel series of pre-nasalised versions of each stop):

m	n	ɲ	
b	d	j	g
p	t	c	k



The most problematic of the consonants was *j, which had been in flux for a century, and Meeussen noted that one might just as well use the notation “/z/ or /y/ instead of /j/”. A generation later, Schadeberg (2003: 146–147) described the continuing uncertainty: “Guthrie (1967–71) distinguishes initial *j from *y, but BLR2 (Coupez et al. 1998) recognises only *j to the exclusion of vowel-initial stems. I regard the two as allophonic but the question needs re-evaluation.” Increasingly, there have been doubts about *j in some lexemes and an inclination to return to at least some vowel-initial stems. This chapter goes further in that direction to argue for reconstructing vowel-initial roots more extensively in PB. After an introduction on the history of the scholarship and some methodological issues, the currently reconstructed *j is systematically examined in the relevant phonological and morphological environments.

1.1 History of the problem

Why did early Bantu scholars reconstruct *j in the first place? The topic was mostly handled in handbooks like Meinhof (1899; 1910) or Homburger (1913), or in discussions of individual languages and words. The stems which are today reconstructed with *j in BLR3 (Bastin et al. 2002) were variously listed by Meinhof et al. (1932: 187–196) with three symbols: *y, *ȳ, and *ø.¹ For example, *yala ‘spread out’, *mu-yaka ‘year’, *yîno ‘tooth’, *yanî ‘leaf’, *yoyû ‘elephant’, and *ato ‘boat, canoe’. Meinhof’s effort to identify which root had which phoneme was complicated by his significant reliance on South Bantu languages where *g > ø is widespread. Homburger sorted out some of these problems but reconstructed PB forms with only initial palatals or velars without much explanation, although her lists of reflexes gave evidence for some vowel-initial roots.

To clarify this situation, in 1954, André Coupez wrote the first article ever focused on the question of PB *j – a mere 3-page note with wordlist. His explicit goal was to correct Meinhof as well as Bourquin (1923). He based his analysis on Yao P21 and Kongo H16, as the only well-attested languages which have regular ‘positive reflexes’ for *g and *j in verb-initial and intervocalic positions. His choice of those languages was unfortunate for elucidating *j because they often introduce hiatus-fillers in those positions. Coupez concluded that at an

¹Approximate orthographic comparisons are: y (Meinhof, Bourquin) ≈ g^w, g’ (Homburger) ≈ g (Greenberg, Guthrie, Meeussen, BLR); ȳ (Meinhof, Bourquin) ≈ g^w, g’ (Homburger) ≈ z (Greenberg) ≈ j, (j) (Meeussen) ≈ j, y (Guthrie) ≈ j (BLR); ng (Meinhof, Bourquin) ≈ ng’ (Homburger) ≈ nj (Meeussen) ≈ nj, ny (Guthrie, BLR). But of course, these authors do not always reconstruct the same series in specific lexemes.

early stage **j* had been lost at the beginning of many nouns,² and there simply were not regularly any verbal stems with initial vowels. He stated his support for Homburger's "conclusions" and gave a wordlist with **g* and **j*, without any vowel-initial PB nouns or verbs.³

At that same time, Malcolm Guthrie (1953) had revived a three-way distinction. In addition to **g* and **j* (e.g. **jàdà* 'hunger'), he added *y* (e.g. **yúdu* 'nose', **yímb* 'sing'). But Meeussen & Tucker (1955: 170, 175–177) writing on Ganda (which has many *j*-initial verbs) rejected Guthrie's **y* as not yet justified and affirmed the unitary **j* of Coupez and Homburger, even adding initial **j* to some of Bourquin's (1923) vowel-initial reconstructions for PB. In *Bantu Grammatical Reconstructions*, Meeussen (1967) generally followed Coupez with **j* as the default (e.g. **jojo* 'life', **jáka* 'year', **júba* 'sun'), but **∅* was allowed to return at the beginning of some verbs (e.g. **igad* 'shut' and **ánik* 'spread in the sun'). The parenthetical consonants in words like **(g)amb* 'speak', **(j)jib* 'know', and **(b)óba* 'fear' further signalled an openness to initial root vowels. This style was continued by Meeussen's (1969) *Bantu Lexical Reconstructions*, already with some changes in particular words, e.g. **(j)áka* 'year', **jji-b* 'know', **ícó* 'your father'. Meeussen did not reconstruct PB semi-vowels but he noted their similarity to contexts with his parenthetical **(j)*.

Guthrie's large dataset (finally published in 1970) continued his approach from the 1950s. He could not confirm a unitary **j*, so he used both PB **j* and **y* (often for the same lexeme) and thought it likely that "there was a mutation **j* » **Y*" and that "**G* » **Y* has to be postulated for most of the **g*/**y* pairs" (1967: 114). This allowed him to have consistent CV 'units' and roots with initial consonants.⁴ Guthrie's idiosyncratic approach with multiple proto-forms made it a difficult path for others to follow – certainly for Meeussen (1973: 10) whose review of Guthrie included: "On the whole, it appears that there is no real ground for setting up **j* and **y* as two distinct correspondences."

BLR2, with a team led by Coupez, maintained his approach with **j* everywhere (without parentheses). As BLR3 (Bastin et al. 2002) notes in the online

²Coupez (1954: 158): "Sans doute **j* s'est-il amui de bonne heure à l'initiale des thèmes nominaux: les thèmes nominaux qui nous sont attestés avec voyelle initiale seraient en réalité des thèmes en **j*." It was also Coupez who introduced a rather vague sense of unspecified allophones (ibid. 157).

³Greenberg (1969: 430) followed this line, pointing out problems in Meinhof's correspondences: "Nor has Meinhof explained any of these deviations in the text of his work. It is now generally accepted that, as first suggested by Homburger 1913, there are two proto-phonemes involved, which are usually symbolized **g* and **y*."

⁴Guthrie (1967: 44, §42.11): "It is from these various unit features that the patterns are made up, and the principal ones involved prove to be C_1V_1 , and C_2V_2 [...]"

legend:⁵ “Guthrie’s *j and *y have been merged into *j. The problems regarding *j/y/zero are far from being resolved.” Subsequently, two standards for reconstruction were in play (Guthrie and BLR), both without initial vowels except in functional morphemes. But scholars periodically pointed out the case for initial vowels in specific roots or specific groups.⁶

In recent years, an increase in knowledge about the North-Western languages has allowed major advances in our understanding of the Bantu family tree. This chapter has taken advantage of these developments to give greater weighting to data from zones A and B. The resulting analysis supports a substantial number of PB reconstructions with an initial vowel or glide rather than a unitary *j.

1.2 Sources, method, and terminology

Reconstructing the phonology of a proto-language at a stage over 4000 years before any record of its descendant languages has significant challenges, and in the case of Bantu there are not even many intermediate reconstructions of late branches. Accordingly, recourse must be made to the primary lexical data in over 400 modern languages (many only partly documented), and then applying a judicious method of sorting out idiosyncrasies, proposing an inevitably simplified starting point, and elucidating the principal developments. One must admire the immense progress made by the early scholars of Bantu, who had developed a respectable grammar and 800-root lexicon of PB by the 1920s. But that was initially based on only a couple dozen languages (eventually becoming over 50), of which some were very closely related and most were from the eastern and southern regions.

Guthrie The first thorough Bantu lexical survey, including substantial attention to North-Western Bantu languages, was the monumental work of Malcolm Guthrie (1967-71). It remains the largest set of comparative data, listing reflexes

⁵The online version of BLR3 is available from: https://www.africanmuseum.be/en/research/discover/human_sciences/culture_society/blr (database last updated on November 6, 2005). Note that BLR3 uses the symbols i and r instead of the pair ÿ and i used by Guthrie and Meeussen. Likewise BLR has u and v instead of y and u.

⁶For example, Creissels (1999: 304): “Tswana data clearly supports the reconstruction of two different types of initials corresponding roughly to Guthrie’s *y and *j”, and he felt that the observed reflexes of one type supported “the hypothesis of the (relatively) ancient absence of any initial consonant.” Bostoen (2019: 311–312): “If one admits the existence of vowel-initial noun stems in PB, it is enough to reconstruct just *j and not *y.” More fully in Teil-Dautrey (2004: 161–192). See also Bulkens (2009: 29–34, written 1997), Bostoen (2009: 115), Bostoen & Bastin (2016: 14–15).

of over 2000 “comparative series” of “Common Bantu” roots and stems from hundreds of languages across all zones, including a systematic sampling of 29 “test languages”. In Guthrie’s system, each “comparative series” (C.S.) is represented by a form with a prefixed asterisk (the usual mark for an artificial construct or reconstruction, although Guthrie is explicit that they are not reconstructions).⁷ In addition to the five test languages from Guthrie’s North-Western zones ABC, there are 70 other languages in those zones which he cites ten or more times. However, Guthrie does not identify sources or informants, which is a problem for determination of speech variety or verification of specific forms since later published sources do not always confirm his data. Nevertheless, Guthrie is currently the only dataset with reflexes for a large number of lexical items in a large number of Bantu languages. Unless otherwise noted, examples below come from his data and are cited using his orthography.

Grollemund Dataset The other lexical dataset to which I will sometimes refer is that accompanying Grollemund et al. (2015), collected from published sources and fieldwork for 409 Bantu and 15 Bantoid languages.⁸ The resulting dataset is notable for its geographical range and depth, including 150 languages in the North-Western zones ABC. Unfortunately, it is limited to up to 100 basic lexical items (meanings), only a few of which concern PB *j.⁹

Bantu Lexical Reconstructions (BLR) The most complete set of lexical reconstructions is provided by the *Bantu Lexical Reconstructions* database at the Royal Museum for Central Africa and is based on a century of work by various scholars. This online database (current version: BLR3) is not a reconstruction of PB but rather a toolkit of reconstructions of lexemes of various Bantu language

⁷Guthrie used the word “reconstruction” occasionally in 1967 regarding Meinhof’s work but avoided it with regard to his own PB X “stems” or “items”. Guthrie takes pains to explain that his “starred forms are in no sense reconstructions of presumed ancestor items” (Guthrie 1965: 43). Rather, they are just “symbolic representations” of “sets of recurrent patterns” (Guthrie 1967: 19, §23.11, 21, §24.11), which become fodder for a process of analysing and attributing related comparative series to PB lects.

⁸This dataset is available from: <http://www.evolution.reading.ac.uk/DataSets.html>. It is an expanded version of Grollemund (2012), a study of about 200 North-Western Bantu and Bantoid languages using a modified version of the wordlist for *Atlas Linguistique du Gabon* (ALGAB).

⁹Another useful dataset is that collected for Bastin et al. (1999), which has 93 meanings from 335 languages, but the Grollemund Dataset often includes the earlier dataset and has fuller zone coverage (with A10, G60, P10, as well as Jarawan). The earlier Bastin et al. (1999) dataset is available from: https://www.africamuseum.be/nl/research/discover/human_sciences/culture_society/lexicostatistic-study-bantu-languages.

groupings and historical stages with varying reliability. For each reconstruction, it provides no reflexes or mentions of specific languages, only zones based on the sources in its bibliography.¹⁰ For our purposes, BLR3 lists over a thousand forms with *j from various time depths, so our focus will be on its 183 “Main” entries with *j in C₁ position and 25 more with *j in C₂ position. I have usually also provided zone information, because only about 2/3 of the “Main” reconstructions have descendants in zones AB, and some which do are not labelled as “Main”. Throughout, I will be using BLR3 reconstructions (which uses only *j), although the comparative data discussed often comes from Guthrie, whose C.S. use *j and *y.

Reconstruction based on parsimony and the Bantu phylogenetic tree Historical reconstruction is based on parsimony (or economy). We propose ancestral states requiring the fewest independent changes needed to derive later reflexes. For this process, we must have languages structured by a reliable family tree, i.e. a phylogeny. One of the major fruits of the half-century since Meeussen’s *Bantu Grammatical Reconstructions* is the determination of a basic family tree for the Bantu language group.¹¹

In Grollemund et al. (2015: Fig. 1 and 2), the evolution of the Bantu languages is graphed in a consensus time tree and a map of migration routes. Although more refinement needs to be done at lower levels, the progressive “backbone” of the tree and major branches is statistically very solid. Node 1 on that tree is the Bantu common ancestor treated here as PB, and then a series of binary splits (repeated 7 to 12 times) leads to a detailed structure with over 400 terminal nodes (the modern languages). In theory, each split is the result of innovations distinctive to one branch or the other, and it is the accumulation of these innovations which marks the divergence from the ancestral language. But the quantity and quality

¹⁰BLR2’s system of *fiabilité* ‘reliability’ had some advantages, but the BLR2 version is no longer supported and the current BLR3 has useful grouping and numerous corrections of details so it was used for this chapter. The history and method of BLR is described by Bostoen & Bastin (2016).

¹¹“[F]rom a purely classificatory point of view, the various trees published over the last 15 years or so by and large agree in their results” (Philippson & Grollemund 2019: 347). Ideally, a family tree classifies languages based on all linguistic changes, both lexical and non-lexical, which are assessed in various ways. Since Bantu is a fairly recent family with much internal contact, lexical and non-lexical innovations sometimes give conflicting isoglosses. The most recent non-lexical analysis (Nurse & Philippson 2003), based on thirty phonological and morphological features, proposes several historical scenarios but does not propose a tree. Accordingly, this chapter follows the most recent and detailed tree based on lexical innovation, being Grollemund et al. (2015).

of innovations vary between splits, and the Bantu phylogenetic tree is scaled for time not divergence, so the *depth* (number of levels a clade is from node 1) is merely a useful approximation of how close or far the clade is to the *root* (the proto-language).

The early stages of this phylogenetic tree can be visualised in Figure 1 with names of major branches and their relevant language zones.¹²

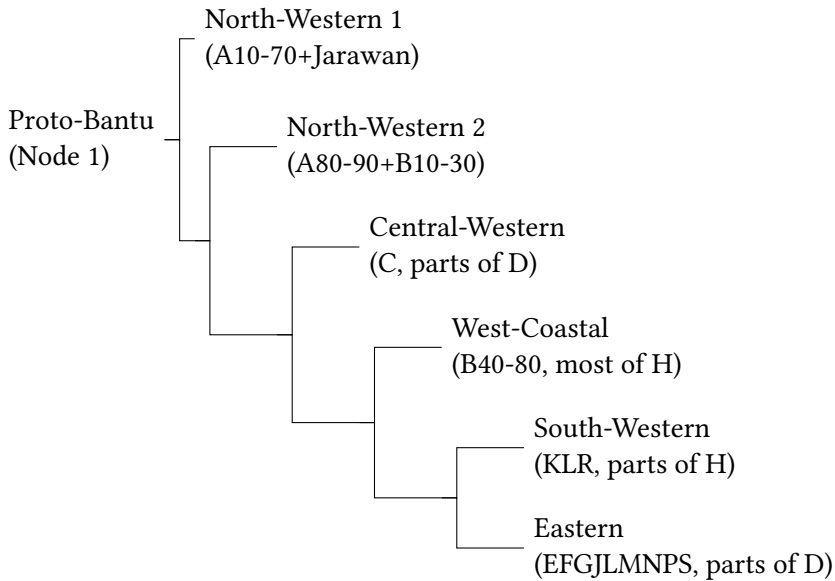


Figure 1: Simplified divisions of the Bantu phylogeny in Grollemund et al. (2015)

Our method is to work back from the modern languages, reconstructing ancestral forms for these major clades, and then proceeding to the nodes closer to the root. In general, we find that these major clades exhibit an internal unity in their reflexes of **j* that allows us to generalise at those levels, despite some inevitable innovations of a few languages among the dozens or hundreds in each branch.

Parsimony (the least number of changes) depends on the placement and distribution of the data in the structure of the tree. If an entire major branch has a distinctive form with **j* contrary to other branches, then it is possibly an innovation but possibly also a relic that escaped early changes in the other major

¹²From the detailed time tree in Grollemund et al. (2015: Fig. 1), I have collapsed three small neighbouring branches into North-Western 2 and three small neighbouring branches into South-Western.

branches, and so it must be studied seriously. But this is rarely the case in this chapter. Usually, the minority form (e.g. *zum* ‘be dry’ instead of the much more common *um*) is dominant in no branch, but rather is distributed across only a few languages in a few branches. So, it is likely to be a sign of innovations at later stages of Bantu development—for if the distinctive form belonged to PB, dozens and dozens of changes would be needed at multiple levels of the tree to account for the much greater number of languages lacking that distinctive form. We will often see that pattern of a few scattered innovations for **j*, indicating fairly recent developments. Of course, it is always theoretically possible that the scattered minority forms preserve an archaic heterogeneity, but then instead of a regional concentration we would require a concentration of the minority forms according to some original allophony or allomorphy.

When the evidence from both the North-Western branches is in agreement, it has great weight because these branches dominate the first splits in the tree. So, strong evidence from the North-Western branches and some currency in other major branches will make a good case for a PB reconstruction. On the other hand, the great majority of documented Bantu languages are in the Eastern branch, a clade which is several levels deep, and any reconstruction at that level must be reconciled with South-Western and West-Coastal (also called West-Western) before it can be given consideration for reconstruction at a higher level. In certain cases, a lexeme is not attested in all major branches, but any reconstruction at one or two levels below PB will be considered to be ‘early’ Bantu, i.e. early enough to be proposed as a candidate for PB but obviously not confirmable as such without support in some other way.

There are, however, two issues that must always be considered along with the phylogenetic approach: contact phenomena and directionality or naturalness of a sound change.

Contact phenomena across branches, which can create changes that are not independent innovations. This is particularly a concern in the North-Western regions of the Bantu domain where dozens of small languages belonging to different branches are geographically adjacent. So, although the North-Western clades have a privileged place in the phylogenetic tree, it is important to support reconstructions in those clades with at least some Bantu branches that are far enough away to discount an areal feature or borrowing of lexemes. Likewise, evidence beyond Narrow Bantu can support PB reconstructions, so relevant Bantoid data from Guthrie and the Grollemund Dataset will be cited.

Directionality or *naturalness* of a sound change, which could lead us to prefer one variant over another. In the case of PB **j*, the most common reflexes are null (\emptyset) or glides (*y*, *w*), but sometimes stops, fricatives or affricates (*j*, *z*, *ʒ*, *dʒ*) are seen.

Weakening (lenition) is the common direction for consonantal sound change, but the strengthening of glides is so common across languages of the world that it has also been argued to be the result of articulatory pressures. In fact, in the systematically compiled AlloPhon database, the strengthening of glides to fricatives is more common than the contrary (12 processes to 8).¹³ Furthermore, the particular strengthening of palatal glides is attested in a dozen language families beyond the database. For example, the initial glide in the Latin month *Januarius* becomes the fricative /ʒ/ in French *janvier* and the affricate /dʒ/ in Italian *gennaio* but disappears in Spanish *enero*.¹⁴ Cross-linguistically, the most common environments for palatal glide strengthening are at a word or syllable onset, and before a high and/or front vowel—both environments where PB *j is most common.

In short, there is some basis for preferring the reconstruction of glides to fricatives, but our default will be to follow parsimony and the usual Comparative Method without assuming a strong natural direction for change one way or the other.

Zones Guthrie’s coding of languages by letter and number, based mainly on geographical zones, has remained standard for identification. But with an increasingly solid family tree, Bantu historical linguists can now group data based on phylogenetic significance, with an emphasis on historical branches rather than geography. Accordingly, the symbol “+” here indicates additional zones, that is to say, “ABDE+” is a shorthand for zones A, B, D, E and some further letter(s). This abbreviation is used partly to save space but also to reduce reliance on Guthrie’s geographical zones as meaningful indicators of a PB ancestry. A lexeme solidly attested in zones AB and E (or any Eastern zone) already implies the first eight or more branchings and 2000 years of geographical spread. An item only attested in zones ABDG is just as likely to have been present in PB as one found in all 16 zones ABCDEFGHJKLMNPRS, although evidence from multiple zones may improve the quality of certain features of the reconstruction or demonstrate the stability of a word in the lexicon.

¹³Bybee & Easterday (2019) describe the data collection and provide examples. For Romance and Basque examples, see Hualde (2011: 2232). For more on Spanish palatal fortition, see Baker & Wiltshire (2003). Meeussen & Tucker (1955: 174–175) noted that the development of Ganda JE15 *ggyá* ‘new’ < **hya* < *PB *pia* “exactly parallels” the glide hardening in Old Norse *tveggja* < Proto-Germanic **twa-jē* ‘of two’. Ganda also has *-jjwa* < **hwa*. In modern German, the initial [j] in words like *Jahr* ‘year’ surfaces as an obstruent in various regional varieties, e.g. [ʒ] in the Mecklenburg dialect and [g] in a variety of Thuringian (Hall 2014: 257–262).

¹⁴Likewise, in medial position, Latin *maior* ‘greater’ > Italian *maggiore*, and Latin *ego* > Vulgar Latin **eo* > Spanish [jo, dʒo]. For initial Indo-European *y > Greek ζ [z, dz], see Sihler (1995: 187–190).

1.3 Outline

The primary method of this chapter is to “sort out” the evidence for PB **j*: first, into different phonological and morphological environments, and then into possible scenarios for reconstruction. Proposals (from BLR, Guthrie) for PB **j* and **y* will be tested using lexical data (from Guthrie, Grollemund, etc.), organised by a phylogenetic tree (from Grollemund et al. 2015).

Procedurally, let us begin by accepting the main reconstructions written with the symbol **j* in BLR3, and then try to elucidate what values they might have had. PB **i* and **n* tend to condition the evolution of subsequent consonants, so three environments can be distinguished:

Group 1: **j* not preceded by **i* or **n*

- Initial **j* in noun stems, e.g. **játò* ‘canoe’
- Initial **j* in verb stems, e.g. **ját* ‘split’
- Medial **j* in noun or verb stems, e.g. **jòjì* 3 ‘belly’

Group 2: **j* preceded by **i*

- Initial **j* in class 5 nouns, e.g. **jàdà* ‘rubbish-heap’, **jìcò* ‘eye’
- Initial **ji-C* and **jij*, e.g. **jìjìb* ‘know’, **jìjì* 6 ‘water’

Group 3: **j* preceded by **n*

- Nouns supporting PB **ny*, e.g. **jóká* ‘snake’
- Nouns supporting PB **nj*, e.g. **jògù* ‘elephant’
- Nouns with mixed classes, e.g. **jìkì* 9/10 ‘bee’ & 14 ‘honey’

So, first to be considered is **j* in the most neutral environment, i.e. at the beginning or middle of roots without a major conditioning factor. Then an examination of the consequences of the two major conditioning factors: a preceding *i* or a preceding nasal. Most roots only occur with **j* in one of these environments, but it will be useful to see what can be learned from those roots with allomorphic variants.

We will go through these environments in order, but in a summary fashion. Our goal is not to be exhaustive but rather to examine a few samples of each category as case studies and consider the issues the category presents.

2 *j not preceded by *i or *n

2.1 Unconditioned initial *j in noun roots

Our first category is one of the easiest: *j reconstructed at the beginning of nominal roots in classes where a CV- prefix does not generally provide an environment conditioning a change.¹⁵

For example, *jâtò 14 ‘canoe’ (BLR 3252) is an old and widespread root, attested in all of Guthrie’s zones, most frequently with the 14/6 (and 14/4) gender. This stem was treated in detail by Bulkens (2009) who lists the previous reconstructions: *gato (Homburger), *ato (Meinhof et al.), *âtò (Greenberg), *yâtò (Guthrie), *(j)âtò (Meeussen). In Bulkens’ collection of 160 reflexes of this stem, only four languages attest a consonant-initial nominal stem and she shows how they developed, mostly due to reanalysis.¹⁶ Otherwise, the stem always begins with a vowel, e.g. Lundu A11 *ádù*, Holoholo D28 *âtò*, Tsonga S53 *àtsò*.

So, the obvious reconstruction at the PB node 1 (and even earlier) is a return to Meinhof’s vowel-initial root *âtò without Guthrie’s *y or BLR’s *j or even Meeussen’s *(j). Bulkens (2009: 58) concludes that the data disproves the hypothesis according to which nominal stems in PB invariably had an initial consonant.

For *jákà 3/4 ‘year’ (BLR 3169, all zones, C.S. 1904), Guthrie gives 33 descendant forms, mostly in the 3/4 gender. Again, the great majority have the class prefix (often with glide formation *mɔ- > mw-) followed by a vowel-initial stem, e.g. Tiene B81 *muáka* (Ellington 1977: 175), Lengola D12 *mwáka* (Stappers 1971: 275), Unguja Swahili G42d *mwaka*. The exceptions are a couple of cases in zone S where the plural class 4 prefix has crept into the singular.

Perhaps most demonstrative is *jéné 1/2 ‘self, same’ (BLR 3296, all zones; C.S. 1970). Not only are there no reflexes in Guthrie with an initial stop, but also the widely occurring variant *méné 1/2adj ‘self’ (BLR 2171 zones ABCK+) suggests that *mò-éné became *méné and was reanalysed as an independent stem at a very early stage, perhaps even by PB. This early development is much harder to imagine with a putative PB *mò-jéné. A similar history of incorporation and reanalysis must be the story with the doublet *jòngó 14 ‘brain’ (BLR 3571, zones BCE+) and *bòngó 14 ‘brain’ (BLR 274, zones ABG+), in this case with the noun prefix of class 14.

¹⁵That is to say classes 1, 2, 3, 4, 7, 14, but not classes 5 or 8 (because of prefix with close front vowel) or 6 (because of class 5 influence), 9 or 10 (because of non-syllabic nasal prefix), or 11 (because of class 10 plural influence).

¹⁶Bulkens’ exceptions are Kota B25 *yàzí* 7/14 (probably not this root), Masaba JE31 *háárò* 5/6, Bukusu JE31c *járò* 5/6, and Pende L11 *wátò* 5/6. Most are due to reclassification of the noun with reanalysis of the former class prefix as part of the new stem.

At the PB stage, in these three roots for ‘canoe’, ‘year’ and ‘self’, there is simply no good evidence in descendant languages that would persuade us to reconstruct an initial stop, spirant, or even glide. There are not too many of these unconditioned **j* nouns, but enough to matter, including several other basic ones, e.g. **ánà* ‘child’ (BLR 3203), **ápà* ‘armpit’ (BLR 3237), **ògà* ‘mushroom’ (BLR 3257), **úmà* ‘thing; bead’ (BLR 3619). Bourquin (1923) listed over a dozen vowel-initial noun roots from earlier scholars and then added a dozen more. Creissels (1999: 305) lists 11 of these nouns where “the languages of subgroup S.30 (and in particular Tswana) demand to accept the possibility of variants of these reconstructions with no initial consonant.”

2.2 Unconditioned **j* in verb stems

We will next look at the important group of verb roots reconstructed with an initial **j*. These 84 verbs account for almost half of the main entries in BLR3 beginning with **j*, and many are widespread through the Bantu area.

2.2.1 Typical reflexes

Following are some of the better attested roots, each with more than twenty languages cited in Guthrie’s (1967–71) comparative series. To simplify the analysis, for each outcome of **j*, I have sorted them into what I have called “weak” outcomes (with no consonant, or with a glide) or “strong” outcomes (with stop, fricative or affricate, especially *j*, *z*). In parentheses, I have put the number of entries in Guthrie with that outcome. Because the strong reflexes are rather rare, occurring only in certain languages, I have explicitly cited those exceptional languages by their Guthrie number (and used Guthrie’s orthography).

- (1) **jác-(am)* ‘open mouth; yawn’ [BLR 3145/6, C.S. 1889(a) **-yác-(am)*]¹⁷
Weak: \emptyset (23), y (4), w (1)
Strong: j (P21, P22)
- (2) **jàd* ‘spread’ [BLR 3147, C.S. 1890 **-yàd-*]
Weak: \emptyset (13), y (10), w (2)
Strong: z (B22b, B82, M63, R24)
- (3) **jáník* ‘spread out (to dry in sun)’ [BLR 3206, C.S. 1924 **-yáník-*]
Weak: \emptyset (20), y (6), ny (2) – Bantoid: Tiv \emptyset
Strong: y/j (B11a), z (M63, R24), j (P21)

¹⁷BLR (following Guthrie) only lists zones CEF+ for this verb, but its presence in zones AB is seen in Proto-Manenguba A15 **sám* ‘sneeze’ (Hedinger 1987: 247) and Bulu A74a *semele* ‘sneeze’.

- (4) **ját* ‘split; separate’ and derivatives [BLR 3242, C.S. 1945-6 *-*yát*-]
 Weak: \emptyset (17), *y* (6) – Bantoid: Tiv \emptyset
 Strong: *z* (B82)
- (5) **jégam* ‘lean against’ [BLR 3291, C.S. 1967ab *-*yégam*-]
 Weak: \emptyset (13), *y* (5) – Bantoid: Ekoid *y*
 Strong: *z* (B82), *j* (P21)
- (6) **jó(o)g* ‘bathe; wash; swim’ [BLR 3525, C.S. 2107 *-*yó(o)g*-]
 Weak: \emptyset (17), *y* (5) – Bantoid: Tiv \emptyset , Ekoid *y*
 Strong: *j* (A74), *y/j* (B11a), *j* (P21)
- (7) **jóm* ‘be dry’ [BLR 3616, C.S. 2161 *-*yóm*-]
 Weak: \emptyset (18), *y* (4) – Bantoid: Tiv \emptyset , Ekoid *y*
 Strong: *y/j* (B11a), *j* (B22b), *z* (B82, M63), *j* (P21)

This data is derived just from Guthrie’s collection and some subclades are more heavily represented than others, but it is a broad survey of Bantu languages and enough to establish a *prima facie* case that the “weak” outcomes are the general rule and “strong” outcomes are the exceptions. According to Guthrie’s data, about 90% of the many modern languages exhibit weak reflexes of **j* in these roots, especially \emptyset but also a fair amount of *y*, which are supported by Tiv and Ekoid cognates. In other words, among about 70 languages tested in the samples above, there are only a few that ever show a consonant /*j*/ or /*z*/ (that is, something stronger than a glide in these roots). From the phylogenetic viewpoint, it is not only the quantity that matters, but also the distribution. These exceptional languages do not form a block supporting a strong reflex preserved in an early branch; rather, they are isolated or in small subclades deep in the phylogenetic tree in Grollemund et al. (2015: Fig. S1). Likewise, an argument that these few strong forms preserve some archaic heterogeneity would need to be based on some original phonological or morphological distinctions (e.g. their concentration in a certain tense), but that is also not the case. Rather, these occasional dispersed drops of *j* or *z* in a Bantu ocean of \emptyset and *y* are a typical pattern for independent innovations in a large dataset.

In addition to Guthrie, we now have the data from the Grollemund Dataset, listing 75 common lexemes in each of 400+ Bantu languages. The only verb relevant for us is PB **jimb* ‘sing’. Analysing all its forms in all zones, one finds that about 140 languages have weak reflexes and 16 have strong reflexes. The strong reflexes mainly come from the few pockets already seen in Guthrie – B11 (3 examples) and N10-P20 (4 examples) – as well as A80 (4 examples) which was sparsely recorded by Guthrie. Although this is only one lexeme and also not a complete

picture (**jimb* is missing a cognate in 200+ languages), the Grollemund Dataset confirms the distributional pattern of Guthrie's data and implies innovations in a handful of recent groups.

So, for the proto-phoneme at the beginning of these verbs it is easiest to posit an original \emptyset from which *y* (or *w*) occasionally arose to resolve a hiatus or various prefixes were reanalysed and incorporated into the onset.¹⁸ Accordingly, our primary interest here in considering Guthrie's exceptional languages does not concern reconstruction, but rather an examination of case studies to see some of the phonetic or phonological paths of development which are possible from PB stem-initial vowels and/or **y*.

2.2.2 Exceptional languages

North-Western Bantu (zones A, B10-30) and Central-Western Bantu (C, parts of D) The North-Western Bantu languages usually show weak onsets in Guthrie, e.g. **jót* 'warm oneself': Duala A24 *ɔl*, Yambasa A62 *ɔt-ɔbɔ*; **jígu* 'hear': Lundu A11 *ɔk*, Bakoko A43b *ɔx*, Bulu A74a *wok*'. Only two of his many languages in these important branches regularly show several strong reflexes, i.e. Mpongwe B11a in the Myene group and Ngom B22b in the Kele group.

For each Mpongwe example, Guthrie gives two forms, one with *y* and one with *j*, e.g. *yemb* & *jemb* 'sing', *yom* & *jom* 'become dry'. In his treatment of the PB reflexes in Nkomi B11e (a related variety of Myene), Rekanga (1994: 157–159) explains the doublets: the usual reflex of **j* is \emptyset but the reflex *dy* (realised [dʒ]) occurs after the nasal prefix in class 9 (see also Grégoire & Rekanga 1994). The infinitive (class 10b) creates this same effect and so is also reconstructed as having once a nasal prefix. In short, the basic verb stem is that seen in the imperative and other forms with *y*, as one would expect. But the effect of a nasal prefix to create an affricate [dʒ] is a topic I will return to in considering class 9 nouns. For Ngom B22b, the reflexes are uniformly *j* (e.g. *jaɔ* 'spread', *jemb* 'sing', *jom* 'become dry'), but Shake B251 *yemp* 'sing' and other forms in closely related languages from the Grollemund Dataset suggest that only a small group was affected by this development. For **jígu* (North-Western **júg*) 'hear', there are over 20 forms from North-Western languages in the Grollemund Dataset, with clearly strong reflexes only in the A80 group. For **jimb* 'sing', there are 7 weak and 3 strong reflexes.¹⁹

¹⁸For example, the irregular Lumbu B44 *yum* and Punu B43 *kum* (< **jóm* 'be dry') reflect the **kɔ* prefix of the cl. 15 infinitive.

¹⁹Weak: Kpe A22 *embà*, Yasa A33a *èhimbà*, Ewondo A72a *yia*, Bulu A74a *yia*, Fang A75 (Bitam and Minvoul) *ɔyie*, Fang A75 (Medouneu) *ɔyee*. Strong: Eton A71 *jà*, Mkaa A15C *jém*, Elung A15C *jé*.

In the Central-Western branch, weak reflexes are the rule in the Grollemund Dataset: Babole C101 *emba*, Mboshi C25 *iyemba*, Bangi C32 *yémbá*, Soko C52 *hamba*, Mongo C61 *émba*, Bushong C83 *yéem*.

In sum, the great majority of the North-Western and Central-Western forms are weak, which supports the testimony of the other early branches for reconstructing a weak stem **imb* or **ymb*. But the mixed evidence in North-Western sub-groups reminds us that there must have been a range of impacts from strengthening (and weakening),²⁰ nasal infinitive prefixes or subsequent front vowels, and analogy to verbal nouns, since some languages use phrases like ‘make a song’. These processes are more clearly seen in other branches.

West-Coastal Bantu (zones B40-80, most of H) Confirming Guthrie, the extensive wordlist of twenty nearby languages (including Teke B70 and Kongo H16) compiled in Koni Muluwa & Bostoen (2015) typically shows initial *y*, *w*, or occasionally \emptyset for these verbs.²¹ The exception in Guthrie is Boma B82 which yields *z* or *j* at the beginning of these words: *zatɔ* ‘split’, *zɪle* ‘get dark’ (< **jid* ‘get dark, black’ BLR 6142), *zɔma* & *zɔmɪ* ‘become hard, dry’, etc. But even Tiene B81, another language with Boma in the Kwa-Kasai North subgroup,²² consistently has *y*, e.g. *yááta* ‘split’, *yíla* ‘get dark’, *yóma* ‘become dry’ (Ellington 1977: 175–176). So, Boma apparently has a language-specific development.

South-Western Bantu (zone R, parts of HKL) Weak reflexes of **j* are the rule. In Guthrie’s data, the only exceptional language in this area is Ngandjera R24 which his inventory describes as “broadly similar” to Ndonga R22 and Kwanyama R21 but with a few distinctive changes including **j* > *z*. Guthrie’s relevant data for Ngandjera was *zar* ‘spread’, *zanik* ‘spread to dry’, *zer* ‘shine’, *zon* ‘spoil’, etc. It is not clear what Guthrie’s source was for Ngandjera and this variety of Wambo is not well attested, so for our purposes I will take the Wambo language R20 as

²⁰In Eton A71, we see the possibility of lenition of fricatives: “the voiced alveolar fricative /z/ is realised by the voiced glottal fricative [ɦ] or simply not realised” (Van de Velde 2006: 28), although that does not affect the verb *jà* ‘sing’ which begins with an affricate.

²¹Nzadi B865 has variation in its reflexes of **j*: *o-yâŋ* ‘spread to dry’, *o-yûm* ‘to dry’, but *o-zwô* ‘bathe (intr)’, *o-zâŋ* ‘to refuse’ < **jâŋ* (zones CJRS), and nouns in *dz*. “There does not seem to be any regularity to this distribution, nor do the reflexes seem to line up consistently with any nearby languages” (Crane et al. 2011: 257). Since Bulu A74a also has an irregular onset in *ɔk* ‘swim’, one avenue to explore is whether some verbs were affected by the reflexive prefix *i-* (‘to wash oneself’), which mutated *y* to *z/j*. For the nouns, *dz* in class 5 is merely a reflex of the prefixes (regularly Nzadi **di/di- > dz-*).

²²New groupings of West-Coastal Bantu can be found in Pacchiarotti et al. (2019).

a whole.²³ Unfortunately, I have not been able to find examples of *j > z in these verbs. Rather, Baucom (1975: 172) reconstructs Proto-Wambo *yoga ‘swim, bathe’ with y/w/ø reflexes of the initial *y in various daughter varieties, e.g. Ngandjera *yoga* ~ *oga*. Likewise, PB *jámu ‘suck’ yields Proto-Wambo *yama, with Ngandjera *ama*. Similarly, Ndonga and Kwanyama only have y as the reflex for initial *j in verbs.²⁴ If, indeed, z reflexes appear in some variety of Ngandjera, they must be a late local innovation.

Eastern Bantu (the broad area of Guthrie zones EFGJMNPS and part of D). Sorting through *all* of Guthrie’s hundreds of entries from *all* of these languages, the only strong reflexes for these *j verbs are found in entries from two subgroups: Ruvuma and Botatwe.

Ruvuma group For these verbs in the closely related languages Yao and Mwera, Guthrie prints a double reflex: j and zero.²⁵

- (8) Yao P21: (j)*asam* ‘open the mouth’, (j)*anik* ‘spread to dry’, (j)*elajel* ‘float’, (j)*egam* ‘lean against’, (j)*oog* ‘bathe, wash’, etc.
- (9) Mwera P22: (j)*aam* ‘open the mouth’

Ngunga (2000: 78–81) explains that in contemporary Yao there are two types of verbal roots: those with a “stable” [j], which is realised in all verb forms, and those with an “unstable” [j], which appears only in some verb forms. He concludes that the infinitive provides the underlying form and that the “unstable” [j] is an insertion in suffix-marked tenses. Ngunga’s analysis is synchronic but it coincides with the obvious diachronic analysis: these *j verb roots historically had a vowel in root-initial position with a later hiatus-filler inserted after some tense markers,²⁶ whereas those verbs with stable [j] should have other origins.

²³Maho (2007: 129): “The entire R20 grouping represents a single language, usually called Wambo or Oshiwambo. Kwanyama R21 plus all varieties coded R211 through R217 correspond to Baucom’s (1975) northern dialect group, while the rest correspond to his southern group.”

²⁴Some examples from Ndonga (Fivaz 1986: 15, 99): *yala* ‘spread’ (*jád), *yela* ‘become bright clear’ (*jéd), *yola* ‘laugh’ (*jòd), *yogá* ‘swim’ (*jóg). From Kwanyama (Turvey et al. 1977): *yala* ‘spread (mat)’ (*jád), *yela* ‘be, become bright’ (*jéd), *yola* ‘laugh, joke’ (*jòd), *yota* ‘warm hands at fire’ (*jót ‘warm oneself’).

²⁵Odden (2003: 529): “Yao and Mwera are very closely related, and might be treated as dialects.” According to Guthrie’s (1967–71: Vol. 2, 59) inventory for Yao: “*C₁: [...] *c, *j > s; *nc, *nj > s; *y > j (in radicals)” and “*C₂: [...] as *C₁ [...] but *y > j in stems”. For Mwera: “Broadly similar to P.21, but *c, *j > ø.”

²⁶Odden (2003: 531): “Avoidance of hiatus is most strict in Yao (and Mwera), which have no V-V sequences within the word. Vowel fusion and glide formation are the rule within the word.”

In short, Yao and Mwera do not provide relics of an early *j but rather support reconstructing a vowel in root-initial position for these verbs.²⁷ The larger lesson is that a palatal stop or even eventually an affricate can develop as one of the options for a hiatus-filler.

Botatwe group Guthrie has examples from two Botatwe languages:

(10) Ila M63: e.g. *zal* ‘spread’, *zambukil* ‘spread’, *zanik* ‘spread to dry’

(11) Subiya K42: e.g. *zimb* ‘sing’, *lu-zimbo* ‘song’ 11/10

Bostoen (2009: 115) gives sample forms from most languages in this group:

(12) **jóm* ‘to be dry’ **jimb* ‘to sing’²⁸

a. Western Botatwe

Shanjo	K36	<i>džúma</i>	<i>ímba</i>
Fwe	K402	<i>zúma</i>	<i>zímba</i>
Totela	K411	<i>yuma</i>	<i>zimba</i>
Subiya	K42	<i>zuma</i>	<i>zimba</i>

b. Eastern Botatwe

Lenje	M61	<i>kú-yuma</i>	<i>kw-ímba</i>
Ila	M63	<i>zuma</i>	<i>ímba</i>
Plat. Tonga	M64	<i>íkú-yuma</i>	<i>ík-w-ímba</i>
Soli	M62	<i>yuma</i>	<i>ímba</i>

As Bostoen (2009: 115) notes, “[t]here is quite some variation in the realization of *j [...]. For most lexical items, certain languages attest a fricative, while others have a zero reflex. The precise languages attesting zero (or glide) may differ, however, from one lexical item to the other.” In short, whatever the source of the variation, the Botatwe data does not clearly lead to any internal reconstruction, even in subgroups.

²⁷ Almost all of the 39 stable-*j* verbs identified by Ngunga lack a clear origin, but many are verbs of noise or movement perhaps connected to ideophones. There are, however, two verbs with ‘stable’ [j] that are derived from PB roots in *j and require another explanation: *juman* ‘quarrel’ and *jiim* ‘to not give’ (which seems to have *j*-less variants and may be influenced by the common Bantu variant **nyim*).

²⁸ Crane (2011: 78) gives *ókúyimbà* ‘(to) sing’ for the Zambian variety of Totela, while Bostoen (2009) mainly reports on the Namibian variety of the language.

2.2.3 Summary of initial *j in verb stems

Overall, the frequency and stability of “weak” forms is quite impressive, and a weak onset of these verbs is to be preferred for PB node 1. It is entirely possible that there are some PB verb roots which begin with *y and some with *ø, and considering the ease with which a glide can be inserted or deleted, further study will be needed to determine the best PB reconstruction for each root along with any allophones (including w before back vowels). Meanwhile, for these verbs I have adopted a convention of writing a parenthetical initial glide, thus: *(y)ác-(am) ‘open mouth’ (BLR 3145/6), *(y)âd ‘spread’ (BLR 3147), *(y)áník ‘spread out (to dry in sun)’ (BLR 3206), *(y)át ‘split’ (BLR 3242), *(y)égam ‘lean against’ (BLR 3291), *(y)ó(o)g ‘bathe’ (BLR 3525), *(y)óm ‘be dry’ (BLR 3616).

In addition to the specific subgroups with apparent strengthening (*y > z, j),²⁹ there are occasional exceptions scattered across other languages. Considering the several hundred forms cited by Guthrie for these verbs with initial *j, it is not surprising to encounter occasional variants or doubtful cases and I will not discuss them all here. Let it suffice to note a few examples of other languages with idiosyncratic forms for *jimb ‘sing’ in the Grollemund Dataset: Kaningi Nord B602 *o-lima*, Soko C52 *hamba*, Bira D32 *nyimbo*, Bembe H11 *kù-giùmbilà* (cf. Vili H12L *kw-imbilá*), Ha JD66 *uku-lilimba*. These are useful reminders that one can always expect exceptions in a large dataset, especially in a category when there are phonological opportunities like hiatus resolution and incorporation of various prefixes (especially nasal and infinitive prefixes) at morpheme boundaries.

A major difference between the vowel-initial nouns and verbs is the frequent presence of glides before the verb stems. Besides the possibility of original glides, one likely reason is the greater range of morphological variation in verbs. For nouns, even with glide formation in the prefix, there are usually only one or two forms, e.g. *bò-átò 14 ‘canoe’, mò-ánà / bà-ána 1/2 ‘child(ren)’. But verbs have a large variety of prefixes of various shapes (ø, CV, V, N) that can lead to allophones in the root-onset. For example, the ‘unstable y’ in some Ganda verbs is so-called because the palatal element appears only at the beginning of the word (in the imperative), after non-high vocalic prefixes (e, a, o) and after n (as nj), e.g. for the stem (y)egeka ‘support’: oyégeka (2SG PRS), yegeka (IMP), njégeka (1SG PRS), but twégeka (1PL PRS), okwégeka (INF) (Meeussen & Tucker 1955: 175–176,

²⁹In fact, Guthrie understood the basic development of these exceptional subgroups (1967: 62–63): “The question of *y is difficult, since in many languages its reflex is zero, although in Boma B.82, Subiya K.42 and Ila M.63 *ya > za, while in Yao P.21 *ya > ja. [...] It is just conceivable that y was the sound in the source-pattern, and if it were, y > j > z is a not impossible development, on the one side, and y > zero on the other.”

also Hyman & Katamba 1999: 369–376). In short, the glide does not appear after the high vowels because the prefixes themselves undergo glide formation, just as seen in PB nouns like **bɔ̀-átò > *bwato*. One can assume that in some languages, in order to preserve morpheme stability, the glide variant of the verb was generalised throughout (and sometimes even strengthened). This development is seen in Ganda in other verbs, where only ‘stable *y*’ (*y* or *nj*) is found, especially before high-vowel stems, e.g. *yíta* ‘call’, *yíga* ‘learn’, *yúmba* ‘sing’. Considering the possibility of cycles of addition and loss of glides and the conditioning factor of preceding prefixes, further study will be needed about the possibility of PB glides in these roots.

2.3 Unconditioned medial **j*

The suspicious paucity of early stems with unconditioned **j* in C_2 position reinforces our doubts about the existence of PB **j* as a standard consonant. There are no solid verbs in this category, but there are three well-attested nouns:

- (13) **kájá* 5/6 (11/10) ‘leaf; tobacco leaf’ (BLR 1736, C.S. 1019 **káyá*, ABCD+)
- (14) **jòji* 3 ‘belly, abdomen’ (BLR 3589, C.S. 2142, ABC) as well as **jòjò* 3 ‘life; spirit; heart’ (BLR 3590, C.S. 2143–44, EF+), **jòjà* 3 ‘life’ (BLR 3588, ps 550, HS)³⁰
- (15) **jòjá* ‘fur, feather, bodyhair’ 14, 3, 11/10 (ACEF+) combining **jòjá* ‘fur’ 14 (BLR 3587, C.S. 2141, FJL+), **jòjá* 5/6, 11/10 ‘feather’ (BLR 3586, C.S. 2140, EGJ), **jòcá* 3 ‘feather; bodyhair’ (BLR 7034, CJ).³¹

Almost all Guthrie’s citations for these roots show *y* or \emptyset in the C_2 position, with a few zone A languages only having one syllable. A few other BLR3 noun reconstructions are marked ‘main’ but without North-Western cognates, e.g. **bɛ̀jɔ̀* 3/4 ‘baobab’ (zones CGM+), *kɔ̀jɔ̀* 3/4, 7/8 ‘fig-tree’ (DE+), **jàjò* 11/10 ‘sole of foot’ (DE+), **káájá* (5, 9a) ‘home village’ (DEFGH+), all almost completely with *y* or \emptyset reflexes.³² It should be noted that in these roots either the vowels are the same

³⁰When Guthrie did not have enough examples for a valid C.S., he created a “partial series”, abbreviated as “ps”. See Guthrie (1967: 42): “Frequently it has not proved possible to complete a valid C.S. but sufficient items have been discovered to make a partial series. Unless there are reasons to the contrary, such series are included in the main catalogue with a separate serial numbering, distinguished by the use of the abbreviation ps.”

³¹An anonymous reviewer kindly added Duala A24 *ɲ-ɔ̀* ‘hair’, Elip A62C *gʸ-ɔ̀yá / bʸ-* ‘feather, hair’.

³²Two of these words (baobab and fig) are flora, possibly added as certain species were encountered during the Bantu Expansion.

or both are low vowels, i.e. the conditions are not favourable for the simple formation of a glide from the first vowel, which is the standard treatment for resolving hiatus in vowel-initial roots. Accordingly, the easiest PB reconstruction here is \emptyset for C_2 with the frequent development of epenthetic elements in various languages or branches but rarely with the strong $*j$ effects seen at morpheme boundaries. Early roots with this structure are rare in PB and, if no glide is reconstructed, one would want to understand their difference from long vowel roots. Other candidates having medial $*ij$ combinations will be treated later, e.g. $*j\acute{i}j\grave{a}$ ‘fire’, $*j\acute{i}j\grave{a}$ 1a ‘mother’, $*j\acute{i}j\grave{i}$ 6 ‘water’.

3 $*j$ conditioned by preceding $*i$

In a significant number of cases, stems reconstructed with $*j$ are conditioned by a preceding $*i$, either as part of the root or in a prefix. There are several ways for this to happen, especially:

- prefix i - before class 5 nouns, e.g. $*j\grave{a}d\grave{a}$ ‘rubbish-heap’, $*j\acute{i}c\grave{o}$ ‘eye’
- $*jij$ in stems, e.g. $*j\acute{i}j\grave{b}$ ‘know’, $*jij$ ‘come’

Here, a distributional pattern appears that is very different from our previous categories. This environment is the major source for the strong reflexes of $*j$ and the tradition of reconstructing some palatal stop or affricate rather than y or \emptyset . But these strong forms result from localised rules mostly in Eastern Bantu. Basically, what I have called weak reflexes (y , w , \emptyset) are regular in the North-Western zones ABC, but strong forms (j , z) are occasional in the north-east and south-west Savannah zones (EFGJKR) and regular in South Bantu (N20-40, P30, S).

3.1 Initial $*j$ in class 5 roots

3.1.1 Typical reflexes

A fair number of class 5 nouns are traditionally reconstructed with $*j$ by both BLR and Guthrie (sometimes with doublets in $*y$), for example:

- (16) a. *jàdà 5/6, (7/8) ‘rubbish-heap’ [BLR 1557, zones ABCDE+, C.S. 918]
 b. *jàná 5 ‘yesterday; tomorrow’ [BLR 1566, ABE+, ps 256]
 c. *jàni 5/6 + ‘leaf; grass’ [BLR 1567, ABCDE+, C.S. 926]
 d. *jóbà 5 ‘sun’ [BLR 1614, ABCDE+, C.S. 955, 2147, ps 508]
 e. *jòì 3/4, 5/6 ‘voice; word’ [BLR 1612, ABCDE+, C.S. 954, ps 260]
 f. *jòdò 3/4, 5/6 ‘nose, nostril’ [BLR 1620, ABCDE+, C.S. 960, 2151]

For PB, the class 5 prefix is reconstructed as the high front vowel **i*, with a pre-prefix (or augment) **dɪ-*, together forming the full template **dɪ-i-ROOT*.³³ It will be seen that these roots are best reconstructed with initial vowels to which the prefixes have attached themselves. Perhaps the strongest evidence for this comes from the fact that class 6 plurals almost never show any strong reflex.

A classic example of this category is **jóbà* ‘sun’, which is attested in all zones and highlights the important evidence from the North-Western branches (sometimes with meanings ‘sky’ or ‘day’):

- from Guthrie: Lundu A11 *d-oba*, Duala A24 *l-oba* cl. 13(?), Mvumbo A81 *duɔ*, Makaa A83 *düawɔ*, Ngom B22b *ɔ-oba* cl. 11, Tsogo B31 *oba* cl. 11.
- other A10-60: Kundu A122 *lobà* (Atta 1993: 89), Batanga A32C *dóbà* (den Besten 2016: 35), Abo A42 *lou*, Dibum A43a *lɔp*, Nen A44 *niɔf*, Kpa A53 *díóó*, Baca A621 *ɲɔ-p* (Mous & Breedveld 1986: 227, 232).

The reflexes of **d* and **di* vary language by language, but all of these forms can be seen as descendants of a vowel-initial root with pre-prefix, **dɪ-(i)-oba*, with an initial *d / d' / l / n* from the conditioning and contraction of **dɪ-/*di-* before the initial vowel of the root. The occasional forms in *j/dz/dj* apparently result from palatalisation before the initial vowel, e.g. **dɪ-V > *dʲ-V > jV*, hence Benga A34 *joba*, Basaa A43 *jɔb*, Bulu A74a *jɔp*.³⁴ In Ewondo A72a, this stem has two forms *yób ~ dzób* ‘sky’, which are apparently the results of the prefix or augment alone:

³³There is possible influence from allomorphs in other classes which lack the *i-* environment (especially the class 6 plurals) or which have N-conditioning. So, in selecting class 5 nouns for analysis, I have excluded any which have class 9 or 10 by-forms, to ensure that there is no influence of those **nj*, **ny*, **nz* forms on the class 5 forms. Accordingly, an analysis of this type would need to be more detailed, especially since the distribution of strong forms varies by lemma.

³⁴There are probably a number of phonological and morphological factors in each language. For example, there are different conditioning factors in Bulu A74a: in *C*₁ unconditioned **d > y*, but **di* (or **dɪ-i*) > *d* (e.g. *dim* ‘extinguish’ < **dim*; *dis/mis* ‘eye(s)’ < **jícò*; *dí/mi* ‘fireplace(s)’ < **jíkò*), and **dɪ-VC > j-VC* (*jal/mal* ‘village(s)’ < **jàdà*) (Yanes & Moise 1987: 10–14).

**i-ǒbà*, **dr-ǒbà* (Essono 2000: 197). A key point is that nowhere do we find forms that reflect an augment plus a consonantal onset like **i-jǒbà* or **dr-jǒbà*. The substantial North-Western evidence for the PB reconstruction of a vowel onset for this class 5 root is matched with straightforward data elsewhere (‘sun’: Nande JD42 *eri-uβa*, Luba-Kasai L31a *di-ūba*, Mwera P22 *li-uβa*, Herero R30 *e-yuβa*).

But it is also important to understand the different changes in certain Eastern branches that led previous scholars to generalise the strong onsets. To understand the general path of development, it is useful to look at a few special nouns reconstructed by BLR3 with initial **ji*, which are also likely to be vowel-initial:

- (17) a. **jícò* 5/6 ‘eye’
 b. **jíkò* 5/6 ‘fireplace; country’
 c. **jínò* 5/6 ‘tooth’
 d. **jínà* 5/6 ‘name’

These class 5 nouns show an unusually wide variety of onsets across the Bantu area. However, if we assume that these were also roots with an initial vowel *i* (as supported by Bantoid forms of ‘eye’: Ekoid *e-yid/a-mid*, Tiv *i-fə/a-fə*), then the variety is quite understandable. The contact of the class 5 prefixes **i-* and **dr-* with the initial vowel inevitably led to certain mergers that blurred the morpheme boundaries. We see three types (examples from Guthrie C.S. 2030 **yícò* ‘eye’, using his orthography):

- Contraction: **dí-(i)-íco* ‘eye’ > **díico* or **díico*. From the full PB augment and prefix, we can expect a contraction of the sequential front vowels. The impact of the vowels on the initial **d* depends on the quality of the contracted vowel, the consonant rules of specific languages, and analogy:
 - languages with a form of *d* conditioned by the vowel *i*, or a pre-vocalic reflex (typically *dʒ*), rather than the unconditioned reflex (typically *l*, *y* or *∅*). Often, we can assume an intermediate **dii*, due to a contraction of the augment and prefix and the root beginning with *i*. For example, Duala A24 *disɔ*, Ngom B22b *dij/mif*, Bali-Teke B75 *dziu*, Bongili C15 *difɔ/mifɔ*, Boloki C36e *dʒiɔ/miɔ*, Bushong C83 *dijf/mijf*, Manyanga H16b *diisu/meeso*, Luba-Katanga L33 *jiiso/meeso*.
 - languages which show the unconditioned reflex of **d*, most likely because the onset was generalised from other class 5 modifiers. For example, Sukuma F21 *liiso/miiso*, Luvale K14 *liso/meso*, Yao P21 *liiso/meeso*, Southern Sotho S33 *lɛjhlɔ/mahlɔ*.

- Inverted augment: **dɪ-i > *i-dɪ* or **ɪ-di* etc. > *ili* etc. Certain Eastern Bantu languages have reversed the order and/or function of the class 5 augment and prefix, probably to create paradigmatic regularity with V-CV structures in other classes.³⁵ Conditioning of the **d* can lead to *r* or *dʒ* instead of *l*. For example, Nande JD42 *erijiso/ameeso*, Nyoro JE11 *eriiso /amaiso*, Luyana K31 *ilito/amiyo*.
- Loss or disuse of the augment: **(dɪ)-i-íco* ‘eye’ > *iso* etc. This development is not uncommon in zones C and D and is characteristic of zone R, e.g. Umbundu R11 *iso/oβaso*, Ndonga R22 *exo/omexo*. It also forms the basis for some South Bantu changes seen below.

Once again, the categories above are explicable by reconstructing the class 5 forms of a vowel-initial PB root **íco* ‘eye’. Likewise, throughout the Bantu languages we see several options in their class 6 plurals based on a vowel-initial PB root:

- *(a)me(e)so*, from a coalescence of **(a)ma-íco*;
- *(a)ma(a)so*, from a contraction of **ma-(i)co* (favouring the first vowel) or a reanalysis of the stem as **co*, perhaps based on a singular form **dico*. This is the standard form in zones AB;
- *(a)mi(i)so*, from a contraction of **m(a)-íco* (favouring the second vowel);
- *(a)ma-iso*, uncontracted, likely an analogical restoration (rare outside JE10).

3.1.2 Eastern cases of class 5 strengthening

In addition to the straightforward development of class 5 vowel-initial roots in most of Bantu above, there are two sub-branches where fricative or other strong onsets developed: South Bantu and North-East Coast Bantu.

3.1.2.1 South Bantu strengthening: class 5 forms with j, z, ž, etc.

In South Bantu languages (zones NPS), we see several types of paradigms in these common nouns:

³⁵For combinations of cl. 5 prefixes in Eastern Bantu, see Kamba Muzenga (1988).

- Vowel-initial stems in singular and plural, e.g. Tswana S31 *li-íná/mà-íná* ‘name’, *li-ínó/mènó* ‘tooth’, *li-ítłó/mâ-tłó* ‘eye’ (Creissels 1999: 325) – notice that there are three different types of plural formation – or Zulu S42 *i(li)so/amehlo*.
- Strong onsets in singular and plural, e.g. Zulu S42 *izinyo/amazinyo* ‘tooth’.
- Mixed onsets in the same paradigm, e.g. Shona group S10 *zísó/mèsó + màzísó + màsó* ‘eye’; *zínó/mènó + màzínó* ‘tooth, teeth’ (Hannan 1974).

Several languages in the region have some mix of types, so analogical processes must be at work. The class 6 plurals (aided by contraction) often preserve vowel-initial stems and we can surmise that the occasional strong onsets in the plural are by analogy to the singular.

What is the source of the several South Bantu strong onsets? An obvious option would be a development from the class 5 augment and prefix **dí-i*, as seen above, and that may be a factor in some languages. But that does not seem to work for languages like Shona where the strong *z* reflex here is not derivable from any version of the prefixes.³⁶ Rather Shona *z* matches the onsets in class 5 forms from PB **g*. In general, PB **g* was lenited to Proto-South Bantu **y* and eventually lost in most languages. After the class 5 prefix **i-* there arose a special set of changes for all the stops, e.g. Shona *dákó/mátákó* ‘buttock’ < **tákò*. For **g*, we see **i-g* > **i-y* > Chewa N31b (*d*)*z*, Shona *z*, Venda S21 *ǰ*, Zulu *z*, Tswa S51 *t*, for example, **gòdò* 5 ‘sky, top’ > Zulu *ízulu* 5 ‘sky, heaven’. This phonological change is also seen inside roots, e.g. **tùìgà* ‘giraffe’ > **tùìyà* > Shona *twìzà*. These are the same reflexes seen for the **j* nouns in class 5. It is for this reason that Meinhof et al. (1932) began many of these class 5 stems with **y* (the graphic predecessor of **g*) rather than **j* (now **j*), and Guthrie had a doublet series in **g* for some of these words: C.S. 831 **gína* and C.S. 2068 **yínà* ‘name’; C.S. 828 **gíkò* and C.S. 2056 **yikò*.

In short, the strong reflexes of **j* in South Bantu nouns appear to reflect stems which had initial *y* at some stage, perhaps because they were the inherited forms in some stems or, more generally, because the glide was inserted to resolve the hiatus between a prefix and a vowel-initial root. In fact, the augment **i-* may have sometimes become that glide and then was reanalysed as part of the root and assigned the root anew, i.e. **i-úba* ‘sun’ > **yúba* > **i-yúba*.³⁷

³⁶In Shona, **dí-C*, **dì-V* > *dz* (**dím-a* ‘extinguish’ > *dzíma*, **dì-ama* ‘sink’ > *dzàma*, **dì-ik-a* ‘bury’ > *dzika*) and **dì-V* > *dy* (**dí-a* ‘eat’ > *dyá*).

³⁷Similar is the development of a glide and then glide strengthening in Ganda JE15, where the class 5 prefix generally causes gemination, e.g. **jìbà* ‘pigeon’ > *ejjìbà* 5 / *amayiba* 6.

3.1.2.2 North-East Coast Bantu strengthening: class 5 forms with j, z, ž, etc.

There are a number of languages in the Sabaki group (E70, G40) and nearby that frequently show strong forms in class 5 (and by analogy in class 6), for example:

- (18) a. Dawida E74a *iziso/meso* ‘eye(s)’ *ijani/mani* ‘leaf/leaves’
 b. Unguja Swahili G42d *jicho/macho* ‘eye(s)’ *jani/majani* ‘leaf/leaves’
 c. Ngazija G44a *dzitso/matso* ‘eye(s)’ *wani/mani* ‘leaf/leaves’

In many North-East Coast Bantu languages, the only class 5 prefix is a single vowel *i-* and often it is deleted, leaving a \emptyset -prefix for polysyllabic consonantal stems, e.g. **pácà* ‘twin’ 5/6 > Swahili *pacha/mapacha*. But for monosyllabic stems, a variety of prefixes are found in the Sabaki languages, e.g. from **bú* come *ivu*, *jivu*, *vuu*, *livu*, *rivu*. A number of hypotheses (including retention of the prefix **di-*, and analogic reformation) led Nurse & Hinnebusch (1993) to reconstruct a series of local changes to explain these monosyllabic stems, as well as our class 5 vowel stems: pre-North-East Coast Bantu *(*i*)*li-* > Proto-North-East Coast Bantu *(*i*)*zi-* > Proto-Sabaki **iji-*.³⁸

3.1.3 Summary of class 5 effects

I have given some attention to the South Bantu and Sabaki groups, because the impact of certain coastal languages (e.g. Zulu and Swahili) on the early Bantuists was high and inclined them to propose some consonantal onset for these stems. But in other branches as well, there are examples of both strong and weak reflexes which suggests that they co-existed for many years, as the form of the class 5 prefixes varied, with possible analogy from class 6 forms in *ma-*. The Kikongo Language Cluster (part of the West-Coastal branch) provides examples of this variety of prefixes and onsets (*y ~ z*) for forms of **jódò* ‘nose’ (with variant **jidò*): Vili H12L *liyilu*, Yombe H16c *yilu*, Soonde H321 *múzulu*, Mbala H41 *muzulu*, Sikongo H16a *zúúnu*, Solongo H16aM *dizunu*, Woyo H16dK *yiilu*, etc. This is paralleled by a variety of class 5 forms in PB **g*: for example, **gòdò* ‘sky, top’: Vili *liyilu*, Yombe *yilu*, Lumbu B44 *diyuuulu*, Yaka H31 *zúlu*, Laadi H16f *zúlù* (from the Grollemund Dataset, itself taken from de Schryver et al. (2015) for the

³⁸The problems of **j* and class 5 forms in the Sabaki group are discussed in Nurse & Hinnebusch (1993: 108–112, 186–196). The process of strengthening in Comorian G44, discussed at pp. 133–145, parallels that found in South Bantu. See also Nurse (1979: 149–153) on Chaga E60 and the North-East Coast.

KLC).³⁹ The various explanations depend on the individual languages and lexemes. For our purposes, it suffices to say that the developments involved are all at intermediate to late stages of Bantu history.

In sum, all the class 5 nouns reconstructed with **j* are best reconstructed with initial vowels for PB node 1.⁴⁰ The general absence of consonantal reflexes in the class 6 plurals of these nouns is a significant problem for reconstructing a consonantal onset.⁴¹ Rather, various phonological processes affected the singular class 5 prefixes **di-* and *i-* before vowels with results that were sometimes reanalysed as strong onsets for the roots, especially in Eastern Bantu. Likewise, there is no need for **j* in Meeussen's (1967: 97) reconstruction of the augments **ju* (cl. 1), **ji* (cl. 9) or **ji* (cl. 10), which were based on Eastern innovations. For class 10, a coronal seems more likely, e.g. **di* (cf. C.S. 2225a).

This is also a convenient time to clarify one important point. Sometimes references are made to Bantu Spirantisation of PB **j*, based on *z* in some of the singular forms of these special words, see for instance, for Kalanga S16, Mathangwane (1999: 82–83, 88, 213). However, these are more easily explained by class 5 effects or reformation. If indeed these PB roots had had an initial **j* and if there had been an effect of the subsequent **i* on it, we should see it in both the singular and plural. But the fact is that we often see some change in the singular but not in the plural. Why would **j* not spirantise systematically before high vowels? Because it is actually zero or a glide.

3.2 Initial **ji-C* and **jij*

Long ago, Meeussen pointed out that his Bantu reconstructions had a surprisingly large number of verb roots beginning with **ji* (Meeussen & Tucker 1955: 177). Perhaps out of deference to tradition, Meeussen (1967: 86, 90) himself later hesitated about **ji-C* structures, reconstructing a parenthetical onset in forms like *(*j*)*ijr*b 'know', and an examination of the specific modern reflexes now shows that the first **j* is not needed.⁴²

³⁹Similar variation can be found under the entries for 'sky', 'fireplace', 'nose', 'eye' or 'tooth' in Koni Muluwa & Bostoen (2015: 72, 99, 127, 130, 181).

⁴⁰The roots were likely vowel-initial at an earlier stage too. Cf. Eastern Grassfields **li-ít* / *mà-it* 'eye', **dínj* 'name' (Elias et al. 1984: 38).

⁴¹There are also nouns like **jáni* 'leaf, grass' (BLR 1567, C.S. 926, 1928) which is commonly cl. 5/6 but its initial vowel is clearly seen in other classes: Lundu A11 *ɛ-anj* 7/8, Bubi A31 *s-anyj* 19/13, Maande A46 *nɔ-any/tɔ-any* 11/13, Luba-Kasai L31a *lw-anyi* 11, Tswa S51 *by-anyi* 14.

⁴²Among dozens of **jij* verb reflexes in the data from Guthrie (1967–71) and the Grollemund Dataset, we find an element before the *i* only in Teke Yaa B73c *yir* 'come', Yao P21 (*j*)*iis*, and Manyanga H16b, where they are resolving the hiatus of vowel-initial stems. Initial *y* is some-

In **ji-C* verbs, the South Bantu consonant changes are similar to what we saw with class 5 reflexes, for which we reconstruct the prefix as **i-* not **ji-*. For example, **ji-kad* ‘dwell, sit’ > Manyika S13 *gara*, Makhuwa P31 *khala*, just as **i-kádà* ‘ember’ 5/6 > Manyika *gara/makara*, Makhuwa *ni-khala*.⁴³ Thus, we would do better to return to the simpler version of Meeussen’s (1967) reconstructions: **ikad* ‘sit’, **igad* ‘shut’, **im* ‘stand (up)’, and *iji* 6 ‘liquid’.

So, if the descendant languages almost never show any consonantal remnant of the proposed first **j*, why was there a reconstruction of **ji* in these verbs instead of simpler **i*, and **jij* instead of simpler **ij*? If I understand the scholarly history, the prefix **ji-* (earlier **yi-*) was reconstructed to explain some verb forms which occasionally show *i* at the beginning of the stem or some consonant mutation. Meinhof et al. (1932: 179) state, “But *yi* can also be what remains of an old infinitive prefix, which has been retained in a few languages only. E.g. **-yikala* ‘sit, remain’, Shambaa *-ikata*, Herero *-kara*, Swahili *-kaa*.” Meinhof’s suggestion that **yi-* is what remains of an old infinitive prefix which later merged with the class 5 prefix has not been accepted. A better source morpheme of the appropriate shape and position is the reflexive pronoun **i-*, which Meinhof et al. (1932: 43) wrote as *yi*. The incorporation of reflexive particles into verb forms is well attested cross-linguistically and seen in Bantu languages in Tswana, Ganda and others.⁴⁴ The fact that many Bantu languages lose or change the reflexive particle allows this particular morpheme to be lost or reanalysed as part of the verb stem. Thus, the initial consonant in **yi-* seems to be due to two factors: Meinhof’s early etymology of the infinitive prefix from a verb *ya*, *ye* or *yia* ‘go’ (ibid. 43), and the occasional forms in *ji/yi* in languages like Sango G61 and Kongo H16.⁴⁵ Accordingly, the reconstruction of the initial **j* in these roots seems to be a relic of Meinhof’s early work and can be removed.

times also found in other **jiC* verb reflexes, e.g. Mpongwe B11a *yir/jir* ‘pour’ < **jit*, Makonde P23 *yigal* ‘open’ (but *id* ‘come’). It is particularly common in the verb **jib* ‘steal’ which has many zone AB reflexes with *yib* or *jib*.

⁴³Botne (1991) gives a wide set of reflexes and an analysis for **jikad* ‘dwell, sit’.

⁴⁴For Bantu reflexives, see Marlo (2015); for a discussion of the lexicalisation of reflexives, especially with **kada*, see Botne (1991: 252).

⁴⁵But certain sample languages dominated. Already in Meinhof’s (1899: 153) *Grundriß*, two of the four reflexes given for **yi-ama*, *yi-ma* ‘stand’ have what looks like a consonantal reflex: Northern Sotho S32 *yema* (*ema*, *yama*) and Sango *jima*. Later Laman’s data for Manyanga H16b had a major role in the sample languages, with **ji-C* reflexes like *yikal* ‘dwell’, *yimit* ‘become pregnant’. Thus, Meinhof et al. (1932: 161) analysed the Kongo *-y-* as a preservation (even though they provided the evidence to show it is actually resolving the hiatus): “**yi* > *yi*, e.g. *yiza* ‘come’ < B. **yiya* [...] In some instances, *y* is completely lost, e.g. *iŋgi* ‘many’ < B. *yŋgi*, *kw-iza* 15 infin. of *yiza* ‘come’. Sometimes *k* appears for *y* [...] e.g. *kima* (dial.) ‘stand fast’ < B. *yima*.”

There is a small but important group of PB nouns and verbs reconstructed with **jij*, with parallel reflexes:

- **jiji* 6 ‘water’ (BLR 3433, ACDE+), **jijà* ‘fire’ (BLR 5884, ACDH+), **jijà* 1a ‘mother’ (BLR 3513, ABCFG);
- **jij* ‘come’ (BLR 3425, ABCE+), **jij-ad* ‘be full’ (BLR 3429, ABCD+).⁴⁶

Obviously, the initial **j* in all these stems can be omitted from the reconstructions. As usual, the noun reflexes are fairly stable: **jijà* 1a ‘mother’ has five forms in zones ABC, all with *iy*. For **jijà* ‘fire’, we see mostly *y* (many *eya*) but also some strong forms in zone A.

In the group of verbs reconstructed with **jij*, the shortness of most roots makes it sometimes hard to be certain of cognates or distinguish other effects. Two of the better documented verbs are **jijad* ‘be full’ (**jijud* ‘become full’) and **jijib* ‘know’.⁴⁷ In the reflexes of these lexemes, we typically see three types of initial sequences with examples of ‘know’ from the Grollemund Dataset:

- Weak (*y*, \emptyset) – frequent everywhere: Yasa A33a *èyíba*, Kuria JE43 *íba*
- Strong (*z*, *j*, *dz*, etc.) – frequent everywhere, regular in South Bantu: Chewa N31b *dziwa*
- *i-* + Strong – frequent in Eastern Bantu: Mwani G403 *idziwa*

These outcomes are somewhat similar to the pattern that was discussed for class 5. Since the South Bantu languages share common reflexes of **jij* with what was reconstructed as class 5 **i*-strengthening of initial *y* (Shona *z*, Southern Sotho *tʰ*, Venda *d*), it seems reasonable to tentatively consider that sort of **iy* structure for these words too. But in this case, **iy* would have to be already present at the PB level.

Let us begin with some examples of **jij* ‘come’ from the North-Western branches: Kundu A122 *iya*, Mkaa A15C *yà*, Kpe A22 *jâ*, Kako A93 *nja*, Tsogo B31 *e-y-a*, with an extended stem *yak/zak* seen in several B20 languages. For **jijib* ‘know’: Wumbvu B24 *u-yiba*. In Central-Western languages, ‘come’ and ‘know’: Mboshi C25 *i-yaa* and *i-yepa*, Bunji C25A *i-jaa* and *i-jépa*, Mongo C61 *yá* and *eb*, Libobi

⁴⁶Cf. also the Eastern compound noun **jijòkòdò* ‘grandchild’ 1/2 (BLR 3435, DEF+).

⁴⁷Cf. C.S. 2047 **yijad* ‘become full’; ‘know’: C.S. 938, 968, 2001. I have not included very reduced forms of ‘know’ like Abo A42 *jī* or Basaa A43a *yi* because of the possible relationship to the stem *yem/jem* ‘know’ seen in A70.

C412 *bo-yéi* and *bo-yebi*. In general, for these branches, there seems to be a majority of weak reflexes but enough strong ones to need more study before making a generalisation. Likewise, West-Coastal and South-Western Bantu have a mixture of weak and strong forms, with much variation even inside subgroups.

In Eastern Bantu, almost all the reflexes are strong, but with such variation (*j*, *c*, *s*, *f*, *ts*, *tf*, *z*, *dz*, *ʒ*, *ɟ*, etc.) that it is not easy to describe a common phonological development for the branch (although perhaps for subgroups like South Bantu). The same can be said for **bàij* ‘work wood’ (BLR 8930, C.S. 32, 86), a rare example of medial **ij* in a verb stem: there are no citations for zones ABC so the reconstruction must be attributed to an intermediate node of the Bantu Expansion, by which time some relevant phonological developments might have taken place.

In short, **jij* has become the traditional reconstruction for several stems regularly showing strong reflexes or *i* + strong in Eastern Bantu and frequently elsewhere. Since there are only a few of these roots (just as with medial **j* in general), this **iX* structure probably arose from the juncture of other elements in the language. At present, I might propose **iy* insofar as it is a common reflex and plausible source for some of the other forms. But one would need to explain the source of the glide, and how to distinguish the evolution of **V-iy-a*, **V-i-ia*, and **V-i-a*.

4 **j* conditioned by preceding **n*

Our final group is reconstructed **j* when pre-nasalised or in nasal combinations. Although BLR3 does not have **y* as a separate phoneme from **j*, it does distinguish **ny* from **nj*.⁴⁸ Altogether, there are several categories we could consider here (each followed by the number of main reconstructions in the BLR3 database):

- **N-j* (stem-initial **j* with the class 9 prefix) – 25 nouns
- **N-ny* (stem-initial **ny* with the class 9 prefix) – 3 nouns
- **ny* (other stem-initial, or final **ny*) – 7 verbs, 3 nouns
- **nj* (stem-medial or final **nj*) – 4 verbs, 8 nouns

⁴⁸I maintain the graphic convention (used by Guthrie and BLR) of writing **ny* in these reconstructions, although **ɲ* may have been the case, as seems more likely in **nyàmà* ‘animal’ below. The emphasis in the discussion is rather on distinguishing reconstructions of **ny*/**ɲ* from those with a stop or fricative under the cover term **nj*.

Whether or not all of these reconstructed categories are correct, there must have been occasional cross-influence and reanalysis. Not surprisingly, BLR3 shows variation between **(N)j* and **ny* in some stems, e.g. Main 7055 **nyóòtà* 5, 9 ‘thirst’ ~ Variant 3580 **jótà* 9 ‘thirst’ and Main 3273 **jéd* ‘shine, be clear’ ~ Variant 2324 **nyényè* ‘star’.

We will mostly look at class 9/10 forms, which have nasal prefixes, but also some forms with nasals in other classes. The patterns are more consistent if we consider them by groups based on reflexes: (1) those with weak reflexes, pointing to **n-y*; (2) those with strong reflexes, pointing to **n-j/z*; (3) those with mixed classes.⁴⁹

4.1 Nouns supporting PB **ny*

There are several nouns reconstructed with **nj* or **ny* that regularly have palatal nasal reflexes in both Bantu and Grassfields languages.⁵⁰

**jókà* 9 ‘snake; intestinal worm’ (Guthrie both **yókà* and **jókà*) is attested in all zones. All citations from zones A and B (which are half of the Bantu family tree) have reflexes with *ɲ* (or occasional *n*) and the preservation of *ɲ* (or *n*) in zones H, L, R and S confirm that **n-yókà* ought to be reconstructed for PB. But in some other zones there frequently arose fricatives, affricates, and palatal stops, e.g. zones C (*ndz, nz, nj, ɲ*), DEF (*nz, nj, ɲ, nc, nf, f, ch*), M (*nz, nj*). This range of mutations shows how **ny* could evolve into strong forms, and the individual variants were probably affected by the developing non-pre-nasalised phonemic inventory in those sub-branches.

**játí* 9 ‘buffalo’ is compiled by Guthrie (and followed by BLR3) in two series: **(n)yátì* (zones ABCEGMNPRS) and **játí* 9/10 (zones BCMN).⁵¹ It is hard to believe that there were really two concurrent stems for a morphologically invariable and semantically stable item (and no single language preserves a doublet). Guthrie’s data has *ny* in all 11 forms from the North-Western branches, and the majority elsewhere – leading us to reconstruct **nyátì* for PB node 1.⁵² Once again it is interesting to note the half-dozen scattered forms in *n-j* or *n-dz* cited

⁴⁹The most extensive study of this category is Bostoen (2005: 182-88) who focuses on **jòngó* ‘cooking pot’, but includes **jòndò* ‘hammer, anvil’, **jénjé* ‘cricket’ and many other relevant lexemes. He assumes these class 9 nouns had a strong **NC* in the *C*₁ position and shows how Meinhof’s Rule plays a significant role in producing weak reflexes in Eastern Bantu.

⁵⁰In this section, unless otherwise specified, Bantu language data comes from Guthrie (1967–71) and the Grollemund Dataset; Grassfields from the Grollemund Dataset.

⁵¹C.S. 927, 1947, ps 495; BLR 1569.

⁵²Frequent nasal-initial weak forms in Grassfields would tend to push the reconstruction back further.

by Guthrie. Several of these are clearly late innovations (distinct even from close neighbours) but useful evidence that the development **ny > nj/ndz/nz* is quite possible in independent languages.

**jùmá 9* ‘back, rear’ is primarily listed by Guthrie under C.S. 2060 **yùmà*.⁵³ For this stem, Guthrie’s data is almost unanimously in favour of a weak onset, with numerous variations on the initial sequence displaying the range that is possible inside what I have called “weak”: *nyi* in zones ADGHJKLR, *ni* DHKR, *ngi* HL, *nyu* BDFGLMP, *nu* DEFM, *nnyu* E. Occasional forms in other classes (e.g. Tikuu G41 *mma* 5, Mbundu H21 *r-ima* 5, Kwambi R23 *oku-nima* 15) show that the initial nasal in class 9 could be perceived as the class 9 prefix or as part of the stem. What is striking is the absence of strong forms (i.e. *n-j*, *n-ʒ*, *n-dʒ*) in Guthrie’s evidence, even in the presence of the high front vowel, which has a spirantising effect in only a few cases, e.g. Sangu B42 *nzìmà* ‘back, behind’ in contrast to *ny* before the back vowels in the Sangu words for meat, god, snake, bird, and body (Idiata-Mayombo 1993: 102).

Guthrie (followed by BLR) considered the basic classes of **jònì* ‘bird’ to be 7/8 or 12/13. However, the zone A and Bantoid evidence shows that the basic classes were 9/10, with the diminutive ‘birdie’ as an alternative formed in Bantu classes 7/8 or 12/13 (class 19 in Grassfields). The Grollemund Dataset lists over sixty forms of this word from zone A, Jarawan and Grassfields languages—all of them with *n*, *ɲ*, or *ny* (likewise Tiv and Ekoid). The later diminutives in other classes sometimes add prefixes to a stem with initial nasal, e.g. Shi JD53 *a-nyonyi* or Oku (Grassfields) *f̄-ńón*, or without, e.g. Luba-Katanga L33 *ky-onyi* (or *koni*), Tumbuka N21 *chi-yuni*.

One of the words most widespread in Bantu languages can be confidently reconstructed at PB node 1 as **nyàmà 9* ‘animal, meat’, with palatal nasals also frequent in Bantoid cognates. But the internal structure of the form is less clear. It might seem simplest to reconstruct the PB root as **yàmà* with a nasal class marker and assume reanalysis led to occasional forms with PREFIX-*nyàmà* in other classes (especially the animate class 1 *mu-*). But several factors argue for treating the palatal nasal as part of the PB root itself, as BLR reconstructs here exceptionally: **nyàmà*. First, it seems there are apparently no strong onsets of this word in Bantu languages. Also, unlike the word for ‘snake’, where some Grassfields and Beoid languages elide the initial nasal, the word for ‘animal’ always maintains an initial nasal in those languages. Possibly a pre-Bantoid proto-form

⁵³BLR 3653 prefers **jùmá*, but the Grassfields, Tiv and Ekoid cognates argue for reconstructing the front vowel for both Proto-Bantoid and PB, which was then sometimes affected by the subsequent bilabial.

had an *i*-prefix in some of these lexemes, e.g. ‘snake’ **i-noka* or **in-oka* or **in-noka* or **ni-oka*, but the persistence of the palatal nasal in ‘animal’ suggests it must have been part of the stem itself before Bantu.

4.2 Nouns supporting PB **nj/nz*

There is also a group of nouns with consistent strong reflexes like *nz*, *ndz*, *ndz*, and *nj* in descendant Bantu languages. Some examples are:

**jògù* 9 ‘elephant’ uniformly has strong reflexes: Mbonge A121 *njeku*, Basaa A43a *ndzɔk*, Mbula (Jarawan) *ñzû*, West Kele B22a *nɔk*, Bangi C32 *nzoku*, Kongo H16 *nza*, Ganda JE15 *enjovu*, Xhosa S41 *indlovu* (all from the Grollemund Dataset).

**jàdà* 9 ‘hunger; famine’ is recorded in all Bantu zones, consistently with strong reflexes: Akoose A15C *nzàà*, Bubi A31 *ecalá*, Mpongwe B11a *ndzana*, Mongo C61 *njala*, Pende L11 *nzala*, Jita JE25 *injara*, Hehe G62 *inzala*, Zezuru S12 *nzara*; as well as Grassfields Fefe *nziè* and Aghem *dziñ*, and Tiv *ijɔn* (all from the Grollemund Dataset).

**jidà* 9 ‘path’ is recorded in all Bantu zones, consistently with strong reflexes: Manenguba A15 *nzè*, Kulung (Jarawan) *njéré*, Eton A71 *zën*, Ngom B22b *nzela*, Punu B43 *nzila*, Rundi JD62 *inzira*, Lenje M61 *nshila*, Tsonga S53 *ndlela*; as well as Grassfields Fefe *má-ñ-zì* and Aghem *dzi* (all from the Grollemund Dataset).

Although our best examples of roots supporting PB **ny* occasionally develop strong forms, roots supporting PB **nz/nj* almost never weaken to *ny*. Accordingly, class 9 roots with mixed reflexes are best reconstructed with **ny*.

4.3 Nouns with mixed classes

So far, we have considered class 9 singular nouns that pair with class 10 plurals, and both classes are reconstructed by Meeussen (1967: 97) with prefix **n-*. But a good way to test the conditioning of **j* is to look at nouns which have allomorphs in different classes, i.e. in the phonological environments of different class prefixes.

Some of the best cases for testing nasal and non-nasal environments are nouns with singular cl. 11 prefix **dɔ-* and plural cl. 10 prefix **n-*. An example is **jàdà* 11/10 ‘fingernail’, for which forms in classes 7/8 and 5/6 are also recorded, often with a semantic difference, e.g. ‘finger’ or ‘hand’ (BLR 1558, C.S. 919-20, 1893-4). In those languages which maintain some form of the cl. 11 prefix (either fully or integrated into another class), we sometimes see the original weak nasal-less stem, e.g. Mbole D11 *lwála*, but also the nasal incorporated, e.g. Wumbvu B24

lipala, Bangi C32 *lónzáli*, and sometimes apparently even the cl. 7 prefix incorporated, e.g. Songola D24 *lù-chálà*. These nasal intrusions into cl. 11 show that analogy played a strong role in paradigm levelling, but the motivation might also be resolving an original hiatus from something like **dɔ-(y)ala*, hence Tetela C71 *lòkálà*.

North-Western class 10 (or 9) forms with palatal nasal reflexes, e.g. Mbuu A15A *nyàn*, Kulung (Jarawan) *nyááli*, Njem A84 *nyâ*, as well as non-nasal forms in other classes, e.g. Abo A42 *tf-ăt*, argue for reconstructing a PB weak stem also for the nasal variants, e.g. **n-(y)ada*. But strengthening of **ny > nz* is seen in certain languages and groups like B50-80 + H16, where almost all the nasal forms are strong. Thus, class 11/10 pairs like Bali-Teke B75 *liyala/ndzala*, Nilamba F31 *lɔala/nzala*, and Zezuru S12 *rwàrà/nzàrà* support PB **(y)àdá*, with some form of post-nasal strengthening (generalised in Nilamba *nzoka* ‘snake’, but not in Zezuru *nyóká*). This post-nasal strengthening or analogy must be localised because a mixture of its presence and its absence is seen among related languages: Kaningi Nord B602 *lepàra* and Atsitsege B701 *lipala*, but “Teke d’Ibali (Congo)” B71a1b *lindzala* and Wuumu B78 *linzál*.

The lexemes **jiki* 9/10 ‘bee’ and 14 ‘honey’ provide another set of allomorphs. Guthrie gives more than thirty forms for ‘honey’ from every zone, yet none of them has a stop or even a glide as an onset to the root: e.g. Bubi A31 *b-ɔɛ*, Bulu A74a *w-ɔɛ*, Mfinu B83 *bɛiɛ*, Kuyu E51 *ɔ-ɔkɛ*, Manyanga H16b *bw-iki*, Luba-Katanga L33 *bu-uki*, Yao P21 *u-uci*, Xhosa S41 *uβ-usi*.⁵⁴ In that sense, the data looks like that of the vowel-initial nominal roots discussed earlier, for example, **bɔ-átò* 14 ‘canoe’. For ‘bee’ (with the nasalising prefix of classes 9/10, and by extension 11), Guthrie provides evidence only for forms in *ny-* in zones A and B: Bubi A31 *lɔ-nyɔɛ*, Mpongwe B11a *nyɔɛ*, Ngom B22b *ɔa-nyɔɛ*, Lumbu B44 *nyosi*, Nzebi B52 *nyux(i)*, Bali-Teke B75 *nnũũ*. Similar forms in *ny-* are found throughout all regions of Bantu. So, the uniform testimony of the North-Western languages, with parallels in other zones, supports a PB weak onset for both words, e.g. **bɔ(v)ki* ‘honey’ and **ni-ɔki* ‘bee’ (or **bɔ-yɔki* or **n-yɔki*).⁵⁵ In that case, the strong forms of ‘bee’ in a number of Bantu languages (e.g. Bangi C32 *lɔ-ndzɔj*, Nande JD42 *en-zuki*, Ila M63 *in-zuki*) must once again be due to some post-nasal

⁵⁴C.S. 962, 2003-4, 2113, 2156-7, 2159 (Guthrie 1967: 124–125, §74.31-4).

⁵⁵The original character of the root’s first vowel is unclear. It could be a front vowel which was affected by the back vowel of the cl. 14 and 11 prefixes, or it could be a back vowel which was affected by the glide *y* or *V₂*. The editors of BLR3 reconstruct the front vowel, but the evidence of most zones (including AB) argues for the back vowel at PB node 1. But cf. Jarawan *i* in ‘honey’: Mbula *nyi*, Jaku *binyi*, Bankal *nyi* (Gerhardt 1982: 92).

strengthening of *ny* > *nz*, occasionally leading to mixed paradigms like Rundi JD62 *uru-yuki/in-zuki* 11/10.

Another example of apparent nasal strengthening would be **jòndò* ‘hammer, anvil’ (C.S. 965, 2171, 706), with weak reflexes in various classes: Pongo A26 *ɛ-ɔndɔ* 7/8 ‘axe’, Ngombe C41 *ɛ-yɔndɔ* 7/8 ‘hammer’, and Ngom B22b *y-ɔndɔ* 19/13 ‘axe’, but strong in Benga A34 *njɔndɔ* 9/6 ‘hammer’. Cf. Tiv *nɔndɔ/ɿ-nyɔndɔ*.

4.4 Summary of **ny* and **nj/nz*

There are two sets of nasal patterns for **j* with distinctly different onsets: palatal nasal (*ny*) and stronger combinations (*nj*, *nz*, *ndz*, etc.). In fact, in the dozens of languages in zones A and B which have reflexes of both ‘snake’ (apparently **yókà*) and ‘elephant’ (apparently **zògù* or **jògù*), none has the same onset for the two words. The same distinction in zones G and S shows that this is not an areal phenomenon and should be reconstructed for PB.⁵⁶

If one wanted to reconstruct both these sets under one proto-phoneme, one would likely start at some pre-Bantu stage with the palatal nasal form and generate the strong nasal form as a conditioned allophone, since that is the directionality seen in the examples above: strong PB **nz/nj* forms (seen uniformly in **jògù* ‘elephant’, **jàdà* ‘hunger’, **jìdà* ‘path’) rarely weaken in Bantu languages, whereas PB **ny* was often strengthened in various ways. This strengthening is seen both in class 9/10 lexemes like **yókà* ‘snake’ and lexemes of mixed classes like *(*y*)*ókí* 9/10 ‘bee’ and 14 ‘honey’. For the lexemes considered in this section, neither the influence of tone nor a subsequent vowel would give us a phonological rule to generate the strong reflexes. A possible rule could be based on C₂:⁵⁷ that voiced C₂ leads to a strong reflex of C₁ after nasal prefixes, e.g. **jàdà* ‘hunger; famine’, **jàdí* ‘lightning’, **jìdà* ‘path’, **jògù* ‘elephant’, **jògù* ‘groundnut’; and the lack of C₂ would also need to qualify, e.g. **jù* ‘house’ and **jái* ‘outside’. But apparent exceptions can be found, and the status and age of each lexeme would need to be studied. Any phonological rule would also need to account for variations in strong and weak reflexes of **nj* in C₂ as well. Even if a rule for allocating allophones could be found, it would have started in some pre-Bantu stage to account for parallels in other Bantoid groups, and it is not clear how long it operated or when the allophones eventually phonemicised.

⁵⁶For nouns maintaining this distinction in Tswana, where the contrast is between weak *n* and strong *tʰ*, see Creissels (1999: 306–307).

⁵⁷This is the approach of Meeussen (1973: 9–10). A phenomenon like Meinhof’s Rule (nasal assimilation of N-C₁ before nasal or nasalised C₂ in nouns) in some Bantu languages supports the consideration of C₂ influence on C₁.

Accordingly, for BLR's nasalised or post-nasal *j, at the stage of PB node 1, it seems simplest to separately reconstruct initial *ny and *nj/nz. For 'animal, meat', one may also maintain a structure like BLR's *n-nyàmà.

5 Conclusions

In looking at environments where PB *j has been reconstructed, we have seen that it is a collection of distinct stories which require separate reconstructions, some clearer than others. Most often *j is really just a placeholder for various effects that occurred at morpheme boundaries and needs to be deconstructed, not reconstructed. To summarise, I have proposed replacing BLR3 *j and *nj with a PB inventory of this sort:

- initial \emptyset (in most nouns and some verbs)
e.g. *átò 'canoe', *(y)át 'split', *ícò 'eye';
- both ny and nz (in cl. 9 or 10 onsets and sometimes medially)
e.g. *nyókà 'snake', *nzògù 'elephant', *nyànzà 9 'lake';
- y (in some onsets and medially)⁵⁸
e.g. *yíb 'steal', *káyá 'leaf', *íyad 'be full'??

This would mean removing *j from the reconstructed consonant chart in Meeussen (1967: 83), and in all his reconstructed forms. Likewise, there is no need for *j in the reconstruction of the pronominal prefixes (augments) of classes 1, 9 and 10 (*ju, *ji, and *ji respectively) nor in the demonstratives built on them (Meeussen 1967: 97, 107).

What are the implications for PB phonology and its evolution?

Vowel-initial roots The reconstruction of vowel-initial roots is an old idea, which was never really refuted. The Homburger-Coupez tradition put initial *g'/j in these roots and led to an expectation of CV-syllable structure in Bantu lexemes, but certain PB inflectional prefixes have always been reconstructed with initial vowels and thus inflected forms are often vowel-initial. It is clearly easier from the phylogenetic viewpoint to explain the exceptional strong (z/j) forms in a few languages than the weak (y, w, \emptyset) forms in the great majority of languages

⁵⁸BLR3 has already addressed other types of stems where Meeussen (1967: 82) considered it "difficult to distinguish VV from VjV, e.g. -béjad/-bé(j)ad/-béad- «plant, sow»" – in this case reconstructing BLR 165 *bíad.

across most major branches. The failure of **j* to undergo Bantu Spirantisation and the extreme rarity of **j* at C₂ in roots is not surprising if reconstructed **j* is understood as a construct based on later effects seen at morpheme boundaries in various languages or groupings. A number of these roots are found in Bantoid languages and further study may justify a reconstruction of some words with a consonant at some pre-Bantu stage, but our goal here was simply to clarify PB node 1.

Distinguishing **ny* and **nz* Whatever the pre-Bantu history, for PB one should make a distinction between **ny* and some other nasal sequence. While [nz], [ndʒ] and [nj] all frequently occur as “strong” reflexes of BLR’s **nj*, the most common is perhaps [nz], so **nz* is a reasonable choice for the PB symbol, and it has the advantage of being detached from the connotations of the current symbol **j*. Of course, the specific phonetic features of any symbol will depend on further study of Bantoid data and directional tendencies in sound changes involving these sorts of fricatives. Since [s] has been seen as the likely phonetic value of **c*, it might be useful to remove the palatal series altogether and follow Greenberg in relabelling both **c* and **j* as **s* and **z*. The presence of **ny* and **nz* in the PB inventory might suggest that independent **y* and **z* were more frequent at some pre-Bantu stage, just as they were later in many Bantu branches.

Is **y* part of the PB phonemic inventory? Many contemporary Bantu or Bantoid languages have semi-vowels, so it would not be surprising to include them in the PB system. Or perhaps the better question is at what stage(s) to reconstruct them.⁵⁹ The strongest cases for an early *y* that we have seen are in medial position in a few nouns, verbs in **iya*, and in the initial position of some verb stems. Also, if we are reconstructing **ny* (**ɲ*) for PB, it would not be a surprise to include a palatal glide. Its initial frequency might not have been high, but various processes have increased its frequency. The extent to which /y/ or /w/ should be reconstructed either as a phoneme or allophone (and at what stages) needs fresh study, free from the legacy of current unitary **j*. One might ask whether PB had rules for vowel contraction or hiatus resolution.⁶⁰

⁵⁹Nurse & Hinnebusch (1993: 61) in their overview of the phonological system of Proto-Sabaki: “the glides *w* and *y* are unchanged from earlier proto periods.” Meinhof et al. (1932: 28) also reconstructed allophonic semi-vowels **y̆* and **w̆* (from **î* and **û*).

⁶⁰Cf. Meeussen (1967: 82): “A closed vowel (i, u; i, u; e, o) followed by a more open vowel (i, u, e, o, a) is sufficient to account for the occurrence of semi-vowels in the present-day languages. It is often difficult to distinguish VV from VjV (which will usually be written here as V(j)V)”

Glide creation and strengthening Several times, we have seen variation between strong (*z, j, ʒ*, etc.) and weak (*∅, y*) reflexes in closely related groups of Bantu languages. This reinforces the cross-linguistic evidence discussed at the beginning that glides can often become fricatives and sometimes vice-versa. Environments that favour strengthening in the history of Bantu are preposed *i* and *n* from a variety of inflectional prefixes, e.g. **nyókà* ‘snake’ > Ngombe C41 *ndʒo*, Chewa N31b *njoka*. But languages can also make changes elsewhere, e.g. Eastern Bantu **kóyò* ‘fig-tree’ > Yao *kuju*.⁶¹ Faytak (2014) presents several examples of “high vowel fricativization” by which front high vowels change to coronal fricatives, i.e. [i] → [z] or [ʒ]. This process “that ends in complete fricativization of reconstructible **i* and **y*” (2014: 60) could be one of the routes of what appears to be strengthening of glides.⁶² Glides, nasals, stops or fricatives could also arise at morpheme boundaries as incorporations of class or infinitive prefixes (**n* or **kw*) or other analogical processes.

Acknowledgements

I am grateful to the participants at the Bantu 6 Conference for their comments and questions, and particularly to Koen Bostoen, Gérard Philippon, and the editors and reviewers of this volume for their many helpful suggestions.

References

- Atta, Samuel Ebongkome. 1993. *The phonology of Lukundu (Bakundu)*. Yaoundé: Université de Yaoundé I. (MA thesis).
- Baker, Gary K. & Caroline R. Wiltshire. 2003. An OT treatment of palatal fortition in Argentinian Spanish. In Ana Teresa Pérez-Leroux & Yves Roberge (eds.), *Romance linguistics: Theory and acquisition. Selected papers from the 32nd Linguistic Symposium on Romance Languages (LSRL), Toronto, April 2002* (Current Issues in Linguistic Theory 244), 33–48. Amsterdam: John Benjamins.

⁶¹See Mortensen (2012) for the emergence of obstruents after high vowels as a recurring sound change, including examples in Grassfields Bantu.

⁶²Following up on Connell (2007), who reviews fricative vowel phonemes in a Mambiloid lect, Faytak (2014: 64–78) discusses fricativisation in Grassfields and other Bantoid languages but mostly with regard to back vowels. Cf. also Hall (2014) for Westphalian German Spirantisation: “the change from an original prevocalic long vowel to the corresponding short vowel plus fricative (i.e. [ʏ]).”

- Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and divergence in the Bantu languages: Perspectives from a lexicostatistic study* (Annals – Human Sciences 162). Tervuren: Royal Museum for Central Africa. https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/lexicostatistic-study-bantu-languages.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Baucom, Kenneth L. 1975. The phonology of proto-Wambo. *African Studies* 34(3). 165–184.
- Bostoën, Koen. 2005. *Des mots et des pots en bantou : une approche linguistique de l'histoire de la céramique en Afrique* (Schriften zur Afrikanistik / Research in African Studies 9). Frankfurt am Main: Peter Lang.
- Bostoën, Koen. 2009. Shanjo and Fwe as part of Bantu Botatwe: A diachronic phonological approach. In Akinloye Ojo & Lioba Moshi (eds.), *Selected Proceedings of the 39th Annual Conference on African Linguistics*, 110–130. Somerville, MA: Cascadilla Proceedings Project.
- Bostoën, Koen. 2019. Reconstructing Proto-Bantu. In Mark Van de Velde, Koen Bostoën, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 308–334. Milton Park, Abingdon: Routledge.
- Bostoën, Koen & Yvonne Bastin. 2016. Bantu lexical reconstruction. In *Oxford handbooks online*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199935345.013.36.
- Botne, Robert. 1991. Variation and word formation in Proto-Bantu: The case of *-YIKAD-. *Afrika und Übersee: Sprachen, Kulturen* 74(2). 247–268.
- Bourquin, Walther. 1923. *Neue Ur-Bantu-Wortstämme, nebst einem Beitrag zur Erforschung der Bantu-Wurzeln* (Beiheft zur Zeitschrift für Eingeborenen-Sprachen 5). Berlin: Dietrich Reimer (Ernst Vohsen).
- Bulkens, Annelies. 2009. Quelques thèmes pour ‘pirogue’ dans les langues bantoues. *Africana Linguistica* 15. 27–58.
- Bybee, Joan L. & Shelece Easterday. 2019. Consonant strengthening: A crosslinguistic survey and articulatory proposal. *Linguistic Typology* 23(2). 263–302.
- Connell, Bruce. 2007. Mambila fricative vowels and Bantu spirantization. *Africana Linguistica* 13. 7–31.
- Coupez, André. 1954. Les phonèmes bantous g et j non précédés de nasale. *Zaire : revue congolaise / Zaire: Congolees tijdschrift* 8. 157–161.

- Coupez, André, Yvonne Bastin & Evariste Mumba (eds.). 1998. *Bantu lexical reconstructions 2 / Reconstructions lexicales bantoues 2*. Tervuren: Royal Museum for Central Africa.
- Crane, Thera M. 2011. *Beyond time: Temporal and extra-temporal functions of tense and aspect marking in Totela, a Bantu language of Zambia*. Berkeley, CA: University of California. (Doctoral dissertation).
- Crane, Thera M., Larry M. Hyman & Simon Nsielanga Tukumu. 2011. *A grammar of Nzadi [B.865]: A Bantu language of the Democratic Republic of Congo* (University of California Publications in Linguistics 147). Berkeley, CA: University of California Press.
- Creissels, Denis. 1999. Remarks on the sound correspondences between Proto-Bantu and Tswana (S.31), with particular attention to problems involving *j (or *y), *j̥ and sequences *NC. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 297–334. Stanford, CA: Center for the Study of Language & Information.
- de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87–162 + supplementary material online.
- den Besten, Margaret G. G. 2016. *A phonology of Batanga (Bantu A30); The Bano's dialect*. Leiden: Leiden University. (MA thesis).
- Elias, Philip, Jacqueline Leroy & Jan Voorhoeve. 1984. Mbam-Nkam or Eastern Grassfields. *Afrika und Übersee: Sprachen, Kulturen* 67(1). 31–107.
- Ellington, John Ernest. 1977. *Aspects of the Tiene language*. Madison, WI: University of Wisconsin–Madison. (Doctoral dissertation).
- Essono, Jean-J. Marie. 2000. *L'ewondo : langue bantu du Cameroun : phonologie – morphologie – syntaxe*. Yaoundé: Presses de l'Université catholique d'Afrique centrale.
- Faytak, Matthew. 2014. High vowel fricativization and chain shift. *UC Berkeley Phonology Lab Annual Report* 10. 52–100.
- Fivaz, Derek. 1986. *A reference grammar of Oshindonga (Wambo)* (African Studies of the Academy 1). Windhoek: Department of African Languages.
- Gerhardt, Ludwig. 1982. Jarawan Bantu: The mistaken identity of the Bantu who turned north. *Afrika und Übersee: Sprachen, Kulturen* 65(1). 75–87.
- Greenberg, Joseph H. 1969. Review of I. Fodor's (1966) *The problems in the classification of the African languages: Methodological and theoretical conclusions concerning the classification system of Joseph H. Greenberg*. *Language* 45. 427–432.

- Grégoire, Claire & Jean-Paul Rekanga. 1994. Nouvelles hypothèses diachroniques sur la cl 10b du myene-nkomi (B11e). *Africana Linguistica* 11. 71–77.
- Grollemund, Rebecca. 2012. *Nouvelles approches en classification : application aux langues bantu du nord-ouest*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guthrie, Malcolm. 1953. *The Bantu languages of Western Equatorial Africa* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1965. Comparative Bantu: A preview. *Journal of African Languages* 4(1). 40–45.
- Guthrie, Malcolm. 1967. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 1: The comparative linguistics of the Bantu languages*. Farnborough: Gregg Press.
- Guthrie, Malcolm. 1967–71. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages (4 volumes)*. Farnborough: Gregg International.
- Hall, T. Alan. 2014. The analysis of Westphalian German spirantization. *Diachronica* 31(2). 223–266.
- Hannan, Michael. 1974. *Standard Shona dictionary*. 2nd revised edition. Salisbury: Rhodesia Literature Bureau.
- Hedinger, Robert. 1987. *The Manenguba languages (Bantu A.15, Mbo cluster) of Cameroon*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Homburger, Liliás. 1913. *Étude sur la phonétique historique du bantou* (Bibliothèque de l'École des hautes études. IVe section, Sciences historiques et philologiques 209). Paris: H. Champion.
- Hualde, José Ignacio. 2011. Sound change. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology* (Blackwell Companions to Linguistics), 2214–2235. Malden, MA: Wiley-Blackwell.
- Hyman, Larry M. & Francis X. Katamba. 1999. The syllable in Luganda phonology and morphology. In Harry van der Hulst & Nancy A. Ritter (eds.), *The syllable: Views and facts* (Studies in Generative Grammar 45), 349–416. Berlin: Mouton de Gruyter.

- Idiata-Mayombo, Daniel-Franck. 1993. Eléments de phonologie diachronique du isangu (B. 42). *Pholia* 8. 67–108.
- Kamba Muzenga, Jean-Georges. 1988. Comportement du préfixe nominal de classe 5 en bantou. *Annales Æquatoria* 9. 89–131.
- Koni Muluwa, Joseph & Koen Bostoen. 2015. *Lexique comparé des langues bantu du Kwilu (République démocratique du Congo): Français – anglais – 21 langues bantu (B, C, H, K, L)* (Grammatical Analyses of African Languages 48). Cologne: Rüdiger Köppe.
- Maho, Jouni F. 2007. *A referential classification of the Bantu languages: Keeping Guthrie's system updated*. Preliminary (online) version # 6 (dated 28 January 2007). Accessed 31 January 2007, but no longer online.
- Marlo, Michael R. 2015. Exceptional properties of the reflexive in Bantu languages. *Nordic Journal of African Studies* 24(1). 1–22.
- Mathangwane, Joyce T. 1999. *Ikalanga phonetics and phonology: A synchronic and diachronic study* (Stanford Monographs in African Language). Stanford, CA: Center for the Study of Language & Information.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeussen, Achiel Emiel. 1969. *Bantu lexical reconstructions*. Unpublished ms., 28 pp. Tervuren: Royal Museum for Central Africa.
- Meeussen, Achiel Emiel. 1973. Comparative Bantu: Test cases for method. *African Language Studies* 14. 6–18.
- Meeussen, Achiel Emiel & Archibald N. Tucker. 1955. Les phonèmes du ganda et du bantou commun. *Africa: Journal of the International African Institute* 25(2). 170–180.
- Meinhof, Carl. 1899. *Grundriß einer Lautlehre der Bantusprachen, nebst Anleitung zur Aufnahme von Bantusprachen. Anhang: Verzeichnis von Bantuwortstämmen* (Abhandlungen für die Kunde des Morgenlandes 11(2)). Leipzig: F.A. Brockhaus.
- Meinhof, Carl. 1910. *Grundriß einer Lautlehre der Bantusprachen, nebst Anleitung zur Aufnahme von Bantusprachen. Anhang: Verzeichnis von Bantuwortstämmen*. 2nd revised edition. Berlin: Dietrich Reimer (Ernst Vohsen).
- Meinhof, Carl, Alice Werner & Nicolaas J. Van Warmelo. 1932. *Introduction to the phonology of the Bantu languages*. Being the English version of *Grundriß einer Lautlehre der Bantusprachen*. Berlin: Dietrich Reimer (Ernst Vohsen).
- Mortensen, David R. 2012. The emergence of obstruents after high vowels. *Diachronica* 29(4). 434–470.

- Mous, Maarten & Anneke Breedveld. 1986. A dialectometrical study of some Bantu languages (A. 40 – A. 60) of Cameroon. In Gladys Guarisma & Wilhelm J. G. Möhlig (eds.), *La méthode dialectométrique appliquée aux langues africaines*, 177–241. Berlin: Dietrich Reimer.
- Ngunga, Armino S. A. 2000. *Phonology and morphology of the Ciyao verb* (Stanford Monographs in African Languages). Stanford, CA: Center for the Study of Language & Information.
- Nurse, Derek. 1979. *Classification of the Chaga dialects: Language and history on Kilimanjaro, the Taita Hills and the Pare Mountains*. Hamburg: Helmut Buske.
- Nurse, Derek & Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history* (University of California Publications in Linguistics 121). Berkeley, CA: University of California Press.
- Nurse, Derek & Gérard Philippson. 2003. Towards a historical classification of the Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 164–181. London: Routledge.
- Odden, David. 2003. Rufiji-Ruvuma (N10, P10-P20). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 529–545. London: Routledge.
- Pacchiarotti, Sara, Natalia Chousou-Polydouri & Koen Bostoen. 2019. Untangling the West-Coastal Bantu mess: Identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica* 25. 155–229.
- Philippson, Gérard & Rebecca Grollemund. 2019. Classifying Bantu languages. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 335–354. Milton Park, Abingdon: Routledge.
- Rekanga, Jean-Paul. 1994. Les réflexes du protobantou en myene-nkomi, langue bantoue du Gabon (B11e). *Africana Linguistica* 11. 149–167.
- Schadeberg, Thilo C. 2003. Historical linguistics. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 143–163. London: Routledge.
- Sihler, Andrew L. 1995. *New comparative grammar of Greek and Latin*. New York, NY: Oxford University Press.
- Stappers, Leo. 1971. Esquisse de la langue lengola. *Africana Linguistica* 5. 255–307.
- Teil-Dautrey, Gisèle. 2004. *Lexiques proto-bantous : étude des cooccurrences segmentales et supra-segmentales*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Turvey, B. H. C., W. Zimmermann & G. B. Taapopi. 1977. *Kwanyama – English dictionary*. Johannesburg: Witwatersrand University Press.

- Van de Velde, Mark. 2006. *A description of Eton: Phonology, morphology, basic syntax and lexicon*. Leuven: Katholieke Universiteit Leuven. (Doctoral dissertation).
- Yanes, Serge & Eyinga Essam Moise. 1987. *Dictionnaire Boulou – français, français – Boulou; avec grammaire*. Sangmelima: P. Monti.

Part II

Proto-Bantu verbal morphology

Chapter 3

Tense in Proto-Bantu

Derek Nurse^a & John R. Watters^b

^aIndependent Scholar ^bSIL International

The focus of this chapter is the appearance of tense in Proto-Bantu (PB). Most Niger-Congo (NC) languages are aspect-prominent, having no tense contrasts, and the same is generally assumed for ancestral Proto-Niger-Congo. PB emerged from part of an eastern subgroup of NC to which we refer as Bantoid. Some 5000 years ago or earlier, tense was innovated at an early stage in a region along and to the east of the Cameroon Volcanic Line. This means that tense is not unique to PB but is inherited by PB from its forebears. We propose two lines of verbal development for Narrow Bantu (NB) based on the verbal phenomena we traced. The data did not always allow us to base our analysis on the strict application of the Comparative Method to the exponents of tense and aspect, but examination of specific systematic features of the verbal systems in NB and parts of Bantoid led us to infer plausible paths of verbal development to explain the data.

1 Introduction

This chapter is organised as follows. §2 deals with what can reasonably be reconstructed for Proto-Bantu (PB). Our reconstruction differs somewhat from that in two earlier works, partly because we took into consideration new evidence from the north-western Narrow Bantu (NB) languages. §3 sets out something of the rich and complicated tense systems that have evolved in NB's eastern Bantoid siblings: Grassfields Bantu, Tikar, Beboid, Yemne-Kimbi, and parts of Mambiloid. In §4 we integrate the first two sections, by juxtaposing the PB reconstructions with what we find in eastern Bantoid.

Reconstruction of tense in these eastern Bantoid languages differs crucially from the reconstruction of tense in other language families, e.g. Romance (Indo-European). Tense categories and their morphological exponents in today's Romance languages can be mostly shown to develop organically from a single set



Derek Nurse & John R. Watters. 2022. Tense in Proto-Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 105–171. Berlin: Language Science Press.

DOI: 10.5281/zenodo.7575819 

of categories and exponents in Latin. That is not the case for the eastern Bantoid languages: while their categories are generally relatable, each has a distinct set of morphological exponents, not derivable from a common ancestral system. We think that tense contrasts developed in two stages. The initial stage saw a single past and maybe a single future developing, most likely at one geographical locus, probably in an early eastern Bantoid lect¹ or a small set of closely related eastern Bantoid lects, in south-western Cameroon. At a later stage, multiple pasts and future contrasts evolved from their respective single earlier tense, probably in Eastern Grassfields. In both cases, we see tense diffusing out from an initial point into adjacent groups, each group imitating the tense category/-ies but using its own morphology, hence the disparity in morphological exponence. Our focus is to identify within the Bantoid variation those exponents of tense that we can relate to reconstructed PB forms.

We would add three caveats. First, any distinctions we may make between groupings within NB on the basis of differing distributions of verbal features, e.g. in §2.2.1 and §2.2.2 below, may or may not correspond to distinctions made by others using different features or criteria. We are not proposing a new classification, but rather we are attempting to account for periods of verbal development within PB, based on specific phenomena.² We think proto-languages are like real languages in having temporal and regional variation. Our distinction might or might not correspond to proposals made by others using different methods.

Second, reconstructing cognitive-systemic-morphological entities such as tense/aspect (TA) differs from the classic Comparative Method (CM). Where the CM has a long and established tradition involving a defined methodology and mostly well-defined results, it will be seen that what we are doing here has no established tradition. It involves some results that few would disagree with, but also several issues for which we have several plausible explanations but no tools to make a definite choice among them. Probability plays an important role in this chapter.

Third, the two foci of this chapter are the Eastern Grassfields languages and the presence of tense and aspect in PB. However, we are mindful that some readers may turn back when faced with the combination of a mass of unfamiliar languages and an unfamiliar topic and/or theory, so we – and our editors – have

¹We use “lect” as a neutral term to cover language, dialect, or other local varieties.

²The latest overall classification of Bantu languages is Grollemund et al. (2015). It is a phylogeny of over 400 Bantu languages relying on basic vocabulary. Despite our reservations about lexicon-based quantitative approaches to language classification, we can identify the present study on the origin of tense in Bantu and Bantoid as primarily concerning nodes 0 and 1 in the tree proposed by Grollemund et al. (2015).

tried to make the content transparent. For definitions of central terminology, see Appendix A. For geographical location, see Figure 3 in the introduction to §3, Figure 5 in §3.4.3, and Figure 6 in §4.4.

2 Reconstructing tense for PB

There have been two previous attempts at reconstructing tense for PB: Meeussen (1967: 112–113) and Nurse (2008: 226–283).³ Their conclusions are quite similar. This is not surprising as their basic assumptions and procedures are similar. They surveyed pre-stem and final vowel (FV) morphemes occurring widely across an array of NB languages and then assembled them to represent categories.⁴ These categories involved drawing on their experience with languages mainly in the east, south, and centre of the NB area. Moreover, they assumed the PB verb had an agglutinating structure. Both scholars worked from morphemes to meaning, because it is easier to work from concrete morphology and structure than from the more elusive semantics.

Following the phylogenetic tree proposed for NB in Figure 1 of Grollemund et al. (2015: 2), we include in this chapter a short but crucial section on tense/aspect categories in NB languages of the North-Western Bantu Cameroon (NWB Cameroon) and Gabon (NWB Gabon), Central-Western Bantu (CWB), and West-Western Bantu (WWB). These include languages of Guthrie zones A, B, C, and D, namely NWB Cameroon (A10-70), NWB Gabon (A80-90, B10-30), CWB (C10-18 and D10-30), and WWB (B40-80 and H10-30-42).⁵ We note that languages of zones D10, D20, and D30 are found in both CWB and Eastern Bantu (EB) in Grollemund et al. (2015). Our concern is with those in CWB. The lower branches in the phylogeny of Grollemund et al. (2015), i.e. EB and South-Western Bantu (SWB), are only of limited relevance to our present purposes.

Of the north-western NB languages, our particular interest is the NWB Cameroon and NWB Gabon languages, partly because Meeussen and Nurse paid

³We do not present the data here, leaving it to readers to consult them. Meeussen's database was (part of) Bastin (1975). Nurse provides his data in Nurse (2019). Previous argumentation is also not repeated but can be seen in Meeussen (1967), Nurse & Philippson (2006), and Nurse (2008).

⁴Meeussen calls them "tense formulae", Nurse "tense-aspect forms". Meeussen uses "tense" as a single cover term for several categories (tense, aspect, focus, etc.) here treated as distinct. Meeussen's formulae "are intended as illustrating guesses rather than as real reconstructions" (Meeussen 1967: 113).

⁵For Guthrie's zones (A, B, ...) and groups (A10, A20, ...), and his referential classification of the Bantu languages in general, see Guthrie (1948; 1971).

less attention to them, and partly because they are involved in what studies to date consider the borderland between NB and other Bantoid language groups. Along with the NWB languages, we engage also with the outgroup Grassfields languages. Of the Bantoid groups along the borderland with NB, the Grassfields is geographically closest to NWB languages and displays behaviour with tense and aspect that indicate a close relationship with NWB.

What follows in Table 1 is a partial comparison, including only pre-stem forms referring partly or exclusively to tense and not primarily to non-tense categories. It includes the FV morphemes **-a*, **-ile*, **-a(n)g-a*, the latter of which Meeussen treats as ‘pre-final’ (see also Sebasoni 1967). Brackets in the second column indicate doubtful status.

Table 1: Tense reconstructions in Meeussen (1967) and Nurse (2008)

Meeussen (1967)	Nurse (2008)
*á-stem- <i>a</i> preterite IPFV	* <i>a</i> -stem- <i>ile</i> past RET
* <i>a</i> -stem- <i>a</i> recent IPFV	—
*á-stem- <i>ide</i> preterite PFV	* <i>a</i> -stem- <i>ile</i> past RET
* <i>a</i> -stem- <i>idé</i> recent PFV	—
—	*∅-stem- <i>ile</i> present RET
*∅-stem- <i>a</i> present 1 IPFV (= CONJ)	*∅-stem- <i>a</i> present
—	* <i>a</i> -stem- <i>ang-a</i> past IPFV
—	*∅-stem- <i>ang-a</i> present IPFV
* <i>da</i> -stem- <i>a</i> present 2 IPFV (= DISJ)	* <i>la</i> -stem- <i>a</i> DISJ (<i>-laa</i> -stem- <i>a</i> future)
* <i>ka</i> -stem- <i>a</i> future	(* <i>ka</i> -stem- <i>a</i> itive/future)

The use of uppercase (e.g. PAST, IPFV) refers to a concrete category in a specific language, but the use of lowercase (past, imperfective) refers to a general category.

Note that Meeussen has a binary contrast for the past between preterite and recent past while Nurse has only one past. See §2.2 for discussion.

These reconstructed morphemes/formulae reflect primarily what occurs in NB outside the NWB languages. However, the NWB languages are crucial to reconstruct PB by identifying what are retentions of PB and what are innovations.

Meeussen (1967) and Nurse (2008) also have in common that they treat PB as the parent language of all current NB languages. They set out mainly to account for the variation they found across NB. Relative to tense they give particular attention up to node 5 in Grollemund et al. (2015), i.e. excluding NWB. This contrasts with our goal. We seek to review PB tense from node 4 up to node 1 and

then bring in node 0. Node 0 involves including Bantoid languages outside NB that may shed light on the development of tense in NB within Bantoid (cf. end of §2.2.1). However, as stated above, we also do consider NB languages from Guthrie’s zones B, C, D, and H, which are north-western geographically speaking, but belong to the CWB and WWB branches in genealogical terms. When we use north-western in a purely geographical sense, we will not abbreviate it. When we refer to the phylogeny of Grollemund et al. (2015), we will use the abbreviations NWB, CWB and WWB.

2.1 The north-western NB languages

Structures expressing TA in north-western NB languages share certain features. Significantly, nearly all have three structures with no pre-stem morpheme reflexes (“pre-stem zero (- \emptyset -)”) and reflexes of the characteristic suffixes in the FV slot. In NB, the pre-stem position typically indicates tense while FV is the dedicated position for aspect. Table 2 displays these recurrent structures.

Table 2: TA structures in north-western NB without tense prefixes

In an aspect system		In a tense-prominent system	
* \emptyset -stem- <i>a</i>	Imperfective	* \emptyset -stem- <i>a</i>	Present
* \emptyset -stem- <i>í</i>	Perfective	* \emptyset -stem- <i>í</i>	Past
* \emptyset -stem- <i>aga</i>	Habitual/Iterative	* \emptyset -stem- <i>aga</i>	Habitual/Iterative

Sebasoni (1967: 131) considers the “Habitual/Iterative” in Table 2 to involve a set of three forms distributed in complementary fashion across NB. Specifically, “-*ag*- prevails in the north-east and east of the NB region, -*ak*- in the north, and -*anga*- in the west and south” [our translation from the original French].

In the perfective * \emptyset -stem-*i* high tone is marked. Where high tone is marked we are fairly confident of the tone. Lack of any tone marking means either low tone or that we are unsure because the data is not conclusive (Nurse & Philippson 2006).

In Table 2, the structures in the left column express aspectual meaning, while those on the right express a mix of aspectual and tense meanings. This is a set of forms which nicely bridge the shift from an aspect-prominent to a tense-prominent system, or thus from Niger-Congo (NC) to NWB. Indeed, the structures in the column on the left occur often across NC (-*ag(a)* in Bantoid, less frequent elsewhere in NC) and they form the skeleton for NWB systems, exemplified in (1).

- (1) Benga A34 (Nurse 2019: Addendum 1)
mbi-a-kal-a 'I talk' (1SG-Ø-talk-a)
mbi-Ø-kal-i 'I talked'
mbi-Ø-kal-ak-a 'I am talking'

Tense-prominent systems in north-western languages also differ in certain ways. For example, most have a small set of tense contrasts, with one/two pasts and one future (Lundu-Akoose A11-15C, Duala A24, Benga A34, Njem A84, Kako A93, Himba-Vove B302-305, Mbuun B87, Babole C101, Mboshi C25, Mbudza C36c, Gesogo C53, etc.), while a few have developed multiple contrasts (Kpe A22, Basaa-Nen-Maande A43a-44-46, Kpa A53, Yangben-Gunu A62A-622, Ewondo A72a, Kwakum A91, Myene-Nkomi B11e, Kota B25, Duma-Nzebi B51-52, Ndumu B63, Teke Yaa B73c, Boma-Yanzi B82-85, Kela C75, Bushong C83, Mbole D11).⁶ To put these on a map gives a haphazard impression as we considered only two languages per Guthrie group (A10, A20, etc.). The picture would probably be more coherent if we included data for all north-western languages. Several morphemes involved in expressing the extra categories in the multiple contrasts in Basaa-Nen, ?Maande, Kpa remain to be investigated. Some of these resemble morphemes in Bantoid languages. For example, a characteristic feature of Bamileke lects is a structure of the shape *N-B*,⁷ which occurs in imperfectives and P1.⁸ It also occurs in Basaa: *a-n-jé* 'he ate P1' and *a-ń-jé* 'he eats'.

2.2 Past tense in PB

2.2.1 One or more pre-stem *a-* 'past' in PB?

Across NB, *a*-⁹ is by far the commonest TA pre-stem marker and the commonest marker of past reference. As can be seen in Table 1, Meeussen postulates a consistent binary contrast between *á-* 'preterite' and *a-* 'recent past'. Nurse has but a single *a-* 'past', based on Nurse & Philippson (2006), which used as its database the same 100 languages as in Nurse (2008). 75% of the languages in that database have a form of *a-* with some past reference, which might mean it is the only past

⁶The referential Bantu language codes seen here, first introduced by Guthrie (1948; 1971), were last updated by Maho (2009).

⁷*N* represents a homorganic nasal which assimilates in place of articulation to the initial consonant of the verb base (B).

⁸P1 stands for "today past"; see the key at Table 3, and in general for abbreviations the section on Abbreviations at the end of this chapter.

⁹In most north-western languages this is prefixed to the verb, so strictly *a-*, while in a few (e.g. A80) it is described as self-standing, so *a*. For the sake of simplification, we describe both here as *a-*.

pre-stem marker, or marks one form of past (near, far) and not another, or combines with another marker to indicate past. It occurs in all 16 of Guthrie's zones, although less frequently in the north-west. There is clear phonetic and phonological evidence for several distinct *a*- morphemes with past reference across NB. Some 22% of the languages examined by Nurse & Philippson (2006) have contrastive *a*-, that is, it is the tone or length of *a*- that distinguishes two tenses, but only a very small number of languages distinguished two pasts on the basis of a suprasegmental contrast alone. Table 1 in Nurse & Philippson (2006: 162) sets out the data for the 53 languages for which they had reliable tonal data. Like Bastin (1994), Nurse & Philippson (2006) conclude that the evidence is good for a contrast in the *a*- involved but not so good in terms of a correlation with meanings. They further conclude that a **a*-stem-*a* form originally had near past and/or retrospective (RET) reference, tonal and length distinctions being later innovations. Nurse & Philippson (2006: 164) finally conclude: "We think [pre-stem] **a* can certainly be reconstructed for Proto-Bantu with past reference [... but] would be reluctant to say more than one past **a*, with different tonal profiles and meanings, can be reconstructed at the level of the proto-language [...] it seems likely that as tense reference, especially past reference, multiplied in Proto- or early Bantu, one of its vehicles was the multiplication of **a*."

We also consider in more detail two factors barely or not at all examined by Meeussen (1967) or Nurse (2008), namely the distribution of *a*- 'past' in the north-western languages, especially zone A, and in the Bantoid languages.

Sifting through Bantoid and even Wider NC (see Williamson & Blench 2000: 18) leads to limited enlightenment. Pre-stem *a* is fairly widespread and scattered in some members of Kordofanian, Mande, Atlantic, Kru, Senufo, Gur, Ubangi, Zande, Kwa, West Benue-Congo (BC) (Yoruba, Nupe), among others, with a considerable range of meanings: past, retrospective, non-past, future (Nurse et al. 2016), and focus. However, a mere listing of the languages and meanings is largely meaningless without being able to systemically link the semantics of the various *a*- and to systematically link *a*- to particular branches and the branches to each other. Bantoid languages are NB's nearest relatives, and some of the 20 Bantoid languages in Watters (2018c) show traces of *a*- 'past' (see Table 10 and its discussion). It is risky to place too much weight on such a short morpheme. There may have been more than one *a*-. Nevertheless, we find it encouraging to find these Bantoid *a*- 'past', and feel they support the hypothesis that a PB *a*- 'past' was inherited from a pre-PB stage.

Table 3 shows that the distribution of *a*- 'past' in zones A, B, C, and bits of D is not as widespread as might be expected. Since *a*- 'past' is so widespread across NB, it should be reconstructable for PB, and was so reconstructed by Meeussen

(1967) and Nurse (2008). Following what is said above, we might expect to find it at least in simple forms, that is, with one past meaning and simple in shape, in the north-western NB languages.

Table 3 can be summarised as follows:

- *a-* ‘past’ is not omnipresent across north-western NB. It is absent from A10-20-30-40, B30-40-60-80, C20 and C60. It occurs in all A50 and also in

Table 3: *a-* ‘past’ in north-western languages with multiple pasts

Language	Code	P3, P4	P2	P1
Kpa	A53			á-
Yambasa	A62		á-	
Ewondo	A72a		á-	
Makaa	A83		a-	
Myene	B11	a-	a-	a
Kota	B25	á-	a-	
Duma	B51	á-	a-	
Nzebi (as per Guthrie 1968)	B52g	á-	a-	
Nzebi (as per Marchal-Nasse 1989)	B52m	á-	á-	
Teke Yaa	B73c	a-		
Babole	C101		a-	
Lingala	C30B		á-	
Ngombe	C41		a(a)-	
Soko-Kele	C50		á-	
Ombo	C76	á-ka-	á-	
Bushong	C83			a-
Mituku	D13	a-P3-6	a-	a-
Enya	D14		a-	á-
Lega	D25	a-	a-	a
Holoholo	D28		á-	á
Nyali	D33		á-	á

Key to the temporal semantics of the categories in this table: languages with four pasts distinguish P1 = today past, P2 = yesterday, P3 = a few days, weeks, or months ago, P4 = remote past. If they only use P1, P2, and P3, then P1 = today past, P2 = recent past, P3 = distant. If they only use P1 and P2, then P1 = recent past and P2 = more distant past. Futures work identically, so if only F1 and F2, then F1 = near future, F2 = distant future, etc. Note: P3-6 in the row for D13 refers to its six past tenses, P1, P2 and P3-6, all using *-a* (see §2.5).

B10-20-50-C10, etc., and in some A60-70-80-90 languages. If we had had access to more languages and better data, this picture might be clearer.

- Parts of north-western NB have a single *a-* ‘past’.
- A few have a binary *a-* ‘past’ contrast.
- Other than B11, D13, and D25 none has a three-way *a-* ‘past’ contrast.
- Not shown for reasons of space is the distribution of this *a-* in the rest of NB, where binary and three-way *a-* contrasts are frequent.
- We can tentatively propose that there is a general development from a single, earlier **a-* to multiple, later *a-* pasts, but it is not a straight line.

We think the best explanation for the absence of reflexes of **a-* in A10-20-30-40 (and the B and C languages above) is to posit that **a-* was part of early PB but subsequently lost in a later PB lect or lects ancestral to A10-20-30-40. This scattered distribution mirrors what we find in Bantoid: *a-* ‘past’ occurs in some Bantoid languages (Ndemli, Ngie, Aghem, Babanki, Mambiloid (Vute)), but not in many others (cf. §3.2.1 and §3.2.2 below). All this suggests that *a-* was once more widespread in Bantoid than it is today, but is now retained in a rather haphazard pattern. We know of no concrete cases where *a-* is lost from synthetic structures – A10-20-30 and most A40 languages are synthetic today – but early PB is more likely to have been analytic (see §2.4) in which situation *a-* could have been more easily replaced, and thus lost, in the ancestral forms of A10-20-30-40 and adjacent Bantoid lects. The ancestors of A10-20-30 and most of A40 subsequently became synthetic.¹⁰

2.2.2 Verb-final *-ile* vs. *-i*

Meeussen (1967) and Nurse (2008) reconstruct for PB verb-final **-ile*, regarded as bimorphemic *-il-e* (cf. Table 1). This is a complicated issue. Closer examination of north-western NB, of Bantoid languages, and of Wider NC suggests a possible different situation. Most zone A, B, C, and some D languages have just *-i*; a few have *-i* and allomorphic variants such as *-ili*, where *-ili* occurs after CV stems, with *-i* after CVC or longer stems.¹¹ Where Bantoid languages have this

¹⁰For zone B and C languages, it also has certain implications, which we prefer to ignore here.

¹¹Lundu A11, Lue A12, Mbo A15, Mbuu A15A, Akoose A15C, Kpe A22, Duala A24, Myene B11, Duma B51, ?Ntomba-Bolia C35a-b, Idakho JE411 (Grégoire 1979; Hedinger et al. 1981: 54, 62 (verbs 8); Bastin 1983; Hedinger 1985: 11; 2008: 111 (verbs 12 and 13); Ebarb & Marlo 2015: 248). Also, consider the discussion in §3.5.1 of this chapter on *-i* and *-ile* in Wider Bantoid. A number of unanswered questions remain about their distribution and origin.

suffix at all, they mostly have it as *-i*; the evidence for *-ile* is sparse and less clear (see §3.5.1.2 below). As far as we know, Wider NC has *-i* and no *-ile* (Nurse et al. 2016). This suggests the original shape was *-i* or *-i//ili*, although we cannot convincingly account for the emergence of *-ile*. It may relate to the notion of suffixal phrasemes in verbal derivation, set out in Bostoen & Guérois (2022 [this volume]).¹² These are historically complex suffixes/extensions which become semantically non-compositional and include the older and shorter simplex suffix with the same meaning (e.g. **-ibv* PASS including **-v* PASS, **-angan* RECP including **-an* RECP, **-idi* CAUS including **-i* CAUS). Could **-ile* also be such a phraseme but in TA marking? Most of these suffixal phrasemes in verbal derivation arise after NWB split off, just as we argue here for **-ile*.

Consequently, we propose a historical scenario with three stages. Stage 1 involves NC and Bantoid¹³ with a basic aspectual contrast between perfective versus imperfective, perfective being widely (not exclusively) represented by *-i*. Stage 2, seen in all languages in zones A, B, C, plus D10, D20, and D30, has *-í*, principally representing ‘past’. A dramatic change then led to Stage 3: in the rest of NB *-ile* came to predominate, with some areal retention of *-i* and some cases of the vowel copy (VC) suffix, where the FV reflects the stem vowel ([CaC]-a, [CeC]-e, [CiC]-i, etc.).¹⁴ In the rest of NB, *-ile* represents primarily retrospective with *a*-taking over the role of ‘past’.

In Figure 1 *-í* is italicised, the VC suffix is underlined, and *-ile* and its many variants are bolded. VC thus occurs when the final vowel is a copy of the first root vowel instead of *-i* and *-ile*. The reconstructions in Meeussen (1967) and Nurse (2008) reflect this large and later (Stage 3) area. We display stages 2 and 3 in Figure 1. So, the shape changes from *-í* in most NC to *-ile* in most NB, the north-west being a transition area, while the meaning shifts from NC ‘perfective’, to north-western NB ‘past’, to ‘retrospective’ in most of NB.

We conclude that *-í* ‘past’ should be reconstructed to PB, rather than *-ile*. That leaves certain unexplained phenomena: why do A10 and A20 languages, Myene B11, Duma B51, Idakho JE411 (and maybe a few others?) have two allomorphs? Why do the *-í* form and meaning change outside the north-west?

An alternative version of Stage 3 would be that while *-ile* widely replaced *-i*, in some languages *-i* and *-ile* coexisted with different meanings. They still do today in a small group of languages based on K10, K20, K40, L10, L30, and

¹²We acknowledge Koen Bostoen’s major contribution to this whole section.

¹³The evidence in Bantoid and north-western languages is obscured by widespread loss of final vowels.

¹⁴The VC suffix is a separate development, with which we do not deal here. See Grégoire (1979).

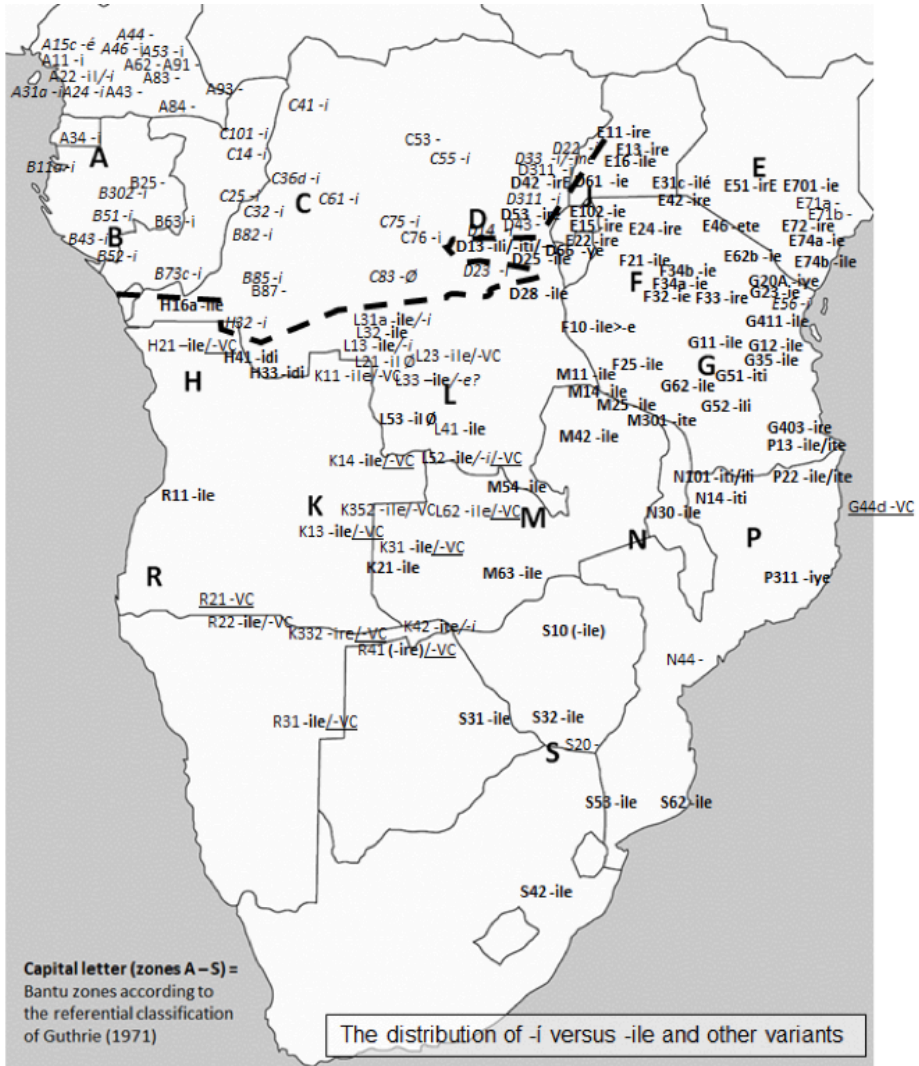


Figure 1: Distribution of -i, -ile, and other minor variants

L50, with some isolates in D10, JE411, E60, and Kongo H16. Where they do co-exist, *-i* represents predominantly ‘near past, anterior, resultant state’, while *-ile* represents a more ‘distant past’.

2.2.3 How did the new pre-stem *a-* fit with *-i*?

We suggested above that the best explanation for the absence of reflexes of **a-* in A10-20-30-40 is to posit that **a-* was part of early PB but subsequently lost in a later PB lect or lects ancestral to A10-20-30-40.¹⁵ Examination of the languages retaining *a-* shows numerous combinations of pre-stem \emptyset and pre-stem *a-* with suffixal *-í*.¹⁶ Common to nearly all is a pair of features: forms with *a-* encode predominantly past reference, and *a-* with past reference predominantly represents a time further removed than pasts without *a-*, so past vs. present, or further past vs. near past, etc.

Hence, *a-* acts as a ‘shifter’ added to another structure.¹⁷ A corollary of this is that *-í* came to be associated with nearer past. This may explain why it finishes up as primarily retrospective (see §2.4). Most retrospectives are associated with events in the more recent past.

2.3 Did PB have future tense(s)?

In contrast to past reference, where one marker predominates, future marking is diverse across NB (Nurse 2008: 85–87). Future morphology is frequently renewed. Nurse’s database has many future markers, all geographically limited and many obviously grammaticalised forms. The only two with any claim to reconstructability are *ka-* and *la(a)-*. Attested in 29% of the languages in the sample of Nurse (2008), *ka-* is the most widespread future marker; *ka-* in general is widespread (71%) in NB in several affirmative functions: itive, narrative, (far) future, (far) past consecutive, if/when/conditional/participial/persistent, subjunctive. It occurs as ‘future’ in all zones, including some zone A languages, though sparse in zones C, G, and S. Nurse & Philippson (2006: 171) hypothesised that *ka-* in its itive function might be the source of many of these other functions,

¹⁵Loss of *a-* in some B and C languages (B30-40-60-80, C20-60: cf Table 3, above) might or might not be related. The pre-stem marker *a(-)* is also lost in most adjacent Bamileke lects.

¹⁶Because of widespread loss of final vowels in zone A, examples from zone B or C are sometimes more transparent.

¹⁷Recall that Mituku D13 has six past tenses. Five of them (P1-5) also have two variants, one with and one without *a-*. Robert Botne (p.c.) has suggested to us that the variants with *a-* may refer to a time further in the past than those without. If so, this would be a remarkable example of the role of *a-*.

including the future. As a future marker, it occurs mainly in SWB and EB languages stretching from the East African Great Lakes region down western Tanzania to Zambia and Namibia: JE30, F10, Kagulu G12, Mbundu H21, Mbala H41, Mpoto N14, K10-20-30, L30-40-50-60, M10, M30-40-50-60, Umbundu-Ndonga R11-22. This distribution does not suffice to reconstruct *ka-* ‘future’ to PB. However, it also occurs as a future marker in some NWB languages (i.e. Benga A34, Basaa A43a, Maande A46, Yangben A62A, and maybe Akoose A15C) and one WWB language (i.e. Nzebi B52). However, some of these futures might derive from an original itive meaning (‘go’) through parallel innovation and others might result from more recent grammaticalisation of auxiliaries or adverbials. All this suggests that reflexes of *ka-* are spread widely enough across NB to warrant its reconstruction for PB, certainly as ‘itive’, possibly in the derived set of meanings, including ‘future’ (cf. Meeussen 1967: 109). A morpheme of the shape *ka* occurs in Wider NC and Bantoid in several functions, i.e. past, (immediate) future, conditional, subjunctive, consecutive, etc. In Bantoid, we only found it as a future in Tikar. This disparate set suggests that while one or maybe more *ka* occurred in NC, no firm statement can be made about the original meaning of *ka* in Wider NC.

Pre-stem *la(a)-*¹⁸ occurs in 17% of the database languages, in a restricted swathe of EB languages from the East African Great Lakes region down western Tanzania to Zambia: Mituku D13, JD60, JE10-20-30-40, maybe E50-60, F20, maybe Rimi F32, Gogo G11, G60, M10-40-50-60, Manda N11. It is maybe also attested in one CWB language (i.e. Kele C55) and one WWB language (i.e. Yombe H16c). The prefix *la(a)-* also occurs in other functions, but is, with 22%, much less frequent than *ka-* and does not occur in NWB. Short vowel morphemes of similar shape, and both future and past reference, occur in some Grassfields languages, but an exact relationship remains to be established. On this basis, we doubt the reconstruction of **la(a)-* as a future tense marker for PB (contra Nurse 2008: 297) and think it is a later innovation. Because *ka-* as a future occurs more widely, including NWB, though sparsely, its reconstruction for PB is more plausible.

2.4 Was the PB verb synthetic or analytic?

Part of the discussion at the end of §2.2.1 involved making a distinction between an analytic and a synthetic verb structure. As discussed more extensively in Nurse (2007), of which this section is a summary, most NC languages have or had an analytic verb structure in which the nucleus [root-EXT-FV] was preceded

¹⁸Larry M. Hyman (p.c.) suggests */laa/* might be bimorphemic, so */la+V/*.

by a variable number of independent items related to the verb. FV is/was the site for expression of aspect. We assume NC had that structure and that such analytic structures today are retentions from early NC, unless it can be proved in individual cases that the opposite happened, i.e. that synthetic structures broke down into analytic ones. In five millennia, much is possible. Our general impression is that early Bantoid inherited analytic structures from NC but that there has been a tendency towards synthetic ones. Outside zone A, almost no NB languages have an analytic verb structure. Within zone A there are different degrees of analyticity as demonstrated in Figure 2.

Basaa	A43a	[H-SBJP] TAM clitics/particles [T-root-EXT-ak-H-na] NEG OP
Makaa	A83	[SP + T H] P1 NEG + CM HAB PROG ADV AUX [OM-root-EXT-FV-H]

Figure 2: Different degrees of analyticity in NB zone A

Although descriptions vary, within zone A, essentially A40 (Maande A46?), A80, and A90 are analytic, while the rest is synthetic.¹⁹ The zone A situation is similar to that in Bantoid (and other NC): most of the few Bantoid languages examined are analytic, but some have tendencies towards becoming synthetic, i.e. Ejagham, Nyang, Jukun, and most Cross River languages. Individual distant NC languages have also become synthetic (Dogon, Kordofanian, Obolo, Zande, etc.). Some analytic languages in Grassfields and zone A show movement to synthetic structures (cf. Nurse et al. 2016: 22). While we need more local detail to better see the overall picture, our general impression is that no coherent synthetic area exists across zone A and Bantoid, so syntheticity seems to have developed among early Bantu lects, around or following the Bantu exodus (cf. Hyman 2004). Since NB languages outside zone A are virtually all synthetic, they must descend from an ancestral lect that was synthetic.

2.5 Our current view of PB tense

Pre-stem morphemes reconstructable to PB are \emptyset ‘vast present’ (interpretable as an absence of marking), *a* ‘past, shifter’, *ka* ‘itive, future’, *ki* ‘persistive, situative’, *a* ‘disjunctive’ (Nurse 2008: 236–257). These are not marked as being prefixal because, as just noted, there was a move from analytic to synthetic status within PB.

¹⁹Also some B70 and B80 languages are partly analytic.

However, they are preverbal and particles, since most are not clearly derivable from auxiliaries.

Suffixes at the end of the verb form reconstructable to PB are *-a* ‘imperfective’, *-í* ‘past/perfective’, *-ag-(a)* ‘habitual/iterative’, *-é* ‘subjunctive’.

Below, because our focus is on tense, we ignore the role of *kí* ‘persistive, situative’, *a* ‘disjunctive’, the itive function of *ka*, *-ag-(a)* ‘habitual/iterative’, and *-é* ‘subjunctive’.

Starting with Table 2, an early or pre-PB, pre-tense stage, and repeating it as a matrix gives Table 4.

Table 4: An earlier pre-PB aspect-prominent stage

Imperfective	Perfective	Progressive
<i>*tò-ø-gòd-a</i> ‘we buy’	<i>*tò-ø-gòd-í</i> ‘we bought’	<i>*tò-ø-gòd-ag-à</i> ‘we are buying, buy HAB’

Adding past and maybe future should give Table 5, possibly a later PB stage.

Table 5: An innovated TA stage

<i>-a</i> Past	<i>*tw-a-gòd-a</i> ‘we bought’	<i>*tw-a-gòd-i</i> ‘we had bought’?	<i>*tw-a-gòd-ag-a</i> ‘we were buying, we used to buy’
<i>-ø</i>	<i>*tò-ø-gòd-a</i> ‘we buy’	<i>*tò-ø-gòd-i</i> ‘we have bought’?	<i>*tò-ø-gòd-ag-a</i> ‘we are buying, we buy HAB’
<i>-ka</i> ? Future	<i>*tò-ka-gòd-a</i> ‘we will buy’	<i>little evidence</i>	<i>*tò-ka-gòd-ag-a</i> ‘we’ll be buying’

The problem here is what stage would Table 5 represent? The significant change between Table 4 and Table 5 is the appearance of pre-stem *a* as ‘past’, beside earlier *-í*, slowly replacing it. Across NC *-í* was primarily a perfective, which most often refers to past time. What Table 5 displays must be unstable because it contains three forms referring to ‘past’ or ‘perfective’: *twagula*, *twagulí*, *tugulí*, so how to label the three columns? Table 5 is a still photo of a slowly changing situation. The evidence shows *a* and *-i* co-existing in north-western languages

and gradually resolving the situation in different ways. As far as we know, *a* and *-í* only co-exist in one zone A language, i.e. Kpa A53. However, more B and C languages combine *a* and *-í* in past reference. Only in zone D, which would be about node 5 in the Bantu phylogeny of Grollemund et al. (2015), does *-í* become *-ile* and ‘past’ become ‘retrospective’. This is much later than PB.

Finally, we think it worth mentioning that a construction consisting of (BE-at) + (locative prefix) + verbal noun occurs widely across NB, Bantoid, and NC; e.g. *tu-li-mu-kugula* ‘we are buying’ (lit. we-are-in-buying) with a progressive meaning or a set of meanings derivable from progressive (Bastin 1989a,b; De Kind et al. 2015). This kind of grammaticalisation is common universally and across Africa (Heine et al. 1993; Heine & Kuteva 2002). We assume it happened often before NB, maybe during PB, and certainly since PB. This is why we do not include it in our reconstruction. This construction could well have co-existed with what is set out in Table 4 and 5.

3 The emergence of tense in Bantoid

In his comprehensive analysis of tense and aspect in NB, Nurse (2008) raises the issue of the origin of tense as a morphological category within NB. From the information available, particularly concerning tense in Bantoid Grassfields Bantu, Nurse (2008) proposes that PB tense likely had a pre-Bantu origin involving ancestor languages of Grassfields and Cross River (CR). At that time, the known distribution of TA systems within Bantoid and CR was limited.

In response to Nurse (2008), Watters (2012) presented the distribution of TA versus aspect-only languages within Bantoid. Aspect-prominent languages appear to the west of the mountain range of the Cameroon Volcanic Line (CVL), while TA languages exist along the CVL and to its east towards the Sanaga River Basin. This present-day distribution points towards a likely origin of tense along the CVL and to its east, in the direction of NB, where tense may have emerged as a morphological category in PB some 5000 years ago. More specifically, the “Grassfields Bantu” group lies along the CVL and to its east, and it is the closest neighbour to the location from which NB is commonly thought to have originated. One implication of eastern Bantoid being involved in the origin of tense in Bantoid and NB is that not all Bantoid languages participated in the innovation of tense, namely, those groups west of the CVL. See the map in Figure 3 for geographical details.

To further clarify the possible presence of the category “tense” elsewhere in East BC, Watters (2018c) expands the coverage of verb systems to include CR and

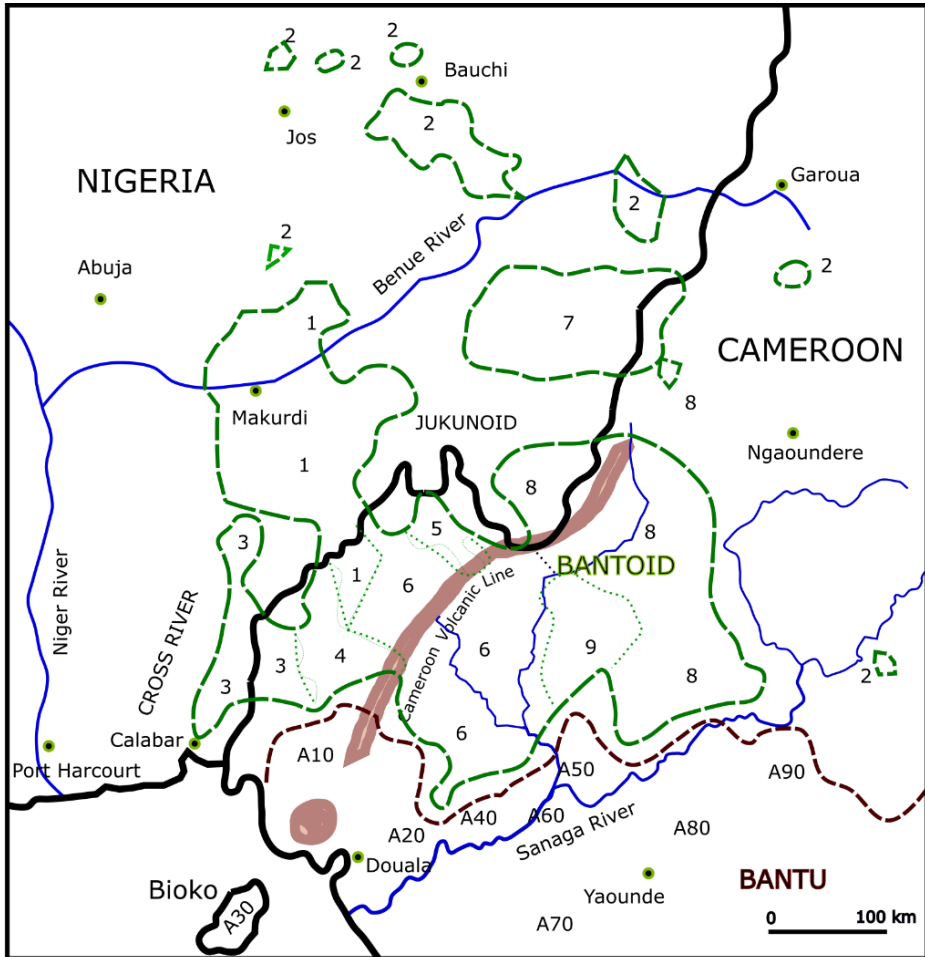
Jukunoid languages within East BC. The evidence from this wider view supports the 2012 conclusion. It also provides additional insight that Proto-CR and Proto-Jukunoid were most likely aspect-prominent and did not participate in the early genesis of tense.

Finally, seeking to test the distribution of tense in the remaining branches of East BC and to review one language claimed to mark tense, Watters (2019) demonstrates that the Plateau and Kainji branches of East BC are essentially aspect-prominent. This conclusion includes the Plateau language Birom that has at times been said to mark tense in its verbal system. Birom is better viewed as aspect-prominent, but if one wants to use the term “tense” for Birom, it only concerns the retrospective and potential aspects using “yesterday/tomorrow”, “today”, and “just now” as degrees of time. Such a system does not resemble the one that led to the TA systems in NB and eastern Bantoid. We can say with fair confidence that we present here a verbal system that developed among lects in southern Cameroon possibly some 5000 years ago or more and nowhere else in BC.

3.1 Position of NB within Bantoid

The NB languages belong to the Bantoid subgroup of East BC. NB languages distinguish themselves linguistically from the other Bantoid groups through their use of passive verb morphology (Watters & Leroy 1989: 445). The passive is absent in the other Bantoid groups with the Sanaga River Basin serving as a boundary. Another distinguishing feature may possibly be NB’s use of the applicative (Hyman 2018: 190; Watters 2018a: 20). Hyman reports that for Bantoid beyond PB he only found Metta and Vute with possible applicative extensions. However, he concluded that the Metta suffix *-ri* is not clearly cognate with the PB applicative **-Id* and that the Vute suffix *-nâ* is a Vute innovation (see also Blench (2022 [this volume])). In contrast to these distinctions between PB and other Bantoid groups, in this section we demonstrate that NB and the eastern region of Bantoid share the verbal category of tense. Ancestors of a subset of Bantoid languages engaged with the PB ancestor to innovate tense as a morphological category.

According to Grollemund et al. (2015: Figure 1), in expanding our focus from NB in §2 to include other Bantoid groups in §3, we move from node 1, i.e. PB at 4000-5000 BP, to node 0 at possibly 5000 BP or older. At node 0 Grassfields, sometimes referred to as “Grassfields Bantu”, and Tiv (Tivoid) represent the other Bantoid groups outside NB. Grassfields and Tiv serve as the outgroups to root the phylogenetic tree.



Key to the codes and numbers on this map: (Narrow) Bantu subgroups identified from A10 to A90; Bantoid subgroups: 1 Tivoid, 2 Jarawan, 3 Ekoid, 4 Nyang, 5 Beboid & Yemne-Kimbi, 6 Grassfields, 7 Dakoid (not included in study – no data), 8 Mambiloid, 9 Tikar, (10 Bendi – if it were included in Bantoid, it lies in the space between 1 Tivoid and 3 Ekoid)

Figure 3: Borderlands of (Narrow) Bantu, Bantoid, Cross River and Jukunoid

3.2 Genealogy of Bantoid outside of NB

In engaging with the Bantoid groups outside of NB, we want to clarify certain relationships within Bantoid and the terminology related to those relationships. First, Bantoid includes the Tivoid, Jarawan, Ekoid, Nyang, Beboid, Grassfields, Dakoid and Mambiloid groups, as shown on Figure 3 above. It also includes the isolate Tikar and possibly Fam. It likely includes the Bendi languages that previously were part of CR. More recently, Good et al. (2011) revised the Beboid group and separated out a new group, i.e. the Yemne-Kimbi languages. Thus, we could say there are ten Bantoid groups and two isolates outside of NB that bear some historical relationship with NB. Dakoid will not figure any further in this study due to a lack of relevant data.

3.2.1 Genealogical relationships

Considering genealogical relationships based on innovations and retentions, Bantoid may appear as a set of scattered groups without much coherence. However, relationships among these languages have gained the attention of linguists over the past fifty years. We consider three of the more recent attempts. One involves a proposed genetic division between Northern Bantoid and Southern Bantoid. Blench & Williamson (1987) proposed this division, and it provided the template for the Bantoid chapters in Bendor-Samuel & Hartell (1989), with Hedinger (1989) presenting Northern Bantoid, and Watters & Leroy (1989) presenting Southern Bantoid. Hedinger (1989: 424, fn. 4) provides the set of thirteen lexical innovations upon which Blench & Williamson (1987) had based their classification of Northern Bantoid as a distinct genealogical subgroup. Northern Bantoid includes Dakoid, Mambiloid, and the isolate Fam. Southern Bantoid includes NB as the major group as well as the seven remaining groups and the isolate Tikar.²⁰

Shifting from lexical innovations to using lexicon-based quantitative methods of genealogical classification, we consider Piron (1995; 1997) and Grollemund et al. (Forthcoming). Piron (1995; 1997) concludes that her lexicostatistic study²¹ does not support a clear division within Bantoid between Northern and Southern Bantoid. Using phylogenetics, Grollemund et al. (Forthcoming) confirm that the Northern-Southern division within Bantoid is not relevant from a genealogical point of view. Figure 4 displays the major Bantoid branches emerging from the

²⁰Compare, however, with Blench (2022: Figure 2 [this volume]), for Blench's current understanding of the sub-classification of BC.

²¹Unlike the newer phylogenetic methods, lexicostatistics builds trees based on lexical similarities and does not distinguish between retentions and innovations (and implicitly assumes a constant rate of lexical change).

new phylogeny of Grollemund et al. (Forthcoming), together with an indication of their tense/aspect systems (for which, see next section, §3.2.2).

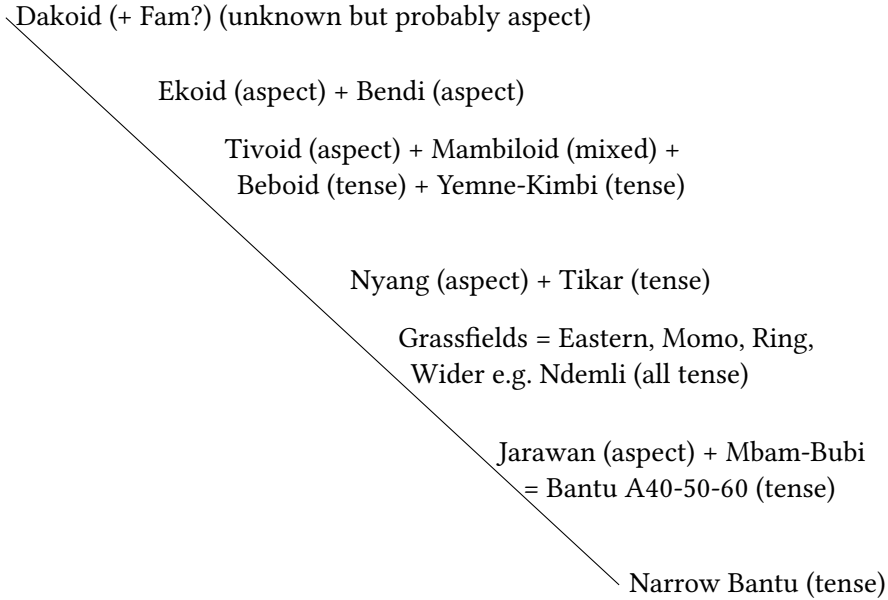


Figure 4: Simplified schema of Bantoid (Grollemund et al. Forthcoming) with an indication of their tense/aspect systems

Note that instead of placing Dakoid and Mambiloid in a separate Northern Bantoid unit, the analysis of Grollemund et al. (Forthcoming) separates them, placing Dakoid as a first group and Mambiloid in the middle of the Bantoid groups as part of a larger group with Tivoid, Beboid and Yemne-Kimbi.

3.2.2 Geographical relationships

Besides genealogical relationships, a more crucial distinction for the study of tense involves the geographical framework for the Bantoid groups. Again, consider Figure 3. Bantoid outside of NB occupies land primarily along the Cameroon-Nigeria border region. A primary feature of the geography are the mountains in Cameroon that originate from the CVL. To the west of the CVL are groups located primarily in Nigeria. To the east are groups located along the CVL and further east into the Sanaga River Basin located primarily in Cameroon. The languages of the western region are aspect-prominent while those of

the eastern region have primarily TA systems. This contrast became clear back in 2011 when preparing the Watters (2018b) manuscript on Ejagham (Ekoid) and its aspect-prominent verb system. All of Bantoid is not like NB when it comes to the matter of tense. The western groups are aspect-prominent. The eastern geographical region is the region that shares tense as a verbal feature with NB. It is from this eastern region that PB emerged. To re-emphasise, “western” and “eastern” Bantoid refer to geographical categories and not to (former) genealogical ones like “Northern” and “Southern”. It is the eastern region that serves as the home of marked tense in their verbal systems similar to NB. The Grassfields group is one eastern group, and it is geographically closest to NB. It displays a TA system like that in NB, yet with some significant differences as well.

One final note, Bantoid languages with tense do not correspond with the phylogenetic units in Grollemund et al. (Forthcoming), as may also be seen from Figure 4. Tivoid is to the west of the CVL and is aspect-prominent. However, it groups with Beboid and Yemne-Kimbi that are along the CVL and have TA systems. Similarly, Nyang and Tikar form a phylogenetic unit, but Nyang is west of the CVL and is aspect-prominent while Tikar is to the east of the CVL and has a TA system. This difference points to tense being developed as an areal feature rather than an inherited feature. The eastern region of Bantoid was the key area for innovating tense.

3.3 Major issues about the origin of tense

We want to focus here on two major issues relevant to the claims about the origin of tense. Nurse (2008) proposed a systematic structure for the PB TA system with a set of exponents for each category. The first issue concerns the systemic structure. Does the proposed PB structure match that of the Bantoid languages that share this possible origin? It appears that general structures do match. This strengthens the claim that tense in NB and other Bantoid languages has a common origin. The concepts “system” and “structure” are illustrated in Table 8 for Bantoid and Table 9 for NB in §3.5.4 below.

The second issue concerns the morphological exponents of tense. Are the exponents of tense that we find in other Bantoid languages cognate with those proposed by Nurse (2008) for PB, and listed in Table 1? The answer to this question is more complicated. The exponents proposed for PB suffixes find some potential matches in Bantoid suffixes in the various Bantoid subgroups along the CVL but fewer in the case of prefixes. There are some possible prefix matches, but many of the Bantoid prefixes differ from PB and even from each other. These Bantoid subgroups present a variety of forms. 5000 years of change no doubt is a

contributing factor. The challenge is explaining the significant variation within the various Bantoid groups including those proposed for PB. At the same time, the critical goal for PB reconstruction is identifying those languages which most closely relate to the PB tense exponents presented in §2 above.

The relevant Bantoid groups in Figure 3 involve more than ninety languages or lects: sixty-seven Grassfields languages (#6), fifteen to twenty Mambiloid lects (Connell 2019) (#8), nine Beboid and five Yemne-Kimbi (#5), and the isolate Tikar (#9). Those Bantoid groups that we have found to date that do not have tense but use aspect-only systems include Tivoid (#1), Jarawan (#2), Ekoid (#3), Nyang (#4), and some Mambiloid lects (#8). Mambiloid is the only group from Northern Bantoid included in this study. Dakoid (#7) is not included because we have no data on its verb systems. The (former) Northern-Southern distinction within Bantoid is not relevant to the discussion about the emergence of tense. Instead, the geographic categories western-eastern are the relevant ones at this point.

Of the Bantoid groups with tense, those in the Grassfields are of the greatest interest since they border on the north-west boundary of what has been referred to as “zone A” (A10 to A90 in Figure 3) of the NB languages, the most north-western NB languages and the closest geographically to the other Bantoid groups with tense. As indicated in Figure 3, the approximate location of the NB groups A10–90 is immediately to the south of the other Bantoid groups.

To represent the details of the more than ninety languages or lects relevant to this topic, we have chosen twenty-four sample languages to represent the five groups. Noni, Nchane and Mungong represent Beboid. Mugbam and Mundabli represent Yemne-Kimbi.²² Sixteen languages represent four subgroups of Grassfields (Eastern, Momo, Ring, and Wider Grassfields). Vute and Ju Ba represent Mambiloid. Tikar represents its own group. Watters (2003) provides an overview and further details about Grassfields. The twenty-four eastern Bantoid languages serving as examples throughout this §3 and the resource(s) used for each of these languages are referenced in Appendix B. Since Bantoid languages in the western region do not mark tense, we are excluding them from the remainder of this study. These involve Tivoid, Jarawan, Ekoid, and Nyang.

Certain morphological categories are important in answering the two questions about structure and exponents. These categories include the distinction between perfective and imperfective aspects, disjoint (+verb focus) and conjoint (+argument focus) forms, and tenses involving past and non-past (present and

²²The languages of Yemne-Kimbi used to be included with Naki as “Western Beboid”. However, Good et al. (2011: 108) argue that there is no substantial evidence to link these languages with Eastern Beboid. They proposed the new name based on two bordering rivers, “Yemne-Kimbi”. Consequently, Eastern Beboid becomes simply “Beboid” with Naki joining this new “Beboid”.

future). Aspects such as retrospective (perfect), habitual, progressive/continuous may also prove helpful, but are not the main concern. In our review of the available literature about these languages, we were not always able to find imperfective forms. In most cases, we were not able to find distinct disjoint and conjoint forms. They may not exist in every language under review. We sought to identify at least perfective forms in all relevant tenses.

3.4 Tenses occurring in a sample of Bantoid

Table 6 presents the number of tenses in the twenty-four sample languages in their TA verbal systems. Appendices C, D and E present them with P0 and F0 included.

Only Vute and Ju Ba represent Mambiloid, both of which mark tense. However, not all of the 15–20 Mambiloid lects have TA systems. While Vute and Ju Ba do have such systems, elsewhere there is variation (Connell 2019). Some lects even seem to vary internally between marking tense and at other times not marking tense. Others only have an aspect-prominent system. These aspect-prominent lects are geographically closer to the western region of Bantoid languages that only have aspect-prominent systems. This indicates a likely areal phenomenon occurring within Mambiloid. It is also probably indicative of how tense diffused among the eastern Bantoid languages as an areal rather than a genetic feature.

3.4.1 Making historical sense of all the past tenses

All twenty-four languages in Table 6 have multiple pasts and all but five (i.e. Nchane, Mungbam, Mfumte, Mengaka, Ngie) have multiple futures. All twenty-four have at least two past tenses, P1 and P2. Four have only two past tenses (i.e. Ngie, Aghem, Obang, Vute). All others in Table 6 have three or four past tenses. These data raise three questions.

The first question concerns the number of past tenses that initially emerged when the Bantoid lects, including the pre-Bantu lects, transitioned from lects with only aspect to lects using tense some 5000 years ago. Some NB zone A languages have one past tense, some two, some three, some four. No language in the eastern Bantoid region has only one past tense. Some have two, but most have three or four. Why is this?

This relates to another issue. Did PB only have one past tense as Nurse (2008: 279, Table 6.4) proposes? Could it be that PB actually marked two degrees of past and NB zone A languages subsequently reduced the number of pasts to one? Hypothetically, it is possible. However, we assume that it is simpler to propose that

Table 6: Tenses in the selected Bantoid groups with TA systems

GROUP / Sub-group / <u>Sub-sub-group</u> / <i>Sub-sub-sub-group</i> / Language	P4	P3	P2	P1	F1	F2	F3	F4
BEBOID								
Noni		✓	✓	✓	✓	✓	✓	
Nchane		✓	✓	✓	✓			
Mungong		✓	✓	✓	✓	✓	✓	
YEMNE-KIMBI								
Mungbam		✓	✓	✓	✓			
Mundabli		✓	✓	✓	✓	✓		
GRASSFIELDS								
Eastern								
<u>North</u>								
Limbum		✓	✓	✓	✓	✓	✓	
Mfumte		✓	✓	✓	✓			
<u>Mbam-Nkam</u>								
<i>Nun</i> : Shupamem		✓	✓	✓	✓	✓	✓	
<i>Ngemba</i> : Bafut		✓	✓	✓	✓	✓	✓	
<i>Bamileke</i> :								
Ngiemboon	✓	✓	✓	✓	✓	✓	✓	✓
Ngomba	✓	✓	✓	✓	✓	✓	✓	✓
Yemba		✓	✓	✓	✓	✓	✓	✓
Mengaka		✓	✓	✓	✓			
Momo								
Mundani		✓	✓	✓	✓	✓	✓	
Ngie			✓	✓	✓			
Ring								
Babanki		✓	✓	✓	✓	✓	✓	
Babungo	✓	✓	✓	✓	✓	✓	✓	
Kom	✓	✓	✓	✓	✓	✓	✓	
Aghem			✓	✓	✓	✓		
Wider								
Obang			✓	✓	✓	✓		
Ndemli		✓	✓	✓	✓	✓		
MAMBILOID								
Vute			✓	✓	✓	✓		
Ju Ba		✓	✓	✓	✓	✓		
TIKAR: Tikar		✓	✓	✓	✓	✓		

the innovation of multiple past tenses begins with a single, general past followed by the addition of one or more pasts. This process is adequate in explaining the presence of languages with single pasts and those with multiple pasts. It is also simpler than positing the development of multiple pasts only to then add another process of losing one or more past tenses until only one is retained. There is no evidence requiring an original two pasts. Zone A indicates a need for only one past tense. In addition, the transition from an aspect-prominent language to a TA language likely begins with the development of one past tense rather than a full array of pasts whether two or more. A transition directly to multiple past tenses is far more complex than an initial transition to one past tense. Furthermore, the natural direction of tense development appears to be from simpler to more complex rather than from more complex to simpler. Is there actual evidence in NB for a language reducing its pasts from two to one, or three to two? Therefore, for reasons of parsimony and current evidence, we posit one past tense for PB.

The second question focuses on the process that led to each of the eastern Bantoid groups developing tense systems. What process was involved? Did each group inherit it from a most recent common ancestor? This is unlikely. It is impossible to identify a common ancestor of all the languages that have TA systems. For example, in the lexicostatistical classification of Piron (1997: 625), Mambiloid (tense) and Tikar (tense) are high on the Bantoid tree and what follows below are both aspect-prominent and TA languages. Tivoid (aspect-prominent) and Beboid (TA) also cluster together based on lexicon. In Figure 4 in §3.2.1 we noted that in the lexicon-based phylogeny of Grollemund et al. (Forthcoming), Mambiloid (tense in some lects), Tivoid (aspect), and Beboid (tense) cluster, while Nyang (aspect) groups with Tikar (tense). In addition, the wide variety of morphological exponents of tense that these languages currently use makes formal morphological inheritance from a common ancestral form doubtful. Therefore, we have no strong basis to conclude they had a common ancestor. They gained tense from another source.

As noted in §3.2.2, the groups that share TA systems also share a geographical region but not a genealogical lineage. Thus, we are left with two choices. Did each group of Bantoid languages innovate tense independently or does a lateral diffusion process account for the spread of tense from a single point of innovation? We think it is very unlikely that all these closely related and geographically close languages would have innovated tense independently. Instead, in some unidentified location among them, the first tense developed and was then inherited or appropriated by related or neighbouring lects. The first step was the innovation of a single past tense. All the lects which invented tense, including the lect that

emerged as PB, must have had this single past tense, despite there being no evidence for a single past in today's Bantoid languages. As PB lects began to separate from the rest of Bantoid, somewhere among the non-NB Bantoid lects a second past tense was innovated, separating "near past" from "more distant past". As Table 6 shows, today all Bantoid languages with tense, apart from NB, have at least two past tenses. Thus, we consider diffusion as the means that led multiple eastern Bantoid groups to gain tense (see §4.4). Later, the spreading of a second past took place among the non-NB Bantoid lects after the PB lect had left the region.²³

The third question concerns the derivation of the multitude of tenses found in the Bantoid languages. Where did they come from? The answer seems to be twofold. The preverbal space allowed for the use of serial verb constructions. The first verb in the sequence gradually took on the role of a tense marker. As these innovations of "verb-as-tense plus verb-root/stem" were shared with neighbouring languages, they used a calque or an analogical formation process to develop their own parallel tense. The variation of tense markers is discussed in §3.5 below.

3.4.2 Making historical sense of the future tenses

The past tenses always involve both a perfective and an imperfective form. Even if in some cases the grammars or briefs have not provided the imperfectives, we assume, by analogy to closely related languages, that imperfectives are available. In the future tenses, however, there is less consistency. Some languages have perfective and imperfective forms. Others have only imperfective forms. In some cases, one future may involve a perfective and the other an imperfective form. These facts point to a less than settled pattern for future forms. In fact, the Ring Grassfields languages Babungo and Aghem only use the imperfective for future time. This is also true for Tikar. For Vute, Thwing & Watters (1987) listed the near future as imperfective and distant future as perfective in form. However, Vute may have formed the morpheme of the perfective from imperfective forms, so Vute may use only imperfective for the future.

In Babungo, Aghem, Tikar, and Vute, the use of the imperfective for future time is essentially a continuation of one of the functions of the imperfective in their earlier aspect-prominent systems. The imperfective in aspect-prominent languages has a default reading as either present or future time. Thus, in these languages today the perfective with its historically default reading as a general

²³See §4.4 below which references Dimmendaal's (2011: 189–194) description of two Nilotic languages that adopted tense distinctions into their inherited aspect-prominent languages.

past temporal reference has transitioned to tense with two innovated past tense forms while their future forms essentially remain unchanged. They maintain their previous imperfective forms to refer to future time as they had done originally.

Of these four languages, if we exclude P0 as we have done in Table 6, Aghem and Vute only have two pasts, as opposed to Tikar and Babungo which have three and four pasts respectively. Aghem and Vute speech communities are not geographically close to one another. Thus, we think Aghem and Vute may represent the simpler process of a Bantoid language transitioning from an aspect-prominent system to a TA one. They expanded beyond the single past to two past tenses (P1, P2) but did not change the imperfective into one or two distinct forms with future reference. Development of future tense was a later expansion that happened independently in different branches.

Working off Anderson's insight (footnote 29 §3.5.3) about the Bamileke languages, these eastern Bantoid speakers first innovated past tense. Then later, perhaps much later, they developed future tenses through the same use of preverbal auxiliaries. The past tense markers are now fully grammaticalised and their history is no longer transparent, but future markers are more recent and tend to be more transparent. See example (2) in §3.4.3 below. So the development of future tenses may have had more than one location of development, either within a given group or sometimes in languages independently. As we have seen, four languages continue to use the imperfective for the future and never developed distinct future tense markers.

3.4.3 Making historical sense of Eastern Grassfields

Eastern Grassfields languages, among all the Grassfields languages, have the largest inventories of pasts and futures (Watters 2003: 246). Considering Table 6 and Appendices C to E where P0 and F0 are included in the tables, several of these languages have up to five pasts and five futures. This is particularly true of the Bamileke subgroup of Eastern Grassfields, except for Mengaka that has only one future. By contrast, three of the Ring Grassfields languages have four pasts as well, but their futures are more limited. Mundani, Ndemli, Noni, and Ju Ba have fewer pasts but still have robust systems. The more limited systems are found in Ngie (Momo), Aghem (Ring), Obang (Wider Grassfields), and Vute (Mambiloid). In all cases, they are definitely TA languages, in contrast to other Bantoid languages to the west.

In considering the Eastern Grassfields, note that they subdivide into a North branch and a Mbam-Nkam branch as referenced in Table 6 and Appendices C

to E. The Mbam-Nkam branch further subdivides into the Nun, Ngemba and Bamileke groups. Of these, the Bamileke is the one that borders on NB. We assume that the ancestor lects of the Eastern Grassfields languages, in particular the Bamileke languages, had a central role in the development of tense in Bantoid. See Figure 5 which displays the eleven Bamileke languages bordering NB as well as the location of the Nun and Ngemba groups.

Anderson (footnote 29 §3.5.3), from his study of Ngiemboon (Anderson 1983) and research on related Bamileke languages, concludes that past tense markers developed before future tense markers. He notes that the future markers behave

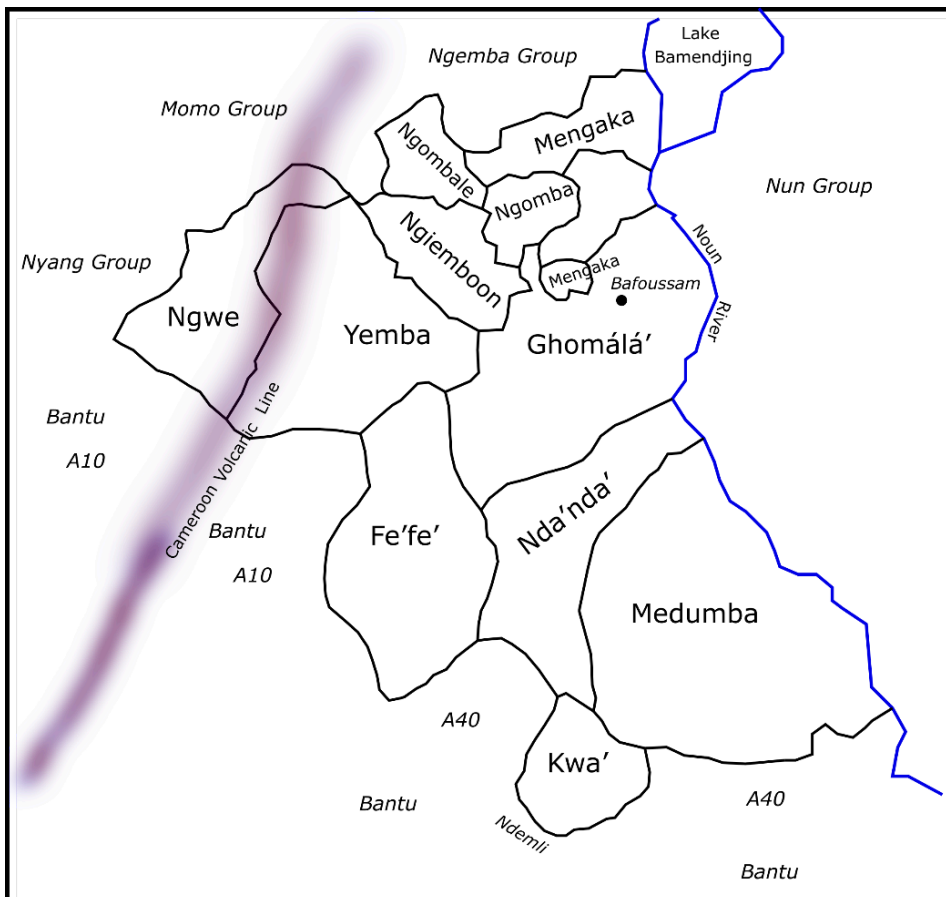


Figure 5: The eleven Bamileke languages, a subgroup of Mbam-Nkam of Eastern Grassfields, bordering NB

as auxiliary verbs with some verbal features while the past tense markers behave as straightforward frozen, verbal morphemes, occurring in a different position than future markers. Thus, we can plausibly conclude that the early lects of Eastern Grassfields languages developed a single past and then eventually developed a second past tense shared among the other eastern Bantoid. However, the Bamileke languages innovated additional past tenses, up to five in some, if P0 is included in the count. From the data in Appendices C to E, it appears that the early forms of Bamileke coalesced around at least three and maybe four past tenses. We assume that these additional past tenses developed after PB had emerged and began expanding.

Thus, we can plausibly conclude that Bamileke developed an initial past tense and later, after separation from the PB lect, developed additional past tenses and tense markers. Later they moved beyond using the imperfective for the future and began developing future tenses using serial verb constructions for which the meaning of the initial tense-marking verb is still transparent today. Hyman (1980: 230) gives the examples in (2) for future tenses in Yemba/Dschang (Eastern Grassfields > Bamileke), using the infinitival prefix *lè-* ‘to’ with the stem.

(2) Future tense derivations in Yemba/Dschang (Hyman 1980: 230)

- a. F1 *píŋ* < *lè-pìŋ* ‘to return’
- b. F2 *lù* / *fù?* < *lè-lù* ‘to get up’ ~ *lè-fù?* ‘to come’
- c. F3 *lá?* < *lè-^hlá?* ‘to spend the night’
- d. F4 *fú* < *lè-^hfú?*

Harro & Haynes (1991: 41–43) compared the Hyman data from the southern/central dialect with their data from the northern dialect. The past tenses were approximately the same, while the future tenses in the northern dialect used *pìŋŋ* (F1), *fù?* (F2), *luū* (F3), and *fú* (F4). Also, in their phonological analysis of these tenses, they posited a floating H tone as the basic marker of past and a floating L tone as basic to the future.²⁴ Lonfo & Anderson (2014: 108–109) report a similar process for future markers in Ngiemboon,²⁵ a closely related Bamileke language.

We therefore attribute the expansion of tenses in various Grassfields and Beboïd languages over the millennia to the grammaticalisation of serial verbs into tense markers.

²⁴In the case of the Yemba/Dschang data, we are treating what Hyman as well as Harro and Haines refer to as P1 and F1 as approximate present tenses P0 and F0. So we have renumbered the tenses changing P2 to P1 and F2 to F1 and so forth.

²⁵Note that here we have omitted the Ngiemboon F0 seen in Appendix E.

3.4.4 Conclusion on merging tense with aspect

In terms of systems, the crucial point concerns the combining of tense and aspect. In the eastern Bantoid languages, the perfective and imperfective aspects form pairs in each of the tenses. Table 7 represents the synthesis of tense and aspect that characterises the eastern Bantoid languages. NB languages share this system as well, suggesting a possible shared history.

Table 7: Systemic structure involving tense and aspect

Tense	Aspect	
	Perfective	Imperfective
Past	Past perfective	Past imperfective
Present	Present perfective	Present imperfective
Future	Future perfective	Future imperfective

The eastern Bantoid languages and many NB languages share the structure in Table 7. The few exceptions are the languages noted above (Babungo, Aghem, Tikar, and Vute) that do not make the perfective-imperfective contrast in their future time reference. They only use imperfectives.

3.5 Exponents of tense

We now examine whether these languages share not only TA categories but also their morphological exponents, one of the issues raised in the introduction to §3.

3.5.1 Exponents of past perfective

Where the data is available, we have expanded Table 7 as Appendix C to include the contrast between disjoint (+verb focus) and conjoint (+argument focus) forms. Both types of perfective may exhibit relevant comparative evidence.

Appendix C shows that innovation in these languages has been entirely pre-verbal, apart from Tikar, demonstrating the recycling of auxiliary verbs that become pre-clitics or prefixes only to be replaced by another auxiliary.

There is a difference between Table 6 and Appendices C to E. Appendices C and D include a column labelled P0, absent from Table 6. Appendix E includes a column labelled F0. The labels P0 and F0 have been a feature of nearly all work on Bantoid languages since the 1980s. However, it is not clear to us that P0 really

is a past tense. It can refer to recent past events, but it has several other functions. It is often, for example, the narrative form in the verbal system. It is typical of aspect-prominent systems to use the least marked, or non-tense-marked, verbal form, i.e. the perfective, to carry the storyline of a narrative, see for instance Watters (1981: 374) for Ejagham (Ekoid) or Paterson (2015) for *Ūt-Ma'in* (Kainji), both East BC languages. However, we leave P0 and F0 in Appendices C to E as part of the relevant data, even though we omitted P0 from Table 6²⁶ as part of the display of past tenses, and do not discuss it further here. For a different treatment, see Sonkoue (2020a) and Sonkoue (2020b).

Appendix C does show some pre-stem *a* possibly cognate with PB **a* ‘past’. Mundani has *a* ‘P2’. Ngie (a Momo language) uses a preverbal *a* [ə] in all its past forms, both (+verb focus) and (–argument focus). Babanki has generalised a preverbal *ə* for all (+verb focus) pasts (and also, see Appendix E, for all (+verb focus) futures) which may derive from an earlier preverbal *á*.

In considering the exponents in Appendix C, we find some morphemes relevant to PB forms in §2 as well as some morphemes that do not have a clear link to such PB forms – see §2.2, §2.4, and §2.5.

3.5.1.1 Preverbal *á* for ‘past’ cognate with PB **á*

Appendix C displays some pre-stem *a* possibly cognate with PB **a*- ‘past’. Evidence is found in all three major divisions of Grassfields. Bamileke languages Yemba/Dschang have *a* in P1. In Momo, Mundani has *a* ‘P2’ and Ngie uses a preverbal *a* [ə] in all its past forms, both (+verb focus) and (–argument focus). In Ring languages, preverbal *á* occurs in the Aghem P2 and P1 (+argument focus) forms, merging with the verbal prefix *mɔ*. Babanki has generalised a preverbal *ə* for all (+verb focus) pasts (but also all (+verb focus) futures), which may derive from an earlier preverbal *á*. In Wider Grassfields, Ndemli uses prefixes *á* and *à* for P2 and P1, respectively. In Yemne-Kimbi, Mundabli uses *a* for P2. See Figure 3 and 5. It is likely that the use of **a* for ‘past’ was more widely present within Bantoid before PB emerged.

3.5.1.2 Postverbal *-i/-ile* possibly cognate with PB **-i/-ile*

As for the NB distinction between *-i* and *-ile* (cf. §2.2.2 and §2.2.3 above), Bantoid data exhibit the following. In the North subgroup of Eastern Grassfields, *-i*

²⁶Just as we omitted F0. Cf. footnote 25.

occurs in Limbum with every verb, perfective and imperfective. Limbum apparently makes no distinction between (+verb focus) and (+argument focus). However, in the Momo group, *-i* occurs in Ngie in its (+argument focus) forms. This contrasts with the (+verb focus) forms which have no suffix. In Ring, Babanki uses *lí* as a post-clitic in its P3 (+argument focus) and P0 (+argument focus) forms, contrasting with no suffix in the (+verb focus) forms. In the isolate Tikar, it occurs in P1, and possibly P2 (*-e*). Like Limbum, Tikar does not distinguish (+verb focus) and (+argument focus). In Beboid, Noni uses a post-clitic *lɔ* in all its past perfective (+argument focus) forms, which may be related. The (+verb focus) forms have no suffix. We think the *-i* forms are probably cognate with PB **-í* and that the *-li* in Babanki may be related to NB **-ile*. Of interest is evidence from farther away in western Bantoid involving Ejagham and Mbe. Ejagham has a suffix *-i* used in the perfective with (+argument focus) that carries three tones, perhaps indicating an earlier disyllabic form like *-ile* and Mbe has a suffix *-le/-li* in the perfective with (+argument focus) (Watters 2017: 941–942).

Thus, across Bantoid groups outside of Bantu, potential cognates of NB **-ile* or of one of its historical components appear to correlate with (+argument focus). They contrast with (+verb focus) forms that have no suffix. Where the vowel *-i* and other vowel cognates occur, the language (e.g. Limbum, Tikar) does not distinguish between (+verb focus) and (+argument focus). The significance of these distinctions is not immediately clear but it may be that earlier **-i* occurred in (+verb focus) contexts and **-lV* or **-le* occurred in (+argument focus) contexts.

Further afield, Emai, an Edoid language in West Benue-Congo, has suffixal *-í* and a postverbal particle *lé* as dual, not co-occurring exponents of anteriority/perfectivity according to Schaefer & Egbokhare (2021). They speculate that “dual exponents of anteriority or perfectivity may have co-existed among the dialectal ancestors of East and West Benue Congo, i.e. Proto Benue Congo, and perhaps late-stage ancestors in Niger-Congo that preceded the Benue Congo split into East and West” (Schaefer & Egbokhare 2021: 5).

3.5.1.3 Forms possibly derived from **màd* ‘finish’ BLR 2143

The Ring language Aghem uses *mɔ* in P1 and P2, Tikar has a suffix *-mɛ*, and Vute P0 has a suffix *-mɛ*. In the North subgroup of Grassfields, Limbum has preverbal *m* in P3 and *mú* in P2. Mfumte has *ma* in both P2 and P0, only distinguished by tone. In the Ngemba subgroup, Bafut has *mə* in P0. These forms possibly derived from **màd* ‘finish’ BLR 2143.

3.5.1.4 Variants of *ka/ke* and *le/la* for past tense

In the Bamileke languages, we note the presence of pre-stem morphemes such as *lè*, *là*, *lɔ̀*, *lò* and *lú*, and *kà*, *kè* and *kà* distributed among the past tense markers P1, P2 and P3. In P4 three use *lá/dá*. In Mundani, *lè* also appears as P3, similar to its Bamileke neighbour Yemba/Dschang. Bafut uses *leŋ* for P3. Kom uses a *læ* in P4 and P2, distinguished from each other only by tone. In some Mungbam dialects *le* occurs in P3 and P2. Ju Ba uses *lo* (P3), *le* (P2), and *la* (P1). The relationship between all of these *IV* past tense markers is uncertain, but it appears that a morpheme *IV* acquired a role in multiple distant pasts. Across NB *la(a)*, infrequent, and *ka*, slightly more frequent, also occur as future markers.²⁷ We have chosen to see past and future *ka* as deriving from an earlier itive ‘go and verb’, whereas Botne sees them as linked through the concept of distal: a distinction in place deixis that indicates location far from the speaker or other deictic centre (cf. Botne 1999). We do not judge here between these two possibilities.

In addition, Mungbam, geographically separated from the Bamileke subgroup by the Ring languages, uses *ka* and *le* in past forms. In two Mungbam dialects, *ka* or *kə/ha* occur in P3, P2, and P1, and in three other Mungbam dialects *le* or *lə* occur in P3, P2, and P1. In Ngiemboon *la* occurs in P3 and *ka* occurs in P2 and Yemba/Dschang has *ke* in P3 and *le* in P2. Ngomba has *ka* in P3 and *la* in P1. These shared exponents point to a particular likely shared history between Yemne-Kimbi languages and the Bamileke languages. It also distinguishes the Yemne-Kimbi languages, once referred to as “Western Beboid”, from the Beboid of today (the old “Eastern Beboid”). Even though various languages have forms possibly related to *-IV*, it is only in Bamileke and Yemne-Kimbi that we see this interplay between *-IV* and *-kV*, suggesting a possible earlier relationship between the two groups despite their current distance from one another.

3.5.1.5 Possible Proto-Beboid forms for P2, P1

Furthermore, the Beboid (old “Eastern Beboid”) forms suggest a possible set of Proto-Beboid forms: *cí* P2, *bé* P1, and *né* ~ \emptyset P0. There may be echoes of these in Grassfields, particularly in Bafut (Eastern Grassfields > Mbam-Nkam > Ngemba) *kì* P2 and *nîŋ* P1, and in Limbum (Eastern Grassfields > North) *bá* P1.

²⁷Thanks to Robert Botne for noting how similar Mungbam exponents are to those in Bamileke and for data on Bamileke lects other than those in our data.

3.5.1.6 Nasal verbal prefix *N-*

In Bamileke and Ring languages, some verb forms, mainly imperfectives and P1, take a nasal verbal prefix *N-*. The two are tonally distinct. It appears in at least P1 in three Bamileke languages, with Ngomba extending it to P2 (and P0) and Mengaka using it in P2. In Babanki, it is present in P3. In the Nun Grassfields language Shupamem, it is present in all tenses of the past and the future imperfective forms. Otherwise, it is not found elsewhere in Bantoid, but it does appear in some nearby NB zone A languages.

3.5.1.7 Summary

In conclusion, given the diversity of exponents for the perfective, it is not currently possible to reconstruct an original, single, full set of tense markers for Bantoid. As for subgroups, Beboid displays a possible set of past tense exponents, and there are strong indications of a set of past tense forms for the Bamileke languages. Otherwise, a few individual forms do stand out across the eastern Bantoid languages: \emptyset as P0/retrospective, *-i* and *-IV* associated with ‘non-near past’, *yV* with ‘past’, and *a* ‘past’. It is plausible that *-i* and *-IV* combined or it may be that the suffix *-iIV* was reduced to *-i* through the loss of *l* and reduction of the resulting long vowel (*-ile* > *-ii* > *i*) or *-IV* through the loss of the initial *-i*.

3.5.2 Exponents of past imperfective

Appendix D displays the various forms of the imperfective aspect, combined with the various tenses where relevant. The matching of the imperfective with the tense categories for each language is not always as uniform as for the perfective.

The imperfective aspect is generally more complex morphologically and semantically than the perfective. Languages find various ways to represent the internal temporal structure of a situation or event. Various category labels capture these differences. The generic label is IMPERFECTIVE (IPFV), but the nuances found often compel researchers to use more specific labels to capture the meanings involved, such as HABITUAL (HAB), PROGRESSIVE (PROG), CONTINUOUS (CONT), DURATIVE (DUR), and INCOMPLETE (INCOMP). We are not sure in some cases of the accuracy of the labels. It is clear that the eastern Bantoid languages had imperfective forms to correspond to the perfectives, and that for each morphologically marked tense category there is both an overtly marked perfective and imperfective aspect.

3.5.2.1 Imperfective suffix *-á* vs. PB **-a(n)g-a*

Two suffixal forms are associated with imperfective marking: one is *-a* and the other involves a velar plosive *g* plus accompanying vowels similar to **-a(n)g-a*.²⁸ Whether these two markers are cognate within Bantoid is unknown at this time.

The suffix *-a* IPFV is the most common marker of the imperfective in Appendix D. It is present throughout the Grassfields, across P1–P4 as *-a*, *-a*, *-e*, or a copy of the verb stem vowel. Tikar, Noni (Beboid), and Vute (Mambiloid) have all also developed CV suffixes for some IPFV forms. The historical relationship between these CV suffixes and the suffix *-a* is not clear. Vute has also developed separate forms for both (+verb focus) and (+argument focus).

The second suffix involves the velar plosive *g*. It only appears in Ndemli among the languages of the eastern Bantoid region. Ndemli uses the suffix *-ŋgè?* IPFV. Relative to the PB form it displays *g* with the optional PB prevelar nasal, but the vowel *ɛ* differs from the PB postvelar *a*.

The Ndemli suffix seems to be unique among the eastern Bantoid languages in its use of the *g* IPFV suffix. However, looking more widely, two languages of the western Bantoid region also use *g* IPFV suffixes. These suffixes appear cognate with PB **-ag-a* in Table 5.

Denya (Mamfe group), in the western Bantoid region, uses a suffix *-gè* IPFV. Western Ejagham (Ekoid), also in the western Bantoid region and in the Cross River basin with Denya, presents a more elaborate relationship with an internally reconstructed Proto-Ejagham **-ág-á* or **-ág*.

Example (3a) displays the suffix *-á* with CV(C)(V) roots; (3b) shows that CV roots use a velar plosive *-g*; (3c) presents the irregular verb root ‘to go’. The imperfective continuous IPFV:CONT, hortative HORT and conditional COND are provided to show that the underlying vowel of the verb root is *i*. However, unlike the other Ejagham CV roots, the historical sequence *-ji-ág* froze into the form *-jǎg*. Rather than deleting the vowel *a* of the suffix it maintained it and deleted the root vowel *i*. This frozen form *a-j-ǎg* gives evidence of an the earlier *-ág* suffix that is now mostly divided into allomorphs *-a* and *-g*. This frozen form *a-j-ǎg* is used for both the perfective and the imperfective. Finally, (3d) shows that CV roots may also use an allomorph *-gá* instead of the simple *-g*. This *-gá* often refers to a general situation. This evidence suggests a Proto-Western Ejagham IPFV suffix **-ágá* or at least **-ág*.

²⁸The presence of the homorganic nasal before the stop occurs spasmodically in Bantu and Bantoid. To date no one has been able to explain its erratic appearance, hence our representation ‘*a(n)g*’.

- (3) Habitual/concomitant forms in Western Ejagham (Watters 1981: 383–389)
[low tone unmarked, all other tones marked]
- a. Roots using -á
 - CVC *a-nâm* ‘she bought’ PFV > *a-nám-á* ‘she buys’ IPFV:HAB
 - CVV *a-sáε* ‘she sharpened’ PFV > *a-sá-á* ‘she sharpens’ IPFV:HAB
 - CVCV *a-káŋε* ‘she fried’ PFV > *a-káŋ-á* ‘she fries’ IPFV:HAB
 - b. Roots using -g
 - CV *a-dí* ‘she ate’ PFV > *a-dí-g* ‘she eats’ IPFV:HAB
 - c. Irregular root -ág
 - CV **a-jí* ‘she went’ PFV
 - a-j-äg* ‘she went/goes’ PFV/IPFV:HAB
 - a-kí-ji* ‘she is going’ IPFV:CONT
 - a-jǐ* ‘she should go’ HORT
 - á-jǐ* ‘if she goes’ COND
 - d. Extended suffix -gá
 - CV *á-dí* ‘they ate’ > *á-^ldígá* ‘they eat’

Even further to the west outside of Bantoid, in Obolo, a Lower Cross River language, one of the imperfective suffixes is *-ga*. This distribution of a *-g* imperfective suffix suggests an origin within wider Bantoid and even beyond (Obolo).

3.5.2.2 Imperfective *shí/sí/tsé* and PB **kí* ‘persistent, situative’

Two other recurring imperfective morphemes are worth noting. One involves the forms *shi* and *si*. In Limbum *shi* is the IPFV, in Bafut *si* marks the IPFV for P2 and P3, and in Yemba/Dschang *si* is one of the variants for the IPFV (PROG). Mengaka uses *tsé* for IPFV. These could be (de)palatalised versions of another morpheme occurring in NB and various BC languages outside NB, i.e. *ki*, but such an analysis needs to be checked against their diachronic phonologies. In Babungo, *yàa kî* ‘- marks the PAST HAB and in the North subgroup of Eastern Grassfields it marks HAB in Limbum and in Mfumte it marks the IPFV (PROG). In fact, in Mfumte *ki* with no tense marker indicates the present. In Bafut, it serves as the F0 present marker. We interpret these as being related to PB **kí*- ‘persistent, situative < imperfective’ (Nurse 2008: 246, 6.2.4(iv)). Looking further afield, *kî* is also found in the western region of Bantoid. In Ejagham *kí*- marks continuous or progressive aspect (Watters 1981: 379–383). In Mbe *-ki* serves as the imperfective or progressive suffix. Even further afield in Obolo in CR we find *kî*- marking the imperfective (Aaron 1999). This morpheme appears to have a long history in CR and Bantoid.

3.5.3 Exponents of future tenses

It appears that 5000 years ago or earlier the innovation of past tense among lects of what is today the eastern region of Bantoid was not matched by a similar innovation of future tense. Future tense appears to be a later development. The earliest form of future reference likely involved the use of imperfective forms from their original aspect-prominent systems, the semantics of which provided a present and a future reading, depending on context. As may be seen from Appendix E, three languages use only imperfective forms for the future even today: Babungo, Aghem, Tikar, and possibly Vute.²⁹

By contrast, while some lects did not participate in the innovation of future tenses, other languages today have developed elaborate combinations of future tense and aspect, as seen in Appendix E. For example, the Bamileke (Eastern Grassfields > Mbam-Nkam), apart from Mengaka, display a full set of future tenses, each with a perfective, imperfective, and a second imperfective (“progressive”) form.

Given this disparity, we ask two questions. First, does the marking of future time reference show signs of developing into a system similar to their past time reference, with each tense realised in both a perfective and imperfective form? Second, do those forms or exponents point to likely shared or proto-forms within Bantoid that relate to Proto-Bantu as discussed in §2.3? Consider the twenty-four languages presented in Appendix E.

What of the systems involving future tenses and perfective/imperfective aspects?

From the data available in the various grammars, it is clear that the development of systems for future tenses was not as systematic as it was for past tenses. There is a spectrum. Some languages have multiple future tenses in both perfective and imperfective aspects. At the other end of the spectrum, some only have one future form or two imperfective forms. From the most elaborate to the least, we find the following.

- One Bamileke lect (Fe’fe’) has five futures.
- Four Bamileke lects (Ngomba, Ngiemboon, Yemba, Nda’nda’) and one NWB language (Nen A44) have four futures, F1 to F4, in both perfective and imperfective aspects. Mengaka only has one future tense, having both a PFV and IPFV form.

²⁹About Grassfields, Stephen C. Anderson (p.c.) says that it is his “[...] belief that Grassfields past tense markers developed before future tense markers, because 1) future markers in Ngiemboon, etc. function as auxiliary verbs, with certain verb characteristics, while past markers do not, and 2) they occur in different slots in the verb phrase.”

- Four Bamileke lects (Ngombale, Ngwe, Ghomala', Medumba), Bafut (Eastern Grassfields > Mbam-Nkam > Ngemba), Mundani (Momo Grassfields), Babanki (Ring Grassfields), Noni and Mungong (Beboid), Wawa (Mambiloid), maybe Tikar, Maande A46, Kpa A53, Gunu A622, Ewondo A72a, maybe Kwakum A91, and some zone B languages (north-western NB) all have three futures, F1 to F3. Bafut and Noni have all four in the perfective and imperfective aspects, but Mundani has it only in the perfective. In the imperfective, Mundani uses a general future form. Babanki has (+FOC) and (–FOC) forms for F1, F2, and F3.
- Many languages have two futures, F1 and F2: Limbum (Eastern Grassfields > North), Shupamem (Eastern Grassfields > Mbam-Nkam > Nun), Babungo, Kom, Oku, Aghem, Mmen (Ring Grassfields), Ndemli, Obang (Wider Grassfields), Mundabli, Koshin (Yemne-Kimbi), Ju Ba, Vute (Mambiloid), and several Bantu zone A and B languages. Limbum and Shupamem have forms for each future in the perfective and imperfective aspects. Obang has one form reported for F1, and F2. Babungo, Kom, Aghem (all Ring Grassfields), Ndemli (Wider Grassfields), Tikar (isolate), Mundabli (Yemne-Kimbi) and Vute and Ju Ba (both Mambiloid) all use two future tenses, F1 and F2. Babungo, Aghem and likely Vute only use forms of the imperfective aspect. Kom, Mundabli, and Ju Ba likely only use imperfective forms also, but it was not possible to verify this likelihood. These languages likely developed a second imperfective out of their earlier aspect-prominent verbal system and never took the second step of developing parallel future perfective forms.
- Mfumte (Eastern Grassfields > North), Ngie (Momo Grassfields), Nchane (Beboid), Mungbam, Ajumbu, Buu (Yemne-Kimbi), and several zone A and B NB languages use a general future. Again, these may represent their inherited imperfective, but we are unable to verify.

In summary, at least five and perhaps nine of these twenty-four Bantoid languages, 20% to 40%, have not expanded their tense system so that it would include perfective and imperfective future time references. Three are using only the imperfective from their original aspect-prominent system to indicate future time reference and another six may be doing the same.

What exponents of future time reference occur in these Bantoid languages, and how do they relate to PB? In §2.3 we stated that of all the various forms for future tense, only two have any claim to possible status as PB forms. They are pre-stem *ka* and *la(a)*. Both of these are present in our data and are the most widespread within the non-NB Bantoid languages:

- The most common exponent is *IV*, especially in Grassfields. Reflexes are most widespread in Grassfields, with one possible cognate in Bebid. Those in Grassfields represent various future tenses from F1 to F4. The Bebid form represents General Future.
- The next most widespread exponent is *ka* and its possible cognates; for which see Appendix F. In terms of Bantoid groups, this form is distributed among a more diverse set of subgroups than *IV*.
- Another five exponents of future tense also occur.

So *IV* and *ka* appear as possible cognates to **la* and **ka* of PB discussed in §2.3. PB did not adopt any of the other future markers, so possibly these were the earliest markers used for the future in the mix of lects in the eastern region of Bantoid.

3.5.4 An alternative representation

Appendices C, D and E are essentially lists of comparative data for the 23 Bantoid languages under discussion, but tense and aspect in real languages are not lists and speakers do not learn lists. They learn systems. Elsewhere up to this point, we have made much mention of structure and system, but have so far not really illustrated them. The verb consists of several interlocking systems, involving tense, aspect, conjunctive vs. disjunctive, focus, positive vs. negative. We cannot include all those here but simply sketch tense and aspect, which we represent as an interlocking system, as in Tables 8 and 9. For Table 8 we choose just one Bamileke (Eastern Grassfields > Mbam-Nkam) language, Ngiemboon, with data from Appendices C, D and E. We opted for Ngiemboon because the data on aspect for it are richer than for the other Bamileke languages.

To clarify similarities between Bantoid Grassfields and north-western NB, we present Table 9, with Mpongwe B11a as the NB language (data from Nurse 2019: Addendum 2). We have simplified the data by including only one-word forms, omitting compounds and the categories represented by them. The original sources of the data are Gautier (1912) and Gérard Philippson (p.c.). Gautier writes all pre-stem morphemes of Mpongwe B11a discretely. Philippson suggests that in Galwa B11c only the 1SG is an independent pronoun.

There are certain obvious differences between Table 8 and 9. One is that between the analytic in Table 8 and the (largely) synthetic structure of the verb in Table 9, mentioned before and dealt with in the next section. Another is the

Table 8: Tense and aspect in Ngiemboon, a Grassfields (Bamileke) language

	PFV, unmarked	IPFV	PROG, <i>nè</i> N-verb- <i>a</i>
P4 <i>là lá?</i>	<i>là lá?</i> N-verb	<i>là lá?</i> N-verb- <i>a</i>	<i>là lá?</i> <i>nè</i> N-verb- <i>a</i>
P3 <i>là</i>	<i>là</i> verb	<i>là-a</i> N-verb	<i>là nè</i> N-verb- <i>a</i>
P2 <i>kà</i>	<i>kà</i> verb	<i>kà-a</i> N-verb	<i>kà nè</i> N-verb- <i>a</i>
P1	<i>ne</i> N-verb	<i>kǒ</i> N-verb- <i>a</i>	<i>kǒ nè</i> N-verb- <i>a</i>
P0 \emptyset	\emptyset verb	\emptyset N-verb- <i>a</i>	\emptyset <i>nè</i> N-verb- <i>a</i>
F0 \emptyset	n.a.	\emptyset verb- <i>a</i>	\emptyset <i>nè</i> verb- <i>a</i>
F1 <i>gè</i>	<i>gè</i> verb	<i>gè</i> verb- <i>a</i>	<i>gè nè</i> verb- <i>a</i>
F2 <i>tó/gyò</i>	<i>tó/gyò</i> verb	<i>tó/gyò</i> verb- <i>a</i>	<i>tó/gyò nè</i> verb- <i>a</i>
F3 <i>lù</i>	<i>lù</i> verb	<i>lù</i> verb- <i>a</i>	<i>lù nè</i> verb- <i>a</i>
F4 <i>lá?/fó</i>	<i>lá?/fó</i> verb	<i>lá?/fó</i> verb- <i>a</i>	<i>lá?/fó nè</i> verb- <i>a</i>

Notes: We have used Anderson (1983) as our basis. Sonkoue (2020b) deals with a second, slightly different, Ngiemboon lect. As a paradigm Table 8 is complete. F0 PFV does not exist. F0 only occurs in the IPFV and PROG. As pointed out above in §3.5.1, we are not happy with the semantics of categories here labelled P0 and F0. They are unmarked for time, as can be seen. There may also be tonal details omitted in those categories (see Sonkoue 2020b). Verb-final /-a/ may rather be a copy of the verb stem vowel.

Table 9: Tense and aspect in Mpongwe B11a, a NWB language

Tense	PFV	IPFV
P3 <i>a-verb-í</i>	<i>my a-yɛn-í</i> 'I saw'	<i>my a yɛn-áy-í</i> 'I was seeing, ...'
P2 <i>a-verb-i</i>	<i>my a-dyên-î</i> as above	<i>my a-dyên-áy-i</i> as above
P1 <i>a-verb-a</i>	<i>my a-dyên-â</i> as above	<i>my a-dyên-áy-a</i> as above
\emptyset -verb- <i>a</i>	<i>mi Ø-dyên-a</i> 'I see, am seeing, will see, I seeing, ...'	
Future <i>bé-verb-a</i>	<i>mi bé-dyên-á</i> 'I will see'	

richness of the Ngiemboon system. A third is the completely different set of morphemes involved – most of the pre-stem morphemes in Ngiemboon appear to derive from auxiliaries.

3.6 Synthetic or analytic verb structure

We can now answer the question as to whether the Bantoid languages with tense outside of Bantu are synthetic or analytic. Of the eighteen NB languages in Nurse (2019), ten are clearly synthetic, six analytic, and two are mixed or unclear, whereas all the 23 non-NB languages above are analytic.

In terms of their internal structure, verbs in non-NB Bantoid languages are synthetic in their use of suffixes but are analytic in their use of preverbal morphemes, particles and auxiliaries. Suffixes mark aspect, inherited from their earlier aspect-prominent stage. The common example is the imperfective suffix *-a(g)* or the perfective suffixes *-IV* (Babanki, Noni) or *-i* (Aghem) involved in the (+/–focus) systems. Suffixes may also include verbal extensions in some languages. The preverbal location is where the innovative work has occurred, where full verbs in serial constructions became auxiliary verbs and, when finally reduced, became particles and prefixes marking tense and modal categories.

4 Tense in PB and its rise in Bantoid

Our primary motivation in this study was to examine tense in Proto-Bantu. In the process, we found it necessary to look more widely. Since (Narrow) Bantu is part of Bantoid and other Bantoid groups border on the north-western region of Bantu, we expanded our search to include the wider Bantoid region. In the process, we identified a set of Bantoid groups in the eastern region of Bantoid immediately bordering north-western Bantu that also have TA verbal systems similar to those in Bantu. These groups are Grassfields, Beoid, Yemne-Kimbi, Tikar, and some Mambiloid lects. It is from a common ancestor with a subset of this group of eastern Bantoid lects that (Narrow) Bantu emerged, assumedly some 5000 years ago. It is reasonable to assume that these groups participated in some way in innovating tense in what would have been a set of aspect-prominent languages. In the innovation of tense, past and future categories were developed. The process, however, was not straightforward, simple, or transparent, and the results are not uniform. Investigating what happened in early Bantoid, especially in past tense development, needs more space and time than are available here.

4.1 Early “past tense”

From the available evidence, tense originated within a set of eastern Bantoid lects. They had inherited a set of verbal suffixes from their original aspect-prominent verbal system. These suffixes encoded aspects: perfective, general imperfective and other more specific imperfective categories (habitual, iterative, progressive). There were no pre-stem affixed morphemes. These suffixed forms shifted semantically into a past perfective and an imperfective present. All of these involved the suffixes already present and the pre-stem zero \emptyset , this playing a role in representing tense (cf. Tables 2, 4, 5). The suffixes continued to mark aspect. Nearly all NB zone A languages, as well as some in B, C, and D10-30, share these features.

This possible shift is repeated graphically in Table 2.

Table 2: TA structures in north-western NB without tense prefixes (repeated from page 109)

In an aspect system		In a tense-prominent system	
* \emptyset -stem- <i>a</i>	Imperfective	* \emptyset -stem- <i>a</i>	Present
* \emptyset -stem- <i>í</i>	Perfective	* \emptyset -stem- <i>í</i>	Past
* \emptyset -stem- <i>aga</i>	Habitual/Iterative	* \emptyset -stem- <i>aga</i>	Habitual/Iterative

From the evidence, we conclude that when tense developed, the first stage would have been a single initial past, contrasting with a present/non-past, with an imperfective used for the future.³⁰ Alternatively, maybe there was a marked “potential” (i.e. future), but more likely the future came later. Given that futures are often renewed, a future marker may have existed at an early stage but was not retained. Multiple contrasts developed later. Most north-western NB languages do not have multiple past contrasts, the exceptions being Kpe A22, A40-50-60, Ewondo A72a, Kwakum A91, and some zone B languages.³¹ The A40-50-60 languages likely developed their multiple pasts from contact with the Eastern Grassfields languages, particularly the neighbouring Bamileke languages, which were prolific in developing multiple tense forms.

The single pre-stem **a* ‘past’ posited for PB (see §2.2 above) existed in the ancestor lect(s) before 5000 years ago and before the Bantu exodus south and east of the CVL, likely in the Sanaga River Basin. However, the ancestor(s) of A10-20-30-40 lost this pre-stem sometime after the Bantu exodus began. Meanwhile, as

³⁰Recall our comments concerning Babungo, Aghem, Tikar, and possibly Vute, and the lack of a perfective form for the future.

³¹Readers should bear in mind that we only examined a sample so there may be more.

other Bantu communities moved away, more ‘past’ *a* contrasts developed. This pre-stem *a* probably first combined with the older *-í* perfective suffix (cf. Table 2) and then slowly but widely replaced it in the representation of “past”. Reconstruction of a future tense for PB is less certain.

We do not pursue here the issue of the kind of contact between NB A40-50-60 languages and Bantoid communities to their north-west that later resulted in the multiple contrasts found in those zone A languages.

Other eastern Bantoid groups also developed tense systems: Beboid, Yemne-Kimbi, Grassfields, and Tikar. Contrary to Watters (2012), we now think they gained their tense from a diffusion process either before or after PB emerged. Some Mambiloid languages also have tense (e.g. Vute and Ju Ba) but not all. This fact suggests that tense was not a feature of Proto-Mambiloid. Instead, Vute and Ju Ba gained tense later as it dispersed into Mambiloid more recently from the south to the north in the eastern Grassfields region.

4.2 System with typological similarities

In the process of innovating multiple tenses, all the Bantoid lects involved, the one that developed into PB and those that developed into other non-Bantu groups, shared a common system inherited from their NC past. The structure involved a contrast between each past perfective and a past imperfective form. The imperfective in non-Bantu Bantoid commonly used an inherited suffix **-a*, probably derived from the fuller imperfective suffix **-a(g)-a*. The fuller form is retained in Bantu.

4.3 Expansion to multiple tenses

There are no traces today of a single past among the non-NB Bantoid languages. All have at least two past tenses. In contrast, NB zone A languages are much more variable in this respect: a few have one, some two, some three, and some four pasts. More diagnostic than the number of pasts is whether they use the new pre-stem *a* in their past formation. The pre-stem *a* occurs in A50 and some A60-70-80 languages, and widely outside zone A (B10-20-50-C10, etc.). It is absent from all A10-20-30-40, A90 languages (also B30-40-60-80, C20, C60). These zone A groups build on the relatively simple morphological structures in the first two lines of Table 2 but in different ways, to create one, and later, several pasts. Languages in zones A50-60-70-80 had a single *a*, later expanded as different *a*. Languages in zones A10-20-30-40, A90 instead add a considerable range of pre-stem morphemes to represent past, which vary from language to language and group to group (cf. Nurse 2019: Addenda 1 and 2).

After the initial development of past tenses, the eastern non-NB Bantoid lects also probably followed different paths in the development of multiple tenses from the NB lects. The development in the non-NB lects involved multilingualism, borrowing, calquing, analogy and recycling. The details of the morphology of past and future tenses in these languages involve significant variation and it is impossible to reconstruct any original morphology with confidence.

Ancestral Eastern Grassfields, especially Bamileke, and possibly Yemne-Kimbi lects appear to have been central in the early development of tense. Early developments then spread to Momo and Ring languages, as well as Beboid, Tikar, and southern and eastern Mambiloid lects. Eastern Grassfields as well as Beboid and some other Grassfields lects continued the development of new tenses beyond the first two by using serial verbs that mutated into tense markers. A few of these lects never created more than a binary past contrast, and never fully developed future tenses. Some Mambiloid lects have not transitioned to a TA system, while others are apparently on the border between aspect-prominent and TA systems. The development of more than two past tenses in most Bantu languages occurred later, separately from the eastern Bantoid lects.

Bantoid languages show limited traces of a 'past', and most groups also encode pasts (and futures) using morphology not clearly or widely found in any of the others. Compare this with the situation in the Romance languages, which share many morphological and systemic similarities, making it possible to reconstruct a proto-system that closely resembles that of their predecessor, Latin (Hewson & Bubenik 1997). Bantoid/NB is not like this, suggesting that the different systems do not derive directly from a single Proto-Bantoid system.

4.4 The dispersal of tense

An alternative model is suggested by what we find in Mambiloid. Vute has two pasts and two futures (Thwing & Watters 1987). A few other Mambiloid varieties have also developed tense; for instance, Ju Ba has three degrees of past remoteness and two futures. Other Mambiloid languages have simple past and future tenses with no degrees of remoteness, while others have no traces of tense (Connell 2019, and p.c.). The geographical location of Vute and Ju Ba suggests that these tense contrasts are not original but may have spread into them from adjacent languages on their western border. That is, they would have adopted the notion of tense distinctions from Grassfields, but encoded it differently, using their own morphology, thus a calque. Such a model is described in some detail by Dimmendaal (2011: 189–194) for two varieties of Nilotic in Kenya. Nilotic languages are aspect-prominent with no inherited tense. However, Southern Nilotic

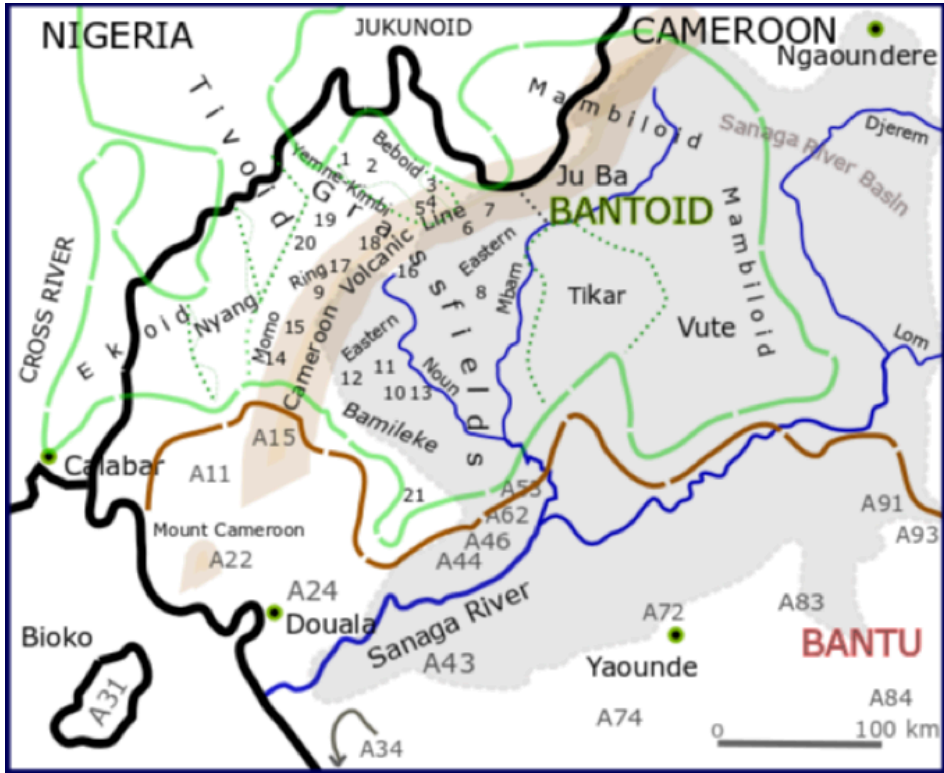
Kalenjin (three past tenses) and Western Nilotic Luo (four pasts) both developed tense contrasts independently along the lines of their Bantu neighbours. The new tense markers are transparently grammaticalised forms of time adverbials. Dimendaal suggests that trade and intermarriage apparently led to shift-induced interference, whereby speakers of a Bantu language introduced these distinctions into Kalenjin and the innovation became the norm. Alternatively, the Bantu-speaking mothers of Kalenjin husbands used them in their speech, which their children then copied.

If we apply this dispersal model here, the following questions arise: In which of the six groups (Beoid & Yemne-Kimbi, Tikar, Bamileke, the rest of Grassfields, Mambiloid, early Bantu) did tense initially appear, how did tense initially disperse, and how did multiple tense contrasts then develop and spread? Because the methodology is not clear, we do not claim to have all the answers, but we can offer some pointers.

Where did tense initially appear? Proto-Mambiloid is unlikely because, assuming Mambiloid is a valid genetic unit, tense is limited to some lects and not others, and in those that have it (Vute, Ju Ba) the encoding is different. In fact, Mambiloid serves as a northern and eastern boundary to the development of tense in Bantoid, and the CVL region to the west serves as a western boundary. Our best hypothesis is that tense initially appeared among the ancestral lects of Eastern Grassfields (Bamileke), Bantu, and maybe Yemne-Kimbi. From these lects it then spread to other regions of Grassfields (Momo, Ring and Wider) and to neighbouring groups Beoid, Tikar and Mambiloid.

How did multiple tense contrasts develop? We think it unlikely that multiple tense contrasts developed in early Bantu in the north-west. Multiple pasts and futures in Bantu are more common outside the north-west, and parts of the morphology encoding the few multiple contrasts that we do find in zone A, especially in A40-50-60, do not occur elsewhere in zone A. Bits of the innovated morphology involved in these multiple tense contrasts in A resemble some morphology in Bantoid. We think it more likely that these multiple contrasts probably intruded into these north-western languages from a Bantoid source, such as Eastern Grassfields. In addition, since we think simple tenses did not develop in Proto-Mambiloid, multiple contrasts did not originate there either. The origin for this activity was towards the Sanaga River Basin rather than the mountains of the Mambiloid region; see Figure 6.

Wider Bantu offers a possible model for the development of multiple tense contrasts. The building blocks for past tense in early Bantu were fairly restricted: *-a-*, *-í*, tone. Combining these, and combining them with pre-stem focus marking,



Key to the codes and numbers on this map:

(Narrow) Bantu languages: Lundu A11, Mbo cluster A15, Kpe A22, Duala A24, Bubi (Bioko) A31, Benga (Equatorial Guinea/Gabon) A34, Basaa A43a, Nen A44, Maande A46, Kpa A53, Yambasa A62, Ewondo A72a, Bulu A74a, Makaa A83, Njem A84, Kwakum A91, Kako A93

Bantoid languages: MAMBILOID: Ju Ba/Mambila, Vute; Isolate TIKAR: Tikar; YEMNE-KIMBI: 1 Mungbam, 2 Mundabli; BEBOID: 3 Nchane, 4 Mungong, 5 Noni; EASTERN GRASSFIELDS, North: 6 Limbum, 7 Mfumte; EASTERN GRASSFIELDS, Mbam-Nkam, Nun: 8 Shupamem; EASTERN GRASSFIELDS, Mbam-Nkam, Ngemba: 9 Bafut; EASTERN GRASSFIELDS, Mbam-Nkam, Bamileke: 10 Ngiemboon, 11 Ngomba, 12 Yemba/Dschang, 13 Mengaka; MOMO GRASSFIELDS: 14 Mundani, 15 Ngie; RING GRASSFIELDS: 16 Babanki, 17 Babungo/Vengo, 18 Kom, 19 Aghem; WIDER GRASSFIELDS: 20 Obang, 21 Ndemli

Other features: Sanaga River Basin = grey area, Cameroon Volcanic Line = wide brown band

Figure 6: NB and Bantoid languages in the region of tense innovation

vowel lengthening, and other tools gave many tense and encoding possibilities over five millennia, as language communities were dispersing.

We assume that multiple tense contrasts developed in the wider Bantoid area beyond NB starting with the common serial verb construction. Serial verb > Auxiliary <aspect> > Auxiliary <tense>, finally becoming integrated as a pre-stem tense marker, is a typical grammaticalisation shift.³² It seems unlikely that multiple contrasts developed independently in the other three language groups, given their adjacency, the small geographical area, and the categories being so similar. We see from the Nilotic example above that a completely calqued mini-system consisting of a set of several tense distinctions can be transferred simultaneously, so we think it plausible that these multiple contrasts dispersed from one source, with each early language group developing its own morphology.

5 Conclusion

We propose that a Bantoid lect or a set of Bantoid lects innovated tense before PB separated from other Bantoid lects – a pre-Bantu stage. Speakers of these lects likely resided on the eastern slopes of the Cameroon Volcanic Line into the Sanaga River Basin. Some 5000 years ago some of these lects emerged from this region as PB forms and other lects formed the beginning of Eastern Grassfields, plausibly the Bamileke and Mbam-Nkam languages. A single initial past probably emerged first, possibly followed by a future. This innovation of this single past tense dispersed to the others, in the circumstances sketched in §4.4 above. Later, multiple pasts developed among the non-NB Bantoid languages. We admit to being unsure exactly where this first developed, but our sense is that the locus was early Eastern Grassfields, and then dispersed to the north and east to the rest of Grassfields, Beboid, Tikar, parts of Mambiloid, and even some NB zone A languages. Later, multiple pasts also developed among the NB languages expanding south and south-east, but that is a separate story.

To conclude, we sketch here an overview of how what we are proposing compares to Meeussen (1967) and Nurse (2008) (cf. Table 1), the focus being on tense. Not surprisingly, our ideas more resemble Nurse's than Meeussen's.

Acknowledgements

We thank Stephen C. Anderson and Robert Hedinger for their comments on an earlier version of this chapter; Larry M. Hyman for his thoughts on certain expo-

³²Thanks to one anonymous reviewer for help here.

Table 10: PB reconstructions by Nurse & Watters (2022 = this chapter), Nurse (2008), and Meeussen (1967)

this chapter	Nurse (2008)	Meeussen (1967)
<i>a</i> -stem- <i>a</i> past	<i>a</i> -stem- <i>a</i> past <i>a</i> -stem- <i>a</i> recent IPFV	<i>á</i> -stem- <i>a</i> PRET IPFV
<i>a</i> -stem- <i>í</i> past RET	<i>a</i> -stem- <i>ile</i> past RET <i>a</i> -stem- <i>idé</i> recent PFV	<i>á</i> -stem- <i>ide</i> PRET PFV
<i>a</i> -stem- <i>ag-a</i> past IPFV	<i>a</i> -stem- <i>ang-a</i> past IPFV	
\emptyset -stem- <i>i</i> RET	\emptyset -stem- <i>ile</i> RET	
\emptyset -stem- <i>a</i>	\emptyset -stem- <i>a</i>	\emptyset -stem- <i>a</i> PRS 1 IPFV (CONJ)
\emptyset -stem- <i>ag-a</i> IPFV	\emptyset -stem- <i>ang-a</i> IPFV	
<i>l(a)</i> -stem- <i>a</i> DISJ	<i>l(a)</i> -stem- <i>a</i> DISJ	<i>da</i> -stem- <i>a</i> PRS 2 IPFV (DISJ)
<i>ka</i> -stem- <i>a</i> itive/future	? <i>ka</i> -stem- <i>a</i> itive/future	<i>ka</i> -stem- <i>a</i> future

nents of tense in Grassfields; two anonymous reviewers for their insights; Robert Botne for comments on facts and interpretations; Bonny Sands, Hilde Gunnink, and Thera M. Crane for various types of help; and the editorial team, especially Koen Bostoen, for their rich and helpful comments and hard work.

Abbreviations

A, B, C, ... ‘zones’ or categories of Bantu languages (Guthrie 1948; 1971; Maho 2009)

ADV	adverb
AUX	auxiliary
B	verb base
BC	Benue-Congo
C	consonant
CAUS	causative
CM	Comparative Method; clause marker
COND	conditional
CONJ	conjunctive
CONT	continuous

CR	Cross River
CV	consonant + vowel
CVL	Cameroon Volcanic Line
CWB	Central-Western Bantu
DISJ	disjunctive
DUR	durative
EB	Eastern Bantu
EXT	(verb root) extension
F	future
F1	near future (see the Key of Table 3 for further detail)
F2	distant future (see the Key of Table 3 for further detail)
-FOC	focus on lexical and phrasal constituents
+FOC	focus including aspect and truth value
FV	final vowel
H	high tone
HAB	habitual
HORT	hortative
INCOMP	incompletive
IPFV	imperfective
IRR	irrealis
ITER	iterative
N	homorganic nasal
NB	(Narrow) Bantu
NC	Niger-Congo; nasal + consonant
NEG	negation
NWB	North-Western Bantu
OM	(prefixal) object marker
OP	object pronoun
P	past
P1	today past (see the Key of Table 3 for further detail)
P2	yesterday past (see the Key of Table 3 for further detail)
P3	days, weeks, months, or distant past (see the Key of Table 3 for further detail)
P4	remote past (see the Key of Table 3 for further detail)
PASS	passive
PB	Proto-Bantu
PFV	perfective
PRET	preterite
PROG	progressive
RECP	reciprocal

RET	retrospective
SBJP	subject pronoun
SG	singular
SP	subject prefix
SUB	subject
SWB	South-Western Bantu
T	tense
TA(M)	tense/aspect/(mood)
V	vowel
VC	vowel copy
WWB	West-Western Bantu

Appendix A Definitions

The following definitions of some basic terms are mostly from Nurse (2008: 308–318).

aspect: verb category expressing the internal temporal constituency of a situation. All finite verbs have aspect, marked or not. Verbs may have more than one aspect. Commonest NC/Bantu aspects are imperfective and perfective. Less common: retrospective, and subdivisions of imperfective (progressive, habitual, iterative).

aspect-prominent (language): language having aspect but not tense.

conjunctive (focus) (CONJ): refers to verb forms focusing on a postverbal constituent (object, adverbial, second verb) or new material.

disjunctive (focus) (DISJ): verb forms indicating no special relationship between that verb and a following constituent. Emphasis is on the verb or one of its categories.

extension (EXT): in NC the verbal base can consist of a root or root followed by one or more productive derivational suffixes known as extensions.

final vowel (FV) or just final: the tenth of eleven positions in Meeussen's template of the Bantu verb, but also used of the verb in NC. See also Pre-final.

focus (FOC): special prominence given to some element to mark it as expressing the most important (new) information or to contrast it with something else.

habitual (aspect) (HAB): refers to a situation characteristic of an extended period of time; the situation is viewed as a characteristic of a whole period.

imperfective (aspect) (IPFV): contrasts with perfective and represents an unbounded situation that lasts over a period of time.

iterative (aspect) (ITER): refers to a repeated situation; an incomplete series of events.

itive (directional): refers to the agent moving away from the current location or time, often using the verb 'go' and may include intention and future time.

perfective (PFV): represents a situation as a single bounded whole, without regard to its constituent phases.

pre-final: the ninth of eleven positions in Meeussen's template of the Bantu verb.

preterite (tense) (PRET): for languages with two degrees of past, some authors, mainly francophone, refer to the near one as recent past, the farther one as preterite.

progressive (PROG): represents an unbounded situation ongoing at or around reference time.

retrospective (RET), also called *perfect or anterior*: refers to a past event with present relevance (with dynamic verbs), or to a situation that started in the past and continued into the present (stative verbs).

tense: the grammaticalised representation of, or verbal inflection for, location in time, relative to some reference point, often the present.

tense/aspect system: we use it to refer to an interlocking morphological/grammaticalised system of tense and aspect in the verb.

vowel copy (VC): a process in which final vowels, in the near past or retrospective, assimilate to those of the verb stem. It applies to a single suffix involving a (more or less) complete copy of the stem vowel into that suffix.

Appendix B Eastern Bantoid languages and their resource(s) serving as examples throughout §3

GROUP/ Sub-group	Sub-sub-group/ <i>Sub-sub-sub-group</i> / Language	Source(s)	
BEBOID	Noni	Hyman (1981)	
	Nchane	Boutwell (2020)	
	Mungong	Boutwell (2014)	
YEMNE-KIMBI	Mungbam	Lovegren (2013)	
	Mundabli	Voll (2017)	
GRASSFIELDS Eastern	<u>North</u>		
	Limbum	Fransen (1995)	
	Mfumte	McClellan (2014)	
	<u>Mbam-Nkam</u>		
	<i>Nun</i> : Shupamem	Nchare (2012)	
	<i>Ngemba</i> : Bafut	Mfonyam (1989)	
	<i>Bamileke</i> :		
	Ngiemboon	Anderson (1983), Lonfo & Anderson (2014)	
	Ngomba	Satre (2004)	
	Yemba/Dschang	Hyman (1980), Harro & Haynes (1991)	
	Mengaka	Sonkoue (2019)	
	Momo	Mundani	Parker (1991), Magba (1995)
		Ngie	Watters (1980)
	Ring	Babanki	Akumbu & Chibaka (2012), Akumbu et al. (2019)
		Babungo	Schaub (1985)
Kom		Shultz (1997)	
Aghem		Anderson (1979)	
Wider			
MAMBILOID	Obang	Asoshi (2015)	
	Ndemli	Ndedje (2013)	
MAMBILOID	Vute	Thwing & Watters (1987)	
	Ju Ba	Perrin (1972), Connell (2019)	
TIKAR	Tikar	Stanley (1991)	

Note: For data on more Bamileke languages than the four included, see Botne (2020).

Appendix C Exponents of past perfective in eastern Bantoid

GROUP .../ Language	P4	P3	P2	P1	P0
BEBOID					
Noni (+FOC)		tò nò verb	cí nò verb	béé verb	ø nò verb
Noni (-FOC)		tò verb ló	cí verb ló	bé verb ló	ø verb ló
Nchane		gē verb	ché SUB verb	bé SUB verb	ø verb
Mungong		ká bé verb	yí verb	bí verb	nó verb
			ká (general PAST)		
YEMNE-KIMBI					
Mungbam		lē ~ ká verb+IRR	lē ~ ká verb	fə ^h ~ ha ^h ~ ká verb	ø verb
Mundabli		kò verb.PFV	á ~ ná verb.PFV	fə ^h verb.PFV	ø verb.PFV
GRASSFIELDS					
Eastern					
<u>North</u>					
Limbum		m verb-i	mú verb-i	bá verb-i	ø verb-i
Mfumte			má verb	nu verb	má verb
<u>Mbam-Nkam</u>					
Nun: Shupamem		kápi ` verb	pi ` verb	pé ` verb	ø ` verb
Ngemba: Bafut		lēŋ verb	kí verb	nāŋ ` verb	mā verb
<i>Bamileke:</i>					
Ngitemboon	là lá? n-verb	là verb	ká verb	ně n-verb	ø verb
Ngomba	nídá'n-verb (gen.)	ka verb	n-verb	lá'n-verb	ø n-verb
	ka lá'n-verb (def.)				
Yemba/Dschang	lè lá? n-verb	lè verb'ŋ	kè verb'ŋ	áá n-verb	á' verb
Mengaka		ká verb	^ n-verb	` verb	ø verb

GROUP .../ Language	P4	P3	P2	P1	P0
Momo					
Mundani		<i>lè</i> verb	<i>á</i> verb (= P2 & general PAST) <i>a</i> verb <i>a</i> verb- <i>í</i>	<i>ghâ/lɔ/lí</i> verb <i>á</i> verb <i>á</i> verb- <i>í</i>	\emptyset verb
Ngie (+FOC)					
Ngie (-FOC)					
Ring					
Babanki (+FOC)		N-verb ()	<i>tá</i> verb	<i>yí</i> verb	` verb
Babanki (-FOC)		<i>á</i> N verb <i>lí</i>	<i>á tá</i> verb	<i>á yí</i> verb	<i>á</i> verb <i>lí</i>
Babungo	<i>ná</i> ` -verb <i>nin læ</i> verb	<i>yáa</i> ` -verb <i>ti</i> verb	<i>sí</i> ` -verb <i>læ</i> verb ` <i>má</i> verb <i>máá</i> verb	<i>níí</i> ` -verb <i>ní</i> verb <i>má</i> verb <i>máá</i> verb	\emptyset verb <i>ní</i> verb
Kom					
Aghem (+FOC)					
Aghem (-FOC)					
Wider					
Obang			<i>yá</i> verb	<i>ká</i> verb	\emptyset verb
Ndemli		<i>ní</i> - verb	<i>á</i> - verb	<i>á</i> - verb	` verb
MAMBILOID					
Vute (+FOC)			<i>yí</i> verb <i>tí</i>	<i>tí</i> ` verb <i>tí</i>	\emptyset verb <i>tí</i>
Vute (-FOC)			<i>yí</i> verb- <i>á</i>	<i>tí</i> ` verb- <i>á</i>	\emptyset verb- <i>á</i>
Ju Ba		<i>lá ná</i> verb	<i>lá ná</i> verb	<i>laná</i> verb	\emptyset verb
TIKAR					
Tikar		verb- <i>ka?</i>	verb- <i>e</i>	verb- <i>i</i>	verb- <i>mε</i>

Appendix D Exponents of past imperfective in eastern Bantoid

GROUP .../ Language	P4	P3	P2	P1	P0
BEBOID					
Noni		<i>tá</i> verb- <i>e</i> <i>lò</i>	<i>cí</i> verb- <i>e</i> <i>lò</i> <i>bé</i>	verb- <i>e</i> <i>lò</i>	∅ verb- <i>e</i> <i>lò</i>
Nchane		<i>gē</i> verb- <i>é</i>	<i>ché</i> SUB verb- <i>é</i>	<i>bé</i> SUB verb- <i>é</i>	∅ verb- <i>é</i>
?Mungong		<i>ká</i> <i>bé</i> verb	<i>yí</i> verb	<i>bí</i> verb	<i>ná</i> verb
		<i>ká</i> (genaral PAST)			
YEMNE-KIMBI					
Mungbam		<i>lē</i> ~ <i>ká</i> verb.IRR	<i>lē</i> ~ <i>ká</i> verb	<i>fə́</i> ^h ~ <i>ha</i> ^h ~ <i>ká</i> verb	∅ verb
Mundabli		<i>ká</i> verb.IPFV	<i>à</i> ~ <i>ná</i> verb.IPFV	<i>fə́</i> verb.IPFV	∅ verb.IPFV
GRASSFIELDS					
Eastern					
<u>North</u>					
Limbum					
PROG		<i>m</i> <i>shí</i> verb- <i>í</i>	<i>mú</i> <i>shí</i> verb- <i>í</i>	<i>bá</i> <i>shí</i> verb- <i>í</i>	∅ <i>shí</i> verb- <i>í</i>
HAB		<i>m</i> <i>kíí</i> verb- <i>í</i>	<i>mú</i> <i>kíí</i> verb- <i>í</i>	<i>bá</i> <i>kíí</i> verb- <i>í</i>	∅ <i>kíí</i> verb- <i>í</i>
Mfumte					
PROG			<i>má</i> <i>kí</i> verb	<i>nu</i> <i>kí</i> verb	<i>má</i> <i>kí</i> verb

GROUP .../ Language	P4	P3	P2	P1	P0
Mbam-Nkam					
Nur: Shupamem		<i>kápú mbú n-verb</i>	<i>pú mbú n-verb</i>	<i>pá mbú n-verb</i>	<i>pá mbú n-verb</i>
Ngemba: Bafut		<i>lêŋ si verb</i>	<i>kí sî verb</i>	<i>^ nîŋ' verb</i>	<i>^ verb</i>
Bamileke:					
Ngiemboon					
HAB	<i>là lá? n-verb-a</i>	<i>là-a n-verb</i>	<i>kà-a n-verb</i>	<i>kǒ n-verb-a</i>	<i>n-verb-a</i>
PROG	<i>là lá? nè n-verb-a</i>	<i>là nè n-verb-a</i>	<i>kà nè n-verb-a</i>	<i>kǒ nè n-verb-a</i>	<i>nè n-verb-a</i>
Ngomba					
NON-PRESENT		<i>míbó n-verb</i>			<i>pǒ</i>
IPFV					
GENERAL IPFV		<i>míbó</i>			
PRESENT PROG		<i>sé n-verb</i>			
HAB		<i>lɔ n-verb</i>			
Yemba					
PROG	<i>ŋ-gǒ é-si n-verb-a</i>	<i>sí n-verb-a</i>	<i>sí N-verb-a</i>	*	<i>sí n-verb-a</i>
PROG	<i>nŋŋ n-verb-a</i>	<i>nŋŋ n-verb-a</i>	<i>nŋŋ N-verb-a</i>	<i>n-nŋŋ N-verb-a</i>	<i>ø-nŋŋ ø-verb-a</i>
Mengaka		<i>kà tsé n-verb-V</i>	<i>^ tsé n-verb-V</i>	<i>bé tsé n-verb-V</i>	<i>ø-tsé n-verb-V</i>
Momo					
Mundani					
PAST IPFV		<i>lé wúá n-verb-á</i>			
PRESENT IPFV		<i>á n-verb-á</i>			
Ngie					
IPFV			<i>mǎǎ verb-á</i>	<i>mǎǎ verb-á</i>	
PAST HAB			<i>bǎǎ verb-á</i>		

GROUP .../ Language	P4	P3	P2	P1	P0
Ring					
Babanki		<i>tǎ N-verb-ə</i>	<i>verb-ə</i>		
Babungo					
IPFV	' - verb	' - verb	' - verb	' - verb	ø' - verb
PAST HAB		<i>yàa kt' -</i>			
PROG		$C_1V_1 > C_1i - C_1V_1 \dots$			
Kom					
INCOMP		<i>verb -a</i> [= IPFV]			
PROG		<i>nin verb</i>			
DUR		<i>nà verb</i>			
Aghem					
Wider			<i>mò verb -à</i>	<i>mò verb-à</i>	ø <i>verb-à</i>
Obang					
PROG		' - verb -(ə)			
HAB		ø			
PERSISTIVE		<i>bé</i>			
Ndemli					
IPFV		<i>verb -ngè?</i>			
MAMBILOID					
Vute (+FOC)			<i>yi' verb ni</i>	<i>ti á verb ni</i>	<i>á verb ni</i>
Vute (-FOC)			<i>yi' verb-ko</i>	<i>ti' verb-ko</i>	<i>á verb-á</i>
Ju Ba			Uncertain		
TIKAR					
Tikar		<i>béka?</i> <i>verb -CV</i>	<i>bé verb -CV</i>	<i>bi verb -CV</i>	<i>tá verb</i>
		(CV = ' -li, -e-, ó, -a, -ni, -bi, -mi, -ø)			

Note: For Yemba/Dschang, Harro & Haynes's (1991) P1–P5 are P0–P4 here.

Appendix E Exponents of future perfective & imperfective

GROUP .../ Language	F0	F1	F2	F3	F4
BEBOID					
Noni					
PFV	<i>nú</i> ∅ verb	<i>nú géé</i> verb	<i>nú bóy</i> verb	<i>nú yíú</i> verb	
IPFV	<i>nú</i> ∅ verb- <i>e</i>	<i>nú géé</i> verb- <i>e</i>	<i>nú bóy</i> verb- <i>e</i>	<i>nú yíú</i> verb- <i>e</i>	
Nchane	<i>lé</i> SUB verb (general FUT)				
Mungong	<i>nā</i> verb	<i>nā fá</i> verb	<i>nā bí</i> verb	<i>nā tá</i> verb	
YEMINE-KIMBI					
Mungbam	<i>á</i> verb (general FUT)				
Mundabli		<i>dí</i> verb	<i>ká</i> verb		
GRASSFIELDS					
Eastern					
North					
Limbum					
PFV	<i>bí</i> verb	<i>bí lò</i> verb	<i>bí fá</i> verb		
IPFV	<i>bí shí</i> verb	<i>bí lò shí</i> verb	<i>bí fá shí</i> verb		
Mfumte					
PFV	<i>kà</i> SUB verb (general FUT)				
IPFV	<i>kà</i> SUB <i>kí</i> verb (general FUT)				
Mbam-Nkam					
<i>Nun:</i>					
Shupamem					
PFV	<i>ná</i> <i>twó</i> verb	<i>ná ló?</i> verb	<i>ná ló</i> <i>twó</i> verb		
IPFV	<i>ná</i> <i>twó píú</i> <i>n</i> -verb	<i>ná ló?</i> <i>píú</i> <i>n</i> -verb	<i>ná ló</i> <i>twó píú</i> <i>n</i> -verb		
Ngemba:					
Bafut					
PFV	<i>ká</i> ∅ verb	<i>ká</i> <i>lě</i> verb	<i>ká</i> <i>lě</i> verb	<i>ká</i> <i>yě</i> verb	
IPFV	<i>ká</i> <i>kí</i> verb	<i>ká</i> <i>lě</i> <i>kí</i> verb	<i>ká</i> <i>lě</i> <i>kí</i> verb	<i>ká</i> <i>yě</i> <i>kí</i> verb	

GROUP .../ Language	F0	F1	F2	F3	F4
<i>Bamileke:</i>					
Ngiemboon					
PFV		gè verb	tó/gyò verb	lù verb	lál/ fò verb
IPFV	verb-V	gè verb-a	tó/gyò verb-a	lù verb-a	lál/ fò verb-a
PROG	nè verb-a	gè nè verb-a	tó/gyò nè verb-a	lù nè verb-a	lál/ fò nè V-a
Ngomba		gɛ/gɛ n- verb	N verb	ɲgɛ/ɲɔ n-verb	ɲɔda' verb
Dschang					
(PFV?)					
Yemba	á'á verb	àà 'piŋ'ɨ verb	àà 'lù'ú/ shuɪ'ɛ verb	á'á lá'ɛ' verb	á'á fú' verb
(PROG)					
(PROG)			é-gó é-sí é-verb-V	← =	← =
Mengaka		nɔŋɔ é-verb-V	← =	← =	← =
PFV					
IPFV		' ní/dé n-verb-'			
Momo		' ní/dé tsé n-verb-V			
Mundani					
PFV	ā verb	ā ghā/ á ló'á ē-verb	ā sà'á ē- verb	ā lí ē- verb	
IPFV	ná á wúá verb				
Ngie					
PFV (+FOC)		mii verb			
PFV (-FOC)		wī verb			
IPFV (+FOC)		mi māā verb-á			
IPFV (-FOC)		wi māā verb -á			

GROUP .../ Language	F0	F1	F2	F3	F4
Ring					
Babanki		á verb	nè verb	lù verb	
(+FOC)		á á verb´	á nè verb´	á lù verb´	
(-FOC)		táa´ - verb	ndó´ - verb		
Babungo		nî verb	læ verb		
IPFV		sì + verb -á	lò´ + verb -á		
Kom		wúlí verb	dúlí verb		
Aghem		i` verb	tá- verb		
IPFV					
Wider					
Obang	mí verb				
Ndemli					
MAMBILOID					
Vute					
IPFV		ba verb ni	kwá´ verb		
Ju Ba		ndē bā verb	ndē né verb		
TIKAR					
Tikar		ka...ka?			
PFV		yè verb -CV	ywime verb -CV		
IPFV					

Appendix F *ka* and its cognates as exponents of future tenses in Bantoid

Bantoid Group	Language	Exponent
Yemne-Kimbi	Mundabli	<i>ká</i> (F2)
Grassfields > Eastern > North	Mfumte	<i>kà</i>
Grassfields > Eastern > Mbam-Nkam > Ngemba	Bafut	<i>ká</i> (F0-F3)
Mambiloid	Vute	<i>kwâ</i> (F2)
Isolate	Tikar	<i>ka...ka?</i> (FUT PFV)

References

- Aaron, Uche E. 1999. *Tense and aspect in Obolo: Grammar and discourse* (SIL International Publications in Linguistics 128). Dallas, TX: Summer Institute of Linguistics & University of Texas at Arlington.
- Akumbu, Pius W. & Evelyn Fogwe Chibaka. 2012. *A pedagogic grammar of Babanki – a Grassfields language of northwest Cameroon* (Grammatical Analyses of African Languages 42). Cologne: Rüdiger Köppe.
- Akumbu, Pius W., Larry M. Hyman & Roland Kießling. 2019. The segmental and tonal structure of verb inflection in Babanki. *UC Berkeley Phonetics and Phonology Lab Annual Report* 15. 57–89.
- Anderson, Stephen C. 1979. Verb structure. In Larry M. Hyman (ed.), *Aghem grammatical structure* (Southern California Occasional Papers in Linguistics 7), 73–136. Los Angeles, CA: University of Southern California.
- Anderson, Stephen C. 1983. *Tone and morpheme rules in Bamileke-Ngyemboon*. Los Angeles, CA: University of Southern California. (Doctoral dissertation).
- Asoshi, Melvice. 2015. *Structural and typological approaches to Obang grammar – a Grassfields Bantu language of Cameroon* (Grammatical Analyses of African Languages 52). Cologne: Rüdiger Köppe.
- Bastin, Yvonne. 1975. *Bibliographie bantoue sélective* (Archives d'anthropologie 24). Tervuren: Royal Museum for Central Africa.
- Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en bantou* (Annales – Série in-8° – Sciences humaines 114). Tervuren: Royal Museum for Central Africa.
- Bastin, Yvonne. 1989a. El prefijo locativo de la clase 18 y la expresión del progresivo presente en Bantú, Parte 1. *Estudios Africanos* 4(6). 35–54.
- Bastin, Yvonne. 1989b. El prefijo locativo de la clase 18 y la expresión del progresivo presente en Bantú, Parte 2. *Estudios Africanos* 4(7). 61–86.

- Bastin, Yvonne. 1994. Essai d'analyse des formatifs de type -a- en bantou central. *Africana Linguistica* 11. 1–8.
- Bendor-Samuel, John T. & Rhonda L. Hartell. 1989. *The Niger-Congo languages: A classification and description of Africa's largest language family*. Lanham, MD: University Press of America.
- Blench, Roger M. 2022. The relevance of Bantoid for the reconstruction of Proto-Bantu verbal extensions. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 235–280. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575823.
- Blench, Roger M. & Kay Williamson. 1987. 'new' Bantoid. Unpublished ms.
- Bostoen, Koen & Rozenn Guérois. 2022. Reconstructing suffixal phrasemes in Bantu verbal derivation. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 343–383. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575829.
- Botne, Robert. 1999. Future and distal -ka-'s: Proto-Bantu or nascent form(s)? In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 473–515. Stanford, CA: Center for the Study of Language & Information.
- Botne, Robert. 2020. The evolution of tense systems in Bamileke (Bantoid) languages. Unpublished ms.
- Boutwell, Richard Lee. 2014. *A sketch grammar of the Mungong language*. Yaoundé: SIL Cameroon.
- Boutwell, Richard Lee. 2020. *A grammar of Nchane: A Bantoid (Beboid) language of Cameroon*. Leiden: Leiden University. (Doctoral dissertation).
- Connell, Bruce. 2019. Tense and/or aspect in Mambila. Paper presented at the 50th Annual Conference on African Linguistics, University of British Columbia, Vancouver, 22–25 May 2019.
- De Kind, Jasper, Sebastian Dom, Gilles-Maurice de Schryver & Koen Bostoen. 2015. Event-centrality and the pragmatics-semantics interface in Kikongo: From predication focus to progressive aspect and vice versa. *Folia Linguistica Historica* 36. 113–163.
- Dimmendaal, Gerrit J. 2011. *Historical linguistics and the comparative study of African languages*. Amsterdam: John Benjamins.
- Ebarb, Kristopher J. & Michael R. Marlo. 2015. Vowel length (in)sensitivity in Luyia morphophonology. *Southern African Linguistics and Applied Language Studies* 33(3). 373–390.
- Fransen, Margo Astrid Eleonora. 1995. *A grammar of Limbum: A Grassfields Bantu language spoken in the North-West province of Cameroon*. Amsterdam: Vrije Universiteit Amsterdam. (Doctoral dissertation).

- Gautier, Jean-Marie. 1912. *Grammaire de la langue mpongwée*. Paris: Procure des PP. du Saint-Esprit.
- Good, Jeff, Jesse Lovegren, Jean Patrick Mve, Carine Nganguép Tchiemouo, Rebecca M. Voll & Pierpaolo Di Carlo. 2011. The languages of the Lower Fungom region of Cameroon: Grammatical overview. *Africana Linguistica* 17. 101–164.
- Grégoire, Claire. 1979. Les voyelles finales alternantes dans la conjugaison affirmative des langues bantoues centrales. *Journal of African Languages and Linguistics* 1(2). 141–172.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Grollemund, Rebecca, Simon Branford, Jean-Marie Hombert & John R. Watters. Forthcoming. Towards a new phylogenetic classification of bantoid languages. In John R. Watters (ed.), *East Benue-Congo: Bantoid classification, Mambiloid, Cross River* (Niger-Congo Comparative Studies). Berlin: Language Science Press.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1968. Notes on Nzébi (Gabon). *Journal of African Languages* 7(2). 101–129.
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Harro, Gretchen & Nancy R. Haynes. 1991. *Grammar sketch of Yemba*. Yaoundé: SIL Cameroon.
- Hedinger, Robert. 1985. The verb in Akɔ̀ɔ̀se. *Studies in African Linguistics* 16(1). 1–55.
- Hedinger, Robert. 1989. Northern Bantoid. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 421–429. Lanham, MD: University Press of America.
- Hedinger, Robert. 2008. *A grammar of Akɔ̀ɔ̀se: A Northwest Bantu language* (SIL International Publications in Linguistics 143). Dallas, TX: SIL International.
- Hedinger, Robert, Joseph Ekandjoum & Sylvia Hedinger. 1981. *Petite grammaire de la langue mboó*. Yaoundé: Association des étudiants mboó de l'Université de Yaoundé.

- Heine, Bernd, Tom Güldemann, Christa Kilian-Hatz, Donald A. Lessau, Heinz Roberg, Mathias Schladt & Thomas Stolz. 1993. *Conceptual shift: A lexicon of grammaticalization processes in African languages* (Afrikanistische Arbeitspapiere 34-35). Cologne: Institut für Afrikanistik.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Hewson, John & Vit Bubenik (eds.). 1997. *Tense and aspect in Indo-European languages: Theory, typology, diachrony* (Current Issues in Linguistic Theory 145). Amsterdam: John Benjamins.
- Hyman, Larry M. 1980. Relative time reference in the Bamileke tense system. *Studies in African Linguistics* 11(2). 227–237.
- Hyman, Larry M. 1981. *Noni grammatical structure: With special reference to verb morphology* (Southern California Occasional Papers in Linguistics 9). Los Angeles, CA: Department of Linguistics, University of Southern California.
- Hyman, Larry M. 2004. How to become a ‘Kwa’ verb. *Journal of West African Languages* 30(2). 69–88.
- Hyman, Larry M. 2018. Common Bantoid verb extensions. In John R. Watters (ed.), *East Benue-Congo: Nouns, pronouns, and verbs* (Niger-Congo Comparative Studies 1), 173–198. Berlin: Language Science Press. DOI: 10.5281/zenodo.1314327.
- Lonfo, Etienne & Stephen C. Anderson. 2014. *Dictionnaire ngiemboon – français – anglais / Ngiemboon – French – English dictionary / tyè’tendù Ngiemboon – Fe-lansé – Ngilisé* (Kay Williamson Educational Foundation – African Languages Monographs 7). Cologne: Rüdiger Köppe.
- Lovegren, Jesse Stuart James. 2013. *Mungbam grammar*. New York, NY: University at Buffalo, State University of New York. (Doctoral dissertation).
- Magba, Elizabeth Ann. 1995. *The INFL constituent in the Mundani language*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Maho, Jouni F. 2009. *NUGL online: The online version of the new updated Guthrie list, a referential classification of the Bantu languages*. https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf.
- Marchal-Nasse, Colette. 1989. *De la phonologie à la morphologie du nzebi, langue bantoue (B52) du Gabon*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- McClearn, Greg L. 2014. *A sketch grammar of the Central Mfumte language*. Yaoundé: SIL Cameroon.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

- Mfonyam, Joseph Ngwa. 1989. *Tone in orthography: The case of Bafut and related languages*. Yaoundé: University of Yaoundé. (Doctoral dissertation).
- Nchare, Abdoulaye Laziz. 2012. *The grammar of Shupamem*. New York, NY: New York University. (Doctoral dissertation).
- Ndedje, René. 2013. *Phonologie, morphologie, syntaxe et classification du ndemli*. Yaoundé: University of Yaoundé I. (Doctoral dissertation).
- Nurse, Derek. 2007. Did the Proto-Bantu verb have a synthetic or an analytic structure? In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 239–256. London: Department of Linguistics, School of Oriental & African Studies.
- Nurse, Derek. 2008. *Tense and aspect in Bantu* (Oxford Linguistics). New York, NY: Oxford University Press.
- Nurse, Derek. 2019. *Data for tense and aspect systems in Bantu* (Documents on Social Sciences and Humanities). Tervuren: Royal Museum for Central Africa. <https://www.africamuseum.be/sites/default/files/media/docs/research/publications/rmca/online/documents-social-sciences-humanities/D.%20Nurse%20Data%20for%20tense%20and%20aspects%20def.pdf>.
- Nurse, Derek & Gérard Philippson. 2006. Common tense-aspect markers in Bantu. *Journal of African Languages and Linguistics* 27(2). 155–196.
- Nurse, Derek, Sarah Rose & John Hewson. 2016. *Tense and aspect in Niger-Congo* (Documents on Social Sciences and Humanities). Tervuren: Royal Museum for Central Africa. <https://www.africamuseum.be/sites/default/files/media/docs/research/publications/rmca/online/documents-social-sciences-humanities/tense-aspect-niger-congo.pdf>.
- Parker, Elizabeth Ann. 1991. Conditionals in Mundani. In Stephen C. Anderson & Bernard Comrie (eds.), *Tense and aspect in eight languages of Cameroon* (SIL International Publications in Linguistics 99), 165–187. Dallas, TX: Summer Institute of Linguistics & University of Texas at Arlington.
- Paterson, Rebecca. 2015. Narrative uses of the Ūt-Ma'in (Kainji) bare verb form. In Doris L. Payne & Shahar Shirtz (eds.), *Beyond aspect: The expression of discourse functions in African languages* (Typological Studies in Language 109), 219–248. Amsterdam: John Benjamins.
- Perrin, Mona. 1972. *Transitivity in Mambila*. Reading: University of Reading. (MA thesis).
- Piron, Pascale. 1995. Identification lexicostatistique des groupes bantoïdes stables. *Journal of West African Languages* 25(2). 3–39.
- Piron, Pascale. 1997. *Classification interne du groupe bantoïde* (LINCOM Studies in African Linguistics 11–12). Munich: LINCOM Europa.

- Satre, Scott A. 2004. *Structure of the verb phrase in Ngomba*. Yaoundé: SIL Cameroon.
- Schaefer, Ronald P. & Francis O. Egbokhare. 2021. Dual exponents of anteriority in Bantu and beyond. Paper presented at the 8th International Conference on Bantu Languages, University of Essex, June 2–4, 2021.
- Schaub, Willi. 1985. *Babungo* (Croom Helm Descriptive Grammars). London: Croom Helm.
- Sebasoni, Servilien. 1967. La préfinale du verbe bantou. *Africana Linguistica* 3. 122–135.
- Shultz, George. 1997. *Kom language grammar sketch, Part 1*. Yaoundé: SIL Cameroon.
- Sonkoue, Eliane Kamdem. 2019. Mengaka tense-aspect system. *Nordic Journal of African Studies* 28(3). 1–17.
- Sonkoue, Eliane Kamdem. 2020a. What grammar for Bamileke languages? A common grammar or a ‘library’ of grammars? *Journal of African Languages and Literatures* 1. 26–61.
- Sonkoue, Eliane Meli Epse Kamdem. 2020b. *Tense-aspect categories and standard negation in five Bamileke languages of Cameroon: A descriptive and comparative study*. Bayreuth: University of Bayreuth. (Doctoral dissertation).
- Stanley, Carol. 1991. *Description phonologique et morpho-syntaxique de la langue tikar (parlée au Cameroun)*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Thwing, Rhonda & John R. Watters. 1987. Focus in Vute. *Journal of African Languages and Linguistics* 9(2). 95–121.
- Voll, Rebecca Maria. 2017. *A grammar of Mundabli: A Bantoid (Yemne-Kimbi) language of Cameroon*. Leiden: Leiden University. (Doctoral dissertation).
- Watters, John R. 1980. Notes on Ngie. Unpublished ms.
- Watters, John R. 1981. *A phonology and morphology of Ejagham, with notes on dialect variation*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Watters, John R. 2003. The Grassfields languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 225–256. London: Routledge.
- Watters, John R. 2012. Towards the reconstruction of the tense-aspect-mood (TAM) system in early Bantoid, with particular attention to the category “tense”. Paper presented at the International Congress “Towards Proto-Niger-Congo: Comparison and Reconstruction”, LLACAN, Paris, 18–21 September 2012.

- Watters, John R. 2017. Tone in Western Ejagham (Etung): Lexical tone on verbs with the perfective constituent focus suffix. In Gratien Gualbert Atindogbé & Evelyn Fogwe Chibaka (eds.), *Proceedings of the 7th World Congress of African Linguistics, Buea, 17–21 August 2012*, 930–947. Bamenda: Langaa Research & Publishing Common Initiative Group.
- Watters, John R. 2018a. East Benue-Congo. In John R. Watters (ed.), *East Benue-Congo: Nouns, pronouns, and verbs* (Niger-Congo Comparative Studies 1), 1–25. Berlin: Language Science Press. DOI: [10.5281/zenodo.1314337](https://doi.org/10.5281/zenodo.1314337).
- Watters, John R. 2018b. Ejagham without tense: Historical implications for Proto-Bantoid. In Eugene Buckley, Thera M. Crane & Jeff Good (eds.), *Revealing structure: Papers in honor of Larry M. Hyman* (CSLI Lecture Notes 219), 252–265. Stanford, CA: Center for the Study of Language & Information.
- Watters, John R. 2018c. Proto-Bantu tense from a Benue-Congo perspective. Paper presented at the International Conference on “Reconstructing Proto-Bantu Grammar”, Ghent University, 19–23 November 2018.
- Watters, John R. 2019. Proto-Bantu tense from a Benue-Congo perspective. Paper presented at the 50th Annual Conference on African Linguistics, University of British Columbia, Vancouver, 22–25 May 2019.
- Watters, John R. & Jacqueline Leroy. 1989. Southern Bantoid. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa’s largest language family*, 430–449. Lanham, MD: University Press of America.
- Williamson, Kay & Roger M. Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*, 11–42. Cambridge: Cambridge University Press.

Chapter 4

Reconstructing the development of the Bantu final vowels

Jeff Good

University at Buffalo

So-called Final Vowel (FV) morphemes are an integral part of the verbal inflectional morphology of most Bantu languages, though relatively little attention has been paid to their historical development in the context of the overall Bantu verbal system. A small set of FVs is reconstructed as appearing at the end of verbs and, along with other morphemes, they play a role in encoding tense, aspect, mood, and polarity. This chapter reconsiders the reconstruction of the Bantu FV system, with the goal of arriving at a better understanding of what the situation may have been like in Proto-Bantu (PB) with respect to these morphemes and how a system of inflectional marking of this kind could have developed. Data is drawn from languages of the north-western Bantu area which have not previously been systematically examined with respect to reconstruction of the FVs. On the basis of data from these languages, it appears that the right edge of the Bantu verb was a more active site for the formation of new morphology than suggested by previous studies and that the standard reconstructions of the FVs may represent a simplification of a more complex PB situation.

1 Introduction

Relatively little attention has been paid to the historical development of so-called Final Vowels (FVs) in the context of the overall Bantu verbal system, although they are an integral part of verbal inflectional morphology in most present-day



Bantu languages.¹ The sentence in (1), from Chewa N31b, illustrates the use of two FVs. The main clause verb *fún* ‘want’ appears with the FV *-a*, which generally has a kind of default status in the verbal system, and the subordinate clause verb *b* ‘steal’ appears with the FV *-e*, which is associated with subjunctive contexts.

- (1) Chewa N31b (Mchombo 2004: 28)
- A-nyaní a-ku-fún-á kutí mi-kángó i-dzá-b-é mi-kánda.*
 2-baboon SP₂-PRS-want-FV that 4-lion SP₄-FUT-steal-SBJV 4-bead
 ‘The baboons want the lions to steal (at a future date) some beads.’

Meeussen (1967: 110) places these FV morphemes in the Final position of the Proto-Bantu (PB) verb form, which is treated as containing both morphemes consisting of a single vowel and longer morphemes, such as the Perfective **-ile*.² He proposes three possible FVs, **-a* (appearing “in most forms”), **-e* (“subjunctive”), and, tentatively, **-I* (“negative”), with no attempt to reconstruct their tones (but see Meeussen 1962; 2014 with regard to the tones of the subjunctive). The first two of these are still proposed as PB reconstructions in more recent work such as Nurse (2008: 261–262), who further reconstructs **-à* with a low tone and **-é* with a high tone, alongside a possible **-I*, associated with past tense encoding rather than negation, whose tone remains uncertain (Nurse 2008: 268). In the reconstructed verbal system, the Final position (see Figure 1) must be filled by some morpheme.³

The presence of FVs in Bantu raises a number of questions beyond their basic reconstruction that have yet to be answered or, in some cases, seriously considered (or reconsidered) from a historical perspective. For instance, what was their

¹In this chapter, single vowel morphemes associated with the Final position in the Bantu verb template proposed by Meeussen (1967: 108–111) (see Figure 1) will be referred to using the capitalised term Final Vowel (FV). This is in opposition to vowels that happen to appear at the end of a verb, which will be referred to using non-technical formulations as the “last vowel” of a stem or the “vowel at the end” of a verb. In some cases, these vowels may also be morphological FVs but are referred to in this way when the focus is the phonological form of the verb rather than its morphological structure. When referring to specific forms, a hyphen will precede vowels when they are being treated as morphemes (e.g. *-a*). Otherwise, they will appear without a hyphen (e.g. *a*). Language-specific and Bantu-specific morphological categories will be referred to using capitalised terms, while general linguistic categories will be referred to using all lower-case letters. Transcription conventions, including those for tone, generally follow those found in the original sources, with minor adjustments made in some cases as indicated.

²See Bastin (1983) for a detailed comparative study of the Perfective, and Nurse (2008: 266–267) for a more recent discussion. Its reconstruction as **-ile* follows Nurse (2008).

³Meeussen (1967) does not explicitly state that the Final slot is obligatory. However, this is implicit in the reconstructed verbal tense system, where all proposed forms are reconstructed with a Final suffix (Meeussen 1967: 111–114).

original source and how did they become an integral part of the expression of the tense, aspect, mood, and polarity marking system found in the Bantu verb? How did they develop their hybrid phonological and morphosyntactic function where they simultaneously play a role in ensuring that verbs adhere to a specific set of prosodic constraints while also taking part in the encoding of verbal semantics? What processes led to the development of the “default” FV *-a, whose semantics appears to be best defined in negative rather than positive terms?

The purpose of this chapter is to reconsider the reconstruction of the Bantu FV system, with the goal of arriving at a better understanding of what the situation may have been like in PB with respect to these morphemes and how a system of inflectional marking of this kind could have developed. The primary data to be examined will be drawn from a survey of languages of the north-western Bantu area, in particular zone A and to a lesser extent zone B, which do not appear to have previously been examined systematically with respect to understanding the origins of the FVs.⁴ While no definitive conclusions regarding the reconstruction of FVs will be reached here, the discussion is intended to serve as a guide for further research in this area. It will be argued, in particular, that the reconstruction of the Bantu FVs is much less straightforward than might be suggested by data from languages like Chewa, as seen in (1), which may very well represent a significant simplification of the PB situation. Instead, the picture that emerges from this study is one where the right edge of the verb is a more active site for the formation of new morphology than suggested by previous studies and where verb roots show greater diversity in their shape than what is implied by the standard reconstructions.

For purposes of reference, a schematic representation of the segmental structure of what will be referred to here as the canonical Bantu verb, based on the adaptation that Güldemann (2003: 184) gives of Meeussen (1967: 108–111), is presented in Figure 1. Verbal positions are assigned a number with respect to the verb root (which is numbered zero), a label, and a list of grammatical functions which morphemes in that slot typically encode. Parentheses indicate that a morpheme does not appear in that position in all tense, aspect, mood, and polarity configurations, while those positions appearing without parentheses are generally always occupied by an affix. Slots that can contain more than one element are noted with an asterisk. Forms in attested languages deviate from this canonical form in various ways, and north-western Bantu languages can be especially divergent (Nurse & Philippson 2003: 5). Nurse (2008: Chapter 6) provides a thorough dis-

⁴A full survey would require a more thorough consideration of zone B as well as zone C, though this was not possible for the present chapter due to time constraints.

cussion of issues surrounding the reconstruction of Bantu verbal structures; see also Nurse & Philippson (2006).

-4	-3	-2	-1	0	1	2	3
(Pre-Initial)	Initial	(Post-Initial)*	(Pre-Radical)*	Radical	(Pre-Final)*	Final	(Post-Final)
negation TAM [†] clause type [†]	subject	TAM negation clause type [†]	object	root	TAM valence	TAM	participant negation [†] clause type [†]

* = more than one element possible; † = local innovation

Figure 1: The Bantu verb template following Meeussen (1967: 108–111) and Güldemann (2003: 184)

In the present study, the most important slot in Figure 1 is slot 2, labelled Final, though the elements in slot 0 and slot 1 are also relevant due to their potential for interaction with FVs in slot 2. Those three slots together also comprise a unit often referred to as the verb stem in comparative Bantu studies, which can have properties that suggest it should be treated as a subconstituent of the larger Bantu verbal structure (e.g. Hyman 1993). Of particular relevance in the present context are two aspects of verb stem phonology. First, the verb stem is often the domain of prosodic phenomena which result in reduced vowel contrasts after the stem-initial vowel (e.g. Hyman 2003a: 45–47). Since these can impact FVs, they are obviously relevant to their reconstruction. Second, the verb stem has a canonical shape of CVC-(VC-)*-V. That is, it is based on a CVC root followed by one or more -VC suffixes and a FV (see Good 2016: 139–141 for an overview). While this is not an exceptionless pattern, it is dominant enough to be relevant for understanding the development of the Bantu verb, as will become clearer in the discussion below of languages like Nen A44 (§4.5) and Kpa A53 (§4.6), among others.

In §2, further information is provided regarding the comparative Bantu data that is relevant to this study, as well as a brief consideration of the connection between Bantu FVs and similar phenomena in other Niger-Congo languages. In §3, an overview of the key descriptive features covered in the survey that forms the core of this chapter is provided. In §4, data is presented on the stem-final verb morphology found in a sample of north-western Bantu languages. In §5 some conclusions are drawn regarding the possible PB FV system and its morphological development.

2 Final Vowels in Bantu and Niger-Congo

This study is based on a sample of fifteen languages which were chosen by reference to Guthrie's classificatory system (Guthrie 1948; 1971), focusing on languages of zones A and B. This sampling choice was made on the assumption that it would provide sufficient representation of FV patterns in north-western Bantu languages to allow for an informed reappraisal of what the PB situation may have been. This approach is in line with a treatment of PB as the reconstructed language associated roughly with node 1 of the phylogenetic tree presented in Grollemund et al. (2015). This essentially includes all languages traditionally classified as Narrow Bantu plus the Jarawan Bantu languages (see Gerhardt 1982 for discussion of this latter group). However, as will be discussed in §5, the results of this survey can be interpreted as suggesting that the reconstructed FV system proposed by Meeussen (1967: 110) may be better associated with a later stage corresponding roughly to their node 2. Bostoen & Guérois (2022 [this volume]) come to a similar conclusion for the long passive form **-ibv*, though they place the innovation of this form at node 3.

The Bantu FVs have not been the subject of extensive comparative study. The most recent work that is comparable in scope to the present chapter is Grégoire (1979), which focused on aspects of FV realisation in languages from the centre of the Bantu domain and specifically excluded the north-western Bantu languages of interest here. Grégoire (1979) did not explicitly consider the origins of the Bantu FV system. Nevertheless, its focus on patterns of FV alternation conditioned by factors other than tense, aspect, mood and polarity parallels this study's close examination of verb-final morphology that deviates from the reconstructed patterns. Unlike the present chapter, Grégoire's (1979) investigation was not designed to reach specific conclusions regarding the general evolution of the PB FV system. However, the fact that it considers the possibility that the system could have developed from a simplification of more complex verb-final morphological patterns through processes of phonological reduction and analogy (Grégoire 1979: 169–170) is in line with proposals that will be made below in §5.2. Grégoire (1979) also provides an overview of relevant work on the topic of the development of the FV to the point of publication (see also van Eeden 1934 for a consideration of earlier work).

Reconstruction of the verb complex in the larger Niger-Congo (NC) context is also an area that has yet to see extensive comparative study, with the partial exception of verbal extensions (Voeltz 1977; Hyman 2007), associated with

slot 1 in the schematic representation of the PB verb in Figure 1.⁵ The broadest and most up-to-date comparative work on this topic is Nurse et al. (2016), which builds on work presented in Nurse (2007). See also Nurse & Watters (2022 [this volume]). Nurse and his colleagues reconstruct a ‘Final Vowel’ category for NC and suggest that it “was originally used for a binary aspect contrast between perfective/factative and imperfective, both indicated by a single vowel” (Nurse et al. 2016: 21). They also tentatively suggest reconstructions of *-i for a vowel coding Factative and *-a for a vowel coding Imperfective. See Welmers (1973: 346–347) for a discussion of the factative category. Their survey appears to have been designed to look for the existence of potential Finals in NC without explicitly considering whether these were obligatory or formed a compact and highly grammaticalised paradigm, as is the case for the reconstructed Bantu FVs.

To pick a relevant example, Nurse et al. (2016: 30–31) treat Bijogo (Segerer 2000; 2002) as making use of two Bantu-like FVs with forms *-ε*, with factative function, and *-i*, with perfective function. Bijogo is classified as part of the Atlantic subgroup of NC, and its verb structure shows striking parallels to what is reconstructed for Bantu despite being quite distant from Bantu both geographically and genealogically (Segerer 2002: 262), which is why it is chosen for comparison here. Bijogo verb forms like those in (2) show FV alternations comparable to those reconstructed for PB.⁶

(2) Bijogo [Atlantic] (Segerer 2000: 226)

- a. *i-booʃi i-tonʃ-i*
I-dog I-jump-IPFV
‘Dogs are jumping.’
- b. *i-booʃi i-tonʃ-ε*
I-dog I-jump-PFV
‘Dogs jumped.’

The vowel alternations exemplified in (2) are reminiscent of the reconstructed Bantu pattern insofar as the last vowel of the verb changes depending on the tense and aspect of the verb. However, the overall Bijogo system differs from the reconstructed Bantu system in crucial ways. First, there is another suffix occupying the Final position, but which has a VC shape. Specifically, the Perfective can

⁵The term *extension* in a Niger-Congo context refers to a verbal suffix attaching to a verb root or stem which derives a new verb stem.

⁶Segerer (2000: 226) categorises the suffixes in (2) using the labels *inaccompli* and *accompli* for (2a) and (2b) respectively. I have translated them as imperfective and perfective following Nurse et al. (2016: 81).

also be formed using an *-ak* suffix rather than *-ε*. Verbs can also lack suffixes entirely in the Perfective, Imperfective, and Infinitive forms. The variants that are found on a given verb appear to be lexically specified, rather than predictable based on other factors.⁷

Based on the summary of evidence for reconstructing a FV in NC provided in Nurse et al. (2016: 30–31), it appears that there is justification for reconstructing verb-final morphemes consisting of a single vowel that encoded aspect (and, perhaps, other categories) at some high level of NC. However, the reconstruction of a Bantu-like system where FVs form a small grammatical paradigm and are obligatory on all verb forms does not appear reasonable for NC as a whole.

A related concern from the perspective of NC is the shape of verb roots in Proto-NC (PNC). Given that most Bantu verb roots are reconstructed as ending in a consonant (Meeussen 1967: 86–92), FVs have a noteworthy prosodic function, alongside their morphological function. Specifically, they allow surfacing verbs to satisfy phonotactic constraints on syllable structure. PB syllable structure was quite restricted, and codas were not allowed (Hyman 2003a: 43). For roots with CVC shape, which Meeussen (1967: 86) labels the “normal type”, the FV allows them to surface as CVCV, thereby satisfying syllable structure constraints.

There does not appear to be much work explicitly on the topic of PNC root structure that would help resolve which aspects of the Bantu situation are archaic and which are innovative, but see Pozdniakov (2016) for a relevant discussion. For instance, if we assume that PNC verb roots were predominantly CVC in shape, then we would model the development of the PB FVs as primarily involving processes through which vocalic morphemes already present in NC became obligatory. By contrast, if we assume PNC verb roots were primarily CVCV shape or could be either CVC or CVCV in shape, then we need to understand how lexical vowels at the ends of verb stems interacted with the development of the FVs.

In the context of the present survey, this is not a completely abstract concern given that some languages, for instance Eton A71 (see §4.8) and Gyeli A801 (see §4.9), show verb roots with CVCV shape where the last vowel is part of the lexical form of the verb root rather than representing a morphological FV. The historical source of these vowels is not obvious, and, if the PNC picture were clearer, this would likely be of value for understanding the origin of such patterns in north-western Bantu languages.⁸ While further consideration of verb stem structure

⁷With respect to the survey of north-western Bantu languages presented below, the Bijogo system is most similar to that of Kako A93 (see §4.11), which has a system of verb-final morphology that is quite distinct from the standard reconstructions.

⁸In Bantu branches from the north-western area such as West-Coastal Bantu, the last vowel of

and verb-final morphology in NC is outside the scope of the present study, future work on PNC verbal form is clearly important for arriving at a complete understanding of the development of the Bantu FVs.

3 Variation in stem-final morphology

3.1 Overview of major points of variation

In order to provide context for the survey presented in §4, a number of general observations about patterns found in the data are introduced in this section. In particular, significant points of variation that have been noted in this survey are: (i) the number of FVs described in a given language (§3.2); (ii) the nature of the morphosyntactic categories that are coded by the FVs (§3.3); (iii) the interaction of FVs with other kinds of verb suffixes, in particular extensions (i.e. verb-to-verb derivational suffixes) (§3.4); (iv) the presence of stems that are lexical exceptions to regular FV patterns (§3.5); and (v) the interaction between stem-final phonological processes and verb-final morphology, in particular stem-controlled vowel harmony and processes affecting environments where two consonants are underlyingly adjacent (§3.6).

The discussion in this section and in §4 looks at suffixes that play a role in encoding tense, aspect, mood, and polarity (TAMP), even in cases where the relevant morpheme is not vocalic, if relevant for understanding the development of FVs. It also looks at patterns in languages which may be indicative of earlier Final position verb morphology or of a proto-language stage before FVs had developed, even if there is no evidence for synchronically active FVs in a given language.

One point that will emerge from the discussion is that the nature of the FV systems found in the north-western Bantu languages suggests that the right edge of the Bantu verb may have been a more active area for the formation of new morphology than indicated in the literature, which has tended to focus on the creation of new morphology in the Post-Initial slot (Güldemann 2003: 185). The issue of whether this is due to processes that have specifically affected north-western Bantu languages or somehow revealing of the situation in PB will be discussed in §5.

infinitive verb forms with a synchronic CVCV shape is often a relic of a former *-VC verbal derivational suffix which underwent phonological erosion, metathesis, and/or was the target of phonological mergers (Guthrie 1967: 60; Rottland 1970; Bostoen & Mundeke 2011; Pacchiarotti & Bostoen 2021: 445). For example, BLR (= Bantu Lexical Reconstructions 3, Bastin et al. 2002) 662 **cón+ík* 'draw a line, write + stative suffix' > Nzadi B865 *ò-sónkà* 'write' (Crane et al. 2011), and BLR 3354 **jímá* 'stand, stop' > Ngwi B861 *yímá* 'be standing' (Sara Pacchiarotti, p.c.).

3.2 Number of Final Vowels

An important point of variation found in the surveyed languages is the number of FVs found within a given language. As will be seen in §4 and as discussed below in §3.6, it is not always straightforward to determine what morphemes should be classified as FVs in systems where the FV position is not as strongly grammaticalised as it is in languages adhering closely to canonical Bantu verb structure, as illustrated with Chewa in (1). However, even in languages where this problem does not arise, there is still significant variation from a comparative perspective. See also Nurse (2008: 47–50) for a relevant overview discussion.

In the present survey, the language with the largest fairly clear-cut system of FVs is Kpe A22 (see §4.3). It shows four segmental forms: *a*, *e*, *i*, and *ɛ*. The first three FVs appear on affirmative main clause verbs in different TAMP configurations, while the last is found only on Past Negative forms in relative clauses and content questions. The largest system of verb-final TAMP-encoding suffixes in the survey, including both vocalic suffixes and other suffixal shapes, is found in Kako A93 (see §4.11), which shows six such elements.

By contrast, there are also languages in the survey which show no evidence for FVs synchronically (even if verb stems may still end in a vowel), namely Yasa A33a (§4.4), Eton A71 (§4.8), Gyeli A801 (§4.9), and, possibly, Makaa A83 (§4.10). These languages are all found in a region encompassing the southern part of Cameroon and northern Equatorial Guinea, suggesting a possible areal pattern, though verifying this would require a more systematic survey of this specific region.⁹ Variation in the inventory of FVs raises important historical questions with respect to the PB situation and how it resulted in these diverging FV patterns. Perhaps the most interesting question is whether the current PB reconstructions, as presented in §1, represent a historical reduction of a more complex system in only one part of the family well after the initial Bantu divergence (see §5.2).

FVs can, of course, appear with different tones depending on the semantics of the verb forms in which they appear. Tone is not considered in any detail here, though it is clear that a full examination of the development of FVs will need to take tone into account. An important question regarding FVs and tone is whether

⁹Nurse (2008: 47) points out that some zone L languages also show FV loss. Pacchiarotti & Bostoen (2021) provide a detailed discussion of an areal pattern of vowel loss at the end of stems found in B80 languages and adjacent groups, while also referencing similar patterns of loss found in other north-western Bantu languages. While this kind of loss does not specifically target the inflectional FVs that are the focus of this chapter, it can obviously impact them due to their stem-final position.

the appearance of different tones on FVs with the same segmental shape should be taken to suggest that there was a historical collapse of oppositions among formerly distinct morphemes.¹⁰ Of particular note from a historical perspective regarding tone and FVs is a general pattern seen across Bantu languages where the tone on non-initial syllables (or, in some cases, moras) in a verb stem is conditioned by the tone of the FV (cf. e.g. Meeussen 1961; Odden & Bickmore 2014: 4).

3.3 Categories encoded by Final Vowels

It can be difficult to assign a clear-cut meaning to specific FVs. However, it is easier to characterise the kinds of categories that they encode, most typically in combination with other verbal markers such as morphemes appearing in the Post-Initial position in Figure 1 (in addition to a tonal melody). Nurse (2008: 42–46) provides an overview of common categories coded on verbs in Bantu languages in general, and specifically describes FVs as having an important role in coding aspect, mood, and tense. In languages where a FV is obligatory, it necessarily plays a role in encoding other kinds of verbal categories even if only appearing in a “default” form of some kind, most typically as *-a* (though see §3.6 for a discussion of harmonising FVs).

In addition, in some Bantu languages, including two of those discussed in §4 – namely, Nen and Kako – FVs have been described as having another, somewhat peculiar, function of encoding whether or not a stem is extended. That is, their form can be conditioned by whether or not a verb stem is longer than two syllables, most typically due to the fact that it appears with a verbal extension (i.e. a verb-to-verb derivational suffix; see Schadeberg & Bostoen 2019), in the Pre-Final position in Figure 1. The presence of these extensions is not necessarily always synchronically transparent, and monosyllabic stems with long vowels can also behave like two-syllable stems with respect to this pattern. Grégoire (1979: 142–143) discusses this phenomenon and describes relevant patterns in a number of Bantu languages outside the north-western area. As the survey presented in this chapter makes clear, it is also found in north-western languages and, in at least one language, Kako, the specific pattern involves a complex interplay between length and the final consonant of a stem.¹¹

¹⁰ Also noteworthy in this context are observations by Grégoire (1979) of languages where the choice of FV can be partly conditioned by the tone of the root synchronically, as in Luba-Katanga L33 (Grégoire 1979: 143), or historically, as in Herero R30 (Grégoire 1979: 157–158).

¹¹ Another north-western Bantu language showing this kind of pattern that is not looked at in detail here is Tiene B81 (Ellington 1977), where the FV is partly phonologically conditioned, partly codes past tense, and partly codes that the stem is extended.

3.4 The interaction between Final Vowels and other suffixal morphemes

Another factor relevant to the realisation of FVs involves their potential interaction with other verbal suffixes. As discussed just above in §3.3, the form of the FV can be impacted by the presence of verbal extensions, and this represents one kind of interaction with suffixes. A different kind of interaction involves cases where the form of a suffix that would be expected to precede a FV affects the FV's realisation on the verb.

A pattern of this kind is found, for instance, in Punu B43 (see §4.15). This language makes extensive use of a default FV *-a*, but the *-a* fails to appear in forms ending in *u*, which include Passive verbs, as seen in a pair such as *lab-a* 'see' vs. *lab-u* 'be seen'. In Bantu languages where verbs adhere more closely to the form depicted in Figure 1, one might expect, instead, a form like *lab-w-a* for the Passive of such a verb, where the appearance of a FV results in the vowel of the preceding suffix becoming a glide.¹²

A difficulty in interpreting such patterns is determining whether these alternations in the realisation of FVs should be analysed as morphologically or phonologically conditioned, an issue that also arises with respect to patterns discussed in §3.3, when alternations are connected to stem length. In particular, are alternations like the one found in Punu morphologically conditioned by the presence of a specific suffix (e.g. a Passive), or should they be seen as a generalisation over stems ending in a specific sound which happens to be the same as the sound of a vocalic suffix (e.g. a *u*)? In Punu, synchronically, the phonological analysis captures the facts of language better, as will be described in §4.15. However, since stems ending in *u* which are not passives do not appear to be common, there is considerable overlap in the predictions of the morphological and phonological analyses, which is why the possibility of morphological conditioning is raised as an issue here. Moreover, given that studies such as Hyman (2003b) and Good (2007) have demonstrated that morphophonological analogy has impacted Bantu verbal morphology, one cannot rule out the possibility that a synchronic phonological pattern has its roots in a morphological generalisation.¹³

¹²van Eeden (1934: 372) argues that FV **-a* must be a relatively recent development due to the nature of its interaction with the vowels found at the end of roots with shape CV. While this is not a generally held view, it emphasises the potential interest of closely examining the interaction between FVs and other elements found towards the end of the verb stem.

¹³Grégoire (1979: 159) describes the case of Kwangali K33 where a class of CVC stems behaves as if they are longer with respect to their choice of FV (see §3.3 for the relevant discussion). These stems end in sounds such as *s* or *z*, which are historically linked to the presence of a Causative **-i* in the stem. This suggests how a generalisation initially tied to morphological structure could be reanalysed as being phonologically conditioned.

In addition to Punu, other languages discussed in §4 where the realisation of FVs is crucially dependent on stem-final morphology are Makaa A83 (§4.10), Myene B11 (§4.12), and Himba B302 (§4.14).

3.5 Lexical exceptions

In some languages of the survey, certain verb roots are described as ending in an invariant vowel and constituting lexical exceptions to regular FV rules. Subminimal roots of shape CV seem especially likely to be exceptional in this way, and this is not limited to the north-western Bantu area. In her study of FV patterns, Grégoire (1979) notes:

Dans le présent article, nous n'avons pas envisagé le cas des radicaux qui sont, en synchronie, de type -CV- ou de type -C- : leur comportement est toujours spécial et mériterait une étude distincte. (Grégoire 1979: 146)

In the present article, we have not considered the case of radicals which are synchronically of type -CV- or of type -C-; their behaviour is always exceptional and merits a separate study. (my translation)

Here, where available, information on such roots will be presented in some cases, especially if they are described as presenting behaviour which is distinct from that of longer roots.

Examples of languages in this survey showing lexical exceptionality in FV patterns are Kpe, Kako, Myene, and Punu. Patterns of lexical exceptionality can overlap with cases of interaction between FVs and other suffixes, as discussed in §3.4, when the exceptionality is connected to specific root-final vowels which have the same form as a verbal suffix, such as *u* in the Punu case.

3.6 Phonological patterns that impact Final Vowels

A final issue that arises with respect to FVs in the surveyed languages is the role of phonological factors impacting the realisation of stem-final morphology, including FVs. Of particular importance are patterns of reduction that result in surface realisations of stem-final morphology where more complex underlying patterns are neutralised. These are sometimes suggestive of possible diachronic sources for FVs from more heterogeneous sets of morphemes.

For instance, as will be discussed in §4.11, in Kako verbs can appear with a past tense suffix *-má*. This suffix appears at the end of the verb where a FV might be expected. This is seen, for instance, in the verb form *βεη-má* 'see-PST'. However, in

verbs ending with certain consonants sequences, including *dy*, the *m* is dropped, producing a surface form like *kwady-á* ‘love-PST’. The form *á* is otherwise associated with a suffix marking Imperative verbs. This suggests a possible pathway for the development of the unusual “default” distribution of the FV **-a*. Two etymologically distinct suffixes may merge in specific phonological contexts, and this merged morphological form could, at a later stage, be generalised across all verbs. This will be discussed further in §5.2.

A more widespread phonological pattern is the presence of vowel harmony affecting FVs. Outside of the north-western area, this is a significant topic in the discussion of Grégoire (1979). It is also relevant in the north-western area as evidenced, for instance, by the FV patterns found in Gunu A622 (§4.7), a language which shows evidence for only a single FV whose form is predictable by vowel harmony rules. This is likely to represent a case where patterns of sound change resulted in the merger of distinctions between FVs which were historically morphologically separate.

In addition to Kako and Gunu, other languages of the survey where phonological patterns are relevant for understanding their FV systems are Akoose A15C (§4.2), Nen A44 (§4.5), Kpa A53 (§4.6), Eton A71 (§4.8), Gyeli A801 (§4.9) and Kota B25 (§4.13).

4 Survey of Final Vowel patterns in north-western Bantu languages

4.1 Introduction to the survey

As a step towards the reconstruction of the development of the Bantu FVs, this section reports on a survey of verb-final morphological patterns, with an emphasis on verb-final morphology that could be classified as a FV or serve as a possible historical source for a FV. A sample of languages across two of the Guthrie zones standardly associated with north-western Bantu (i.e. zones A and B) was examined. Nurse (2008) was used as an initial guide in the selection of these languages, and other sources were located as needed with the goal of having one language for each of the nine high-level subdivisions of zone A and the lower-numbered subdivisions of zone B. Two languages are discussed below from group A80. While the availability of a detailed description for one A80 language, i.e. Gyeli (see §4.9), made it ideal from the perspective of a survey like the one presented here, another language, Makaa (see §4.10), was found to show an interesting pattern involving the realisation of an apparent reflex of a FV under particular tonal and syntactic conditions which seemed important to include in the discussion.

In some cases, the choice of a language within one of these subdivisions was relatively opportunistic by virtue of being based on a source that was readily available. In other cases, the choice was more or less dictated by the lack of any other appropriate and available source for the languages of that subdivision. While it seems likely that the FV patterns found in the survey provide a reasonable sense of the overall north-western Bantu picture, it is almost certainly also the case that important comparative evidence could be uncovered by examining languages not considered here (see also §5.2). Moreover, as indicated in §2, the languages that were surveyed were chosen by reference to Guthrie's referential classificatory system rather than any particular genealogical proposal for the internal structure of the Bantu family, such as the one derived from phylogenetic analyses found in Grollemund et al. (2015). Future work in this area would likely benefit from the development of an expanded sample that takes such proposals into account, for instance, by the inclusion of Jarawan Bantu languages or by targeted sampling across groups which are not believed to form low-level genealogical units.

At the same time, it should be noted that this approach resulted in a sample that is relatively genealogically diverse in the context of north-western Bantu. Akoose A15C, Kpe A22, Yasa A33a and Eton A71, on the one hand, and Nen A44, Kpa A53 and Gunu A622, on the other, are part of two distinct groupings, under Grollemund et al.'s (2015) node 1. Makaa A83 and Kako A93 are part of another grouping placed under their node 2.¹⁴ Kota B25 is part of a grouping under node 3, while Myene B11 and Himba B302 are placed in a low-level group under node 4. Finally, Punu B43 and Nzebi B52 are both placed under Grollemund et al.'s (2015) node 6, a group also known as West-Coastal Bantu (see Vansina 1995; de Schryver et al. 2015; Pacchiarotti et al. 2019).

In the rest of this section, the basic descriptive facts of the FV patterns across the fifteen surveyed languages from zones A and B will be presented, going from lower Guthrie numbers to higher numbers within each of the two zones. In §4.17, the overall results of the survey are summarised.

4.2 Akoose A15C

In Table 1, the basic TAMP forms of verbs in Akoose A15C are presented in schematic form, with “...” representing where the verb stem appears. The asso-

¹⁴One of the surveyed languages, Gyeli A801, is not included Grollemund et al.'s (2015) study. However, two languages that are reported to be closely related to Gyeli by Grimm (2015: 108), Shiwa A803 and Kwasio A81, are part of that study. Both are placed within the same group as Makaa A83 and Kako A93 in Grollemund et al. (2015).

4 Reconstructing the development of the Bantu final vowels

ciated surfacing forms of the verb for ‘wash’, coded with a third singular (class 1) subject marker appearing before the TAMP markers, are presented in Table 2 (Hedinger 2008: 100–101).¹⁵

The Akoose system can be described with reference to two verbal aspects, Perfective and Imperfective, coded via suffixes, and two tenses, Past and Future, coded via prefixes. Both Affirmative and Negative verb forms are included. Verb forms not appearing with prefixes in the Affirmative forms or only a single *e*-prefix in Negative forms can have a Perfect or Present Imperfective interpretation, to which I have applied the label *Factative* here, adapting the term as used by Welmers (1973: 346–347).¹⁶

Table 1: Schematic representation of verbal forms in Akoose

	PFV	IPFV	PFV.NEG	IPFV.NEG
Past	\tilde{N} -...-’	...-é’-áá	<i>e-nkêN</i> -...-é	<i>e</i> -...- <i>e</i> -áá
Factative	...-é	...-é’	<i>e</i> -...- <i>e</i> -’é	<i>e</i> -...-é’-’é
Future	â-...-’	â-...-é’	^ <i>e</i> -â-...-’-’é	^ <i>e</i> -â-...-é’-’é

Table 2: Third singular subject forms of verb ‘wash’ in Akoose

	PFV	IPFV	PFV.NEG	IPFV.NEG
Past	<i>anwóg</i>	<i>awógáá</i>	<i>enkênwógké</i>	<i>ewógáá</i>
Factative	<i>awógé</i>	<i>awógé’</i>	<i>ewógéé</i>	<i>ewógéé</i>
Future	ǎ’ <i>wóg</i>	ǎ’ <i>wógé’</i>	ê <i>wógké</i>	ê <i>wógéé</i>

¹⁵The schematic representation of the Future Imperfective in Hedinger (2008: 100) represents the prefix as *dâ*-. The *d* appears to be a typographic error and, therefore, is not included here (see Hedinger 1985: 33). In the Negative forms, the appearance of the *e*-prefix overrides the appearance of the third person subject prefix *â* due to a vowel deletion rule (Hedinger 1985: 20). The presentation in Table 1 follows the source where an apostrophe is used to represent a glottal stop. The relationship between the proposed underlying forms for Akoose inflected verbs in Table 1 and surface forms is not always straightforward, and the reader is referred to Hedinger (2008) for a full treatment.

¹⁶Welmers (1973) uses this label to refer to verbal constructions associated with past tense semantics when applied to verbs expressing events, and present tense semantics when applied to verbs expressing states. The use of that label here extends the term to apply to constructions where the same verb root is used but its temporal reference is connected to its aspectual encoding.

Various features of the Akoose FV system stand out from the forms presented in Table 1 and Table 2. For instance, the Past and Future Perfective forms lack a FV entirely. Second, phonological simplification at the right edge of the verb results in complex underlying patterns surfacing in ways that adhere more closely to the canonical Bantu verb form than might be expected from their morphological composition.¹⁷ This is seen most directly in the Past Imperfective form which appears with a final sequence of *-áá* and where the Imperfective *é'* does not surface at all.¹⁸ To a lesser degree, it can also be seen in the simplification of the *-e'-é* and *-é'-'é* sequences to *éé* and *éé*, respectively, in the Factative Perfective Negative, Factative Imperfective Negative, and Future Imperfective Negative forms.

While the forms in Table 2 do not make clear what the source is for the abstract analyses presented in Table 1, evidence for them can be found in dialectal variants as well as in verb forms whose stems do not have canonical CVC shape (Hedinger 2008: 101, fn. 9). For instance, in the Past Imperfective forms of CVV stems, the glottal stop of the Imperfective suffix is found, as seen in a form like *abóó'áá* 'it was breaking', based on the root *bóó* (Hedinger 2008: 123). Moreover, the analysis in Table 1 abstracts away from at least one complication of clear comparative interest, namely the fact that the Imperfective suffix in stems with CV shape has the form *-ag*, as seen in a verb such as *a-dy-ág-áá* 'he was eating' based on a root *dyé* 'eat' (Hedinger 2008: 123). This form is readily identifiable with a reconstructed form **-ag* generally associated with imperfective semantics (Sebasoni 1967; Nurse 2008: 262–264) and suggests that Akoose verbs have been affected by processes of phonological reduction in the portion of the verb stem between the initial CVC sequence and the FV. This portion of the verb stem is identified as the prosodic trough in Hyman (1998), i.e. a domain characterised by reduced possibilities of phonological contrast in comparison to other parts of the verb stem. In this case, such processes appear to have resulted in **-ag* developing into *-'é*, though the details of such a process remain to be worked out. In CV stems, the suffix would have been protected from such effects due to the fact that their short lexical forms would allow **-ag* to appear before the trough position. From a general diachronic perspective, these Akoose patterns suggest that new FV patterns can arise due to phonological reduction affecting the end of the stem (see §3.6).

¹⁷In the Perfective Negative form *enkênwógké* in Table 2 a *k* appears before the last vowel due to a process where a glottal stop appearing immediately after a consonant is partly assimilated to the preceding consonant.

¹⁸Hedinger (2008: 6) states that glottal stops are frequently dropped between vowels, connecting that aspect of this reduction to more general processes in Akoose.

Another aspect of the Akoose FV system of interest here is the specific surface segments seen at the end of verbs, namely \emptyset , *e*, *aa*, ε , and ε' (see Table 1). The fact that they appear with different lengths and tone patterns and that one of these ends in a glottal stop means that they should not be directly equated with the reconstructed FVs of Meeussen (1967: 110). They do, however, potentially point to the kinds of historical processes of reduction and fusion at the end of the verb stem that could have resulted in the canonical pattern. If we assume that the verbal template in Figure 1 represents an earlier stage of Akoose, then present-day Akoose would appear to be a language where a new FV system is emerging as the earlier one is breaking down. Alternatively, we could treat Akoose as representing a branch of Bantu where the template in Figure 1 never developed in the first place (see also Güldemann (2022 [this volume])). The results of this survey do not clearly indicate which analysis is to be preferred, but this is clearly an important issue for PB reconstruction (see §5.2).

4.3 Kpe A22

In Table 3, the major verbal patterns of Kpe A22, also known as Bakweri or Mokpwe, are presented following the description of Marlo & Odden (2007: 20).¹⁹ Three segmental forms of the FV are found in Table 3: *a*, *e*, and *i*. The overall verb structure largely follows a canonical pattern, and the forms of these vowels are in line with the reconstructions proposed by Nurse (2008: 268) (see §1). The different FVs are not associated with straightforward semantics, but the *-a* FV appears to fulfil the expected default function. The column *SPT* in Table 3 indicates if a high tone appears on the subject prefix of the verb (in these examples having segmental shape *na*). The ' is used to represent downstep.

Marlo & Odden (2007: 21) describe two deviations from the system exemplified in Table 3. First, monosyllabic stems and stems longer than CVC often have final vowels that do not vary. Relevant examples are provided in Table 4, which compares the monosyllabic root *và* 'divide' and the trisyllabic root *lakízé* 'forgive', with the CVC root *zoz* 'wash'.²⁰ These patterns are placed here under the

¹⁹Marlo & Odden (2007: 20) discuss two Perfective forms in Kpe, simply labelling them Perfective1 and Perfective2, and they do not appear to discuss how they are semantically or functionally distinct.

²⁰Marlo & Odden (2007: 21) do not indicate the source of the long vowel in the verb 'divide' in Table 4, which differs from the short vowel they present in the citation form. It could presumably be due to the morphological presence of a FV that has assimilated to the vowel of the root or due to a lengthening effect connected to a minimality constraint of some kind (cf. e.g. Downing 2006: 54–55 and Hyman 2008 for discussions of minimality constraints in Bantu).

Table 3: Forms of verb ‘wash’ in Kpe

Verb	Translation	Category	SPT	TAMP	FV
<i>na-Ø-zoz-a</i>	‘I will wash’	FUT	Ø	-Ø-	-à
<i>na-zí-zoz-e</i>	‘I did not wash’	PST NEG	Ø	-zí-	-è
<i>na-zá-zoz-á</i>	‘I will not wash’	FUT NEG	Ø	-zá` -	-á
<i>na-ma-zoz-á</i>	‘I washed’	PST	Ø	-ma-	-á
<i>na-zí-zóz-í</i>	‘I have not washed yet’	INC	Ø	-zí	-í
<i>na-Ø-zoz-î</i>	‘I have washed’	PFV1	Ø	-Ø-	-î
<i>ná'-má-zoz-a</i>	‘I have washed’	PFV2	H	-` má`	-a

heading of lexical exceptions discussed in §3.5. It is not obvious how to interpret such stems in historical terms. Could the lack of distinct segmental FVs in forms based on roots like *và* and *lakízá* be conservative and representative of a historical stage where FVs were not obligatory? Or, could they be innovative and have resulted from FVs having been lost in some contexts that are yet to be determined? These questions will be considered again in §5.2.²¹

Table 4: FV variation by stem type in Kpe

Infinitive	Gloss	PFV1	Gloss
<i>li-zoz-á</i>	‘to wash’	<i>na-Ø-zoz-î</i>	‘I have washed’
<i>li-vaá</i>	‘to divide’	<i>na-Ø-vaâ</i>	‘I have divided’
<i>li-lakízá</i>	‘to forgive’	<i>na-Ø-lakízê</i>	‘I have forgiven’

A final important aspect of the Kpe system is that the verbal encoding of relative clauses and content questions involving Past Negative forms makes use of a fourth final vowel with the segmental shape *-ε*. Examples are provided in (3).²² This pattern raises two questions. First, what is the historical source of this vowel? Given the fact that it appears with a complex (rising) tone, the most likely

²¹The *ize* ending of *lakízá* ‘forgive’ in Table 4 is formally identical to the Causative suffix in the language, as described by Atindogbé (2013: 100), who transcribes it as *-izre*. This suggests that the presence of an invariant FV in this stem may be comparable to cases found in other surveyed languages involving the interaction between verb-final morphology and FVs, as discussed in §3.4, at least in historical terms.

²²The glossing in the examples in (3) is my own on the basis of information in Hawkinson (1986: 243), Marlo & Odden (2007), and Atindogbé (2013), as well as Michael R. Marlo (p.c.).

possibility would seem to be that it represents a fusion of two formerly morphologically distinct vowels, in a manner comparable to the morphological fusions seen in Akoose forms like the Factative Perfective Negative, as exemplified in Table 2 in §4.2. If that was the case, then this suggests that the Final position of the verb has been an active site of morphological formation in Kpe. Second, the presence of this fourth vowel opens up the question of just how large the FV inventory can be in Bantu languages. I am not aware of any work systematically exploring this topic, though Kpe represents the upper limit of verb-final TAMP-encoding morphemes with a V shape in this survey.

(3) Kpe A22 (Marlo & Odden 2007: 25)

- a. *emó a-zí-mo-zoz-ě*
 1.PRO SP₁-PST.NEG-OP₁-wash-FV
 ‘the one who didn’t wash him’
- b. *njé 'a-zí-zoz-ě*
 who SP₁-PST.NEG-wash-FV
 ‘Who didn’t he wash?’

On the whole, the Kpe FV system largely follows the canonical pattern in that most verbs appears with one of three FVs with similar forms to the FVs presented in Meeussen (1967: 110), though the ways that it deviates from that pattern suggests interesting historical possibilities regarding when FVs became obligatory and how new FVs may have developed. From an areal perspective, Kpe is somewhat unusual in this survey. In terms of its FV patterns, Kpe behaves more like languages of zone B, to be discussed below, than the other languages of zone A surveyed here, none of which show such canonical behaviour (see also §5.1). This is despite the fact that Kpe is associated with the south-west region of Cameroon and is geographically separated from zone B languages by other zone A languages surveyed here.

4.4 Yasa A33a

The Yasa A33a verbal system presents an example of a language where all verbs end in vowels, but where there is no evidence that they are morphologically independent. Instead, they appear to be part of the verb stem. Examples of verb stems, drawn from Bôt (2011: 90), are provided in Table 5. The vowels appearing at the end of a stem are restricted to *ɛ*, *ɔ*, or *a* in a seven-vowel system.

In Table 6, a number of Causative verb forms in Yasa are presented (Bôt 2011: 92), in Table 7, a number of Passive forms (Bôt 2011: 95), and, in Table 8, a number

Table 5: Yasa verb stems

Stem	Gloss
<i>lɔndɛ̀</i>	‘sew’
<i>bámɛ̀</i>	‘scold’
<i>péwà</i>	‘weigh’
<i>ànà</i>	‘fight’
<i>óbò</i>	‘fish’
<i>òkò</i>	‘curse’

of Reciprocal forms (Bôt 2011: 97).²³ What is important in this context is that these forms provide no evidence for the presence of a distinct FV morpheme.

Table 6: Yasa Causative stems

Stem	Gloss	CAUS	Gloss
<i>nɔ́ǎnɔ́ǎ</i>	‘work’	<i>nɔ́ǎnɔ́ǎjɛ̀</i>	‘make work’
<i>bétà</i>	‘show’	<i>bétíjɛ̀</i>	‘make show’
<i>làpà</i>	‘speak’	<i>làpíjɛ̀</i>	‘make speak’
<i>lùngà</i>	‘get angry’	<i>lùngùwɛ̀</i>	‘annoy’
<i>kúd^wà</i>	‘go out’	<i>kúd^wúwɛ̀</i>	‘bring out’
<i>wùpà</i>	‘become wet’	<i>wùpùwɛ̀</i>	‘make wet’

The Causative suffix is analysed as having a VCV shape where it invariably ends in ϵ , the initial V harmonises as i or u depending on whether the vowel preceding it is rounded, and the intervening C appears as j after i and as w after u . The presence of this suffix is associated with the loss of the last vowel of the stem. This can be accounted for via a general elision rule where the first vowel in a VV sequence arising as the result of morphological concatenation is deleted (Bôt 2011: 93). Due to the fact that Causative forms end in an invariant vowel, they provide no evidence for the presence of morphologically active final vowels in Yasa.

²³Based on the translations provided for the Passive forms in Yasa by Bôt (2011), the Yasa Passive appears to be used as a marker of both passive and middle verbs. Its form suggests it can be associated with the PB positional *-am (see Dom et al. 2016: 135–137 for a relevant comparative discussion).

4 Reconstructing the development of the Bantu final vowels

Table 7: Yasa Passive stems

Stem	Gloss	PASS	Gloss
<i>lòndè</i>	‘sew’	<i>lòndèmè</i>	‘be sewn’
<i>bàdè</i>	‘add’	<i>bàdèmè</i>	‘be added’
<i>tókólò</i>	‘collect’	<i>tókólòmò</i>	‘be collected’
<i>ò:kò</i>	‘slander’	<i>ò:kòmò</i>	‘be slandered’
<i>pútà</i>	‘hunt’	<i>pútàmà</i>	‘be hunted’
<i>ḡá</i>	‘smoke’	<i>ḡámà</i>	‘be smoked’

Table 8: Yasa Reciprocal stems

Stem	Gloss	RECP	Gloss
<i>búsèmè</i>	‘reunite’	<i>búsèmènè</i>	‘meet’
<i>vítèmè</i>	‘pursue’	<i>vítèmènè</i>	‘pursue each other’
<i>ò:kò</i>	‘slander’	<i>ò:kònò</i>	‘slander each other’
<i>tìlà</i>	‘write’	<i>tìlàrà</i>	‘write to each other’
<i>lówà</i>	‘insult’	<i>lówàrà</i>	‘insult each other’
<i>lúmà</i>	‘prick’	<i>lúmàrà</i>	‘prick each other’

The Passive and Reciprocal are both formed with the addition of CV syllables at the right edge of the verb, where the vowel fully harmonises with the preceding vowel. This pattern, too, does not provide any evidence for a morphologically active final vowel. It would be logically possible to analyse these stems as having a CVC-VC-V morphological structure, following what is reconstructed for PB extensions (see §1), with a vowel harmony rule affecting the non-initial vowels. However, there is no synchronic evidence for this in Yasa, making an analysis involving a suffix with CV shape the most straightforward one for this language. What is crucially lacking in Yasa, as compared to languages showing a more canonical pattern such as Kpe, just discussed in §4.3, are FV alternations that justify treating the last vowel of a verb stem as a distinct morpheme.

Tense forms for *nḡánḡà* ‘work’ and *tìlà* ‘write’ are presented in Table 9 (Bôt 1998).²⁴ The tense labels presented in Table 9 are not specifically found in Bôt

²⁴The F2 form presented in Bôt (1998: 55) for ‘work’ appears to be an error, where the P3 form was inadvertently repeated. This is why ‘work’ is replaced with ‘write’ for the F2 tense in Table 9.

(1998) but are used to reflect the fact that the different past and future tenses are characterised as encoding remoteness distinctions.

Table 9: Yasa tense forms for *nɕánɕà* ‘work’ and *tilà* ‘write’

Tense	Form	Coding
PRS	<i>nɕánɕá-ndí</i>	-’ <i>ndí</i>
P1	<i>nɕánɕà</i>	Ø
P2	<i>bé-nɕánɕà-kà</i>	<i>bé ...-KV̂</i>
P3	<i>nɕánɕ-é</i>	-é
F1	<i>mú-è-nɕánɕà</i>	<i>mú-è-</i>
F2	<i>nɕè-tilà-kà</i>	<i>nɕè-...-KV̂</i>

As is the case for data involving verb extensions, the tense forms of Yasa also do not provide evidence for a morphologically active FV. Instead, they can be analysed as simply taking TAMP suffixes appearing after a vowel-final verb stem. The only form which, superficially, appears to provide evidence for a FV is the P3 form, where the last vowel of the verb changes to *ɛ*. However, Bôt (1998: 51–52) analyses this as involving a process of VV reduction comparable to what was analysed for Causative forms in Table 6 above. Specifically, the P3 form, consisting solely of a vowel, unlike the other tense suffixes, triggers deletion of the preceding vowel. Moreover, in stems ending in *ɔ*, a glide epenthesis process is found where the P3 suffix is preceded by a *j* and the last vowel of the root is not deleted, as seen in a form such as *tɔkɔjé*, based on the root *tɔkɔ* ‘boil’. This further suggests that P3 -é is a suffix appearing at the end of a vowel-final verb stem rather than a FV that morphologically alternates with other FVs. Therefore, while it would clearly be possible to see these patterns as historically connected to the reconstructed FVs, there is no good reason to analyse them as FVs from a synchronic perspective.

A natural interpretation of the Yasa patterns is that former FV morphology became lexically incorporated into verb roots. A complication for such an analysis is determining the source of the specific vowels found at the end of verb forms such as those given in Table 5 above since they are not fully predictable. This issue will be discussed further in §5.2.

4.5 Nen A44

Following the description of Dugast (1971), Nen A44 is a language lacking a system of obligatory FVs (see also Mous 2003: 288). Dugast (1967) includes verbs which appear to lexically end in a vowel (e.g. *hékè* ‘remove’), though a casual inspection of this dictionary suggests that such monomorphemic verbs are not especially common. At least three extensions end in a vowel, a Direct Causative *-ì*, an Indirect Causative with allomorphs *-asi* and *-osi*, and a Neuter with allomorphs *-ε*, *-i*, *-o*, and *-u*. The use of these suffixes results in the appearance of numerous other vowel-final verbs (Dugast 1971: 167–168). These verb-final vowels cannot be readily associated with the reconstructed FVs. However, there is a class of verbs which appears with a partly harmonising vowel suffix after the root in a set of environments that Dugast (1971: 166) characterises as involving the encoding of the past tense or the imperative mood. Examples of verbs appearing with this suffix are provided in Table 10, in some cases alongside formally similar verbs that do not appear with this suffix in the relevant contexts and which are included for purposes of comparison (Dugast 1971: 230–231).²⁵

Table 10: Nen verbs taking vocalic suffixes

Root	PST	Gloss	Comparison
<i>liŋ</i>	<i>liŋə</i>	‘be angry’	<i>lêŋ</i> ‘be fat’
<i>bùl</i>	<i>bulə</i>	‘hunt’	
<i>hès</i>	<i>hesa</i>	‘tilt a container’	<i>ndés</i> ‘limp, be wobbly’
<i>lënd</i>	<i>lenda</i>	‘realise’	<i>lánd</i> ‘be desiccated, be dry’
<i>nək</i>	<i>nəkə</i>	‘break’	<i>kók</i> ‘be pure, guard against’
<i>tàŋ</i>	<i>təŋa</i>	‘be early, arrive first’	<i>nyáŋ</i> ‘be spicy’

What conditions whether or not a verb appears with this vowel in the relevant semantic environments is not clear (Dugast 1971: 229). It seems likely that there is

²⁵Dugast (1967) does not appear to give specific information about which verbs appear with this suffix, and Dugast (1971: 229–232) does not systematically present verbs which do not appear with it but provides examples of a number of stative verbs of this kind, since many of the verbs which do take the suffix are stative (though there are also non-stative verbs which appear with the suffix as made clear by the data in Table 10). Because of this, all of the comparison verbs are stative, even though the description makes clear that there are non-stative verbs which do not take the suffix. In Table 10, the suffixed form of the verbs is characterised as being associated with past tense contexts, following Dugast (1971: 230), though, as mentioned above, they appear to also be used in at least some imperative contexts.

at least some degree of lexical conditioning even if further analysis could partly account for which verbs appear with this vowel. Based on the examples provided in Dugast (1971), it seems that this vowel is only found on low-tone monosyllabic roots ending in a consonant, providing a possible parallel to developments in central Bantu languages noted by Grégoire (1979: 167). Mous (2003: 291–292) offers an additional discussion of these vowels including a historical analysis of them as reflexes of FVs which did not undergo processes of reduction affecting other FVs.²⁶

The unusual distribution of this harmonising vocalic suffix suggests that it is a relic of what was a once more productive FV system, since it is otherwise hard to envision a pathway through which such a system of marking would develop only on some verbs. A good candidate for the source of this vowel may be **-a*, given that one of the surfacing shapes of this suffix is *a* and the unusual semantic distribution of the suffix in past and imperative contexts, which fits the reconstruction of **-a* as a default FV. If some reflex of **-a* alternated with a FV like **-ɪ* in certain TAMP contexts, with **-a* retaining reflexes in Nen while **-ɪ* was lost entirely, this could have produced the two classes of verbs found today, i.e. one class appearing with the vocalic suffix and another class not appearing with it. These two reconstructed FVs are associated with past meaning, and there is also evidence that their appearance may have been conditioned by the properties of the stem that they attached to, for instance, whether or not it was extended (see Nurse 2008: 271–276, and §3.3). If this interpretation is correct, then Nen represents a case where phonological changes at the end of the verb are relevant for understanding the development of the FVs. Nen is also a case where the historical situation may have involved the use of FVs to encode unusual categories, such as a combination of tense and prosodic features of the stem like specific tone patterns and the presence of a verbal derivational suffix (see §3.3 and §3.6), assuming that factors such as these might have conditioned the source patterns for the split in Nen verbs seen today, where some appear with this vocalic suffix and others do not.

4.6 Kpa A53

The system of verb prefixes and suffixes found in Kpa A53, also known as Bafia, for encoding various TAMP functions in the affirmative is presented in Table 11 (based on Guarisma 2003: 320), where a plus sign indicates that a combination

²⁶Mous (2003: 292) also discusses vowels appearing at the end of verbs which are not sentence final in Nen. He treats these as being epenthetic due to the fact that their quality, tone, and appearance is predictable, unlike the other verb-final vowels discussed here.

4 Reconstructing the development of the Bantu final vowels

is possible and a minus sign that it is impossible. As indicated in Table 11, Kpa has a relatively developed system of verb prefixes, but only two segmental final suffixes. One of these, coding perfective semantics, is vocalic, and the other, labelled retrospective, codes something along the lines of anterior semantics and has a CV shape. The third ‘suffix’ is tonal in nature, classified as an instance of metatony by Guarisma (2003: 319).²⁷ It involves a high tone appearing after the root when the verb is not phrase final. As can be seen in Table 11, there are also verb forms lacking any suffix. Guarisma (2003: 319) places the prefixes into two sets on the basis of the divergence found in the prefixes that they can combine with.

Table 11: Kpa TAMP encoding

			Suffixes			
			-∅	-í	metatony	-yà
			NTR	PFV	IPFV	RET
Set 1	∅-	NTR	+	+	-	+
	á-	P0	-	+	-	-
	ń-	P1	-	+	-	-
Set 2	ê-	PROG	-	-	+	-
	rí-	HYPO	-	-	+	-
	má-	INCEP	-	-	+	+
	ká-	ANT	-	(+)	+	+
	kí-	CONC	-	(+)	+	+

Examples of the use of the two segmental suffixes are provided in (4) for the Perfective suffix and (5) for the Retrospective suffix. Where the relationship between the surfacing verb form and the underlying morphological pattern is obscured by phonological processes, an underlying representation of the verb is provided. Of note here are phonological processes affecting verbs which produce surface patterns that are fairly distinct from the underlying patterns. In (4b), a deletion process results in the surfacing form *yéé*, where a long vowel, in effect,

²⁷Metatony is a term used to describe phenomena in specific Bantu languages where, in certain TAMP configurations, verbs appear with high tones in syllables following the root if they are not phrase final. See Guarisma (2003: 320) for her specific use of this term as it applies to Kpa, and Hyman (2017: 108–112) for a more general discussion of metatony in Bantu.

marks the Perfective. In (5b), a process of coalescence results in a single consonant appearing when a consonant-final root is followed by a consonant-initial suffix. Notably, the resulting form *tékà* has a shape that formally matches the canonical CVCV form of a CVC root followed by a FV.

(4) Kpa A53 (Guarisma 1992: 218, resp. Guarisma 2003: 321)

- a. *à-kós-í* *bì-dùn*
SP₁-gather-PFV 8-fruit
'He gathered the fruits.'
- b. *ɓà-ɣéé* *zòʔ* /-ɣén-í/
SP₂-see.PFV 9.elephant
'They saw an elephant.'

(5) Kpa A53 (Guarisma 1992: 219, resp. Guarisma 2003: 321)

- a. *à-tékà* *mòní* /à-téʔ-ɣà/
SP₁-take.RET 6.money
'He had taken the money.'
- b. *bèl* *ì-bá-ɣà* *rì* *ɓón* *bíí* *ɓá'ráá*
9.ancestor SP₉-be-RET with 2.child 2.3SG.POSS 2.three
'God had three children ...'

While it would be premature to come to strong generalisations based on only a few forms, Kpa is a language where phonological processes appearing at the right edge of the verb stem create surface patterns which are suggestive of possible processes for the development of the Bantu FVs from a system involving suffixing morphemes with more diverse shapes, as most clearly illustrated by the verb form in (5b) (see §3.6). The *ɣéé* verb form in (4b) is also indicative of how assimilatory processes can reduce vowel distinctions in ways which could result, under the right conditions, in a collapse of morphological distinctions.

4.7 Gunu A622

The Gunu A622 FV system, as described in Orwig (1989), is in some ways reminiscent of what is seen in Yasa (§4.4) insofar as a vowel at the end of a verb has no clear semantic function. Unlike Yasa, however, there is evidence that it is morphologically active. To the extent that this vowel is coding anything, this would simply be that the word in which it is found is a verb. The form of this FV is largely predictable via rules of vowel harmony provided by Orwig (1989:

4 Reconstructing the development of the Bantu final vowels

288–293) that assume an underlying form of the FV as *-a*. Examples of Gunu verbs with the FV are provided in Table 12 (Orwig 1989: 288–289). The divisions in Table 12 represent the different forms of the FV as conditioned by the root vowel. The last form in the table, *dɔmb-a*, is specifically treated as exceptional with respect to vowel harmony.

Table 12: Examples of Gunu verbs illustrating FV patterns

Stem	Gloss
<i>lab-a</i>	‘enjoy’
<i>báan-a</i>	‘serve food’
<i>fá-a</i>	‘give’
<i>fɛm-a</i>	‘hate’
<i>léɛg-a</i>	‘talk’
<i>bɛ-a</i>	‘have’
<i>fon-a</i>	‘bless’
<i>bid-e</i>	‘ask’
<i>mi-e</i>	‘bury’
<i>déb-e</i>	‘contain’
<i>fug-e</i>	‘mix’
<i>buig-e</i>	‘close’
<i>sós-ɔ</i>	‘suck’
<i>ʃb-ɔ</i>	‘feel’
<i>dɔmb-a</i>	‘leave, pass’

In Table 13, examples of verbs appearing with various extensions are presented (Orwig 1989: 290–293). Forms carrying a Causative suffix are separated from the others due to the fact that this suffix triggers vowel harmony patterns affecting both the root and the FV. While FVs are not used to code morphological distinctions in Gunu, the fact that they are separated from the root in the presence of extensions indicates that they are still analysed as morphologically distinct from the root.

The Gunu system is an instance where phonological processes affecting vowels at the end of verb stems are relevant for understanding its FV system (see §3.6). In this case, it appears that patterns of vowel harmony, as well as other potential changes that are harder to recover historically, have completely neutralised any morphosyntactic distinctions that may have been encoded by the

Table 13: Gunu verbs with various extensions

Stem	Extended stem	Form	Gloss
<i>fól-a</i>	<i>fól-εd-a</i>	DIM	‘sweep’
<i>bɔl-ɔ</i>	<i>bɔl-εn-ɔ</i>	APPL	‘borrow’
<i>mag-a</i>	<i>mag-εn-a</i>	APPL	‘try’
<i>nog-a</i>	<i>nog-εn-a</i>	APPL	‘braid’
<i>ʒb-ɔ</i>	<i>ʒb-ɔn-ɔ</i>	ITER	‘feel’
<i>ság-a</i>	<i>ság-an-a</i>	ITER	‘dry’
<i>dɔmb-ɔ</i>	<i>domb-i-o</i>	CAUS	‘be tired’
<i>nod-a</i>	<i>nud-i-e</i>	CAUS	‘vomit’
<i>húm-e</i>	<i>húm-i-e</i>	CAUS	‘go out’

FVs. At the same time, the synchronic patterns provide good evidence that Gunu once made use of a more canonical FV system since it would otherwise be difficult to understand how a morphologically independent FV could develop on its own without any semantic function.²⁸

4.8 Eton A71

Van de Velde (2008: 114) describes Eton A71 as lacking FVs. Verbs can end in a vowel, but these are not identified as associated with the reconstructed FVs. Common shapes for underived verb stems are CV, CVC, and CVCV, with CVC forms comprising around sixty percent of the collected verbs, CVCV forms around twenty-five percent, and CV forms around fifteen percent (Van de Velde 2008: 115). Longer stems are found either because they are derived from shorter stems

²⁸For a language with a minimal FV system like Gunu, it might also be reasonable to analyse extensions as infixes appearing before the last vowel of a CVCV verb root. However, since the last vowel of the verb is largely predictable, a FV analysis is also possible, despite its minimal semantic function. Which synchronic analysis might be adopted has relatively little bearing on the historical concerns of this chapter given that the last vowel of Gunu verbs is transparently a reflex of at least one of the reconstructed FVs (and possibly more than one depending on the precise historical details). The language Cicipu of the Kainji subgroup of Benue-Congo offers an interesting contrast to Gunu since its verbs show a largely similar pattern except for the fact that the second vowel of verbs with CV₁CV₂ structure is unpredictable. Extensions still appear before the last vowel producing a CV₁C-VC-V₂ pattern. In the Cicipu case, an infixation analysis can more straightforwardly account for the unpredictability of the last vowel of a derived verb (McGill 2009: 209–210).

4 Reconstructing the development of the Bantu final vowels

via verbal extensions or appear with a limited set of expansions, some of which are identical to extensions (Van de Velde 2008: 116–118).²⁹ The second vowel of roots with the shape CVCV is restricted to underlying *i* and *a*, the latter of which is subject to vowel harmony, and the same restriction holds for the vowels found in verbal extensions and expansions. The harmony affecting *a* is triggered by preceding mid vowels (Van de Velde 2008: 31). Examples of underived verb roots are provided in Table 14 (Van de Velde 2008: 115).

Table 14: Examples of monomorphemic Eton verb stems

Verb	Gloss
<i>jà</i>	‘sing’
<i>vé</i>	‘give’
<i>tùg</i>	‘rub’
<i>pún</i>	‘be afraid’
<i>cilà</i>	‘forbid’
<i>kómô</i>	‘admire’
<i>bémî</i>	‘warn’

The fact that the last vowel in CVCV forms is restricted to two underlying vowels with forms that are similar to those of the reconstructed FVs, namely **-a* for *a* and potentially either **-é* or **-i* for *i*, might suggest that they should be treated as FVs. Moreover, there are morphological constructions where the vowel disappears in the presence of other suffixes in a way that is reminiscent of what is seen for FVs in languages whose verbs adhere more closely to canonical Bantu verb structure. Specifically, the vowel is lost when the verb appears with the causativising suffix *-là*, as seen in the verb pair *yéǵî* ‘learn’ vs. *yéǵlê* ‘teach’ (Van de Velde 2008: 121).³⁰ However, Van de Velde (2008: 115) makes clear that the properties of these vowels can be predicted based on general prosodic patterns in Eton (Van de Velde 2008: 19), and there is no evidence for analysing them as

²⁹Expansions are similar to extensions in that they are suffixal and have comparable phonological behaviour to extensions. However, they cannot be associated with any specific meaning, and they appear after roots which are not found without an expansion. See Schadeberg & Bostoen (2019: 172–173) for further discussion.

³⁰See also Van de Velde (2008: 123, 129) for other morphological constructions showing similar patterns. The realisation of the *-là* suffix with a front mid vowel in the word ‘teach’ appears to be due to a process of harmony affecting stem-final open syllables (see Van de Velde 2008: 38–39).

separate morphemes. Even if they were analysable as such, there would still be the problem of explaining why they only appear on some verbs.

The patterns found in Eton raise a number of historical questions. On the one hand, the lack of FVs can, in principle, be viewed as resulting from an innovation where historical FVs were lost as morphologically active elements. Under this scenario, some roots would have lost any trace of the FV, while in other roots a former FV would have become part of their lexical form. If this was the case, what processes would have governed which stems would have developed CVC shapes and which would have developed CVCV shapes? On the other hand, if Eton is somehow seen as representing a state of Bantu before the FV system had morphologised, there is still the same problem of understanding why some verbs have ‘final’ vowels and others do not. In principle, one could simply say that this was due to variation in the lexical forms of different verbs, though that would raise important issues for the reconstruction of PB verb roots, suggesting that their possible shapes may have been more heterogeneous than generally assumed (cf. e.g. Meeussen 1967: 89). This issue will be discussed further in §5.2.

More generally in the present context, Eton is another language where phonological restrictions affecting the right edge of the verb are relevant for understanding the realisation of vowels at the end of the stem (see §3.6). In particular, the presence of prosodic constraints limiting possible vowel oppositions in the second syllable of stems and also limiting stem size suggest a route through which a more heterogeneous system of verbal suffixes could, in effect, be reduced to result in something like the Bantu FV pattern. If these restrictions were subsequently “relaxed” at some stage of Bantu, verbal suffixal morphology could then allow stems to be expanded beyond two syllables. However, at that stage, the reduced FV pattern would have already morphologised and, in some sense, still attest to the presence of earlier prosodic restrictions.³¹

4.9 Gyeli A801

Grimm (2015: 215–216) discusses vowel patterns found at the end of Gyeli A801 verbs. She provides arguments as to why, even though all Gyeli verbs end in a vowel, these should not be considered FVs, but are rather present due to syllable structure constraints. While there are restrictions on which vowels can appear

³¹I leave open here the question of the timing of this proposed set of changes on the assumption that this would need to be considered in light of a more careful examination of NC verb structure (see §2). An additional complication is the possibility of cyclical change in NC and Bantu verb structures where periods of morphological and phonological reduction may have alternated with periods of morphological and phonological expansion (see Hyman 2011).

4 Reconstructing the development of the Bantu final vowels

in non-initial syllables, there is no evidence that these restrictions are tied to a limited number of morphological FVs and, instead, these seem to be prosodic in nature. Furthermore, extensions do not have the canonical -VC shape, where they appear between a stem and a FV. Rather, they have the shape -V or -VCV (Grimm 2015: 219), and they override the last vowel which would otherwise be found on the verb. Finally, the quality of the last vowel of a verb is not predictable and is, therefore, best analysed as part of the lexical specification of the verb, unlike canonical FVs.

In Table 15, examples of Gyeli verbs are provided. These are mostly drawn from Grimm (2015: 223), with forms also taken from Grimm (2015: 217, 218, 224). Gyeli verb extensions are illustrated in Table 16, with forms based on the disyllabic roots *lúndɔ* ‘fill oneself’, *vidɛ* ‘turn something’, and *kɛ̀ɛ* ‘hang something’. While the presence of extensions is associated with the loss of the root-final vowel in disyllabic stems, this can be accounted for straightforwardly in phonological terms as the result of a deletion connected to hiatus resolution given the invariant vowels found in the extensions (Grimm 2015: 216–217).

Table 15: Examples of monomorphemic Gyeli verb stems

Verb	Gloss
<i>bà</i>	‘smoke’
<i>dè</i>	‘eat’
<i>dji</i>	‘open’
<i>dyû</i>	‘kill’
<i>liè</i>	‘cede, let’
<i>bàwe</i>	‘carry something’
<i>kɛ̀ɛ</i>	‘hang something’
<i>kwàɛ</i>	‘love’
<i>vidɛ</i>	‘turn something’
<i>bvúɔ</i>	‘break something’
<i>djiwɔ</i>	‘steal’
<i>lúndɔ</i>	‘fill oneself’
<i>gyàga</i>	‘buy’
<i>kòla</i>	‘add’

Monosyllabic verbs have a different behaviour when appearing with extensions. They generally show an epenthetic consonant whose form is not synchronically predictable and which appears between their single vowel and the vowel

Table 16: Verbal extensions on disyllabic verb roots in Gyeli

Extension	Stem	Gloss
Reciprocal	<i>lúnd-ala</i>	‘fill one another’
Passive	<i>lúnd-a</i>	‘be filled’
Causative	<i>lúnd-εεε</i>	‘make something full’
Applicative	<i>lúnd-ελε</i>	‘fill something’
Autocausative	<i>vid-εga</i>	‘turn (by itself)’
Positional	<i>kèl-ɔwɔ</i>	‘assume hanging position’

of the extension. This means that they do not lose their lexical final vowel. For the few verbs that do not appear with these epenthetic consonants, their vowels still do not drop, creating exceptional hiatus environments between the root and the extension. Relevant examples are provided in Table 17 (Grimm 2015: 217–218). Forms in the first half of the table appear with epenthetic consonants (which are bolded), and forms in the second half take the extensions without the addition of an epenthetic consonant.

Table 17: Verbal extensions on monosyllabic verb roots in Gyeli

Root	Gloss	Derived stem	Gloss	Extension
<i>sâ</i>	‘vomit’	<i>sâŋgala</i>	‘vomit together’	Reciprocal
<i>bà</i>	‘smoke’	<i>bàyaga</i>	‘smoke (by itself)’	Autocausative
<i>dvùð</i>	‘hurt’	<i>dvùgεεε</i>	‘make hurt’	Causative
<i>bû</i>	‘destroy’	<i>bûla</i>	‘be destroyed’	Passive
<i>sóð</i>	‘continue’	<i>sósεεε</i>	‘continue with’	Applicative
<i>bâ</i>	‘marry’	<i>bânala</i>	‘marry one another’	Reciprocal
<i>dyû</i>	‘kill’	<i>dyûwala</i>	‘kill one another’	Reciprocal
<i>vèè</i>	‘try on’	<i>vèʔεεε</i>	‘make try on’	Applicative
<i>dè</i>	‘eat’	<i>dèβa</i>	‘be eaten’	Passive
<i>dyâ</i>	‘lie down’	<i>dyâaala</i>	‘lie down together’	Reciprocal
<i>kwê</i>	‘fall’	<i>kwêεεε</i>	‘make fall’	Causative

As suggested by Grimm (2015: 218), the appearance of the epenthetic consonants could be historically explained via loss of consonants in roots when they appeared in word-final position, while the consonants were protected from such

a process in the presence of an extension. Even if that is the ultimate historical source of the pattern, some degree of synchronic restructuring must have taken place given that the form *dè* ‘eat’ has an apparently straightforward PB etymology which has been reconstructed as a CV root, namely **dí* (BLR 944) and, therefore, had no historical consonant which could have been lost.³² Regardless as to the precise historical analysis, Gyeli appears to provide another example, alongside Eton, just discussed in §4.8, of the importance of prosodic size constraints for understanding Bantu morphological patterns given the differing behaviour of monosyllabic and disyllabic stems (see also §3.6).

Since Gyeli lacks a class of CVC verb roots, which is the canonical shape of monomorphemic verb roots in Bantu, in a pattern similar to what was seen for Yasa in §4.4, this again raises the questions of how stems in languages of this kind developed the particular lexical final vowel that they are found with.

4.10 Makaa A83

Makaa A83 is a language that appears to largely lack FVs. However, it does employ a verb-final high tone that is associated with the appearance of a vowel in the expected FV position under specific circumstances. This high tone generally appears in non-progressive constructions except when coding the Distant Past, in which case it only occurs in non-progressive constructions coded for polar focus (Heath 1991: 6). This high tone does not appear on the verb but can appear on the following word, where it replaces the tone found in the word’s first vowel, or it can be realised on an epenthetic vowel. The epenthetic vowel is found when this high tone would otherwise be placed before “a pronoun, a preposition, another verb, or an object without a prefix with a [low-toned] root” (Heath 1991: 6).³³ Relevant examples are provided in (6).

³²As was also seen in footnote 8, the reconstructed roots referred to here are drawn from Bastin et al. (2002) and an identifier is provided for their specific reconstruction in the online version of the Bantu Lexical Reconstructions 3 (BLR) database.

³³Heath (1991) does not appear to explicitly indicate how this high tone is realised in clauses where the verb is final. However, two contrasting examples that are provided by Heath (1991: 12–14), one with a transitive verb and one with an intransitive in a present perfect construction, suggest that the high tone does not appear when the verb is in final position.

(6) Makaa A83 (Heath 1991: 6–7)

- a. *Má ámà wííng ó-mpyâ.*
Mà ´ ` ámà wííng ´ ò-mpyâ
 1SG H1 P1 chase_away H2 2-dog
 ‘I chased away the dogs.’
- b. *Má ámà wííng ú ncwòmbè.*
Mà ´ ` ámà wííng ´ Ø-ncwòmbè
 1SG H1 P1 chase_away H2 7-sheep
 ‘I chased away the sheep.’³⁴
- c. *Mà á gù ú gwòó.*
Mà a ´ gù ´ Ø-gwòó
 1SG P2 H1 pick H2 7-mushroom
 ‘I picked the mushroom.’

All of the examples in (6) appear with two floating high tones, one at the left edge of the verb (glossed H1) and one at the right edge of the verb (glossed H2). The high tone of interest here is the one at the right edge. In (6a), the verb appearing with this high tone is followed by the noun *òmpyâ* ‘dog’ which begins with a low-tone noun class prefix. This prefix can serve as a host for the high tone, and the noun surfaces with an initial high tone, as indicated in the example. In the other two sentences, there is no postverbal host for the high tone, and, instead, an epenthetic vowel appears. This vowel is analysed as having the basic segmental form *ɥ*, which is the form that is found when it appears between two consonants, as in (6b). If it follows a verb ending in a vowel, it assimilates to that vowel, as in (6c), where its segmental form is *u* rather than *ɥ* due the fact that it follows the root *gù*.

One possible historical explanation for the appearance of these epenthetic vowels is that they represent traces of FVs that were lost in most phrasal contexts. If this is the case, it raises the question of the extent to which specific phrasal environments need to be considered in historical accounts of the development of the FVs and related patterns. The restricted distribution and quite specific conditioning of these vowels in Makaa would seem to make their analysis as relic forms more likely than their being innovations specifically to host a floating tone.

The Makaa case is not like any other language surveyed here, though the overall pattern fits into the general set of questions connected to the relationship between FVs and other suffixes, in this case a tonal suffix (see §3.4). While serious

³⁴The translation in (6b) has been adjusted from what was provided in Heath (1991: 6) to match the gloss, since the original translation had a pronominal object rather than a nominal one.

consideration is not given to the role of tone in FV formation in this survey, the Makaa epenthetic vowel makes clear that a full account of their development will need to take tonal patterns closely into account.

4.11 Kako A93

Kako A93, described by Ernst (1991; 1995) and Yukawa (1992), makes use of a system of verb-final suffixes that follows a pattern comparable to that associated with canonical FVs, but with a number of complications. Example data involving forms of the verb *bɛ̃ŋ* ‘see’ is provided in Table 18 (Ernst 1995: 11, 13). As can be seen, for this verb, one verbal category, the Subjunctive, is coded by the lack of a suffix and another category, the Past, is coded by a CV suffix. With five segmentally different Finals (including the \emptyset final for the Subjunctive), Kako has the largest system of Finals for any language in this survey, though the realisation of these Finals can include consonants, as seen for the Past form in Table 18, which means that these cannot all be considered FVs.

Table 18: Verbal suffixes in Kako

Category	Form
Infinitive	<i>bɛ̃ŋ-ɛ̃</i>
Atemporal	<i>bɛ̃ŋ-a</i>
Imperative	<i>bɛ̃ŋ-á</i>
Subordinate	<i>bɛ̃ŋ-ē̃</i>
Subjunctive	<i>bɛ̃ŋ-∅</i>
Past	<i>bɛ̃ŋ-má</i>

As discussed below, there are a series of complicated positional restrictions on non-initial vowels in Kako which limit the range of contrasts at the end of the verb. While it is not explicitly described as such, the FV appearing on infinitives seems to be lexically specified, though, as can be seen, it still appears to be morphologically active insofar as it can be replaced by other suffixes and its absence can be used to encode a verbal category. Ernst (1995: 15) presents at least one near-minimal pair of verb roots which seems to illustrate the lexical nature of FVs in infinitives: *kít-ɛ̃* ‘to advise’ and *kít-ɔ̃* ‘to style (hair)’. Comparable to what is found in some of the other surveyed languages (see §3.5), CV roots lack a FV in infinitive forms (Ernst 1995: 3). (This pattern will be discussed in more detail in §4.15.)

There are a number of morphophonological complexities involved in the realisation of the Kako verb-final suffixes. This can be exemplified by considering the Past forms, as illustrated in Table 19 (Ernst 1995: 18–22).³⁵ The basic form of this suffix appears to be *-má*, but its form varies depending on the verb's last consonant.³⁶ The *-má* variant is found after CV roots and roots whose Infinitive ends in *e*, *ɛ*, and *o* and whose last consonant is a sonorant, as seen in, for example, forms like *womá* 'kill.PST' and *kelmá* 'do.PST' in Table 19. Stems ending with these vowels in the Infinitive and whose last consonant is a stop that is not palatalised, labialised, or *r* appear with a high vowel before *-má* that appears to represent a raised variant of the FV found in the Infinitive. This is seen in the forms *kitimá* 'advise.PST' and *wokumá* 'hear.PST' in Table 19. Stems whose FV in the Infinitive is *a* retain the vowel when *-má* is added, as seen in a form like *sanamá* 'work.PST'. Stems ending in palatalised or labialised consonants or extended stems whose Infinitives have FVs *e* or *ɛ* take *-á* in the Past, as in forms like *kwad'ýá* 'love.PST' and *njesá* 'send.PST'. Ernst (1995: 21) only provides an example of an extended verb that independently ends in a palatalised consonant, which is why this is the only extended form included in the fifth set of forms provided in Table 19. Finally, there are a number of CV stems which are irregular, forming the past by replacing their last vowel with *á*, as seen, for example, in the opposition between the Infinitive and Past forms of 'hear' where *gwé* alternates with *gwá*.

Other verbal forms show similar complications in their phonological realisation. For instance, the Atemporal, which takes the suffix *-a* in Table 18, has allomorphs with a CV shape where the initial consonant is a velar followed by a harmonising vowel, resulting in forms such as *wo-ku*, based on *wó* 'kill'. There are also forms where the Atemporal is coded by the lack of a suffix, as in *kel*, based in *kélb* 'do', among various other realisations (Ernst 1995: 24–28). The Imperative also shows fairly complex patterns of allomorphy as well as segmental overlap with the Atemporal in a number of cases (Ernst 1995: 37). The Subordinate form has a simpler segmental realisation, which is often the same as the Infinitive, except for verbs whose Infinitive ends in an *-ɔ*, in which case it changes to *-ɛ* (Ernst 1995: 38). The Subjunctive also often takes on the same form as the Infinitive, except for verbs whose FV is *-ɛ* or *-ɔ*. These become *-i* and a harmonising high vowel respectively after non-palatalised and non-labialised stops and delete after nasals (Ernst 1995: 41). This last variant is the one seen in Table 18.

³⁵The transcription of the Infinitive verb of 'arrive' in Table 19 has been adapted to represent nasalisation with a tilde rather than the original diacritic found in the source.

³⁶A potential etymology for this suffix is that it represents a morphologisation of the reconstructed verb **màd* 'finish'. See Nurse (2008: 252–253) for discussion of verbal prefixes with shape *ma-* found in some Bantu languages which can be traced to the same reconstructed verb.

4 Reconstructing the development of the Bantu final vowels

Table 19: Variation in past tense formation in Kako

Infinitive	Past	Gloss
<i>wó</i>	<i>wo-má</i>	‘kill’
<i>d̥yà</i>	<i>d̥yaŋ-má</i>	‘arrive’
<i>kèl-ɔ̃</i>	<i>kel-má</i>	‘do’
<i>kít-è</i>	<i>kit-i-má</i>	‘advise’
<i>wōk-ɔ̃</i>	<i>wok-u-má</i>	‘hear’
<i>sán-à</i>	<i>san-a-má</i>	‘work’
<i>jámbin-à</i>	<i>jambin-a-má</i>	‘cook’
<i>kwàdy-è</i>	<i>kwady-á</i>	‘love’
<i>bèŋw-è</i>	<i>beyw-á</i>	‘follow’
<i>njès-è</i>	<i>njes-á</i>	‘send’ (s = [ʃ])
<i>dók-ìdy-è</i>	<i>dok-idy-á</i>	‘add’
<i>kè</i>	<i>ká</i>	‘go’
<i>njè</i>	<i>njá</i>	‘come’
<i>gwé</i>	<i>gwá</i>	‘die’

Taken together, the pattern that emerges is one where there is evidence for six distinct verbal patterns of suffixation associated with the expression of TAMP categories, as exemplified in Table 18, though these distinctions can be neutralised in specific verb forms and their realisation can be controlled by a number of complicated factors sensitive to root-final phonology (see also §3.6).

An added layer of complication to the Kako system are restrictions on the allowable vowel qualities in final syllables of the Infinitive forms of verbs. These are summarised in Table 20 (Ernst 1995: 8, see also Ernst 1991). The table provides information on what vowels are allowed, separated by the final consonant of the root and whether the stem, including the FV, is two or three syllables (in the column labelled σ). This means that, in effect, FVs can simultaneously encode a morphological category (e.g. that a verb is in the Infinitive form) while also partly encoding aspects of the prosodic properties of the verb that they appear with (e.g. whether it is two or three syllables in length), though these categories are not necessarily uniquely encoded by a given vowel (see also §3.3). The FV $-ɔ̃$, for instance, is only associated with a subset of verbs that have two syllables

including the FV, while the FV *-e* is mostly restricted to three-syllable verbs, excluding those whose final consonant is a palatal. A vowel like *-a*, by contrast, is less prosodically restricted and cannot be seen as encoding any information about the phonological properties of the verb other than the fact that it does not end in a labial-velar.

Table 20: Prosodic restrictions on final vowels in Kako Infinitives

Final consonant	σ	Final vowel
Sonorant (<i>m, n, ŋ, l, y, w</i>)	2 3	ϵ a \circ e a
Palatal (<i>dʲ, ny, nj</i> ; some <i>s, w</i>)	2 3	e ϵ a e ϵ a
Labial-velar (<i>kʷ, gʷ, ŋgʷ</i>)	2	ϵ
Occlusive (<i>p, b, t, k, mb, nd, ŋg, ŋgb, gb, r</i> ; some <i>s, nj</i>)	2 3	ϵ a \circ e ϵ a

The Kako patterns are perhaps the most interesting among the surveyed languages in the present context since they suggest a historical model for the development of the Bantu FVs involving an interplay between stem-final processes resulting in suffixal consonant loss (e.g. the *m* of the Past suffix as exemplified in Table 19) alongside prosodic processes affecting vowels, including not only vowel harmony patterns but also positional restrictions of the sort seen in Table 20. Taken together, these processes could result in phonological reductions that would cause formerly distinct morphemes to become segmentally homophonous. While these patterns of homophony are limited to specific phonological contexts in Kako, if they were to become extended to a wider range of contexts, it is possible to imagine the resulting system being limited to a small set of vowels, along the lines of what we see in Bantu languages showing the canonical Bantu FV pattern. Of course, this is speculative, and many details would need to be filled in to map out how a Kako-like system could develop into a canonical Bantu one. However, given that the ultimate source of the Bantu FVs is otherwise unclear, the phonological alternations associated with TAMP suffixes in Kako can potentially be seen as a model for the initial steps of a pathway for their development.

4.12 Myene B11

The Myene B11 FV system is closer to the canonical Bantu system than the languages considered to this point except for Kpe (§4.3), and, in general, the surveyed languages of zone B show more canonical FV patterns than the surveyed languages of zone A. In Myene, most verb stems end in a FV *-a* which can change in different TAMP configurations, as would be expected. There is, however, a class of verbs ending in *o* and *e* which have an invariant last vowel, meaning that they constitute lexical exceptions to the usual patterns (see §3.5). Verbs ending in *o* also include those which have been passivised via the replacement of FV *-a* with Passive marker *-o* (Gautier 1912: 82). The historical significance of this pattern will be discussed in more detail in §4.15. Examples of verbs showing FV *-a* in their Infinitive forms and *o* or *e* as their last vowel in their Infinitive forms are provided in Table 21 (Gautier 1912: 82).³⁷ The form *ke* ‘go’ in Table 21 is synchronically treated as a short form of the regular verb *kēnda*, though these forms go back to two distinct PB forms, namely **gì* (BLR 1371) and **gènd* (BLR 1363).

Table 21: Myene Infinitive forms

Verb	Gloss
<i>bag-a</i>	‘bring’
<i>ben-a</i>	‘plant’
<i>but-a</i>	‘look for’
<i>pon-a</i>	‘watch’
<i>dēnd-a</i>	‘do’
<i>tāng-a</i>	‘count’
<i>dyɔgo</i>	‘hear’
<i>avoro</i>	‘know’
<i>ko</i>	‘be able’
<i>re</i>	‘be’
<i>zele</i>	‘not be’
<i>ke</i>	‘go’

³⁷The transcription of the data in Table 21 has been adjusted to replace an original use of *è* and *ò* with *ɛ* and *ɔ* respectively, following my interpretation of Gautier (1912: 3). The verb *dyɔgo* ‘hear’ in Table 21 is presumably a reflex of the reconstructed root **jigu* (BLR 3423), whose CVCV structure would already have been irregular in PB.

Myene otherwise appears to make use of two FVs, *-i* and *-e*, in addition to *-a* in the canonical way. Gautier (1912) does not appear to make an explicit statement regarding the existence or specific semantics of these two vowels, but they can be found in the verb conjugations provided, and the *-i* is associated with the past and the *-e* with subjunctive and subordinate contexts. Representative forms are provided in Table 22, based on the verb root *dyen* ‘see’ (Gautier 1912: 84–85).³⁸ Morphological segmentation in Table 22 is my own. Some forms are coded with an Imperfect suffix *-ag* appearing between the stem and the FV. Except for the Imperative, the forms are provided with a first person singular subject pronoun, as given in the source. Spaces in the forms are also those provided in the source. Verbs whose stems end in *o* and *e* do not participate in these suffixal alternations, and verbs ending in *e* are additionally characterised as defective, i.e. lacking certain expected inflectional forms (Gautier 1912: 82).

Table 22: Verb forms in Myene

Verb form	Function
<i>yen-a</i>	Imperative
<i>mi dyen-a</i>	Present
<i>my a-dyen-a</i>	Immediate Past
<i>my a-dyen-ag-a</i>	Immediate Imperfect
<i>my a-dyen-i</i>	Near Past
<i>my a-dyen-ag-i</i>	Near Imperfect
<i>mi be dyen-a</i>	Future
<i>mi ga yen-e</i>	Necessitative Subjunctive

While the presence of lexical exceptions in Myene raises historical questions regarding the path through which the FV system developed, its FV system is otherwise quite recognisable as a canonical Bantu system.

4.13 Kota B25

The Kota B25 FV system appears to be largely canonical in form. Piron (1990: 129) describes three different segmental final suffixes, with forms *a*, *e*, and *εε*.

³⁸Although not of direct relevance to the current study, Myene verbs can also show an interesting pattern of initial consonant mutations, resulting in two stem forms, one with a “weak” consonant and one with a “strong” consonant, as seen, for instance, in the Infinitive/Imperative pair *pona* ‘to watch’ vs. *wona* ‘watch!’ (Gautier 1912: 81–82). This is seen in Table 22 where the initial consonant of the root alternates between *dy* and *y*.

4 Reconstructing the development of the Bantu final vowels

The *-a* can appear with a high tone or a low tone, while the *-e* and *-eɛ* are both described as only appearing with a high tone. The FV *-a* is used with a wide range of verb forms in affirmative and negative contexts, as well as past, present, and future contexts, and clearly is best understood as the default FV. The *-e* is described as appearing only with the Negative Present and the *-eɛ* only in a Present Affirmative form associated with semantics involving an action that is being done for the first time, and it can presumably be historically associated with the Perfective **-ile* (see §1).³⁹ Examples of Kota verbs with these suffixes are given in Table 23 (Piron 1990: 131–139).

Table 23: Verb forms in Kota

Verb form	Gloss	Translation	Function
<i>jók-ák-á</i>	listen-IPFV-FV	‘listen!’	Imperative
<i>mà-hút-á</i>	1SG.TAMP-pay-FV	‘I pay’	Present
<i>já-mò-làp-á</i>	3SG.TAMP-TAM-disappear-FV	‘he disappeared’	Recent Past
<i>já-ká-làp-á</i>	3SG.TAMP-TAM-disappear-FV	‘he has not disappeared’	Recent Past Negative
<i>mé-kón-àk-à</i>	1SG.TAMP-plant-IPFV-FV	‘I will plant’	Immediate Future
<i>mà-jók-été</i>	1SG.TAMP-listen-FV	‘I am listening’	Present (“first time”)
<i>má-bép-é</i>	1SG.TAMP-carry-FV	‘I don’t carry’	Present Negative

Piron (1990) does not appear to contain an explicit statement regarding how roots which end in a vowel, such as verb roots with form CV, behave with respect to the presence of FVs and whether there may be lexical exceptions to the normal patterns. However, a partial paradigm of forms is provided for the CV verb root *dì* ‘be’, and it does appear to take FVs.⁴⁰ For instance, there is a Present form *àdjé* (with a class 1 subject prefix), analysable as *à-dì-é*, which irregularly takes a FV with form *é*, rather than the expected *-a*. This is unlike Myene (§4.12) where

³⁹Piron (1990: 129) labels this suffix as coding a present tense, but Piron (1990: 133) implies it codes a past tense. The translations suggest it codes a present tense form, which is why I am using that label here.

⁴⁰Piron (1990: 62) contains an abstract analysis of the conjugation of a CV verb *sò* ‘say’ that includes a FV at an underlying level of representation, though it is difficult to assess the extent of the evidence for this analysis. Piron (1990: 68) also gives derivations for CV and CVCV roots where they take the usual FV verb morphology. However, an intervening Imperfective suffix in these examples makes them not ideal for establishing their overall morphological behaviour with respect to FVs.

comparable exceptional verbs do not take a FV of any kind and end in invariant vowels regardless of the TAMP configuration in which they appear. Similarly, there is a past form with a class 1 subject prefix *àdjàsá*, analysable as *à-di-à-sá* (with a Post-final suffix *-sá*), which takes a FV *-a* in the canonical way (Piron 1990: 142–143).

Piron (1990: 62) describes a vowel harmony rule that can affect FV *-a* causing it to surface as *ɔ* after *ɔ* and *ɛ* after *ɛ* (see also §3.6). The latter change is of interest here due to the fact that this could result in a partial formal segmental overlap between FV *-a* and the Final suffix *-εε* in certain phonological contexts. This could lead to partial formal conflation, for instance with the second vowel of *-εε* being reanalysed as a surfacing form *-a*. Nevertheless, overall, the Kota system is in line with the canonical Bantu system.

4.14 Himba B302

Himba B302 has four segmental FVs, an *-i* and *-e*, both appearing with low tones, an *-a* appearing with high and low tones, and an *-o*, appearing with a high tone (Rekanga 2000: 468). This makes it, along with Kpe, one of the languages with the highest number of FVs encountered in the survey. The first three FVs follow typical patterns. The *-i* and *-e* are relatively restricted in the contexts in which they occur, with *-i* found in recent past contexts and *-e* found in some present and future contexts. The FV *-a* appears in a wide range of other contexts, following its general pattern as a default FV. The FV *-o* is restricted to a specific infinitive construction that is also coded with a prefix. Relevant examples are provided in Table 24 (Rekanga 2000: 468–472).⁴¹

The FV *-o* only appears in infinitive forms with a specific prefix and also, apparently, requires the presence of an additional *-ay* suffix, which is found in the two examples provided in Rekanga (2000: 472). The restricted distribution of FV *-o* is an indication that it is a recent innovation. This is also suggested, of course, by the comparative picture. The source of this *-o* is not clear, but it does at least point to the potential for the development of new FVs in languages that otherwise appear to have a relatively stable canonical FV system.

While the behaviour of verbs that might present potential lexical exceptions to FV patterns does not appear to be discussed in a general way, Rekanga (2000:

⁴¹Rekanga (2000: 469) appears to label the form *òndéhù:mè* in Table 24 as a Recent Past but provides a Present translation, which is the category in which the verb is placed here. The identification of the *-ay* suffix as an Imperfective is my own, on the assumption that it is a reflex of **-ag*. The identification of the *mo-* prefix in the last form of the table with class 18 is also my own.

Table 24: Verb forms in Himba

Verb form	Gloss	Translation	Function
<i>à-hù:p-i</i>	3SG-TAMP-throw-FV	‘he threw’	Recent Past
<i>ò-ndé-hù:m-è</i>	2SG-TAMP-descend-FV	‘you descend’	Present
<i>à-ndé-bé-pù:p-á</i>	3SG-TAMP-TAMP-move-FV	‘he moves first’	Present Precessive
<i>à-má-kin-áy-à</i>	3SG-TAMP-refuse-IPFV-FV	‘he had refused’	Remote Past
<i>ò-hómb-à</i>	2SG-buy-FV	‘that you buy’	Present Subjunctive
<i>ò-há-nà-d^y-à</i>	2SG-TAMP-TAMP-come-FV	‘that you do not come’	Negative Subjunctive
<i>mó-γènd-á:γ-ò</i>	18-go-IPFV-FV	‘to go’	Infinitive

471) does provide the example *òhánàd^yà*, presented in Table 24 and based on a verb root analysed as underlyingly having the form *dyè* ‘come’. This suggests that there are no lexically-conditioned exceptions to the FV patterns, given that this inflected form of the verb ends in *a* rather than *e*. It does appear, however, that the presence of the Passive suffix overrides the presence of a FV, as evidenced by forms like *nómàitsú* ‘it (cl. 11) was given’, based on the verb root *its* ‘give’, and *àndéhèβónó* ‘he was chosen’, based on the verb stem *hévón* ‘choose’ (Rekanga 2000: 321) (see §3.4). In both cases, the last vowel of the verb can be associated with a Passive suffix analysed as being underlyingly *o* that is affected by vowel harmony.⁴² A full range of passivised verbs, in particular across the different possible FVs, is not presented in Rekanga (2000). So, there may be complications in the realisation of the Passive that were not reported.

Overall, the Himba FV system is largely in line with the canonical Bantu system with the major points of difference being the development of the FV *-o* in one of the language’s infinitive constructions and the fact that the Passive suffix apparently overrides the appearance of FVs that would otherwise be expected.

4.15 Punu B43

Punu B43, following the descriptions of Bonneau (1956: 44–45) and Fontaney (1980: 75), makes use of FVs that are in line with the canonical Bantu pattern,

⁴²For this study, I had access only to the second volume of a multivolume work, which focuses on the morphology of Himba. Because of this, I was not able to examine the part of the work discussing processes of vowel harmony.

though with some noteworthy differences. Most inflectional verb forms end in a FV *-a* which has a default status. There is also a final *-i* that appears in a small set of inflected forms, namely the Affirmative and Negative Present, Imperative verbs which are also marked for an object prefix, and the Affirmative Subjunctive. The latter two domains are associated with FV suffixes whose segmental form is reconstructed as **e* by Meeussen (1967: 112), and the Punu *-i* in these contexts is presumably connected to the same pattern that prompted this reconstruction. The use of *-i* in negative verbs can also be connected to a form (tentatively) reconstructed by Meeussen (1967: 110) as **-i*. The use of an *-i* in the present affirmative would appear to represent some kind of innovation. Example verb forms, based on the stem *dibíg* ‘close’, which appears with an Impositive extension *-ig* [iy] (Fontaney 1980: 59), are presented in Table 25.⁴³ Those verbs with subject marking appear with the 1PL marker *tu-*. The forms and category labels are drawn from Fontaney (1980: 78–80).

Table 25: Regular verb forms in Punu

Verb form	Function
<i>u-dibíg-a</i>	Infinitive
<i>tu-kí-dibíg-a</i>	Near future
<i>tu-gó-dibíg-a</i>	Future Negative
<i>tú-tsi-dibíg-a</i>	Perfective
<i>tu-má-dibíg-a</i>	Imperfective
<i>díbig-a</i>	Imperative
<i>tú-í-díbig-i</i>	Present
<i>tu-gé-díbig-i</i>	Present Negative
<i>tu-dibíg-i</i>	Subjunctive
<i>ji-díbig-i</i>	Imperative (with object marker)

There are two general classes of exceptions to the patterns exemplified in Table 25. Verbs which end in vowels other than *-a* in their citation form have invariant vowels when conjugated. Some of these have the form CV, but Bonneau (1956: 44) also indicates that there are underived longer verbs, such as *ulu* ‘hear’ and *gufi* ‘be small’. Example forms from two CV verbs of this class, *ji* ‘eat’

⁴³See Schadeberg & Bostoen (2019: 178–179) for a discussion of the Impositive in a comparative Bantu context.

4 Reconstructing the development of the Bantu final vowels

and *nu* ‘drink’ are presented in Table 26 (Fontaney 1980: 95–96). Regular FV patterns can re-emerge in these stems in the presence of extensions. For instance, a causativised form of ‘drink’ has the form *nu-ís-a*, which then follows the pattern seen in Table 25.

Table 26: Invariant final vowel forms in Punu

Verb form (‘eat’)	Verb form (‘drink’)	Function
<i>tú-í-ji</i>	<i>tú-í-nu</i>	Present
<i>tú-ú-ji</i>	<i>tú-ú-nu</i>	Future
<i>tú-tsí-ji</i>	<i>tú-tsí-nu</i>	Perfective

The other class of exceptions to the patterns exemplified in Table 25 are passivised verbs, which appear with a final *-u*. The Passive *-u* overrides any other expected FV (Fontaney 1980: 75). For instance, a verb like *lab-a* ‘see’ would have the form *lab-u* when passivised and behave like the verb *nu* ‘drink’, seen in Table 26 (Bonneau 1956: 45).

The Punu patterns are interesting due to the presence of lexical and morphological exceptions to canonical FV patterns (see §3.5). The primary question that this raises from a comparative perspective is how such a pattern could have developed. While the available sources do not provide comprehensive lists of exceptional verb roots such as those seen in Table 26, three that are presented have apparent PB etymologies. For the verbs in Table 26, it is presumably the case that *ji* ‘eat’ is a reflex of PB **dí* (BLR 944) and *nu* ‘drink’ a reflex of **nyó* (BLR 7047), while a third verb *fu* ‘die’ Fontaney (1980: 96) is a reflex of **kú* (BLR 2089). See also Nsuka-Nkutsi (1980). The exceptionality of these forms cannot be seen, therefore, as attributable to their being borrowed or connected to some other irregularity due to contact. The same holds for the Passive suffix.

If we assume that the canonical Bantu FV pattern can be associated with PB, we would need to propose a process whereby FVs were lost in these forms in Punu, perhaps accompanied by a vocalisation process if their vowels had surfaced as glides when followed by a FV. The conditions that would allow such a change to take place are not obvious. However, in the Punu case, the fact that the vowel *a* would typically have appeared as [ə] in final position (Kwenzi-Mikala 1980: 10) may have made it more likely to be lost in that context when preceded by another vowel. Alternatively, if we assumed that the exceptional verb forms

in Punu represent an archaism before the canonical FV system had completely developed, then that raises questions about the timing of the emergence of FVs in Bantu and specifically suggests that they may not yet have been fully morphologised in PB.

4.16 Nzebi B52

Nzebi B52 is described as making use of two FVs, *-a*, in most forms – i.e. serving as a default – and *-i* in forms encoding the Perfect (Marchal-Nasse 1989: 119). The *-a* FV can be affected by a rule of vowel harmony causing it to assimilate to a preceding *ε*, *ɔ*, or *u*. It can also appear as *ɔ* in some cases (Marchal-Nasse 1989: 113). The FV *-i* can trigger patterns of regressive harmony raising preceding vowels (cf. e.g. Marchal-Nasse 1989: 121–123, 130–131).⁴⁴ None of these processes appears to create any ambiguity with respect to which FV is appearing on a verb. Example verbs are provided in Table 27 (Marchal-Nasse 1989: 461–489).⁴⁵

Table 27: Verb forms in Nzebi

Verb form	Gloss	Translation	Function
<i>bà:-só:mb-á</i>	SM ₂ -TAMP-buy-FV	‘they buy’	Neutral Present
<i>bi-sa:-díb-əg-á</i>	SM ₈ -TAMP-close- IPFV-FV	‘they don’t close’	Negative Present
<i>ní-ba-bó:ng-ɔ</i>	TAMP-8-take-FV	‘they will take’	Indeterminate Future
<i>bi-ma-díb-əg-a</i>	SM ₈ -TAMP-close-FV	‘they have already closed’	Near Past
<i>by-á-só:mb-ás-á</i>	SM ₈ -TAMP-buy- CAUS-FV	‘they had sold’	Distant Past
<i>bi-dib-ig-i</i>	SM ₈ -TAMP-close- IPFV-FV	‘they have closed’	Perfect
<i>a-sa-bé:mb-i</i>	SM ₁ -TAMP-touch-FV	‘he has not touched’	Negative Perfect

Verb roots with shape CV appear with an expansion *-ad*, which is then followed by the regular FV. So, while these have exceptional behaviour, they fol-

⁴⁴This pattern is also discussed in Guthrie (1968: 102–103).

⁴⁵The identification of the *-Vg [Vx]* suffix with an imperfective is my own, on the assumption that it is a reflex of **-ag*.

low the regular FV patterns. The verb *kú* ‘die’, for example, has the infinitive *u-kw-á:d-a* (Marchal-Nasse 1989: 439–440).

While not specific to FVs, vowels found at the end of words longer than one syllable, which includes FVs, are subject to various reduction processes, which present possible historical models for the loss of FVs in other languages (Marchal-Nasse 1989: 42–43, see also Guthrie 1968: 119).

Overall, from a formal perspective, Nzebi’s FV system more or less follows the canonical pattern. The system is the smallest logically possible size of just two vowels, and it has a limited functional load, since the *-i* vowel appears in a fairly narrow set of forms. It shows some phonological complications, though not any that appear to shed particular light on the development of FVs generally.

4.17 Overview of survey results

Table 28 places the languages covered in this survey into five broad categories based on the nature of their FV systems. This categorisation is intended to complement the information provided in §3, which focused on grammatical phenomena relevant to understanding FV patterns rather than the languages themselves. The five categories are as follows:

1. Languages that show verb-final morphology that may show some similarities to FVs as found in canonical Bantu verbs (as described in §1), but which also have features that cannot be straightforwardly connected to the reconstructed FVs of Meeussen (1967: 110). These are classified as having *non-canonical Final morphology*.
2. Languages where vowels are found at the end of verbs in ways that show phonological parallels to verbs that appear with FVs but where there is no evidence that these vowels should be treated as distinct morphemes. These are classified as languages with *vowels found at the end of verbs, no distinctions encoded*.
3. The one language in the survey that did not show any evidence of regular segmental FVs, Makaa (§4.10), is categorised under the label *no regular segmental Final Vowels*.
4. Languages which showed evidence for a relatively canonical FV system but where there were also exceptional verbs where FVs were not found. These are classified as having *canonical Final Vowel morphology, with exceptions*.

5. Languages whose FV systems follow the canonical pattern, which are classified as having *canonical Final Verb morphology*.

The results of the survey are further discussed in §5.

Table 28: Overview of the results of the survey

Category	Language	Summary of system
Non-canonical Final morphology	Akooe A15C (§4.2)	A complex system of verb-final TAMP encoding potentially involving multiple suffixes
	Nen A44 (§4.5)	A FV appears on some verbs in a narrow set of semantic contexts
	Kpa A53 (§4.6)	A complex system of verb-final TAMP encoding with only one vocalic suffixal element
	Kako A93 (§4.11)	A complex system of verb-final TAMP encoding with both vocalic and non-vocalic suffixal elements
Vowels found at the end of verbs, no distinctions encoded	Yasa A33a (§4.4)	Verbs end in lexically unpredictable vowels which do not exhibit clear evidence of having morphological status
	Gunu A622 (§4.7)	Verb stems end in a morphologically separable vowel, but it does not encode TAMP contrasts
	Eton A71 (§4.8)	Some verbs end in lexically unpredictable vowels which do not exhibit clear evidence of having morphological status
	Gyeli A801 (§4.9)	Verbs end in lexically unpredictable vowels which do not exhibit clear evidence of having morphological status
No regular segmental Final Vowels	Makaa A83 (§4.10)	Verbs do not generally end in vowels, but epenthetic vowels can appear after verb stems where a Final Vowel would be expected under specific phonological circumstances

4 Reconstructing the development of the Bantu final vowels

Category	Language	Summary of system
Canonical Final Vowel morphology, with exceptions	Kpe A22 (§4.3)	Four FVs are found, but some verbs exceptionally end in invariant vowels
	Myene B11 (§4.12)	Three FVs are found, but some verbs end in invariant lexical vowels and Passive verbs are derived via an invariant verb-final vocalic suffix
	Himba B302 (§4.14)	Four FVs are found, but Passive verbs appear to be derived via an invariant verb-final vocalic suffix
	Punu B43 (§4.15)	Two FVs are found, but some verbs end in invariant lexical vowels and Passive verbs are derived via an invariant verb-final vocalic suffix
Canonical Final Vowel morphology	Kota B25 (§4.13)	Two FVs are found without any apparent exceptional behaviour, along with one segmentally more complex Final element
	Nzebi B52 (§4.16)	Two FVs are found without any apparent exceptional behaviour

5 Conclusion

5.1 When did the Final Vowels develop?

Given the variability found with respect to verb-final morphology in this survey, an important concern that arises in the PB context is determining the stage of PB at which the reconstructed system developed. In order to consider this more closely, the map provided in Figure 2 provides the location of each of the surveyed languages while also indicating how they were categorised with respect to the five broad classes introduced in §4.17 and summarised in Table 28.⁴⁶

In this map, a clear pattern emerges, where the languages in the southern part of the survey area in Gabon all have FV systems that are in line with the standard reconstructions, though some of these have lexically exceptional verb stems. In the northern area, with the exception of Kpe, the languages deviate from the reconstructed patterns in various ways, and they do not provide clear evidence for the reconstructed system.

⁴⁶This map was created using tools developed by Moroz et al. (2022). The background map was produced by Thunderforest (see <http://www.thunderforest.com>) using data from OpenStreetMap (see <https://www.openstreetmap.org/>).

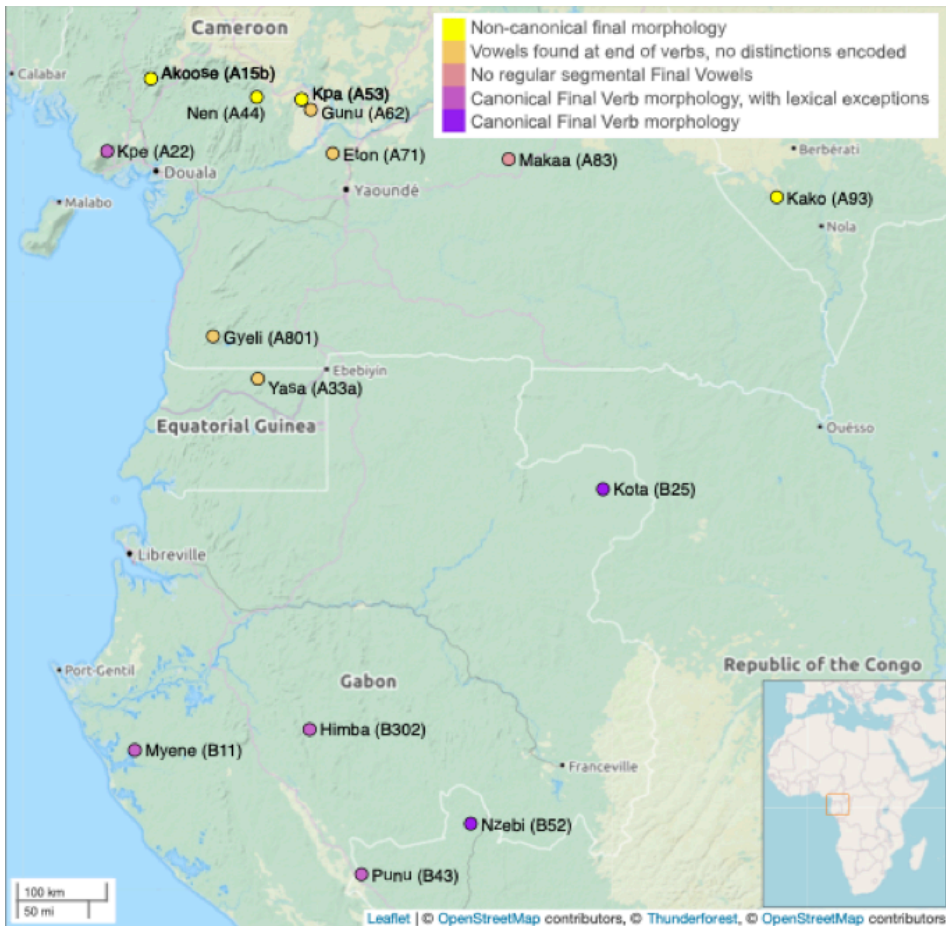


Figure 2: Locations of languages and overview of FV patterns found in this survey

The languages whose FV systems are in line with the reconstructed PB system, with the exception of Kpe, would appear to correspond most closely to those belonging to node 2 of the phylogenetic tree in Grollemund et al. (2015). This node includes most of the languages in zone B and higher, as well as some languages from the A80 and A90 groups. This could suggest that the canonical FV system was a relatively late development. Alternatively, a contact explanation could be given wherein FV patterns in line with the reconstructions were earlier present in the languages in the northern part of the north-western area. These might have become reduced due to areal processes of morphological reduction by virtue of

the fact that the languages in zone A are in the “buffer” region between isolating Kwa-type languages to their north and west and more canonically agglutinating languages to their south – cf. e.g. Hyman (2004), Good (2012; 2017: 476–484) for relevant discussions, and Stilo (2005) for a discussion of the concept of a buffer zone. Under this interpretation, Kpe could be viewed as a relatively conservative language within the region.

At this stage, I believe it would be premature to argue that either of these scenarios is more likely than the other, or that they should be seen as mutually exclusive since the full historical picture almost certainly involved an interplay between genealogical and areal factors. The most important point to emerge from this study is that the north-western Bantu data does not obviously point to the reconstructed FV system having been fully in place before the diversification of the Bantu languages, at least given the conventional understanding of which languages comprise the family.

5.2 How did the Final Vowels develop?

The survey presented in §4 does not point to any clear answers with respect to the development of the FVs in Bantu. However, it does reveal patterns of complexity suggesting that FV systems in the north-western Bantu area cannot simply be understood as a straightforward reduction of the canonical Bantu pattern that has served as the basis for the existing reconstructions of FVs, as presented in §1. The most noteworthy questions raised by this survey in my view are those listed in (7).

- (7) Questions raised by this survey:
- a. Should languages like Akoose (§4.2) and Kako (§4.11), which make use of non-vocalic verb morphology in the Final position, be seen as having innovated non-vocalic Final morphology after losing FVs or as representing precursor systems to those languages showing canonical FV patterns?
 - b. In languages which show no synchronic evidence for FVs, such as Yasa (§4.4) and Gyeli (§4.9) where all verb roots end in vowels, or Eton (§4.8), where many do, what is the source of the last vowel in those roots?
 - c. In languages with canonical FV patterns on most verbs, but with a class of verbs exceptionally not taking FV morphology, such as Kpe (§4.3), Kako (§4.11), Myene (§4.12), and Punu (§4.15), does the lack of FVs on these roots represent an archaism or an innovation?

- d. How interconnected are processes of prosodification and the development of the canonical FV system?
- e. Can FVs be used to establish isoglosses that separate different historical stages of Bantu?

At this stage, I believe that all of the questions in (7) must be considered open and without obvious answers. However, the languages of the survey do, at least, point to a historical path for the development of the canonical Bantu FV system along the following lines, though its time depth is not clear. These steps could be viewed roughly as follows, with languages from the survey that can be seen as models for each of the steps indicated:

- a. An initial stage along the lines of what is found in Eton (§4.8), where there is no system of FVs, and verb roots may or may not end in a vowel;
- b. The rise of a system of prosodic constraints on verbs placing restrictions on which vowels can appear in non-initial position, as seen in languages like Yasa (§4.4) or Eton (§4.8), and even requiring the presence of a vowel at the end of the verb, as seen in a language like Gyeli (§4.9);
- c. The integration of postverbal elements coding TAMP categories into the end of the verb which, in turn, became subject to prosodic constraints on verb shape, as seen in languages like Kpa (§4.6) and Kako (§4.11);
- d. The reduction of these elements to a vowel if they previously had a more complex shape, though only on a subset of verb forms, thus creating some classes of verbs appearing with FVs and some verbs appearing without them, as found, for example, in Myene (§4.12) and Punu (§4.15); and
- e. The analogical extension of these vocalic morphemes to all verb forms, even if some had previously not been part of the full set of historical processes outlined above, along the lines of the system described for Nzebi (§4.16) in this survey.⁴⁷

Even if the historical sketch just presented is a reasonable representation of the development of the canonical Bantu FV system, this does not mean that all of

⁴⁷ Although not part of the formal survey found in this chapter, the comparative description of Duma B51, Mbede B61 and Ndumu B63, provided in Adam (1954: 72), shows languages apparently representing the transition between stages 4 and 5, with Mbede showing CV roots without FVs, Ndumu showing these roots appearing with a FV, and Duma showing them appearing with an expansion suffix that hosts a FV.

the languages referenced above represent direct reflexes of this development. It could also very well be the case that some of the languages once had the canonical pattern and lost it due to phonological change connected to the prosodification of the verb, with a language like Gunu (§4.7) having a morphologically active FV, but without any morphological oppositions in the FV system, possibly representing a near-final stage of this process. Another significant concern is what the source would be of the postverbal elements which would have morphologised into FVs, but proposals for this would require a separate study.⁴⁸

Of the open questions listed in (7), the one where I think the historical situation is most easily reconstructed is the third one. The presence of lexical exceptions to Final patterns, in particular on CV stems, some of which are clearly quite old, as well as forms marked with the Passive, as discussed in §4.15, is most likely an archaism in my view. A historical scenario where FVs developed in longer verb systems, in part due to the fact that suffixes in these stems would have been subjected to more restrictive prosodic constraints, and were then extended to CV verbs, as well as verbs marked with the Passive, seems relatively historically plausible. By contrast, a scenario where CV verbs selectively lost FVs and then vocalised the glides that were part of the stem that would have appeared before a FV seems much less likely. If this is the right interpretation, it would mean that, even if the canonical Bantu FV system was largely in place for roots with CVC shape in PB, it had yet to fully extend to all verbs, suggesting a possibly interesting isogloss to look for carefully somewhere in the border between zone A and zone B languages.

Of the other questions listed in (7), the one that strikes me as most difficult to resolve is the second. In the languages which show CVCV roots where the last vowel is lexically specified (at least partially), where did the last vowel come from? There are various imaginable sources. They could be former FVs, with different FVs lexicalising across different verbs, perhaps due to varying frequency patterns for the use of each verb in different TAMP configurations. They could represent vowels derived from other sources, such as elements associated with the Pre-Final slot in Figure 1. They could also have arisen from postverbal elements beginning with a vowel, where the following vowel was reanalysed as belonging to the verb due to a reparsing, though developing this analysis would require determining what those elements might have been. In principle, they could also represent archaic elements which were lost in languages which developed

⁴⁸In this regard, Akoose (§4.2) suggests an interesting possibility that the source of FVs could be non-Final verbal suffixes such as *-ag, which took on reduced forms in longer verbs and then were reanalysed as morphologically distinct from their longer forms.

a canonical FV system due to elision effects, or some other phonological process along those lines. This last scenario strikes me as unlikely, given the comparative Bantu context, since it would require a major alteration in the reconstruction of PB verb roots, and I simply point it out here as a logical possibility.

Overall, I think the most significant result of this survey is that it points to the need for a more thorough consideration of the development of the canonical FV system in Bantu, since it is not obvious how such a system could have developed and, as the survey makes clear, patterns in north-western Bantu suggest that it may not have even fully developed within PB, i.e. ancestral node 1 in Grollemund et al. (2015), even if the seeds of the system must have already been in place. I should also stress that this is a domain where expanded data collection is likely to reveal interesting new facts, especially in zone A. The historical picture outlined above, for instance, is strongly influenced by the description of Finals in Kako, and to a lesser extent Kpa and Akoose, and it would be especially worthwhile to have a better sense of how many other languages in the north-west show systems like those described for these languages, which are, in my view, promising models for either an early stage of PB or some pre-PB variety from which the Bantu languages emerged. More broadly, this study underscores the need for revisions to PB reconstructions which might be biased towards central and eastern Bantu patterns (cf. e.g. Schadeberg 2003: 156). This appears to have been the case for Meeussen's (1967) reconstruction of the FV system as well.

This conclusion should also be considered in the context of the ongoing debate about the historical time depth of many of the features of the Bantu verbal system presented in Figure 1, in particular the verbal prefix system – see Güldemann (2011: 123–129) and Hyman (2011: 29–40) for relevant discussions, as well as Güldemann (2022 [this volume]) and Nurse & Watters (2022 [this volume]). The key question is whether the verbal prefixes, in particular those associated with slots -3 to 0 in Figure 1, should be treated as having already fully morphologised in PB or whether they were still expressed by morphosyntactically independent elements such as pronouns and auxiliary verbs, which would later develop into agreement markers and TAMP markers, respectively. If the FV system was not fully developed in PB, this would seem to be more in line with the position of Güldemann (2011) that the prefixes also represent a post-PB innovation insofar as both reconstructions point to a verbal system that was less morphologically elaborated and involved greater use of elements with some degree of syntactic independence than implied by Meeussen's (1967) reconstruction. If that is the case, however, it leaves open the important historical question of what might have triggered the processes of morphologisation that resulted in the development of

what has long been viewed as the canonical structure of the Bantu verb in such a large part of the Bantu-speaking area.

Acknowledgements

I would like to thank participants at the *International Conference on Reconstructing Proto-Bantu Grammar*, organised by the Ghent University Centre for Bantu Studies and the Service of Culture & Society of the Royal Museum for Central Africa in Tervuren, as well as attendees of a meeting of the Philological Society on May 1, 2020, for their feedback on the work discussed in this chapter in addition to two anonymous reviewers, Koen Bostoen, Rozenn Guérois, Larry M. Hyman, Dmitry Idiatov and Sara Pacchiarotti for more detailed comments.

Abbreviations

ANT	anterior
APPL	applicative
CAUS	causative
CONC	conclusive
DIM	diminutive
F1, F2	future tenses with remoteness distinctions, from more to less recent
FUT	future
FV	final vowel
H1, H2	floating high tone morphemes (Makaa)
HYP	hypothetical
I	noun class with shape <i>i</i> (Bijogo)
INC	incompletive
INCEP	inceptive
IPFV	imperfective
ITER	iterative
NC	Niger-Congo
NEG	negative
NTR	neutral
OP _N	object prefix with subscript indicator of noun class
P0, P1, P2, P3	past tenses with remoteness distinctions, from more to less recent
PFV	perfective
PB	Proto-Bantu
PL	plural

PRON	pronoun
PROG	progressive
PRS	present
PST	past
RET	retrospective
SBJV	subjunctive
SG	singular
SP _N	subject prefix with subscript indicator of noun class
TAMP	tense/aspect/mood/polarity marker
1 ... 18	noun classes
1/2/3 SG/PL	person and number

References

- Adam, Jean-Jérôme. 1954. *Grammaire composée : mbede, ndumu, duma* (Mémoires de l'Institut d'études centrafricaines 6). Montpellier: Imprimerie Charité.
- Atindogbé, Gratien Gualbert. 2013. A grammatical sketch of Mòkpè (Bakweri), Bantu A20. *African Study Monographs, Supplementary Issue* 45. 5–163.
- Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en bantou* (Annales – Série in-8° – Sciences humaines 114). Tervuren: Royal Museum for Central Africa.
- Bastin, Yvonne, André Coupeze, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bonneau, Joseph. 1956. *Grammaire pounoue et lexique pounou – français* (Mémoires de l'Institut d'études centrafricaines 8). Montpellier: Imprimerie Charité.
- Bostoën, Koen & Rozenn Guérois. 2022. Reconstructing suffixal phrasemes in Bantu verbal derivation. In Koen Bostoën, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 343–383. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575829.
- Bostoën, Koen & Léon Mundeke. 2011. The causative/applicative syncretism in Mbuun (Bantu B87, DRC): Semantic split or phonemic merger? *Journal of African Languages and Linguistics* 32(2). 179–218.
- Bôt, Dieudonné Martin Luther. 1998. Temps verbaux et aspects du yasa. *Afrikanistische Arbeitspapiere* 53. 47–65.

4 Reconstructing the development of the Bantu final vowels

- Bôt, Dieudonné Martin Luther. 2011. Les verbes dérivés du yasa. *Journal of West African Languages* 38(1). 89–98.
- Crane, Thera M., Larry M. Hyman & Simon Nsielanga Tukumu. 2011. *A grammar of Nzadi [B.865]: A Bantu language of the Democratic Republic of Congo* (University of California Publications in Linguistics 147). Berkeley, CA: University of California Press.
- de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87–162 + supplementary material online.
- Dom, Sebastian, Leonid Kulikov & Koen Bostoen. 2016. The middle as a voice category in Bantu: Setting the stage for further research. *Lingua Posnaniensis* 59(2). 129–149.
- Downing, Laura J. 2006. *Canonical forms in prosodic morphology* (Oxford Studies in Theoretical Linguistics 12). Oxford: Oxford University Press.
- Dugast, Idelette. 1967. *Lexique de la langue tûnen (parler des Banen du Sud-Ouest du Cameroun)* (Langues et littératures de l’Afrique noire 2). Paris: Librairie C. Klincksieck.
- Dugast, Idelette. 1971. *Grammaire du tûnen* (Langues et littératures de l’Afrique noire 8). Paris: Librairie C. Klincksieck.
- Ellington, John Ernest. 1977. *Aspects of the Tiene language*. Madison, WI: University of Wisconsin–Madison. (Doctoral dissertation).
- Ernst, Urs. 1991. Temps et aspect en kakɔ. In Stephen C. Anderson & Bernard Comrie (eds.), *Tense and aspect in eight languages of Cameroon* (SIL International Publications in Linguistics 99), 17–45. Dallas, TX: Summer Institute of Linguistics & The University of Texas at Arlington.
- Ernst, Urs. 1995. *Les formes verbales du kakó*. Yaoundé: SIL Cameroon.
- Fontaney, V. Louise. 1980. Le verbe. In François Nsuka-Nkutsi (ed.), *Éléments de description du punu* (Linguistique et sémiologie. Travaux du centre de recherches linguistiques et sémiologiques de Lyon), 51–114. Lyon: Université Lumière – Lyon 2.
- Gautier, Jean-Marie. 1912. *Grammaire de la langue mpongwée*. Paris: Procure des PP. du Saint-Esprit.
- Gerhardt, Ludwig. 1982. Jarawan Bantu: The mistaken identity of the Bantu who turned north. *Afrika und Übersee: Sprachen, Kulturen* 65(1). 75–87.
- Good, Jeff. 2007. Slouching towards deponency: A family of mismatches in the Bantu verb stem. In Matthew Baerman, Greville G. Corbett, Dunstan Brown & Andrew Hippisley (eds.), *Deponency and morphological mismatches* (Proceedings of the British Academy 145), 203–230. Oxford: Oxford University Press.

- Good, Jeff. 2012. How to become a 'Kwa' noun. *Morphology* 22(2). 293–335.
- Good, Jeff. 2016. *The linguistic typology of templates*. Cambridge: Cambridge University Press.
- Good, Jeff. 2017. Niger-Congo languages. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics* (Cambridge Handbooks in Language and Linguistics), 471–499. Cambridge: Cambridge University Press.
- Grégoire, Claire. 1979. Les voyelles finales alternantes dans la conjugaison affirmative des langues bantoues centrales. *Journal of African Languages and Linguistics* 1(2). 141–172.
- Grimm, Nadine. 2015. *A grammar of Gyeli*. Berlin: Humboldt University of Berlin. (Doctoral dissertation).
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guarisma, Gladys. 1992. *La bafia (rì-kpā?) : langue bantoue (A53) du Cameroun*. Paris: Université Paris V 'René Descartes'. (Doctoral dissertation).
- Guarisma, Gladys. 2003. Kpā? (A53). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 307–334. London: Routledge.
- Güldemann, Tom. 2003. Grammaticalization. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 182–194. London: Routledge.
- Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas: With special reference to Africa* (Tokyo University of Foreign Studies, Studies in Linguistics 2), 109–141. Amsterdam: John Benjamins.
- Güldemann, Tom. 2022. Predicate structure and argument indexing in early Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 387–421. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575831.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1967. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 1: The comparative linguistics of the Bantu languages*. Farnborough: Gregg Press.
- Guthrie, Malcolm. 1968. Notes on Nzébi (Gabon). *Journal of African Languages* 7(2). 101–129.

4 Reconstructing the development of the Bantu final vowels

- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Hawkinson, Ann Katherine. 1986. *Bakweri verb morphology*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).
- Heath, Daniel. 1991. Tense and aspect in Makaa. In Stephen C. Anderson & Bernard Comrie (eds.), *Tense and aspect in eight languages of Cameroon* (SIL International Publications in Linguistics 99), 3–15. Dallas, TX: Summer Institute of Linguistics & The University of Texas at Arlington.
- Hedinger, Robert. 1985. The verb in Akɔɔse. *Studies in African Linguistics* 16(1). 1–55.
- Hedinger, Robert. 2008. *A grammar of Akɔɔse: A Northwest Bantu language* (SIL International Publications in Linguistics 143). Dallas, TX: SIL International.
- Hyman, Larry M. 1993. Conceptual issues in the comparative study of the Bantu verb stem. In Salikoko S. Mufwene & Lioba Moshi (eds.), *Topics in African linguistics. Papers from the XXI Annual Conference on African Linguistics, University of Georgia, April 1990* (Current Issues in Linguistic Theory 100), 3–34. Amsterdam: John Benjamins.
- Hyman, Larry M. 1998. Positional prominence and the ‘prosodic trough’ in Yaka. *Phonology* 15(1). 41–75.
- Hyman, Larry M. 2003a. Segmental phonology. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 42–58. London: Routledge.
- Hyman, Larry M. 2003b. Sound change, misanalysis, and analogy in the Bantu causative. *Journal of African Languages and Linguistics* 24(1). 55–90.
- Hyman, Larry M. 2004. How to become a ‘Kwa’ verb. *Journal of West African Languages* 30(2). 69–88.
- Hyman, Larry M. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville, MA: Cascadilla Press.
- Hyman, Larry M. 2008. Directional asymmetries in the morphology and phonology of words, with special reference to Bantu. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 46(2). 309–350.
- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.
- Hyman, Larry M. 2017. Disentangling conjoint, disjoint, metatony, tone cases, augments, prosody, and focus in Bantu. In Jenneke van der Wal & Larry M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu* (Trends in Linguistics. Studies and Monographs 301), 100–121. Berlin: Walter de Gruyter.

- Kwenzi-Mikala, Jérôme T. 1980. Esquisse phonologique du punu. In François Nsuka-Nkutsi (ed.), *Eléments de description du punu* (Linguistique et sémiologie. Travaux du centre de recherches linguistiques et sémiologiques de Lyon), 7–18. Lyon: Université Lumière – Lyon 2.
- Marchal-Nasse, Colette. 1989. *De la phonologie à la morphologie du nzebi, langue bantoue (B52) du Gabon*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- Marlo, Michael R. & David Odden. 2007. The exponence of TAM in Bakweri. In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 19–31. London: Department of Linguistics, School of Oriental & African Studies.
- McGill, Stuart. 2009. *Gender and person agreement in Cicipu discourse*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Mchombo, Sam A. 2004. *The syntax of Chichewa* (Cambridge Syntax Guides). Cambridge: Cambridge University Press.
- Meeussen, Achiel Emiel. 1961. Le ton des extensions verbales en bantou. *Orbis, Bulletin international de documentation linguistique* 10(2). 424–427.
- Meeussen, Achiel Emiel. 1962. De tonen van subjunctief en imperatief in het Bantoe. *Africana Linguistica* 1. 57–74.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeussen, Achiel Emiel. 2014. Tones of the subjunctive and the imperative in Bantu. *Africana Linguistica* 20. 15–38.
- Moroz, George, Kirill Koncha, Mikhail Leonov, Anna Smirnova & Sasha Shakhnova. 2022. *Lingtypology: Linguistic typology and mapping*. <https://CRAN.R-project.org/package=lingtypology>.
- Mous, Maarten. 2003. Nen (A44). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 283–306. London: Routledge.
- Nsuka-Nkutsi, François. 1980. Quelques réflexes du proto-bantou en punu. In François Nsuka-Nkutsi (ed.), *Eléments de description du punu* (Linguistique et sémiologie. Travaux du centre de recherches linguistiques et sémiologiques de Lyon), 129–178. Lyon: Université Lumière – Lyon 2.
- Nurse, Derek. 2007. Did the Proto-Bantu verb have a synthetic or an analytic structure? In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 239–256. London: Department of Linguistics, School of Oriental & African Studies.

4 Reconstructing the development of the Bantu final vowels

- Nurse, Derek. 2008. *Tense and aspect in Bantu* (Oxford Linguistics). New York, NY: Oxford University Press.
- Nurse, Derek & Gérard Philippson. 2003. Introduction. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 1–12. London: Routledge.
- Nurse, Derek & Gérard Philippson. 2006. Common tense-aspect markers in Bantu. *Journal of African Languages and Linguistics* 27(2). 155–196.
- Nurse, Derek, Sarah Rose & John Hewson. 2016. *Tense and aspect in Niger-Congo* (Documents on Social Sciences and Humanities). Tervuren: Royal Museum for Central Africa. <https://www.africamuseum.be/sites/default/files/media/docs/research/publications/rmca/online/documents-social-sciences-humanities/tense-aspect-niger-congo.pdf>.
- Nurse, Derek & John R. Watters. 2022. Tense in Proto-Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 105–171. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575819.
- Odden, David & Lee S. Bickmore. 2014. Melodic tone in Bantu: Overview. *Africana Linguistica* 20. 3–13.
- Orwig, Carol. 1989. Les extensions verbales en nuguunu. In Daniel Barreteau & Robert Hedinger (eds.), *Descriptions de langues camerounaises* (DELAN: Description systématique des langues nationales), 283–314. Paris: SIL & ACCT & ORSTOM.
- Pacchiarotti, Sara & Koen Bostoen. 2021. Final vowel loss in Lower Kasai Bantu (DRC) as a contact-induced change. *Journal of Language Contact* 14(2). 438–476.
- Pacchiarotti, Sara, Natalia Chousou-Polydouri & Koen Bostoen. 2019. Untangling the West-Coastal Bantu mess: Identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica* 25. 155–229.
- Piron, Pascale. 1990. *Éléments de description du kota, langue bantoue du Gabon, B.25*. Brussels: Université libre de Bruxelles. (MA thesis).
- Pozdniakov, Konstantin. 2016. From Proto-Atlantic to Proto-Niger-Congo: The root structure. Paper presented at the Second International Congress “Towards Proto-Niger-Congo: Comparison and Reconstruction”, LLACAN, Paris, 1–3 September 2016.
- Rekanga, Jean-Paul. 2000. *Essai de grammaire himba (langue bantoue du Gabon, B36), Tome II : morphologie*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- Rottland, Franz. 1970. *Die Verbformen des Yanzi*. Leiden: Leiden University. (Doctoral dissertation).

- Schadeberg, Thilo C. 2003. Historical linguistics. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 143–163. London: Routledge.
- Schadeberg, Thilo C. & Koen Bostoen. 2019. Word formation. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 2nd edn. (Routledge Language Family Series), 172–203. Milton Park, Abingdon: Routledge.
- Sebasoni, Servilien. 1967. La préfinale du verbe bantou. *Africana Linguistica* 3. 122–135.
- Segerer, Guillaume. 2000. *Description de la langue bijogo (Guinée Bissau)*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Segerer, Guillaume. 2002. *La langue bijogo de Bubaque (Guinée Bissau)* (Afrique et Langage 3). Louvain: Peeters.
- Stilo, Donald L. 2005. Iranian as a buffer zone between the universal typologies of Turkic and Semitic. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion. Case studies from Iranian, Semitic and Turkic*, 35–63. Abingdon: RoutledgeCurzon.
- Van de Velde, Mark. 2008. *A grammar of Eton* (Mouton Grammar Library 46). Berlin: Mouton de Gruyter.
- van Eeden, Bernardus Izak Christiaan. 1934. The terminating vowel of the Bantu verbal stem. *Bantu Studies, A Journal Devoted to the Scientific Study of Bantu, Hottentot and Bushmen* 8(1). 367–375.
- Vansina, Jan. 1995. New linguistic evidence and ‘the Bantu expansion’. *Journal of African History* 36(2). 173–195.
- Voeltz, Erhard F. K. 1977. *Proto Niger-Congo verb extensions*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Welmers, William E. 1973. *African language structures*. Berkeley, CA: University of California Press.
- Yukawa, Yasutoshi. 1992. A tonological study of Kaka verbs. In Ryohei Kagaya (ed.), *Studies in Cameroonian and Zairean languages* (Bantu Linguistics 3), 47–66. Tokyo: Institute for the Study of Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.

Chapter 5

The relevance of Bantoid for the reconstruction of Proto-Bantu verbal extensions

Roger M. Blench

Kay Williamson Educational Foundation

In this chapter the relevance of Bantoid for the reconstruction of verbal extensions in Proto-Bantu (PB) is assessed. The Bantoid or Wide Bantu languages are a body of some 150–200 languages positioned geographically between Nigeria and Cameroon. They do not form a genetic subgroup, but all are in some way related to Narrow Bantu, i.e. Bantu as referentially classified by Guthrie (1948; 1967–71), more closely than other branches within Benue-Congo. The most well-known subgroups are Dakoid, Mambiloid, Tivoid, Beoid, Grassfields, and Mbe-Ekoid. The chapter discusses the characteristics of verbal extensions in Bantoid and their possible relation to extensions attested in Narrow Bantu on the one hand, and in other branches of Benue-Congo on the other hand. Based on a review of the literature on verbal extensions in the various branches of Bantoid and on case studies of individual languages, the chapter concludes that a rich system similar to Narrow Bantu can be reconstructed for Proto-Grassfields, while in other Bantoid subgroups, it is now lost or much reduced. Only the causative *-si* is attested in a substantial number of subgroups. Some Bantoid extensions show significant segmental similarities to certain extensions in Narrow Bantu zone A languages, which have never been reconstructed for PB. It is argued that these extensions shared between the highest branches of the Bantu family tree warrant a revision of PB verb derivation suffixes.

1 Introduction

The Bantoid languages are a body of some 150–200 languages positioned geographically between Nigeria and Cameroon. There is no evidence they form a



genetic subgroup, although they are all in some way related to Narrow Bantu more closely than to the rest of Benue-Congo. The most well-known Bantoid subgroups are Dakoid, Mambiloid, Tivoid, Beboïd, Grassfields and Ekoid. Bendi, formerly classified as Cross River, may also be Bantoid. Jarawan is sometimes claimed to be Narrow Bantu instead of Bantoid (or Wide Bantu). The division between (Narrow) Bantu and Bantoid used in this chapter considers that (Narrow) Bantu consists of the subgroups as defined in the referential classification of Guthrie (1967–71).

Both (Narrow) Bantu and Bantoid are characterised by systems of nominal affixes and alliterative concord, although these are highly eroded in some languages. However, Bantoid noun morphology is not that of classic Bantu, despite its prefixes being often ascribed the same class numbers in a somewhat misleading way. Bantoid does not represent a genetic group, although the languages are related. It is simply a cover term for those subgroups which split away from Benue-Congo before the genesis of Narrow Bantu (Blench 2015). Even the division between Bantu and Bantoid is now often questioned, as some authors have observed that much of Bantu A, with its highly reduced noun classes, would perhaps be better treated as Bantoid.

Apart from noun classes, one of the characteristic features ascribed to Proto-Bantu (PB) is its system of verbal extensions (Schadeberg 2003). These are (V)(C)V elements which are (usually) suffixed to the verb stem, and in some languages can be stacked in complex strings. They can transform the semantics and syntax of the verb, marking number, directionality, or reflexivity and bring about other changes, as well as denote some types of aspectual marking. Verbal extension morphology can almost certainly be traced back considerably further in Niger-Congo (e.g. Voeltz 1977; Trithart 1983; Hyman 2007; 2014). Such suffixes are present in some form in many Niger-Congo branches, though not in Mande, some branches of Kordofanian, Dogon and Ijò. Ijò, intriguingly, does have a small repertoire of verbal extensions synchronically, but these show no segmental cognacy with other branches of Niger-Congo (Kay Williamson, p.c.). Whether these should be reconstructed to Proto-Niger-Congo depends on what internal structure is claimed for the phylum. Similarly, the state of scholarship is not such that we can easily assert that particular segmental features can be reconstructed. Hyman (2014) discusses the uneven distribution of verbal extensions in the different branches of Niger-Congo and the extent to which they reflect those found in Bantu.

In any event, it is reasonable to assume that several of the extensions reconstructed for PB go back to Proto-Benue-Congo. Benue-Congo is of considerable importance, because some languages exhibit features which resurface in Bantu,

but which are only attested in fragmentary form or not at all in Bantoid. In most branches of Benue-Congo these have become unproductive, becoming incorporated in roots. Nowhere in Bantoid are these systems wholly functional, but their former presence can be detected from the presence of “frozen” morphemes. Hyman (2017a) addresses this issue in what he terms *from syntheticity to analyticity* and discusses the way in “which [Bantoid] languages compensate for the loss of valence-adding extensions, e.g. the applicative, which has multiple functions in Common Bantu”. He identifies periphrasis, unmarked double objects, adpositions and nominal constructions as strategies for dealing with the loss of verbal extensions. Table 1, adapted from Hyman (2017a: Table 3), summarises the sort of contrasts which can be expected.

Table 1: Canonical Bantu compared with Bantoid (Hyman 2017a: Table 3)

Feature	Canonical Bantu	Bantoid
phonology	minimum word = 2 syllables	maximum stem = mostly 2~3 syllables
morphology	highly synthetic, agglutinative	less so, gradual move towards analyticity
verb extensions	many, mostly marking valence	few, mostly marking aspect
unmarked objects	multiple	at most two, ultimate limitation to one per verb
object marking	head marking on verb	various prepositions and/or serial verbs
ditransitive verbs	a few (<i>*pá</i> ‘give’)	few or none

The concern of this chapter is primarily with identifying the trail of evidence that links segmental evidence for existing or former extensions in Bantoid with those in Bantu. Although a standard list of proposed reconstructed verbal extensions exists for PB, comparative data from Guthrie’s zone A and closely related Bantoid languages provide only limited support for the proposed forms.

The definition of extensions varies from author to author, and in the maximal interpretation it is any suffix on a verb, including tense/aspect markers. In

the Bantoid region, many languages have verbs with unproductive suffixes that have no assignable meaning. The hypothesis is that these are the traces of now fossilised extensions, although this claim would need to be supported by the semantics of synchronic verb forms. Note that in some languages, changes in meaning similar to those brought about by extensions occur through tonal change. It seems reasonable to include these in a list of extensions (Hyman 2017b). In certain languages, such as Vute, innovative extension-like suffixes originate in Serial Verb Constructions. Over time, these forms may be lexicalised to merge with the set of authentic extensions. Productive extensions are those for which there is evidence that they have an assignable semantics and can be suffixed to roots as part of the derivational process in speech.

The verbal extensions of PB have generated a considerable literature. The first discussions of these go back to Meinhof (1899; 1910) and the *Bantu Grammatical Reconstructions* of Meeussen (1967). The literature on this is summarised in Schadeberg (2003: 72) whose list of proposed reconstructions, reproduced in Table 2, is still the most widely cited (see also Schadeberg & Bostoen 2019: 173).

Table 2: Proto-Bantu verbal extensions

Proto-Bantu	Semantics
*-i/-ici	causative
*-id	applicative
*-ik	impositive
*-ik	neuter
*-am	positional, stative
*-an	associative, reciprocal
*-a(n)g	repetitive
*-al	extensive
*-at	tentative, contactive
-ɔ/-ibɔ	passive
*-ɔl/-ɔk	reversive

This system is relatively rich and has the potential for stacking. In certain Bantu languages, up to four extensions can be added to the stem to generate very specific subsets of meaning. The analytic question is the extent to which these can be linked to extensions attested for Bantoid, or further back, for Benue-Congo. Since Bantoid is a key element in understanding the genesis of Bantu verbal extensions, this chapter summarises the presence or absence, morphology and semantics of extensions in the Bantoid languages. Hyman (2018) has

reviewed verb extensions in some Bantoid branches with a view to reconstruction, although the coverage is far from comprehensive. Hyman (2018: 176) divides these into three classes: (1) productive extensions; (2) unproductive extensions often restricted to post-radical position or specific combinations; and (3) frozen, mostly unidentifiable -VC- expansions. He also suggests lists of allomorphs of the forms cited in Table 2.

Since extensions preserved in some branches strongly resemble Bantu, this chapter also considers briefly the relationship of Bantoid to Benue-Congo (§2). Overall, Bantoid languages are poorly documented, so in §2.3 time is given to discussing the question of internal classification and data sources. The ancestry of the characteristic extensions in Bantoid can be traced in Benue-Congo languages as discussed in §3.1. Existing information about the presence or absence of extensions in the established branches of Bantoid is summarised in §3.2. Case studies of synchronic extensions are presented in §3.3 which includes a section on Bantu zone A languages. The conclusion summarises the evidence presented and considers this evidence for the historical origin of attested Bantu extensions.

2 Classification of Bantoid

2.1 Bantoid (Wide Bantu) vs. (Narrow) Bantu opposition

Sigismund Koelle (1854) and Wilhelm Bleek (1862–69) noted that many languages of West Africa also showed noun classes marked by prefixes, and Bleek went so far as to include a “West-African” division in the family he named Bantu. According to Jungraithmayr & Möhlig (1983), the term “Bantoid” was introduced by Krause (1895), but it seems to have been subsequently forgotten. It re-appears in Guthrie (1948; 1967–71) to describe what he called “transitional” languages, replacing the vague term “Semi-Bantu”, which goes back to Johnston (1919–22). The modern sense of the term Bantoid to refer to Bantu-like languages of the Nigerian-Cameroon borderland may have first appeared in Jacquot & Richardson (1956). This includes summary sketches of Nyang, Ekoid, Tikar and Grassfields languages, although the volume as a whole also incorporates material on Narrow Bantu and a variety of Adamawa and Ubangian languages so it is rather unspecific.

Despite the discussion in Johnston (1919–22) and Guthrie (1967–71) of the place of Bantoid languages with apparent correspondences to Bantu, it was Greenberg (1963, 1974) who first emphasised the issue of genetic classification as opposed to typology. He treated Bantu as one branch of Benue-Congo, i.e. the adjacent languages of southern and eastern Nigeria and Cameroon. He says:

“the Bantu languages are simply a subgroup of an already established genetic subfamily of Western Sudanic [i.e. Niger-Congo, broadly speaking]” (Greenberg 1963: 32). Figure 1 shows Greenberg’s classification.

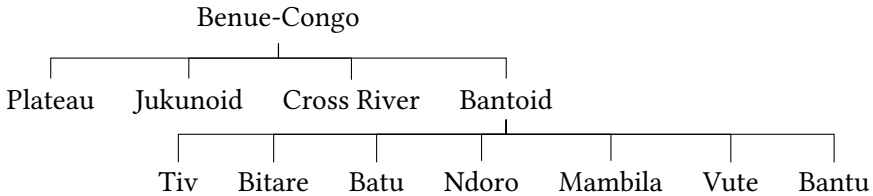


Figure 1: Greenberg’s (1963) situating of Bantu

Greenberg (1963: 35) also clearly stated that “supposedly transitional languages are really Bantu”. In other words, many languages lacking some features typical of Bantu are nonetheless related to it. This approach to Bantu was refreshing and made historical sense in a way that Guthrie’s views never had. But since the 1960s, data has gradually accumulated on the vast and complex array of languages in the “Bantu borderland”, i.e. the region between southern Cameroon (where Guthrie’s Bantu begins) and eastern Nigeria. The next step in the evolution of our understanding of Bantoid was the formation of the Grassfields Working Group in the early 1970s. Many of these findings were summarised in overview articles from this period, including Hedinger (1989) and Watters & Leroy (1989a,b).

Bantoid and Bantu represent nested subsets of Benue-Congo, a large and complex group of languages, whose exact membership remains disputed. Originating with Westermann’s (1927) *Benue-Cross-Fluss*, it took shape in Greenberg (1963), Williamson (1971) and de Wolf (1971). The name “Benue-Congo” was introduced by Greenberg (1963) who proposed a division into four branches: Plateau, Jukunoid, Cross River, and Bantoid. For a period in the 1980s and 1990s, it was considered that all the languages in the former “Eastern Kwa”, i.e. Yoruboid, Igboïd, Nupoid etc. were part of Benue-Congo, i.e. Western Benue-Congo. However, the evidence for this was never published and, in my view, it seems easier to revert to Benue-Congo as in Greenberg’s original, with the potential addition of Ukaan, a small cluster of languages spoken south-west of the Niger-Benue confluence. Ukaan has alternating prefixes (i.e. those which change on a predictable basis), marking number and concord, as well as some segmental cognates, hence its likely affiliation with Benue-Congo, but its exact position remains to be determined. With this in mind, Figure 2 provides a schematic representation of my current understanding of the sub-classification of Benue-Congo languages as the result of numerous years of research on many of those languages.

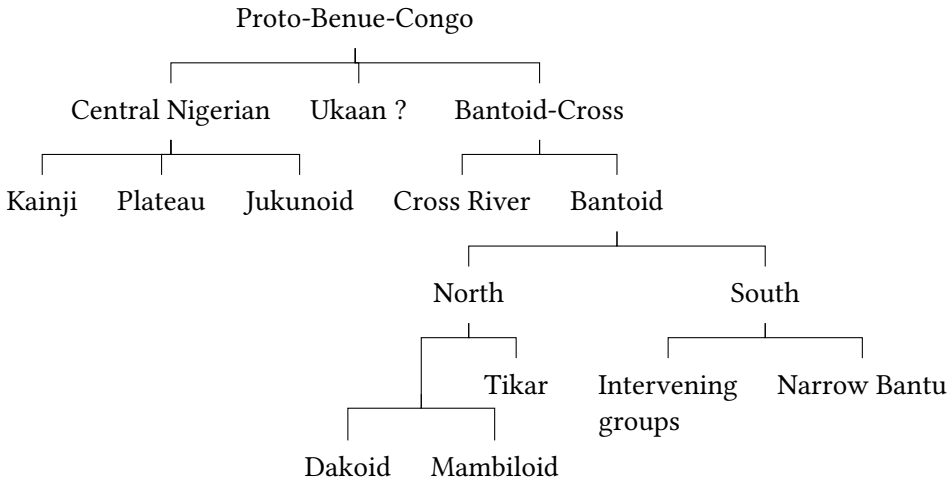


Figure 2: Revised sub-classification of Benue-Congo languages

It is emphasised strongly that no claim is made for Bantoid as a genetic group; it is rather a referential term covering all languages with a discernible relationship to Narrow Bantu. Bendi, previously considered part of Cross River, has been shifted to Bantoid, a change of affiliation proposed by Blench (2001).

2.2 Membership of Bantoid

Although (Narrow) Bantu has been treated as a genetic unity since the middle of the nineteenth century, it is unlikely there is any distinctive boundary between Bantu and the languages related to it. As Bostoen & Van de Velde (2019) note, no lexical or morphological isoglosses have been identified that clearly demarcate Bantu from its closest relatives. Figure 2 shows the subgroups that “stand between” Bantoid-Cross and Narrow Bantu. The languages represented are very numerous (150 ~ 200) and also highly diverse morphologically. New languages are likely to be discovered and more work in historical reconstruction will improve our understanding of how these languages relate to one another. This section lists the major Bantoid subgroups as presently understood. Most of these groups are uncontroversial, although the genealogical validity of poorly documented isolate branches, such as Buru, which may either be Tivoid or an independent branch, need more study. A more complete list of the languages which Bantoid includes is given in the Ethnologue (Eberhard et al. 2022) and Glottolog (Hammarström et al. 2021). In the absence of more in-depth historical-linguistic

research, I assume that individual groups split away from a common stem, and developed their own characteristics. The order in which this took place remains controversial, and will take considerable further work to resolve in a satisfying manner.

One particular aspect of Figure 2 requires further consideration, namely the division of Bantoid into North and South. Dakoid, Mambiloid and Tikar represent language groups with either no noun classes, or relics of a divergent system, as in Tikar. I believe that these three should be classified together as “North Bantoid”. However, the lack of data for some languages and convincing reconstructions of their historical morphology makes this at best a speculative hypothesis. The other side is “South Bantoid”, which is not a discrete branch in itself, but just a convenient cover term for Narrow Bantu and its closest relatives that do not belong to Dakoid, Mambiloid and Tikar. A proposal for the stepwise branching of different “Southern Bantoid” subgroups is presented in Figure 3. Narrow Bantu is depicted here conventionally as a separate subgroup, although several lexicon-based classifications, such as Piron (1997; 1998), Grollemund et al. (2015) and Grollemund et al. (2018), point out that it is genealogically not discrete from Grassfields and Jarawan Bantu.

Table 3 lists the major subgroups of Bantoid following the order in which I believe them to have diverged from Benue-Congo.

It is important to flag some caveats. Not all authors agree that Dakoid is Bantoid (e.g. Boyd 1994; 1997) and the placing of Ndoro in Mambiloid remains doubtful. Bendi has long been treated as Cross River following Greenberg (1963) and Williamson (1989), but without good evidence. The data on Furu is too uncertain to be sure whether it has been correctly classified; a Jukunoid affiliation is possible. Jeff Good and his colleagues have argued convincingly that Beboïd is not a unity, and even that the languages within Yemne-Kimbi (= formerly West Beboïd) may not constitute a genetic group (Good et al. 2011). Ambele and Menchum are treated as co-ordinate with Grassfields, but the evidence remains sketchy. Momo has been split up into Momo proper and South-West Grassfields. The evidence for the placing of Jarawan, treated in previous texts as Bantoid, remains controversial. Lexically, it is more closely related to Narrow Bantu languages, perhaps Guthrie’s A60 group (cf. Piron 1997; Grollemund et al. 2015), but the loss of both verbal and nominal morphology makes its integration into Narrow Bantu uncertain.¹ An alternative interpretation could be that this loss is a later areal feature.

¹A striking disagreement over the classification of Jarawan Bantu was aired at the First Bantoid Conference held in Hamburg in March 2022. Contrary to the present author’s claim of an A60 affiliation, Van de Velde & Idiatov (2022) argued for A80-A90, while Jeffrey Wills and Rebecca Grollemund (p.c.) assign Jarawan to Bantoid. Clearly this argument has some way to go.

5 *Reconstruction of Proto-Bantu verbal extensions with Bantoid*

Table 3: Major subgroups of Bantoid

Group	Country	Location	Representative language(s)
Dakoid	NGA	around Ganye	Daka, Taram, Tiba
Mambiloid	NGA/CMR	around Gembu	Mambila, Kwanja, Vute, Ndro
Tikar	CMR	NE of Fouban	(three dialects)
Bendi	NGA	around Ogoja	Bokyi, Bekwara, Alege
Tivoid	NGA/CMR	around Obudu	Tiv, Iyive, Ugara
Buru	NGA	Buru	Buru
Furu	NGA/CMR	Furu Awa	Furu
East Beboid	CMR	around Nkambe	Noone, ^a Ncane
Yemne-Kimbi	CMR	NE Grassfields	Fungom, Mundabli
Nyang	CMR	Mamfe	Kenyang
Ekoid	NGA/CMR	Mamfe	Ejagham, Etung
Mbe	NGA	Ogoja	Mbe
Ambele	CMR	Grassfields	Ambele
Menchum	CMR	Grassfields	Menchum
Grassfields	CMR		
• Ndemli	CMR	Nkam, Littoral region	Ndemli
• Ring	CMR	Grassfields	
– Centre	CMR	Grassfields	Babanki, Kom, Mmen, Oku
– East	CMR	Grassfields	Lamnso?
– South	CMR	Grassfields	Bamunka
– West	CMR	Grassfields	Aghem, Isu
• Momo	CMR	Grassfields	Moghamo
• South-West	CMR	Grassfields	Manta
• Eastern	CMR	Grassfields	
–Bamileke	CMR	Grassfields	Bamileke, Ngiemboon, Ngomba
–Ngemba	CMR	Grassfields	Bafut, Mankon, Ngemba
–Nkambe	CMR	Grassfields	Limum, Mfumte, Yamba
Jarawan	NGA/CMR	East-Central Nigeria	Jar, Mbula-Bwazza, Mama
Bantu zone A	CMR	Southern Cameroon	Akoose A15C

^aThis language name is spelt in various ways (Noni, Nooni) in bibliographic references and even within the Noone community.

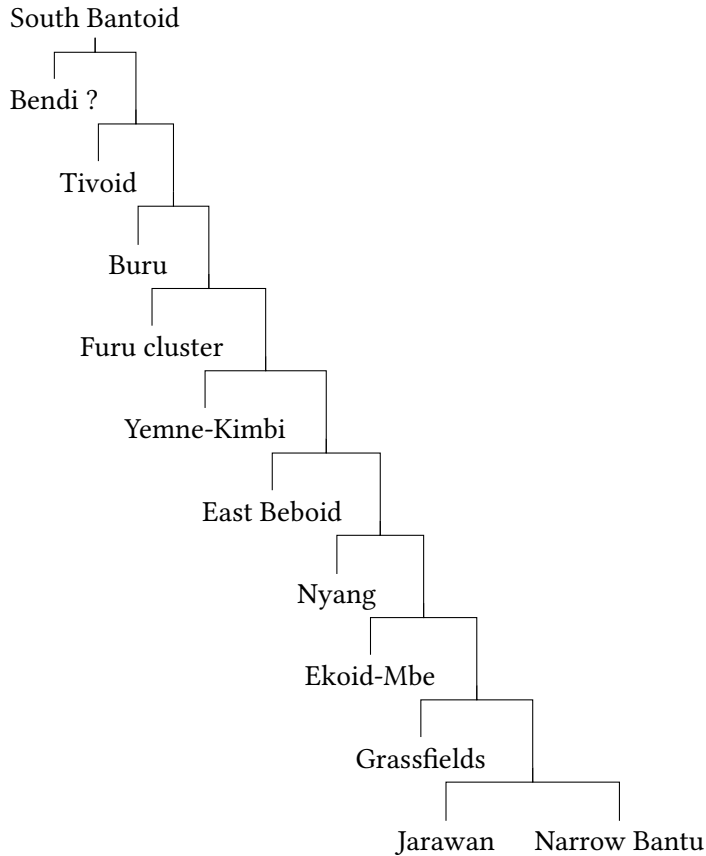


Figure 3: Proposal for the stepwise divergence of Bantoid languages

Common to this body of work is that the classifications were presented with limited justification. This is perhaps unsurprising as the number of languages is very large and many were poorly known, then and still today. Piron (1997) and Bastin & Piron (1999) represent classifications of Bantoid using lexicostatistics. The PhD thesis of Grollemund (2012) applies more recent statistical techniques to basic vocabulary for the classification of Bantu and Bantoid, but its focus is on Narrow Bantu with a random sample of South Bantoid languages. Blench (2015) is the only overview of all families which, in my view, can be assigned to Bantoid.²

²Overviews of the major Bantoid branches, together with wordlists of isolates such as Buru, and arguments for their coherence, can be found on the relevant page of my website: <http://www.rogerblench.info/Language/Niger-Congo/Bantoid/BantdOP.htm>.

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

Maps of the main Bantoid groups are provided in the relevant sections below. These are in the main based on those available on the relevant Wikipedia pages which are in turn redrawn from the Ethnologue. However, where errors were spotted, for example in the Beboid and Dakoid maps, these have been redrawn to reflect current understanding. Tivoid is shown in Figure 4, together with the unclassified Esimbi and Buru.

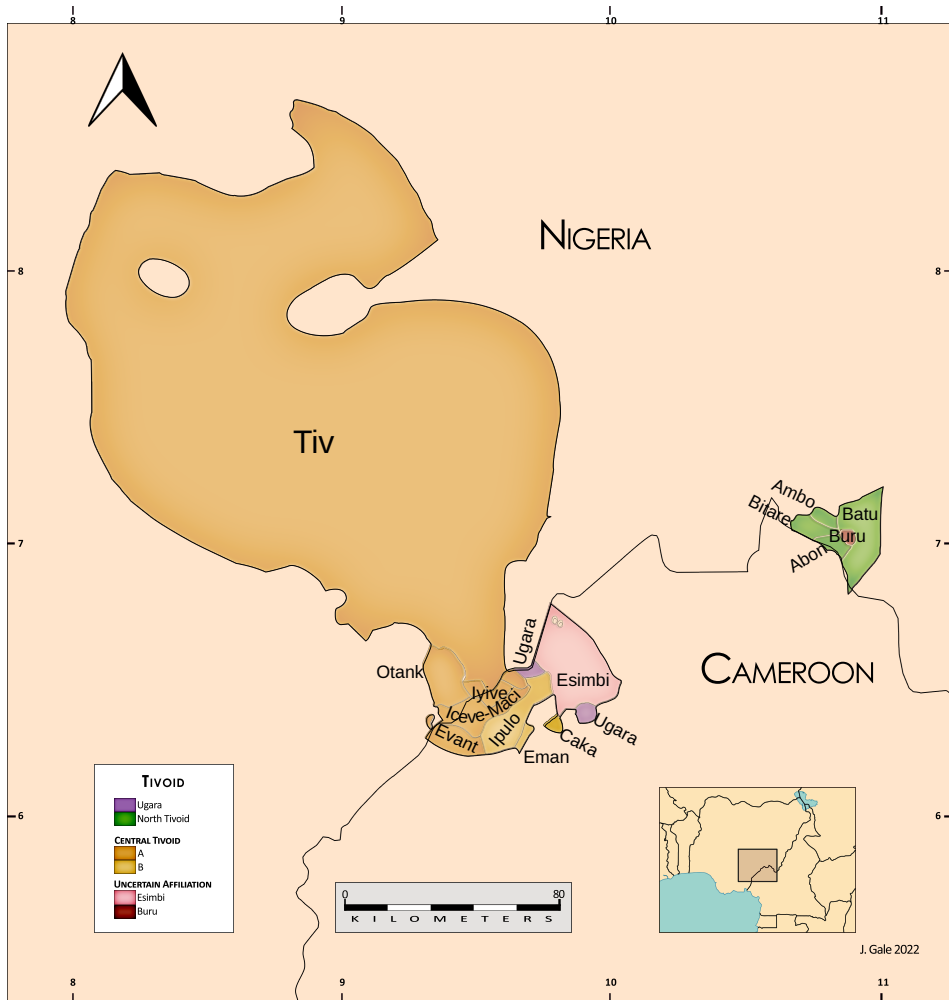


Figure 4: Map of the Tivoid languages, together with Esirobi and Buru

A feature of the Bantoid area is intensive borrowing, both between closely related languages and between different branches of Bantoid. Bantoid languages are largely found in an area of high density settlement, linked by complex trade networks and long noted for extensive multilingualism. Warnier (1979) analysed this in respect of another grammatical feature, viz. noun classes, noting their extensive borrowing and consequent morphological re-analysis. More recently, Di Carlo et al. (2018), Di Carlo et al. (2019) and Di Carlo et al. (2020) have reviewed multilingualism in Africa in general, but also focused on the Lower Fungom area of the Grassfields, where the details of language interaction can be analysed at the micro-level. This type of multilingualism, which involves borrowing grammatical features as well as vocabulary, goes a long way to explaining why verbal extensions in Bantoid do not form tidy patterns.

2.3 Overview of the data sources

The descriptive data required to characterise Bantoid languages in ways which would satisfy historical linguists and typologists is not available for many branches. The literature on many subgroups is sparse, to say the least, and many important sources are unpublished. Because so much of the material has focused on an ultimate goal of orthography and literacy, phonology and noun classes remain much better understood than, for example, verbal extensions.

There are two key caches of unpublished and mainly electronic data, the files of SIL International – which incorporates much of the data collected for ALCAM, the *Linguistic Atlas of Cameroon* (Dieu & Renaud 1983) – and the student dissertations supervised at the University of Yaoundé I. Part of the legacy material is available on the SIL Cameroon website (<https://www.silcam.org/>) although much material, especially fieldwork lexicons, remain in the hands of its members.³ Wycliffe Nigeria has recently undertaken surveys of the Bantoid languages on the Nigerian side of the border, resolving numerous queries about the extent and classification of particular branches.⁴ Jeff Good has facilitated the scanning of University of Yaoundé I theses in linguistics up to 2006, and these are now available electronically.

³Thanks to Robert Hedinger for making this material available.

⁴Materials from Nigeria created by SIL survey staff are available on personal application.

3 Bantoid verbal extensions

3.1 Verbal extensions in Benue-Congo

To assess the time depth of possible verb extensions in Bantoid, their historical origin can be explored within Benue-Congo. However, much of Benue-Congo, including Plateau, Jukunoid and Cross River, retains only traces of a verbal-extension system. Only the Kainji languages in north-west Nigeria have elaborate Bantu-like systems, analysed in McGill (2009) for the Cicipu language, part of the Kambari cluster, and in Mort (2012) for tiCind, a Kamuku language. Cicipu (McGill 2009: 227ff.) has the extensions listed in Table 4; the labels are copied from the author.

We cannot reconstruct forms for extensions in Proto-Kainji, due to the limited number of grammatical descriptions (though see Paterson 2019), and it is therefore not possible to discriminate between older segmental patterns and those which may be innovative.

Extensions have either disappeared or been reduced to unproductive segments in most branches of Kainji, Plateau, Jukunoid and Cross River. However, it is possible to infer likely extensions from synchronic verb forms. Table 5 lists three recurrent suffixes identified in the lexicon of Tarok (Plateau).

However, these are unproductive today and do not clearly resemble any of those reconstructed for Bantu. Nonetheless, their fragmentary survival leads to the conclusion that a system of verbal extensions has to be reconstructed back to the level of Proto-Benue-Congo, and must therefore have been present in early Bantoid. However, their segmental forms can no longer be identified.

Table 4: Verbal extensions in Cicipu (McGill 2009)

Segment	Interpretation
<i>-is- ~ -sV</i>	causative, intensive
<i>-wA</i>	valence-increasing, anti-causative, separative
<i>-nA</i>	ventive
<i>-nu</i>	resultative, intensifier, de-intensifier
<i>-il</i>	pluractional ^a

^aAlthough suffixed after the root, it can be followed by tense/aspect markers and then another extension.

Table 5: Fossilised verbal extensions in Tarok

Segment	Interpretation
<i>-fi</i>	singulative, do something once
<i>-dar</i>	do something completely, intensively
<i>-ri/-li</i>	unassigned

3.2 Synchronic distribution of verbal extensions in Bantoid

A primary question in analysing Bantoid verbal extensions is accounting for their absence in some branches, especially in those more remote from Narrow Bantu, where they have disappeared without leaving obvious segmental traces. Table 6 summarises the situation for the different Bantoid subgroups identified in the literature. It should be emphasised that there are no specific publications on extensions in many of them. Those marked functional have been identified in the literature as in active use, whereas inferred suffixes are those which I have extracted from lexical data. The claim for their presence or absence has to be based on inferences from the lexicon or incidental data. Some of the more diverse subgroups, such as Mambiloid, may include languages with no remaining extensions and those where they are evidently present. Key references are given for individual languages.

Hyman (2018) is a survey of Bantoid verb extensions which includes Grassfields, Mbe (Ekoid), Tikar, Noone, Kemezung (Beboid) and Vute (Mambiloid) in his comparative tables. To throw light on the ancestry of Bantu verbal extensions we must create a basic tabulation of the presence of extensions in individual Bantoid branches, although some may eventually be discarded as not relevant to Bantu.

3.3 Case studies

3.3.1 Dakoid: Sama Mum

The Dakoid languages represent one of the least-described subgroups of Bantoid and were previously classified as Adamawa by Greenberg, presumably because of their cultural relationship with the Samba Leko. They are spoken in eastern Nigeria around the Shebshi mountains, see Figure 5.

There are no specific publications on extensions, so these must be inferred from lexical data. The main resource is a dictionary of Sama Mum or Samba Daka,

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

Table 6: Identifying verbal extensions in major subgroups of Bantoid

Group	Verbal extensions		Language	Reference
	Functional	Inferred		
Dakoid		✓	Daka	Boyd & Sa'ad (2010)
Mambiloid	✓	–	Nizaa	Kjelsvik (2002: 19ff.)
Mambiloid	✓	–	Vute	Thwing (2006)
Tikar	✓	–	Tikar	Stanley (1991)
Bendi	–	–	Bekwara	Stanford (1967)
Tivoid	–	–	Tiv	Arnott (1958)
Buru	?	–	Buru	Koops (s.d.)
Furu	?	–	Furu	Breton (1993), Kießling (2007)
East Beoid	✓		Noone	Hyman (1981)
East Beoid	✓		Mungong ^a	Boutwell (2014)
East Beoid	✓		Nchane	Boutwell (2020)
Yemne-Kimbi	–	–	Mundabli	Voll (2017)
Nyang	–	–	Denya/Kenyang	Unpublished lexicons
Ekoid	–	–	Ejagham	Watters (1981)
Mbe	✓	–	Mbe	Bámgbósé (1967)
Ambele	?	?	Ambele	Nganganu (2001)
Menchum	?	?	Befang	Gueche Fotso (2004)
Grassfields				
•Ndemli	✓		Ndemli ^b	Ndedje (2013)
•Ring	✓		Lamnso?	Grebe & Siiyaatan (2015) ^c
•Momo	✓		Meta'	Spreda (1995), Hyman (2018)
•South-West	?	?	Manta	Ayotte & Ayotte (2003)
•Eastern				
– Bamileke	✓	–	Ngiemboon	Lonfo & Anderson (2014) ^d
– Ngemba	✓		Bambili	Ayuninjam (1998)
– Nkambe	–	–	Mfumte	McLean (2014)
Jarawan	✓	✓		Van de Velde & Idiatov (2022)
Bantu zone A	✓	✓	Akoose A15C	Hedinger (1992; 2008)

^aHowever, the verbal extensions for Mungong consist only of a multiple action extension and an extremely rare causative in *-si*.

^bAlthough Ngoran (1999: 73) states that “[i]n this language, we have been unable to uncover any vestiges of suffixal extensions”, they are identified in Ndedje (2013).

^cSee Table 15 in §3.3.7.

^dSee Table 14 in §3.3.6.

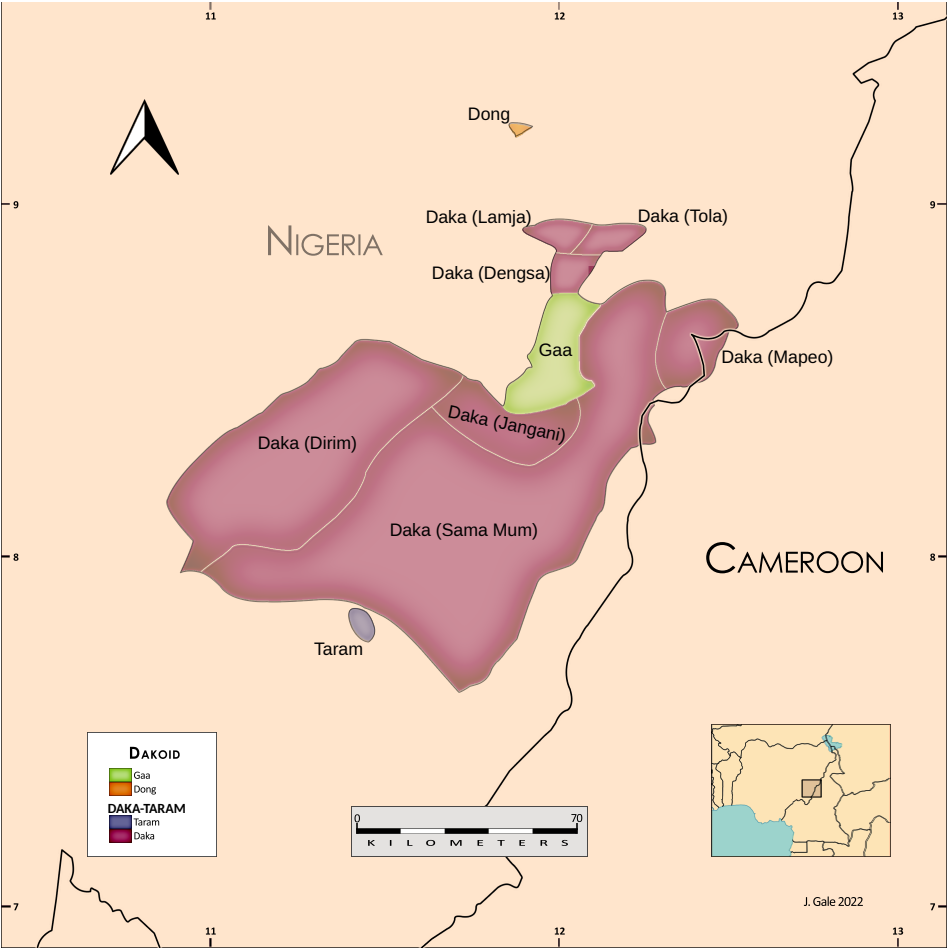


Figure 5: Map of the Dakoid languages

which has a list of the semantic categories of verbal derivations in the introduction, but without any information on their segmental form (Boyd & Sa'ad 2010). I therefore had to infer the extensions and their semantics from the dictionary entries. I have given an example of each verb with these extensions, but for two categories listed in the text, no examples are apparent. The proposed extensions are shown in Table 7.

Table 7: Sama Mum verbal extensions (inferred from Boyd & Sa'ad 2010)

Segment	Interpretation	Example
-kì, -sì	pluractional	<i>bāl-kì</i> 'move many things about', <i>bīp-sì</i> 'twist many'
-sì	causative	<i>bāā-sì</i> 'recall'
-rì	causative II	<i>bōb-rì</i> 'spot, stain'
-èn, -sèn, -kèn, -mèn	resultative I, II	<i>bāl-èn</i> 'move about', <i>dāk-sèn</i> 'be walking stealthily', <i>būū-kèn</i> 'be inhaled', <i>bān-mèn</i> 'be farmed'
-kèn, -sèn	reciprocal	<i>bān-kèn</i> 'farm for each other'
?	applicative	not listed in text
	diminutive	only one case known

Since the authors do not always mark their lexical examples, it is not always clear where some segments are to be found. A striking aspect of Sama Mum is the allomorphy of /s/ and /k/ and the absence of extensions indicating motion, which is characteristic of other branches of Bantoid. The CVn structures which characterise Sama Mum recur in several Bantoid branches and Akoose A15C, which argues either for a genetic connection or the repeated fusing of two extensions (see Bostoen & Guérois (2022 [this volume])).

3.3.2 Mambiloid

The Mambiloid languages are a very internally diverse family spoken in Nigeria and north-west Cameroon (Blench 1993). Figure 6 shows their approximate distribution.

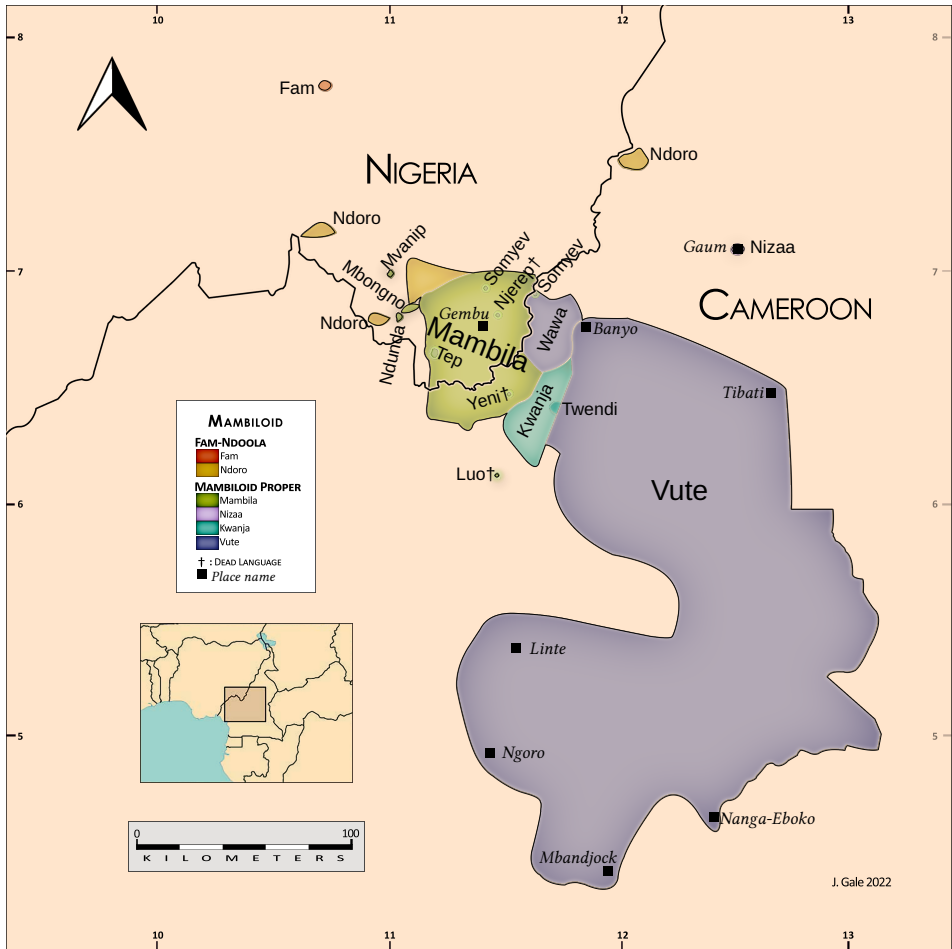


Figure 6: Map of the Mambiloid languages

3.3.2.1 Nizaa

The Nizaa language preserves verbal morphology far better than some other languages in the group, in contrast to Mambila itself, which has lost virtually all nominal and verbal morphology. The main summary of verbal extensions in Nizaa is Kjelsvik (2002: 18). Table 8 outlines the forms she identifies, although she does not provide examples for the directional.

Table 8: Nizaa verbal extensions (Kjelsvik 2002: 18)

Function	Segment	Interpretation	Example
Number marking	-r suffix, vowel lowering	verbal plurality	<i>ki</i> ‘cut once’, <i>kir</i> ‘cut many times’, <i>njúb</i> ‘strike once’, <i>njwáb</i> ‘strike many times’
Directional	-a	‘illative’, motion into an enclosure	
	-ri	‘allative’, motion towards a location, often the deictic centre of the sentence [perhaps corresponds to Bantu applicative]	
	-wa	‘distantive’, motion away from a location, or from the deictic centre [perhaps corresponds to Bantu separative]	
	-sa	‘down’, motion towards a lower location	
Completive	-ki	‘totality’	<i>kibkirá</i> ‘wean’, <i>kagkirá</i> ‘attach’

Kjelsvik (2002) notes that stacking of up to three suffixes is allowed, highly unusual for Bantoid.

3.3.2.2 Vute

Vute, also part of the Mambiloid group, is spoken in north-west Cameroon around Banyo (Guarisma 1978). The only published description of Vute verbal extensions is Thwing (1987), but Thwing (2006) can be downloaded and provides a more complete overview. Vute has either developed or retained a rich repertoire of extensions, in contrast to other languages in its group. It is notable because, like Nizaa, it allows strings of up to four suffixes on the verb root (Thwing 2006: 28). Thwing (2006: 29) summarises the extensions and these are presented in Table 9.

Thwing (2006) also includes a long list of adverbial extensions, which are omitted here. One of these, *-kí* for ‘completely’, resembles Nizaa *-ki* marking ‘totality’. She also notes “phasal” extensions, essentially marking inceptive and completive, both of which have transparent etymologies. The benefactive *-nà* and the directionals are undoubtedly innovative, as Thwing (2006) proposes language-internal etymologies for them. She calls the last three ‘additive/conjoining extensions’, which function to join two clauses or sentences.

Note also that, although Nizaa and Vute are related, there are no clear segmental cognates between the extensions identified for the two languages. It is possible that Nizaa *-sa* and Vute *-sé/-só*, both meaning ‘downwards’, are cognate. However, they could equally be independently innovated, possibly from a cognate language-internal source, such as the reflex of PB **cí* ‘ground; country; underneath’ (BLR 562) (Bastin et al. 2002). This suggests that even within an identified genetic group there must be significant innovation.

3.3.3 Tikar

The Tikar language is spoken on the Tikar Plain in the Adamawa Province of Cameroon (Hagège 1969).⁵ In her lengthy grammar of Tikar, Stanley (1991: 355–384) treats verbal extensions under derivation. Table 10 is extracted from the FLex database of Tikar (Jackson 1988) as well as the PhD thesis of Stanley (1991). Tikar extensions are characterised by very extensive allomorphy.

Note that although Blench (2015) has classified Dakoid, Mambiloid and Tikar in a putative North Bantoid grouping based on lexical and phonological correspondences (see also Figure 2), verbal extensions provide little or no evidence to support this.

⁵For an indication of where this language is spoken *vis-à-vis* the other Bantoid languages, see the note at Figure 11.

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

Table 9: Vute verbal extensions (Thwing 2006: 29)

Function	Segment	Comment
Causative	-tì	could also be interpreted as a transitivity marker; e.g. 'become black' → 'blacken'
Valence change	-lì	can either raise or lower valence depending on verb root
Valence raising	verb root vowel lengthening / -hì	e.g. 'be lost' → 'lose', also 'bite' → 'bite many times'
Valence lowering	change in the quality of the root vowel	
Benefactive (?)	-nà	indirect object marker; innovative from the verb 'give'
Modal	-ná	cohortative or imperative plural form of the verb
Directional	-wò -sò -tè(è) -wú -sé/-só -hó -lé	towards centre of reference away from centre of reference around centre of reference up, upwards down, downwards out (often combined with -sò) in, into
Participant additive	-bwê	adds a participant to an event
Sequential events	-cé	
Simultaneous events	-cè	

Table 10: Tikar verbal extensions (Jackson 1988; Stanley 1991)

Segment	Interpretation
(N)s-zi	causative
-li, -ni, -mi	reflexive/intransitive; nearly bleached of meaning
-ì, -bì, -mì, -nì, -lì	allative; marks action in the direction of the speaker The initial consonant is determined by the final consonant of the verb stem.
-ḍ, -bḍ, -mḍ, -nḍ, -lḍ	ablative; marks action away from the speaker. The initial consonant is determined by the final of the verb stem, infixes where final stem consonant is alveolar.
-ɔʔ plus reduplication of final stem consonant, -lɔʔ, -nɔʔ	iterative ~ <i>répétitif</i>
-(k)aʔ, -ŋga, -a, [-saʔ, -naʔ, laʔ-] ^a	iterative ~ <i>répétitif</i>
-ki	habitual; suffixed to previous extensions
-ti, -ndi	denotes physical action on an object especially with hands ~ <i>alteratif</i> ; some semantics not identifiable

^aThese forms in square brackets are rare in the data.

3.3.4 East Beoid: Noone, Mungong and Nchane

The Beoid languages are spoken in the northern Grassfields of Cameroon, with an extension into Nigeria (Hamm et al. 2002). They are conventionally divided into East and West, although Jeff Good (p.c.) has argued that West Beoid cannot be shown to be a coherent genetic group. He uses the label ‘Yemne-Kimbi’ for West Beoid. Figure 7 shows the distribution of the Beoid languages.⁶

⁶Thanks to Jeff Good for assistance in updating the Beoid map with recent community-preferred names.

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

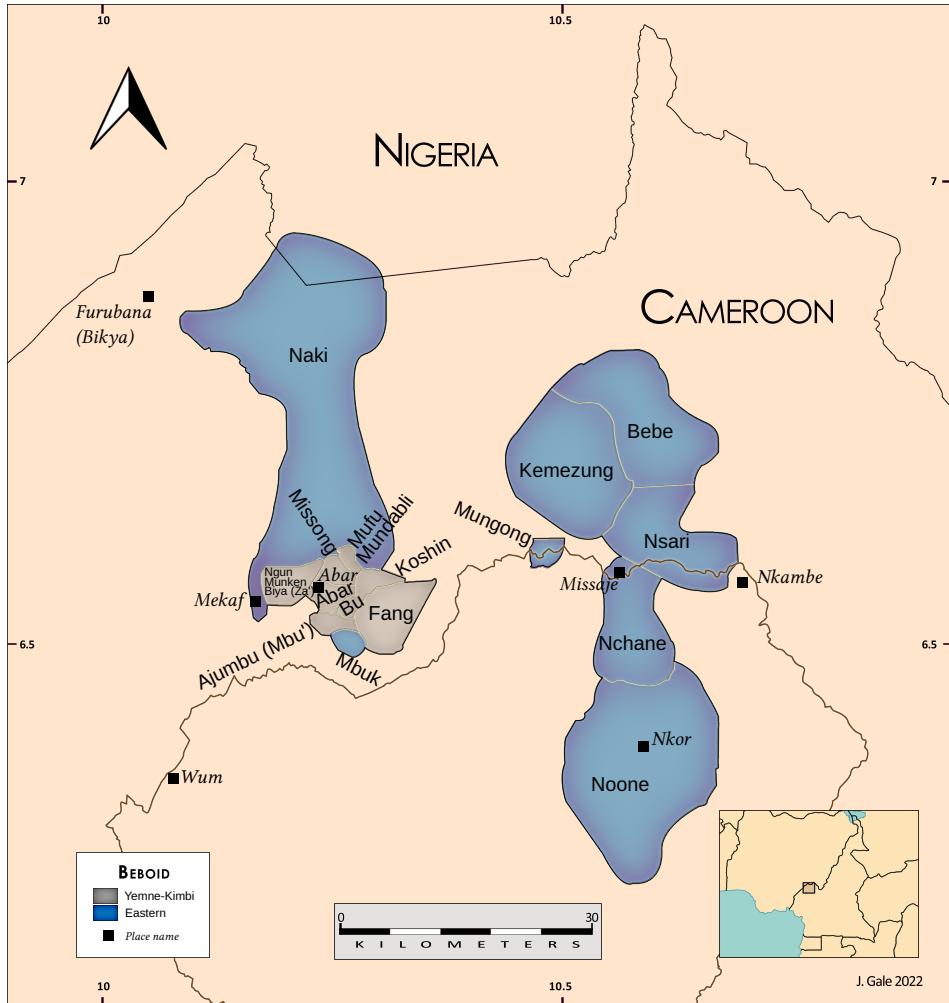


Figure 7: Map of the Beboid languages

Noone is an East Beboid language, first described in Hyman (1981). Table 11 summarises the extensions listed for Noone. Whether ‘reduplication’ should be considered an extension is doubtful.

Table 11: Noone verbal extensions (Hyman 1981)

Category	Segment	Interpretation
Aspectual	<i>-cɛ</i>	attenuative
	<i>-yɛ</i>	distributive
	<i>-kɛn</i>	iterative
	<i>-tɛn</i>	bifurcative
	<i>-RED(uplicated)</i>	frequentative
Relational	<i>-se</i>	causative
	<i>-ke</i>	transitive
	<i>-m</i>	positional
	<i>-n</i>	reciprocal

The aspectuals form quite a restricted set and it is problematic to link these segments with other Bantoid branches. However, some of the relational suffixes are clearly cognate with those in PB (cf. Table 2), for example the positional *-m* (PB *-*am*) and the reciprocal *-n* (PB *-*an*). The causative *-se* is similar to the forms occurring across Bantoid.

However, Mungong, also East Beboid and described in Boutwell (2014), is quite different in that the extensions of Noone are absent, and only one inferred extension *-fə* is identified, a plural or iterative.

Nchane, also East Beboid and described in Boutwell (2020), is still more surprising, since the typically suffixed elements have become preverbal. For example, the iterative *ká-* precedes the verb; judging from form and meaning, it is perhaps cognate with the Noone suffix *-kɛn*. Boutwell (2020) identifies a durative and sequential marker *tú*, a resultative *mɔ* and a habitual *tɔ* in addition to other TA marking. Nchane also has a wide range of postverbal adverbials, but these do not function like usual extensions. As with Mambiloid, East Beboid seems to be very diverse internally, with considerable innovation in individual languages.

3.3.5 Mbe and Ekoid

Mbe is a single language, related to Ekoid, spoken on the Cross River in south-east Nigeria. Figure 8 shows the location of Mbe and the Ekoid languages.

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

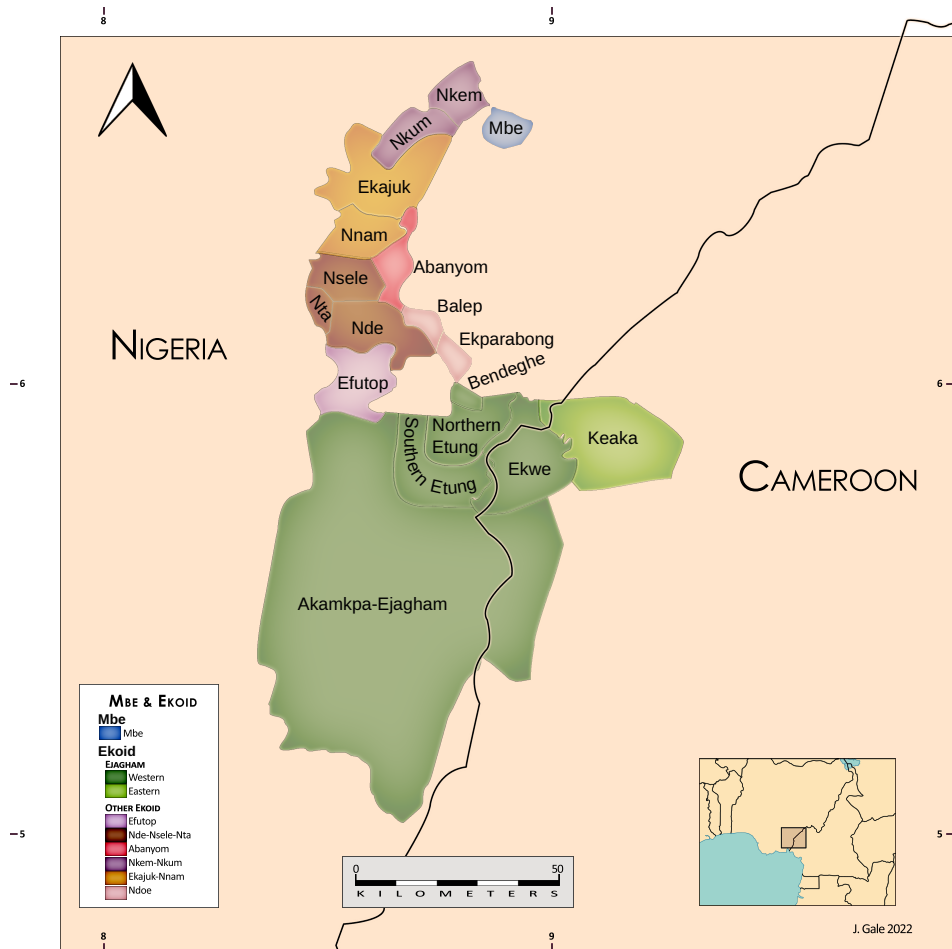


Figure 8: Map of Mbe and the Ekoid languages

In contrast to Ekoid, Mbe seems to have a significant repertoire of verbal extensions (Gerhardt 1978, Blench 2013). The main source for Mbe is Bámbgǒsé (1967) whose paper describes the morphology of Mbe verbs in some detail but gives little or nothing on the interpretation of the forms listed. However, it is clear that almost all verbal extensions in Mbe involve either valence change or plurality (both marking plural subjects and multiple and iterative action). Reduplication is a common strategy and is sometimes combined with the extended forms. Mbe permits multiple plurals on individual verb roots (Bámbgǒsé 1967). Hyman (2018: Table 5) lists only *-li*, *-ri* as separative and intransitive, but clearly the Mbe system is richer than this. Table 12 shows the main Mbe extensions, together with my inferences as to their interpretation.

Table 12: Mbe verbal extensions (Bámbgǒsé 1967)

Number	Operation	Segment	Interpretation
	Transitiviser	<i>-ô, -î</i>	
	Transitiviser	falling tone	
Plural	Ubiquitiser	<i>-nî</i>	do something all over the place
Plural	Reversive (?)	<i>-lî</i>	‘close’ → ‘open’ etc.
Plural	Intensifier	<i>-rî</i>	
Plural		<i>-î</i>	
Plural	Complete reduplication		
Plural	Reduplication of first syllable		

An unpublished dictionary of Mbe, by Pohlig (s.d.), lists forms from which other unproductive extensions can be inferred; see Table 13.

The ubiquitous *-lí* and *-rí* are presumably allomorphs of *-nî*.

3.3.6 Eastern Grassfields: Ngiemboon and Yemba

Grassfields languages are spoken in Cameroon, with a few isolated communities in Nigeria. They constitute a large and complex group, divided into Wide and Narrow Grassfields; see Figure 9. Momo and South-West within Wider Grassfields remain extremely poorly known and the internal configuration of Grassfields is yet to be demonstrated convincingly.

Ngiemboon is spoken in the Grassfields of Cameroon and is a Grassfields language in the Western Bamileke subgroup. Ngiemboon no longer has a productive

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

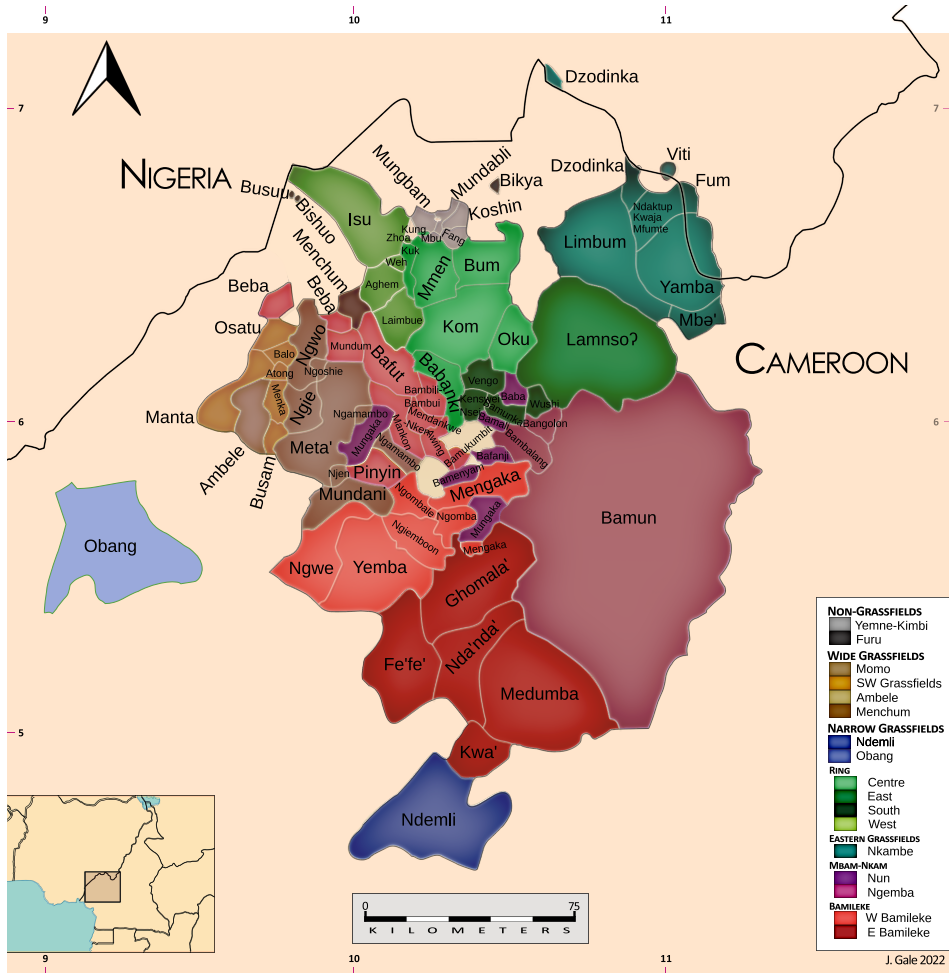


Figure 9: Map of the Grassfields languages

Table 13: Mbe verbal extensions (inferred from Pohlig s.d.)

Operation	Segment	Example
Detransitiviser	<i>-ló ~ -ró</i>	<i>fuéló</i> ‘be brushed off’, <i>duéló</i> ‘become deep’, <i>sháro</i> ‘come to resemble’
Ubiquitiser (‘do x all over the place’)	<i>-lí ~ -rí</i>	<i>shébrí</i> ‘cut into pieces’, <i>shárí</i> ‘scatter things’, <i>yéblí</i> ‘turn over’
Stativiser	<i>-nú</i>	<i>nénú</i> ‘be happy’, <i>lénú</i> ‘reside in a place’, <i>rénú</i> ‘be almost enough’

system of extensions, but the numerous pairs and triplets of verb roots plus (C)V segments show that a rich system must have existed in the recent past. An early sketch of its extensions is contained in Mba & Djiafeua (2003). However, a very large lexical database exists, published as a dictionary (Lonfo & Anderson 2014). Table 14 shows the likely extensions which can be extracted from that database, together with their proposed interpretations (Blench & Martin 2010). Included are segments which appear to be present segmentally but have no obvious semantics.

It is very difficult to map any of these clearly to other attested Bantoid evidence, and the extensive potential meaning-sets suggests that Ngiemboon has undergone extensive mergers and reanalysis.

Harro (1989) and Mbanji et al. (2007) describe the extremely limited extension system of Yemba, another Bamileke language in the same subgroup as Ngiemboon. There are just two segmental extensions, *-tí* and *-ní*: *-tí* is a pluralising extension marking distributive and iterative; *-ní* is more opaque, but there are examples of stativising, reciprocal marking. Surprisingly, these do not resemble other documented Grassfields languages. Mankon as described by Leroy (2007: 225–232) has examples of several extensions, e.g. *-ni* (detransitiviser, often reflexive, comparable with Tikar *-ni*), *-ki* (detransitiviser, iterative, comparable with Tikar *-ki*), *-ti* (diminutiviser, also found in A60 languages) and *-si* (causative).

3.3.7 Ring: Lamnso?

Lamnso? is a Ring language spoken in the Grassfields of Cameroon; see Figure 9. An extensive dictionary of Lamnso? has been published (Grebe & Siyaatan 2015) and from the associated fieldwork database it is possible to infer plausible verbal

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

Table 14: Evidence for verbal extensions in Ngiemboon (Blench & Martin 2010)

Candidate segment	Plausibility	Semantics
-a	conclusive	valence-changing
-bE	evidence inconclusive; some cases clearly final -e	highly varied, perhaps intensification?
-e	conclusive	valence-changing
-le	present but rare	reversive, reflexive
-me	no conclusive evidence for a -me verb extension	all examples valence-changing final -e
-ηV	limited evidence for a valence-changing final -ηv	valence-changing
-ɔ	inconclusive	
-tE	conclusive	iterative, reversive, reciprocal, plurative, cessive, intensive, valence-changing
vowel doubling	conclusive	reversive, reflexive, cessive, valence-changing
tone reversal	conclusive	valence-changing, reversive, iterative, intensive

extensions. Table 15 summarises all the probable extensions in Lamnso? with their meanings. For almost all extensions, there are words that do not ‘fit’ either because the simplex form of the verb is missing or because the semantics do not lend themselves to any unambiguous analysis.

These verbal extensions for Lamnso? do not resemble those for Ngiemboon (Table 14), the language assumedly more closely related to Lamnso?, but there are striking similarities with Akoose A15C (for which, see Table 18 in §3.3.9 below).

Table 15: Lamnso? verbal extensions (inferred from Grebe & Siiyaatan 2015)

Form	Subset	Semantics
CVC	-kir	distributive, plural subject, iterative, continuous action, reflexive
	-nen	reciprocal, plural subject, valence-changing
	-nin	reciprocal, excessive
	-sin	completive
	-tir	paucal, diminutive
	-tin	plural subject, completive, valence-changing, intensification
	-rin	resultative
CV	-si	completive, causative
	-ri	multiple action
	-ti	multiple action, action creating plural objects, intensification
	-fi	process
	-ne	intensification [greater speed, intensity of behaviour]
	-vi	reductive
VC	-Vm	inchoative, inceptive
	-Vy	completive
	-Vr	causative
	-Vn	isiautonomic [i.e. indicating self-initiated action]
V	-V	extensive

3.3.8 Jarawan Bantu

The Jarawan Bantu languages are spoken in scattered communities in eastern and central Nigeria and formerly also in northern Cameroon (Rueck et al. 2007). Figure 10 shows the distribution of Jarawan Bantu. Maddieson & Williamson (1975) remains the only overview of Jarawan Bantu. Many languages have very few speakers, and those recorded in Cameroon in the early twentieth century have apparently become extinct. The extinct Jarawan Bantu languages of northern Cameroon (Dieu & Renaud 1983) are marked with the symbol † in Figure 10.⁷

⁷We do not know when these became extinct, but when the region was surveyed for the *Linguistic Atlas of Cameroon* in the 1970s (Dieu et al. 1976; Dieu & Renaud 1983), no more speakers could be found.

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

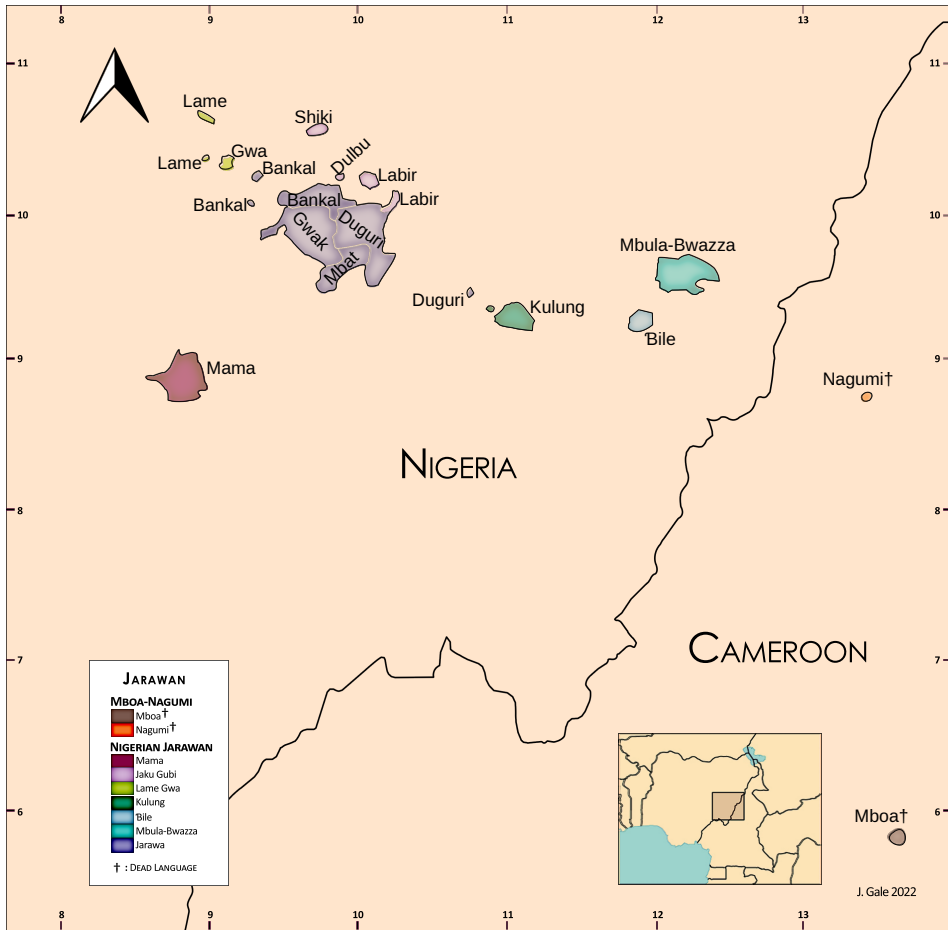


Figure 10: Map of the Jarawan Bantu languages

Jarawan Bantu remains poorly described, with no complete grammar of any individual language. The first published analysis of verbal extensions in Jarawan Bantu is Gerhardt (1988), who points out that the remaining ones are generally interpreted as perfectives; see Table 16. Otherwise, Jarawan Bantu has lost, along with the loss of noun classes, all the usual functions of extensions, including iteratives and plurals, as well as valence-changing extensions.

However, a fresh field study of Mbula (Van de Velde & Idiatov 2022) has revealed a more complex picture. Table 17 shows the verbal extensions of Mbula.

These Mbula verbal extensions align Jarawan Bantu more obviously with the other Bantoid branches described here, but do not clearly establish its nearest genetic neighbours.

Table 16: Jarawan Bantu verbal extensions (Gerhardt 1988)

Segment	Interpretation
stem-vowel lengthened, with occasional inserted glottal	intensifier
stem-vowel lengthened, final vowel neutralised to <i>-a</i>	perfective
<i>-m ~ -Vm ~ -mV</i> (where V is commonly <i>-a</i>)	perfective

Table 17: Mbula verbal extensions (Van de Velde & Idiatov 2022)

Segment	Extension
<i>-sə</i>	causative
<i>-k^H</i>	applicative
<i>-ki</i>	pluractional
<i>*ji-</i>	petrified: reflexive
<i>*-id</i>	petrified: applicative

3.3.9 Bantu: Akoose and Mbonge Oroko

Akoose A15C is a Narrow Bantu language spoken in south-west Cameroon. One might expect its extensions to be close to the forms which have been attributed to those reconstructed for PB given its membership of Narrow Bantu. Since this is not the case, then either Akoose has been significantly transformed by borrowing or has undergone idiosyncratic local development. Akoose verbal extensions have been described in detail by Hedinger (1992; 2008) and are summarised in Table 18.

This should be compared with the proposed PB extensions set out in Table 2. If *s*→*t*, then the causative might be cognate. There are very limited correspondences between the synchronic extensions in Akoose and the PB reconstructed forms and it is notable that Akoose shows more resemblances with Lamnso? (Table 15) and Noone (Table 11), particularly the prevalence of CVN forms, and parallels such as the reciprocal in (n)-Vn, which is part of the PB reconstructed set.

Perplexingly, a study of the Mbonge dialect A121 of Oroko A101 reveals a system quite different from Akoose, despite the fact that both languages are rather close lexically. In some cases, Oroko extensions match the forms reconstructed for PB more closely (Friesen 2002). Table 19, adapted from Friesen (2002: Table 7),

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

Table 18: Akoose verbal extensions (Hedinger 1992; 2008)

Form	Segment	Interpretation
V(C)	<i>-ed/t</i>	causative
	<i>-en/n</i>	instrumental, reciprocal, comitative
	<i>-e/-²/-d</i>	applicative
	<i>-el</i>	unexplained
CVC	<i>-led</i>	unexplained
	<i>-len</i>	unexplained
	<i>-med</i>	unexplained
	<i>-ned</i>	instrumental, comitative, applicative
	<i>-nen</i>	unexplained
	<i>-ted</i>	applicative, causative
	<i>-ten</i>	instrumental, reciprocal, comitative
	<i>-sen</i>	unexplained
	<i>-gen</i>	unexplained

shows the extensions identified in Mbonge Oroko compared with those in PB proposed by Meeussen (1967) and Schadeberg (2003). Friesen adds four extensions for which she can identify no parallel.

In the case of extensions like *-isele*, certain combinations of extensions can become fused with specific functions. Narrow Bantu has many examples of verbs with frozen expansions, some of which indeed look like existing extensions (cf. Bostoen & Guérois (2022 [this volume])). Akoose and Oroko are expected to be close to one another, but they only have a small number of resemblances in terms of extensions except for the reciprocal, applicative and instrumental. This is unlikely to be a consequence of weak description as both publications are the result of long-term study.

4 Discussion and conclusion

The use of verbal extensions was evidently a feature of early Niger-Congo (Voeltz 1977; Hyman 2014) and they remained part of the morphological system at the time of the diversification of Benue-Congo, as strongly suggested by the evidence from West Kainji—see the debate on this topic between Güldemann (2011) and Hyman (2011). The remarkable verbal extensions in the Katloid languages

Table 19: Mbonge Oroko verbal extensions (Friesen 2002: Table 7)

Mbonge	Description	Meeussen	Schadeberg	Proto-Bantu label
-isε	causative – default	-jç-	-ici-	causative
-elε	causative – indirect agent	-ud-	-ʊd-	separative transitive
-isεlε	causative – indirect effector			
-ε	causative – lexicalised	-j̄/-ik-	-i/-ɪk-	causative/impositive transitive
-eε	applicative	-id-	-il-	applicative
—	—	-ik-	-ik-	impositive
-eε	anti-causative	-ik-	-ik-	neuter
-am	stative	-am-	-am-	stative
-an	instrument, accompaniment	-an-	-an-	reciprocal
-en with a-	reciprocal			
—	—	-ad-	-al-	(function unidentified)
—	—	-at-	-at-	contactive
-o	inversive	-uk-	-ʊk-	separative intransitive
-ab	passive			passive (cf. Bostoën & Guérois (2022 [this volume]))
Additional Mbonge extensions				
-ε (with a-)	reflexive	-ik-	-ik-	neuter
-en	intensity			PB *-an? (cf. Bostoën et al. 2015)
-i	lexicalised			
-el	lexicalised			

in Kordofanian (e.g. Hellwig 2013: Table 4) illustrate the importance of this morphosyntactic feature at an earlier stage of Niger-Congo (see also Hyman 2020). However, verbal extensions are now preserved only in fragmentary form in individual Plateau and Cross River languages and have largely disappeared in many branches of Bantoid. Few studies have analysed verbal extensions specifically, but where substantial lexicons exist their former presence can sometimes be inferred. The outcomes of this loss remain to be more fully explored, but clearly an expansion of the verbal auxiliary system, verb serialisation and adverbs are typical replacement strategies (see also Hyman 2017a). Kießling (2004) and Kießling & Wung (2011) have written about the evolution of verb serialisation in Ring languages, which has essentially replaced functional verbal extensions.

Where languages preserve extensions, many are very restricted (i.e. they only occur on a few verbs, as in Yemba, Nizaa, Vute or Mungong). Only some Eastern Grassfields languages have complex, if now unproductive, systems. From the point of view of historical reconstruction, there are few correspondences even within Grassfields, as a comparison of Table 14 and Table 15 makes plain. Languages such as Ngiemboon and Lamnso? would be expected to be more closely related to one another than to Narrow Bantu, but this is not apparent from the data. This is not to say that more conservative languages such as Mankon (Leroy 2007) do not preserve more elements that correspond to elements outside Grassfields. Comparison with Bantu (Table 2) is hardly more illuminating. As Hyman (2018) observes: “[t]he forms or functions of the extensions may not correspond to those in Narrow Bantu”. Indeed the only extension which is clearly preserved from the remoter branches of Bantoid is the causative in *-si*, which is also widespread in Niger-Congo. The degree to which the other extensions are cognate is contentious, and will not be resolved until group level reconstructions are available.

Another major difference with Narrow Bantu is the rareness of stacked extensions. Given the productive nature of this process in Bantu, it is perhaps surprising that hardly any Bantoid languages, except Vute and Nizaa, can be demonstrated to permit strings of extensions. Other languages exhibit strong maximality constraints. It is plausible to suggest that the *-CVN* forms which are attested in Dakoid, Grassfields and Beboid represent two originally distinct extensions now fused, or reanalysis of the final C of the root, but this has yet to be actually demonstrated. An important element in the loss of extensions, is the imposition of a maximum size constraint on stems (root + suffix) which leaves little room for two extensions except for the fused *-CVN* forms.

Despite this lack of obvious cognates, there are strong similarities in semantics. Valence change, iteratives, plural, reciprocal, reflexive, and instrumental are

often present, which suggests that concepts are transmitted, in the absence of (easy to establish) inherited segments. Given the relative conservatism of noun class prefixes, this variability is quite surprising. To explain it, we must invoke metatypy, the notion that ideas are conserved more than segments, that verbal plurality, iteratives, directionals and transitivisers effectively need to find expression but are constantly re-encoded, perhaps because of continuing segment merger and subsequent splitting. Ngiemboon represents this situation, where some extensions with a consistent segmental form encompass a whole variety of semantics. Such systems are very dynamic and probably change on a generational scale, while the underlying parameters are conserved. Semantic similarities are, of course, in the eye of the beholder; the extent to which the meanings can be bleached and repurposed varies from one researcher to another.

The comparison between Akoose A15C (cf. Table 18) and the proposed reconstructed forms for Proto-Bantu (PB, cf. Table 2) reveals a significant analytic problem. Akoose presumably represents Bantu shortly after the split from Bantoid and, as such, its extensions should either resemble those reconstructed for PB or there should be evidence from fossil morphology of a wholesale replacement process. Akoose forms manifestly do not resemble the proposed PB forms, whether semantics or segments are considered. Akoose is similar to Lamnso? in terms of its -CVN segments, although the difficulties of assigning meaning to many of these makes semantic matches more difficult. The explanation for this is unknown; either Akoose has come under areal influence from Grassfields or possibly parallel developments have led to convergent surface forms. Oroko A101 (Table 19) has more similarities to PB, but is also quite different from Akoose.

The proposed verbal extensions of PB are reconstructed forms. In other words, they would ideally be supported by lengthy data tables and sound correspondences to account for the synchronic forms, especially for zone A languages. It is more likely they represent a synthesis of forms evident from inspection of a range of languages across Bantu, which would not necessarily reflect the forms of PB. Akoose shows that Bantu retained significant segmental matches with languages outside Bantu, in Grassfields, and perhaps also with Dakoid, which is far from Ring, making contact-induced change unlikely (Table 7). This suggests that at the very least the repertoire of extensions in PB should be extended. Some Bantu extensions can plausibly be traced outside Narrow Bantu, as suggested in Table 20.

The similarities between Dakoid, Beboid and Grassfield's Ring are striking, since Dakoid is quite geographically remote from the others and contact is a less plausible explanation. Indeed, the relative distances between the different Bantoid groups discussed in this chapter may best be appreciated from the synthesis map shown in Figure 11.

5 Reconstruction of Proto-Bantu verbal extensions with Bantoid

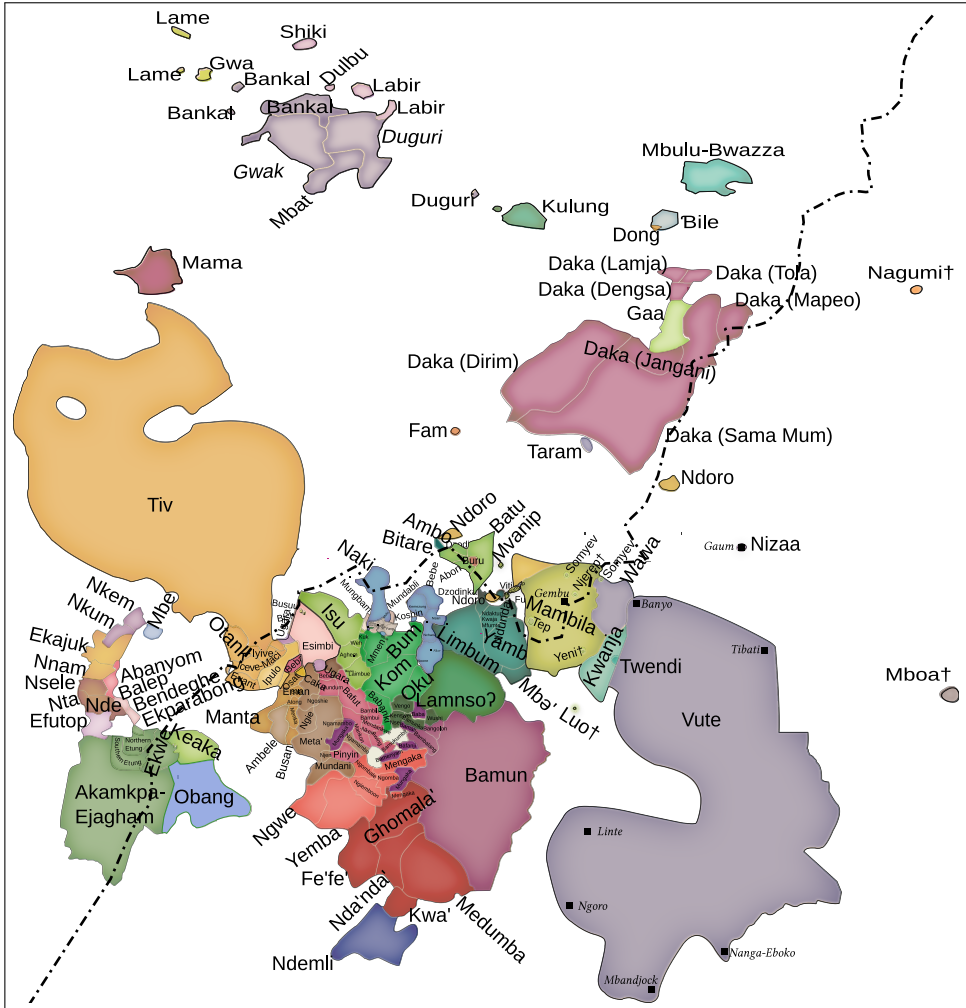
Table 20: Proposed cognates of Proto-Bantu verbal extensions outside Narrow Bantu

Family	Language	Form	Semantics	Proto-Bantu
Kordofanian	Tima	<i>-ik</i>	productive causative suffix functions as a neutro-passive marker	*-ik
Kainji	Cicipu	<i>-sV</i>	causative	*-i/-ici
Dakoid	Sama Mum	<i>-si</i>	causative	*-i/-ici
Dakoid	Sama Mum	<i>-(k)èn, -(s)èn</i>	reciprocal	*-an
East Beoid	Noone	<i>-m</i>	positional	*-am
East Beoid	Noone	<i>-n</i>	reciprocal	*-an
East Beoid	Noone	<i>-se</i>	causative	*-i/-ici
Jarawan	Mbula	<i>-sə</i>	causative	*-i/-ici
Ring	Lamnso?	<i>-nen</i>	reciprocal	*-an

Longer term, however, a major review of the evidence for Bantu, focusing on zone A languages, is required, conforming to the principles of the Comparative Method.

Acknowledgements

I would like to take this opportunity to thank SIL members in both Nigeria and Cameroon, who have always been willing to share material and to observe that our knowledge of Bantoid would be markedly impoverished without their contributions. Special thanks to Robert Hedinger and Stephen C. Anderson for sharing documents and assisting with field trips in Cameroon, and Luther Hon and Michael J. Rueck from the SIL Survey Team in Jos, Nigeria. My thanks also go to the reviewers for picking up errors and omissions in earlier versions. Thank you also to the editors, especially Koen Bostoen and Gilles-Maurice de Schryver, for financing the maps found in this chapter.



All the languages of the Bantoid groups discussed in this chapter are shown, except for Tikar, spoken on the Tikar Plain, which lies to the east of the Ring language Lamnso?.

Figure 11: Synthesis map of the Bantoid language groups

References

- Arnott, D. W. 1958. The classification of verbs in Tiv. *Bulletin of the School of Oriental and African Studies* 21(1). 111–133.
- Ayotte, Michael & Charlene Ayotte. 2003. *Sociolinguistic language survey of Ngwe: Fontem, Alou, and Wabane subdivisions, Lebalem division, South West province* (SIL Electronic Survey Reports 4). Yaoundé: SIL Cameroon.
- Ayuninjam, Funwi F. 1998. *A reference grammar of Mbili*. Lanham, MD: University Press of America.
- Bámgbòsé, Ayò. 1967. Verbal classes in Mbe. *Afrika und Übersee: Sprachen, Kulturen* 50(3). 173–193.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bastin, Yvonne & Pascale Piron. 1999. Classifications lexicostatistiques : bantou, bantou et bantöide (de l'intérêt des « groupes flottants »). In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 149–163. Stanford, CA: Center for the Study of Language & Information.
- Bleek, Wilhelm H. I. 1862–69. *A comparative grammar of South African languages*. London: Trübner & Co.
- Blench, Roger M. 1993. An outline classification of the Mambiloid languages. *Journal of West African Languages* 23(1). 105–118.
- Blench, Roger M. 2001. *The Bendi languages: More lost Bantu languages?* Paper presented at the Benue-Congo Workshop, 32nd Annual Conference on African Linguistics, Berkeley, 26–27 March 2001.
- Blench, Roger M. 2013. *Ekoid: Bantoid languages of the Nigeria-Cameroun borderland*. Unpublished ms. Jos: Kay Williamson Educational Foundation.
- Blench, Roger M. 2015. The Bantoid languages. In *Oxford handbooks online*. DOI: 10.1093/oxfordhb/9780199935345.013.17.
- Blench, Roger M. & Marieke Martin. 2010. *Ngiemboon verbal extensions: A new analysis*. Unpublished ms. Jos: Kay Williamson Educational Foundation.
- Bostoën, Koen, Sebastian Dom & Guillaume Segerer. 2015. The antipassive in Bantu. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 53(4). 731–772.
- Bostoën, Koen & Rozenn Guérois. 2022. Reconstructing suffixal phrasemes in Bantu verbal derivation. In Koen Bostoën, Gilles-Maurice de Schryver, Rozenn

- Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 343–383. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575829.
- Bostoën, Koen & Mark Van de Velde. 2019. Introduction. In Mark Van de Velde, Koen Bostoën, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 1–13. Milton Park, Abingdon: Routledge.
- Boutwell, Richard Lee. 2014. *A sketch grammar of the Mungong language*. Yaoundé: SIL Cameroon.
- Boutwell, Richard Lee. 2020. *A grammar of Nchane: A Bantoid (Beboid) language of Cameroon*. Leiden: Leiden University. (Doctoral dissertation).
- Boyd, Raymond. 1994. *Historical perspectives on Chamba Daka* (Westafrikanische Studien 6). Cologne: Rüdiger Köppe.
- Boyd, Raymond. 1997. Chamba Daka and Bantoid: A further look at Chamba Daka classification. *Journal of West African Languages* 26(2). 29–43.
- Boyd, Raymond & Isa Sa'ad. 2010. *A Chamba – English dictionary*. Lagos: Malt-house Press.
- Breton, Roland. 1993. Is there a Furu language group? An investigation on the Cameroon-Nigeria border. *Journal of West African Languages* 23(2). 97–118.
- de Wolf, Paul P. 1971. *The noun class system of Proto-Benue-Congo* (Janua Linguarum. Series Practica 167). The Hague: Mouton.
- Di Carlo, Pierpaolo, Angiachi Demetris Esene Agwara & Rachel Ojong Diba. 2020. Multilingualism and the heteroglossia of ideologies in lower Fungom (Cameroon): Use and attitudes. *Sociolinguistic Studies* 14(3). 321–345.
- Di Carlo, Pierpaolo, Jeff Good, Ling Bian, Yujia Pan & Penghang Liu. 2018. Socio-spatial networks, multilingualism, and language use in a rural African context. In Paolo Fogliaroni, Andrea Ballatore & Eliseo Clementini (eds.), *Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory (COSIT 2017)* (Lecture Notes in Geoinformation and Cartography), 35–39. Cham: Springer.
- Di Carlo, Pierpaolo, Jeff Good & Rachel Ojong Diba. 2019. Multilingualism in rural Africa. In *Oxford research encyclopedia of linguistics*. DOI: 10.1093/acrefore/9780199384655.013.227.
- Dieu, Michel & Patrick Renaud. 1983. *Situation linguistique en Afrique centrale : inventaire préliminaire : le Cameroun* (Atlas linguistique de l'Afrique centrale; Atlas linguistique du Cameroun). Paris; Yaoundé: ACCT; CERDOTOLA & DGRST.
- Dieu, Michel, Patrick Renaud & Michka Sachnine. 1976. L'atlas linguistique du Cameroun. *Bulletin de l'ALCAM* 1. 1–32.

- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2022. *Ethnologue: Languages of the world*. 25th edn. Dallas, TX: SIL International.
- Friesen, Lisa. 2002. *Valence change and Oroko verb morphology (Mbonge dialect)*. Grand Forks, ND: University of North Dakota. (MA thesis).
- Gerhardt, Ludwig. 1978. Die Beziehungen des Mbe (Cross River State, Nigeria) zum Bantu. *Afrika und Übersee: Sprachen, Kulturen* 61(3–4). 229–241.
- Gerhardt, Ludwig. 1988. A note on verbal extensions in Jarawan Bantu. *Journal of West African Languages* 18(2). 3–8.
- Good, Jeff, Jesse Lovegren, Jean Patrick Mve, Carine Nganguép Tchiemouo, Rebecca M. Voll & Pierpaolo Di Carlo. 2011. The languages of the Lower Fungom region of Cameroon: Grammatical overview. *Africana Linguistica* 17. 101–164.
- Grebe, Karl H. & Patrick Siyaatan. 2015. *Lamnso' – English dictionary / Iḵwà' nsàv* (Kay Williamson Educational Foundation – African Languages Monographs 8). Yaoundé; Cologne: Cameroon Association for Bible Translation & Literacy; Rüdiger Köppe.
- Greenberg, Joseph H. 1963. *The languages of Africa* (Indiana University Research Center in Anthropology, Folklore and Linguistics 25). Bloomington, IN; The Hague: Indiana University; Mouton.
- Grollemund, Rebecca. 2012. *Nouvelles approches en classification : application aux langues bantu du nord-ouest*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Grollemund, Rebecca, Jean-Marie Hombert & Simon Branford. 2018. A phylogenetic study of North-Western Bantu and South Bantoid languages. In Rajend Mesthrie & David Bradley (eds.), *The dynamics of language: Plenary and focus lectures from the 20th International Congress of Linguists*, 120–133. Cape Town: UCT Press.
- Guarisma, Gladys. 1978. *Etudes vouté (langue bantoïde du Cameroun) : phonologie et alphabet pratique, synthématique, lexique vouté – français* (Société d'études linguistiques et anthropologiques de France 66–67). Paris: SELAF.
- Gueche Fotso, Hugues Carlos. 2004. *Noun morphology of Befang*. Yaoundé: University of Yaoundé I. (MA thesis).
- Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas: With special refer-*

- ence to Africa* (Tokyo University of Foreign Studies, Studies in Linguistics 2), 109–141. Amsterdam: John Benjamins.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1967–71. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages* (4 volumes). Farnborough: Gregg International.
- Hagège, Claude. 1969. *Esquisse linguistique du tikar (Cameroun)* (Société d'études linguistiques et anthropologiques de France 11). Paris: SELAF.
- Hamm, Cameron, Ferdinand Assako a Tiati, Jason Diller & Kari Jordan-Diller. 2002. *A rapid appraisal survey of Western Beoid languages (Menchum division, Northwest province)* (SIL Electronic Survey Reports 14). Yaoundé: SIL International.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank (eds.). 2021. *Glottolog 4.5*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://glottolog.org/>.
- Harro, Gretchen. 1989. Les extensions verbales en yemba (bamiléké-dschang). In Daniel Barreteau & Robert Hedinger (eds.), *Descriptions de langues camerounaises* (DELAN : Description systématique des langues nationales), 239–269. Paris: SIL & ACCT & ORSTOM.
- Hedinger, Robert. 1989. Northern Bantoid. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 421–429. Lanham, MD: University Press of America.
- Hedinger, Robert. 1992. Verbal extensions in Akɔɔse: Their form, meaning and valency changes. *Afrika und Übersee: Sprachen, Kulturen* 75(2). 227–251.
- Hedinger, Robert. 2008. *A grammar of Akɔɔse: A Northwest Bantu language* (SIL International Publications in Linguistics 143). Dallas, TX: SIL International.
- Hellwig, Birgit. 2013. Verbal morphology in Katla. In Thilo C. Schadeberg & Roger M. Blench (eds.), *Nuba Mountain language studies*, 237–250. Cologne: Rüdiger Köppe.
- Hyman, Larry M. 1981. *Noni grammatical structure: With special reference to verb morphology* (Southern California Occasional Papers in Linguistics 9). Los Angeles, CA: Department of Linguistics, University of Southern California.
- Hyman, Larry M. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville, MA: Cascadilla Press.

- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.
- Hyman, Larry M. 2014. Reconstructing the Niger-Congo verb extension paradigm: What’s cognate, copied or renewed? In Martine Robbeets & Walter Bisang (eds.), *Paradigm change: In the Transeurasian languages and beyond* (Studies in Language Companion Series 161), 103–126. Amsterdam: John Benjamins.
- Hyman, Larry M. 2017a. Multiple argument marking in Bantoid: From syntheticity to analyticity. In Walter Bisang & Andrej Malchukov (eds.), *Unity and diversity in grammaticalization scenarios* (Studies in Diversity Linguistics 16), 67–95. Berlin: Language Science Press. DOI: 10.5281/zenodo.823230.
- Hyman, Larry M. 2017b. On reconstructing tone in Proto-Niger-Congo. In Valentin F. Vydrin & Anastasia V. Lyakhovich (eds.), *В жёлтой жаркой африке... к 50-летию Александра Желтова / in the hot yellow Africa... in honor of Alexander Zheltov on the occasion of his 50th birthday*, 175–191. Saint Petersburg: Nestor-Istoria.
- Hyman, Larry M. 2018. Common Bantoid verb extensions. In John R. Watters (ed.), *East Benue-Congo: Nouns, pronouns, and verbs* (Niger-Congo Comparative Studies 1), 173–198. Berlin: Language Science Press. DOI: 10.5281/zenodo.1314327.
- Hyman, Larry M. 2020. A note on Nuba Mountain verb extensions. *Faits de langues* 51(1). 29–36.
- Jackson, Ellen W. 1988. *Lexique français – tikar*. Unpublished ms. Yaoundé: SIL Cameroon.
- Jacquot, André & Irvine Richardson. 1956. Report of the western team: Atlantic coast to Oubangui. In Malcolm Guthrie & Archibald N. Tucker (eds.), *Linguistic survey of the Northern Bantu borderland*, vol. 1, 9–62. London: Oxford University Press (for the International African Institute).
- Johnston, Harry H. 1919–22. *A comparative study of the Bantu and Semi-Bantu languages, Volumes I & II*. Oxford: Clarendon Press.
- Jungraithmayr, Herrmann & Wilhelm J. G. Möhlig. 1983. *Lexikon der Afrikanistik: Afrikanische Sprachen und ihre Erforschung*. Berlin: Dietrich Reimer.
- Kießling, Roland. 2004. “The giraffes burst throw emerge climb pass through the roof of the hut”: Verbal serialisation in the West Ring languages (Isu, Weh, Aghem). In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 225–235. Cologne: Rüdiger Köppe.
- Kießling, Roland. 2007. Furu lexical data. Unpublished ms.

- Kießling, Roland & Bong M. Wung. 2011. *Verbal serialisation in Isu (West-Ring): A Grassfields language of Cameroon* (Grammatische Analysen afrikanischer Sprachen 39). Cologne: Rüdiger Köppe.
- Kjelsvik, Bjørghild. 2002. *Verb chains in Nizaa*. Oslo: University of Oslo. (MA thesis).
- Koelle, Sigismund W. 1854. *Polyglotta Africana, or a comparative vocabulary of nearly three hundred words and phrases in more than one hundred distinct African languages*. London: Church Missionary House.
- Koops, Robert. s.d. Buru wordlist. Unpublished ms.
- Krause, Gottlieb A. 1895. Die Stellung des Temne innerhalb der Bantusprachen. *Zeitschrift für afrikanische, ozeanische und ostasiatische Sprachen: mit besonderer Berücksichtigung der deutschen Kolonien* 1. 250–267.
- Leroy, Jacqueline. 2007. *Le mankon: Langue bantoue des Grassfields (Province Nord-Ouest du Cameroun)* (Société d'études linguistiques et anthropologiques de France 437). Paris: Peeters.
- Lonfo, Etienne & Stephen C. Anderson. 2014. *Dictionnaire ngiemboon – français – anglais / Ngiemboon – French – English dictionary / tyè'tendù Ngiemboon – Felansé – Ngilisé* (Kay Williamson Educational Foundation – African Languages Monographs 7). Cologne: Rüdiger Köppe.
- Maddieson, Ian & Kay Williamson. 1975. Jarawan Bantu. *African Languages / Langues Africaines* 1. 124–163.
- Mba, Gabriel & Prosper Djaifeua. 2003. Les extensions verbales en ngiemboon. In Daniel F. Idiata & Gabriel Mba (eds.), *Studies on voice through verbal extensions in nine Bantu languages spoken in Cameroon, Gabon, DRC and Rwanda* (LINCOM Studies in African Linguistics 57), 113–166. Munich: LINCOM Europa.
- Mbanji, Bawe E., Albert A. Tsomejio, Ann Wester & Mary Annett. 2007. *Anu ŋkūŋfák: aŋwa'ne wə tswí-ázi'ne / L'histoire de Kande : livre de l'apprenant / Kande's story: Comité Oeucumenique de Suivi et de la Traduction de la Bible en Yémba*.
- McGill, Stuart. 2009. *Gender and person agreement in Cicipu discourse*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- McLean, Greg L. 2014. *A sketch grammar of the Central Mfumte language*. Yaoundé: SIL Cameroon.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meinhof, Carl. 1899. *Grundriß einer Lautlehre der Bantusprachen, nebst Anleitung zur Aufnahme von Bantusprachen. Anhang: Verzeichnis von Bantuwortstämmen*

- (Abhandlungen für die Kunde des Morgenlandes 11(2)). Leipzig: F.A. Brockhaus.
- Meinhof, Carl. 1910. *Grundriß einer Lautlehre der Bantusprachen, nebst Anleitung zur Aufnahme von Bantusprachen. Anhang: Verzeichnis von Bantuwortstämmen*. 2nd revised edition. Berlin: Dietrich Reimer (Ernst Vohsen).
- Mort, Katherine. 2012. *Kamuku grammar outline*. Unpublished ms. Orlando, FL: Wycliffe Bible Translators.
- Ndedje, René. 2013. *Phonologie, morphologie, syntaxe et classification du ndemli*. Yaoundé: University of Yaoundé I. (Doctoral dissertation).
- Nganganu, Kenfac Lucy. 2001. *The phonology of Ambele*. Yaoundé: University of Yaoundé I. (MA thesis).
- Ngoran, Loveline Lenaka. 1999. *A sketch outline of the phonology of Ndemli*. Yaoundé: University of Yaoundé I. (MA thesis).
- Paterson, Rebecca Dow Smith. 2019. *Nominalization and predication in Ut-ma'in*. Eugene, OR: University of Oregon. (Doctoral dissertation).
- Piron, Pascale. 1997. *Classification interne du groupe bantoïde* (LINCOM Studies in African Linguistics 11–12). Munich: LINCOM Europa.
- Piron, Pascale. 1998. The internal classification of the Bantoid language group, with special focus on the relations between Bantu, Southern Bantoid and Northern Bantoid languages. In Ian Maddieson & Thomas J. Hinnebusch (eds.), *Language history and linguistic description in Africa. Selected papers of the Twenty-Sixth Annual Conference on African Linguistics, held March 23–25, 1995, in Santa Monica, California* (Trends in African Linguistics 2), 65–74. Trenton, NJ; Asmara: Africa World Press.
- Pohlig, James N. s.d. Mbe dictionary. Unpublished ms.
- Rueck, Michael J., Zachariah Yoder & Katarína Hannelova. 2007. Sociolinguistic survey of the Mbula, Tambo, Bakopi, Gwamba, Bwazza, Kulung and Bille people, of Adamawa and Taraba States, Nigeria. Unpublished ms. Jos.
- Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 71–89. London: Routledge.
- Schadeberg, Thilo C. & Koen Bostoen. 2019. Word formation. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 2nd edn. (Routledge Language Family Series), 172–203. Milton Park, Abingdon: Routledge.
- Spreda, Klaus W. 1995. *Drafts and notes towards a sketch grammar of Meta'*. Yaoundé: SIL Cameroon.
- Stanford, Ronald. 1967. *The Bekwarra language of Nigeria: A grammatical description*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).

- Stanley, Carol. 1991. *Description phonologique et morpho-syntaxique de la langue tikar (parlée au Cameroun)*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Thwing, Rhonda. 1987. *The Vute noun phrase and the relationship between Vute and Bantu*. Arlington, TX: University of Texas at Arlington. (MA thesis).
- Thwing, Rhonda. 2006. *Verb extensions in Vute*. Yaoundé: SIL International.
- Trithart, Mary Lee. 1983. *The applied suffix and transitivity: A historical study in Bantu*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Van de Velde, Mark & Dmitry Idiatov. 2022. A first analysis of Mbula. Paper presented at the First Bantoid Conference, Hamburg, 24–26 March 2022.
- Voeltz, Erhard F. K. 1977. *Proto Niger-Congo verb extensions*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Voll, Rebecca Maria. 2017. *A grammar of Mundabli: A Bantoid (Yemne-Kimbi) language of Cameroon*. Leiden: Leiden University. (Doctoral dissertation).
- Warnier, Jean-Pierre. 1979. Noun-classes, lexical stocks, multilingualism, and the history of the Cameroon Grassfields. *Language in Society* 8(3). 409–423.
- Watters, John R. 1981. *A phonology and morphology of Ejagham, with notes on dialect variation*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Watters, John R. & Jacqueline Leroy. 1989a. Bantoid overview. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 401–429. Lanham, MD: University Press of America.
- Watters, John R. & Jacqueline Leroy. 1989b. Southern Bantoid. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 430–449. Lanham, MD: University Press of America.
- Westermann, Diedrich. 1927. *Die westlichen Sudansprachen und ihre Beziehungen zum Bantu* (Mitteilungen des Seminars für Orientalische Sprachen an der Friedrich Wilhelms-Universität zu Berlin 29). Berlin: de Gruyter.
- Williamson, Kay. 1971. The Benue-Congo languages & Ijò. In Thomas A. Sebeok (ed.), *Linguistics in Sub-Saharan Africa* (Current Trends in Linguistics 7), 245–306. The Hague: Mouton.
- Williamson, Kay. 1989. Benue-Congo overview. In John T. Bendor-Samuel & Rhonda L. Hartell (eds.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 246–274. Lanham, MD: University Press of America.

Chapter 6

Causative and passive high tone in Bantu: Spurious or proto?

Larry M. Hyman

University of California, Berkeley

In this study I address Meeussen's (1967: 92) tentative proposal to reconstruct a H tone on the Proto-Bantu causative **-i* and passive **-u* suffixes, contrasting with all other Proto-Bantu extensions, which are reconstructed as toneless (Meeussen 1961; 1967). After surveying the phenomenon, I conclude that the causative-passive H (CPH) is almost exclusively limited to certain of the interlacustrine Bantu languages (JD40-60 and JE10-40) and should not be reconstructed. I exemplify the CPH tone effects in several of these languages and consider other cases of H tone extensions outside of the interlacustrine area which I argue to be unrelated. Although still requiring further investigation, I conclude by considering different morphological and phonological scenarios by which the CPH effects might have evolved.

1 Introduction

The purpose of this study is to survey and evaluate the tonal effects of the two Bantu vocalic verb extensions **-i* 'causative' and **-u* 'passive' in order to determine whether they carried a H in Proto-Bantu (PB), as Meeussen (1967: 92) considers in his *Bantu Grammatical Reconstructions* (BGR): "The high tone of (the Proto-Bantu) suffixes *-i-* and *-u-* is set up tentatively, and in any case its manifestations seem to have been very much limited." Meeussen was quite clearly concerned about this issue which he referred to in work both prior and subsequent to BGR:

Dans quelques langues, les extensions vocaliques remontant à -i- (caus.) et -u- (passif) ont un ton haut dans certaines formes verbales à finale basse. Ce



(2) Ombo C76 (Meeussen 1961: 425)

folot ‘pull’ *kóngɔl* ‘gather’

- | | | | | | |
|----|-----------------|------------------|-------------------|----------------------------|-----------|
| a. | <i>to-/tɔ-</i> | <i>folotak-a</i> | <i>kóngɔlak-a</i> | ‘we pull/gather’ | (PRS.HAB) |
| | <i>tá-</i> | <i>folotak-i</i> | <i>kóngɔlak-i</i> | ‘we used to pull/gather’ | (PST.HAB) |
| b. | <i>tátándá-</i> | <i>folót-á</i> | <i>kóngól-á</i> | ‘we are pulling/gathering’ | (PRS) |
| | <i>to-/tɔ-</i> | <i>folót-í</i> | <i>kóngól-í</i> | ‘we pulled/gathered’ | (HOD.PST) |

As seen, the post-radical vowels are L before the L final vowel (FV) in (2a), but H before the high FV in (2b), which I will refer to as an inflectional suffixal H (ISH). The general case of no tonal contrast on verb extensions continues in most Bantu languages, sometimes violated only by CPH, as in Fuliiru JD63: “Verbal extensions are all toneless with the exception of Passive (PS), Causative (CS) and CS+PS, any of which contributes a single floating H tone.” (Van Otterloo 2014: 386)

Unravelling the tonal properties of **-i* and **-ɔ* thus potentially requires an understanding of the relation between the FV and inflectional stem (aka melodic) tones—in fact, even beyond tone, as we will see. There are two logical explanations for the exceptional CPH: (i) CPH occurred in PB; (ii) CPH was innovated subsequent to PB. Of the two hypotheses, the first is the easy way out. All instances of CPH would be from PB. Where not attested, the CPH has been lost. If adopting the second hypothesis, we have the more difficult task of explaining how CPH came into being, i.e. why would only **-i* and **-ɔ* acquire a H tone—and, as it turns out, only in certain inflectional contexts?

In what follows I will first present arguments in §2 that CPH was innovated (cf. Hyman & Katamba 1990). Then, in §3 I present other cases of H tone extensions which are unrelated to CPH. I conclude in §4 by considering morphological and phonological scenarios by which **-i* and **-ɔ* might have acquired H tone.

2 Arguments in favour of the innovation of CPH

In this section I outline six facts about CPH that would seem best to be explained if we assumed that the H tone was not originally a property of the causative and passive extensions themselves.

2.1 Limited geographical distribution

The first argument is that CPH is found on both *-i* and *-ɔ* only in (some) inter-lacustrine (JD, JE) languages. Those that have been so identified to show H tone

effects from both extensions are listed in (3) along with their revised Guthrie referential classification (Maho 2009).¹

(3)	a. Nande	JD42	(Mutaka 1994)	HTA
	b. Shi	JD53	(Polak-Bynon 1975)	HTA (+HTI)
	c. Tembo	JD531	(Shigeki Kaji, p.c.)	HTA (+HTI)
	d. Fuliiru	JD63	(Van Otterloo 2011)	HTI
	e. Ganda	JE15	(Hyman & Katamba 1990)	HTA
	f. Soga	JE16	(Hyman, personal notes)	HTA (+HTI)
	g. West Nyala	JE18	(Marlo 2007, Ebarb & Marlo 2010)	
	h. Wanga	JE32a	(Christopher Green, p.c.)	HTI
	i. Marama	JE32C	(Kristopher J. Ebarb, p.c.)	HTI (?)
	j. Kabarasi	JE32E	(Ebarb 2016: e136)	HTA
	k. Marachi	JE342	(Marlo 2007)	
	l. Idakho	JE411	(Ebarb 2014)	HTA
	m. Isukha	JE412	(Kristopher J. Ebarb, p.c.)	
	n. Tiriki	JE413	(Marlo 2013: 182)	HTA

While CPH is widespread in zone J, it is not present in all interlacustrine languages. Thus, CPH is absent in Rwanda-Rundi JD61-62, Nkore-Kiga JE13/14, Lamogi (misclassified under Soga JE16) and presumably mutually intelligible Gwere JE17, the Haya-Jita group JE20, Bukusu JE31c and Logooli JE41. While not a knock-out argument, it would seem more likely that CPH was innovated in this area rather than inherited from PB and lost multiple times everywhere else.

The above-cited zone J languages are those that have H tone on both causative and passive extensions. Although non-interlacustrine Holoholo D28, Lega D25, and Herero R30 have been cited in this context, these do not show the same CPH phenomena. Coupez' (1955: 30) brief discussion of *-y* and *-w* concerns Holoholo

¹Also indicated are languages which have H tone anticipation (HTA) and can be analysed with H tone inversion (HTI), changing the original *H to /L/. I return to this below. Thanks to Michael R. Marlo for help in identifying the Luyia languages and providing additional information concerning their CPH, HTA and HTI properties. Although Michael R. Marlo has provided additional suggestions that HTA or HTI appear in more of the above (and other) Luyia variants, my own characterisation is quite restrictive: HTA in (3) is intended to indicate which languages have a regular process of realising Proto-Bantu *H tone on the preceding mora, e.g. Soga **bón* 'see' (BLR 266, Bastin et al. 2002) > *ò-kú-bòn-à* 'to see', realised *ò-kú-bòn-á* with a final H% boundary tone. He also adds: "Tiriki does not have H associated with causative *-its-*, just with passive *-w-*" (p.c.). See also Marlo (2008; 2009) for tonal analyses of Tura JE32G and Khayo JE341, respectively, which he says do not appear to have CPH effects.

6 Causative and passive high tone in Bantu: Spurious or proto?

infinitives, where H tone shifts one mora to the right, then spreads an additional mora, forming a HL contour on the FV, as shown in (4).²

- (4) Holoholo D28 (Coupez 1955)
- | | underlying | H tone shift | H tone spreading | |
|----|----------------|----------------|------------------|------------------------|
| a. | /kù-kálik-à/ | → kù-kàlik-à | → kù-kàlik-â | ‘cut oneself’ (p. 72) |
| | /kù-tégelel-à/ | → kù-tègélel-à | → kù-tègélel-â | ‘hear’ (p. 24) |
| b. | /kù-món-à/ | → kù-mòn-à | | ‘see’ (p. 22) |
| | /kù-món-i-à/ | → kù-mòn-i-à | → kù-mòn-y-â | ‘make see’ (p. 74) |
| c. | /kù-kúat-à/ | → kù-kúát-à | → kù-kwàt-â | ‘hold’ (p. 24, 30, 32) |
| | /kù-kúat-i-à/ | → kù-kúát-i-à | → kù-kwàt-y-â | ‘make hold’ (p. 30) |

Since the above tonal differences are not specifically tracked in the different tense-aspects, I conducted a search through all of the examples in Coupez (1955) and did not find CPH in any of his 18 ‘*tiroirs*’ (i.e. tense/aspect/mood constructions). Regarding Lega, Meeussen (1971: 21, 25, 27) only mentions that certain tenses with a final H are instead realised with final HL in the presence of *-i-* or *-u-*. If this is all, then H on final CVV is realised HL, which Meeussen writes either *-í-ê* or *-y-ê* as in (5), e.g. *ku-lí-â ~ ku-ly-â* ‘to eat’ (p. 3).

- (5) Lega D25 (Meeussen 1971)
- | | | | | |
|----|-----|-----------------------|------------------|---------|
| a. | i. | <i>tw-â-kúbúl-é</i> | ‘we will pour’ | |
| | ii. | <i>tw-â-súbán-i-ê</i> | ‘we will return’ | (p. 21) |
| b. | i. | <i>tú-kubul-é</i> | ‘let us pour’ | |
| | ii. | <i>tú-suban-y-ê</i> | ‘let us return’ | (p. 23) |

What Meeussen apparently had in mind was that the FV H shifts onto the *-V* extension:

Les suffixes vocaliques -i- (caus.) et -u- (pass.) ont, à l'imminent, au subjonctif et à l'impératif, un ton haut comptant comme ton de finale et suivi de ton bas.
Meeussen (1971: 27)

The vocalic suffixes *-i-* (caus.) and *-u-* (pass.) have, in the near future, in the subjunctive and in the imperative, a high tone counting as tone of the final and followed by a low tone. (my translation)

²There is a regular tone absorption rule by which LH-H is simplified to L-H, as in (4c).

However, it is not enough to show that the presence of *-i* and *-v* cause a tonal difference, since this may simply be due to the extra tone-bearing unit (mora) they add. In Lungu M14, tenses which have H tone spreading to the penult realise the H on the FV if the word ends in CwV or CyV (Bickmore 2007: 165), as in (6). This however also applies if the final syllable is from a CV root + FV, as in (6c).

(6) Lungu M14 (Bickmore 2007)

- | | | | | | |
|----|---------------------------|---|------------------------|--|----------|
| a. | <i>/tú-ku-ful-a/</i> | → | <i>tú-kú-fúl-a</i> | ‘we are washing’ | (p. 150) |
| | <i>/tú-ku-ful-il-a/</i> | → | <i>tú-kú-fúl-íl-a</i> | ‘we are washing for’ | |
| b. | <i>/tú-ku-ful-u-a/</i> | → | <i>tú-kú-fúl-w-á</i> | ‘we are being washed’ | (p. 167) |
| | <i>/tú-ku-ful-is-i-a/</i> | → | <i>tú-kú-fúl-ísh-á</i> | ‘we are washing a lot’ (< <i>-is-i-a</i>) | |
| c. | <i>/tú-ku-lu-a/</i> | → | <i>tú-kú-lw-á</i> | ‘we are fighting’ | (p. 171) |
| | <i>/tú-ku-zu-a/</i> | → | <i>tú-kú-zw-á</i> | ‘we are bleeding’ | |
| d. | <i>/tú-ku-lu-il-a/</i> | → | <i>tú-kú-lw-ííl-a</i> | ‘we are fighting for’ | (p. 171) |
| | <i>/tú-ku-zu-il-a/</i> | → | <i>tú-kú-zw-ííl-a</i> | ‘we are bleeding for’ | |

The contrast between H vs. L on the final syllable clearly correlates with final vowel shortening which converts the penultimate H assigned to Cwáa, Cyáa to Cwá, Cyá (Bickmore 2007: 232). Many other Bantu languages exhibit final CGV-conditioned tonal differences which have nothing to do with CPH, including Mwanga M22 (Bickmore 2000), Yao P21 (Hyman & Ngunga 1994), and Makonde P23 (Liphola & Odden 1999), among others. However, in order to be considered a case of CPH, there has to be a H present somewhere that can only be attributed to *-i* and *-v*, as was seen in West Nyala above in (1).

2.2 Absence of tone reversal

The second argument for innovation is that the CPH does not invert to L in languages which have inverted PB *H to L. This is seen in the Tembo JD531 examples in (7), where the H and L verb roots are reconstructed as **dimb* ‘catch’ (BLR 9693) and **kód* ‘work’ BLR 1875, respectively (Bastin et al. 2002).

(7) Tembo JD531 (Shigeki Kaji, p.c.)

- | | | | | |
|----|-------------------------|---|-----------------------|--|
| a. | <i>/n-á-símb-a/</i> | → | <i>n-á-símb-a</i> | ‘I am catching (sth./s.o.)’ |
| | <i>/n-á-símb-í-a/</i> | → | <i>n-á-símb-y-á</i> | ‘I am making (s.o.) catch (sth./s.o.)’ |
| | <i>/n-á-símb-ú-a/</i> | → | <i>n-á-símb-w-á</i> | ‘I am being caught’ |
| b. | <i>/n-á-kol-a/</i> | → | <i>n-á-kol-a</i> | ‘I am working’ |
| | <i>/n-á-kos-í-a/</i> | → | <i>n-á-kos-á</i> | ‘I am making (s.o.) work’ |
| | <i>/n-á-kol-el-ú-a/</i> | → | <i>n-á-kol-el-w-á</i> | ‘I am being worked for
(assisted in my work)’ |

(10) Nande JD42 (Hyman & Valinande 1985)

- a. **mó-tu-a-mu-hum-ir-u-a*
 b. H H *H on FV à la Meeussen
 c. H HHH copying of H from FV to second stem mora
 d. H H HHH H tone anticipation (HTA)
 e. H H L L L Meeussen's Rule (MR): H-H → H-L
 f. H H L HL insertion of CPH
 g. H H H H(L L)⁵ second HTA (doubling)
 → *mó-tw-a-mú-húm-ír-w-â* ... 'we were hit for him ...'

The two questions are: (i) If the PB forms were H tone *-í and *-ú in PB, why do not they undergo MR? (ii) Failing to undergo MR, how did the H of the FV spread through the CPH? As seen in the next section, there are more anomalies concerning the FV.

2.5 Unexpected effects on the final vowel

The fifth argument concerns unexpected tonal and segmental effects on the FV. Continuing with Nande, I have put the final L-L in parentheses in (10g) because of an irregularity discovered by Mutaka (1994: 115–116). After H tone copying and anticipation in (10c,d), all but the first H becomes L by MR in (10e), exactly as in (9e). This L on the FV blocks the assignment of the phrase-final H% boundary tone, as seen in (11a).

(11) Nande JD42 (Hyman & Valinande 1985)

- a. /*mó-tu-a-mu-hum-ir-á*/ → *mó-tw-a-mú-húm-ì:r-â*
 'we hit for him'
 b. /*mó-tu-a-mu-hum-ir-u-a*/ → *mó-tw-a-mú-húm-í:r-w-â*
 'we were hit for him'
 c. /*mó-tu-a-mu-hum-is-i-a*/ → *mó-tw-a-mú-húm-í:s-y-â*
 H%L//
 'we caused him to be hit'

⁵The reason for putting the LL of the final syllable in parentheses is explained in the next section.

In (11b) and (11c), however, where the CPH is anticipated onto the preceding *-í:r* and *-í:s*, respectively, the FV fails to show the L derived from *H on the FV. Instead the phrase-final H% and declarative L// boundary tones are both realised on the last syllable. We know this from the fact that the final HL is lacking if something follows, e.g. *mó-tw-á-húm-ís-y-à Váliná:ndè* ‘we caused s.o. to hit Valinande’ (Mutaka 1994: 115). For further discussion, see Mutaka (1994: 115–116), who proposes a rule of final L tone deletion, as well as Jones (2014: 227–233).

While Nande only shows a tonal irregularity on the FV, both Ganda JE15 and Soga JE16 show additional effects of the CPH. First, the H tonal effects of the CPH are found only when there is an ISH (as in other languages), but also requires that the tense/aspect be marked by the perfect(ive) *-il-e* ending, realised post-consonantly as */-i-e/*, i.e. with deletion of the */l/*. Second, when co-occurring with the causative or passive, the FV is surprisingly realised with the FV *-a* instead of *-e* (Hyman & Katamba 1990: 145–146). This is seen in Soga in (12), which has undergone a diachronic process of HTA shift onto the preceding mora which is most visible in (12a) (cf. PB **-tí-* ‘fear’).

(12) Soga JE16 (Hyman 2018)

- a. *ò-kú-ty-à* ‘to fear’ (cf. Ganda *ò-kù-ty-â* < */-tí-a/*)
- b. *tù-tì-il-é* ‘we have feared’ (*-il-e* ending after a CV- root)
- c. *tù-tì-is-ííz-à* ‘we have caused to fear’ (< *-iiz-y-a* < *-iiz-i-a*)
- d. *tù-tì-il-ííb-w-à* ‘we have been feared’ (< *-iib-u*)
- e. *tù-tì-is-ííb-w-à* ‘we have been made to fear’

In (12b), the perfect(ive) ending *-il-e* is seen to occur after a CV verb. Since all the tone-bearing units were originally H, MR applies to all but the first, hence intermediate *tù-tì-il-é*, followed by the initial H undergoing HTA shift, and the FV receiving a H% tone. In (12b) and (12c), the CPH is anticipated from the mora of the *-i* and *-u* which glide (with the intermediate *-y* being absorbed into the preceding *z*). The CPH is realised all H on *-ííz-i* and *-ííb-w* rather than LH, since rising tones are prohibited in Soga (and Ganda). The long vowel itself is attributed to an extra ‘imbricated’ (fused) *-il* morph, and the *z* in (12c) due to the degree 1 **-i: -ilil-i-a* > *-iil-i-a* > *-iiz-i-a* > *-iiz-y-aa* (by gliding and compensatory lengthening) > *-iiz-a* (with absorption of the *-y* and final vowel shortening). While there are many cases in Bantu where the FV is different after a bare vs. extended root, e.g. ROOT-*i* or *-e* vs. ROOT-EXT-*a* (cf. Grégoire 1979), such variation of the FV on

perfective *-*il-e* is (almost) unique to Ganda and Soga.⁶ Note in (12e) that the CAUS-PASS sequence introduces only one extra H. We would expect two H tones if *-*i* and *-*ɔ* each carried an independent *H tone in PB.

2.6 Local realisation

The last argument for innovation is the fact that the CPH is always realised locally, either on the syllable with *-*i* or *-*ɔ* or on its neighbour (e.g. when HTA has applied). If this H has been there from the days of PB, why has it not migrated or been subject to the major changes that the root and inflection stem tones have undergone (other than loss, of course)? It would only have to have been innovated before HTA shift in Nande and Soga. At least in the case of Soga, HTA shift is extremely recent, since the language is near-mutually intelligible with Ganda, which lacks HTA shift.

In fact, besides being closely tied to the *-*i* or *-*ɔ*, the general assumption has been that the CPH always originates in the final syllable of the verb stem.

The causative -*y-* and passive -*w-* suffixes surface in pre-final position in Lumarachi, immediately before the FV. As a result, when the H of the causative or passive suffix surfaces, it surfaces on the FV. (Marlo 2007: 255)

[...] I propose that there is a morphologically conditioned rule that assigns a H on the causative or passive vowel in the penultimate position [in Nande]. (Mutaka 1994: 115)

However, there are two environments that allow *-*i* and *-*ɔ* to occur earlier than the final syllable in the verb stem in some CPH languages, but have not been tonally scrutinised: (i) verb stem reduplication; (ii) sequences of *-*i* or *-*ɔ* + reciprocal -*an/-agan* + FV. While variations in total vs. partial verb-stem reduplication and a widespread requirement that reciprocalised causatives be realised -*i-an-i* limit the interest of these two contexts in some CPH languages, they both potentially occur in Ganda and Soga. Let us begin with reduplication.

In languages that truncate and prepose the “frequentative” reduplicant, tone is usually not copied. Typically the root H is only in the reduplicant while an ISH

⁶Michael R. Marlo indicates that this is also found in Luyia. One intriguing similarity is Mwiini G412, which replaces the -*e* of *-*il-e* with -*a* in the passive (Kisseberth & Abasheikh 1975: 251), but without -*u* appearing: *bush-il-e* ‘s/he hit’, *bush-il-a* ‘s/he was hit’ (Charles W. Kisseberth, p.c.).

is assigned to the whole reduplicated stem, as in Nande in (13), where the source moras of the H tones are underlined.⁷

- (13) Nande JD42 (Mutaka & Hyman 1990; Philippe Mutaka, p.c.)
- a. * \emptyset FV (e-ri) *hum-ir-an-a* ... → (e-ri) *hum-a+hum-ir-an-a* ...
'to hit for e.o.'
 - b. * \emptyset FV (e-rí) *tùm-ir-an-a* ... → (e-rí) *tùm-a+tum-ir-an-a* ...
'to send to e.o.'
 - c. *H FV (mó-tw-a-mú-) *húm-ir-à* ... → (mó-tw-a-mú-) *húm-a+hum-ir-à* ...
'we hit for him'
 - d. *H FV (mó-tw-á-mú-) *túm-ír-à* ... → (mó-tw-á-mú-) *tùm-a+túm-ír-à* ...
'we sent to him'
 - e. *H FV (mó-tw-á-) *húm-w-a* ... → (mó-tw-á-) *húm-w-a+húm-w-a* ...
'we were hit'
 - f. *H FV (mó-tw-a-) *túm-w-á* ... → (mó-tw-a-) *tùm-w-a+túm-w-á* ...
'we sent to him'

(13c) has the output of the derivation in (9). The root *túm* 'send' has an historical H, i.e. **tóm* 'send' (BLR 3055), which is anticipated onto the prefixes in (13b) and (13d), but deleted after the tense marker *a-* in (13f). (13e) crucially shows the suffixal *H shifting to *á-húm* plus another H on the second *húm* from passive *-w*. The CPH causes the FV *-a* to lose its L tone in (13e) (Mutaka 1994: 116), cf. (8b).

However, a different picture emerges from Soga in (14), which has full verb stem reduplication.⁸

- (14) Soga JE16 (Hyman, personal notes)
- a. (ò-kú-) *ty-à* 'to fear'
(ò-kú-) *ty-àà-ty-á*
 - b. (tù-) *tì-il-é* 'we have feared'
(tù-) *tì-il-è-tì-il-é*
 - c. (tù-) *tì-ìs-ííz-à* 'we have caused to fear'
(tù-) *tì-ìs-ííz-à-tì-ìs-ììz-á*

⁷For discussion of some of the tonal variation occurring in Bantu reduplication, see Downing (2003). This variation can depend on the type of reduplication even within the same language, as in Tonga N15 (Mkochi 2017).

⁸Again, the final H on the base form of (14b) and all of the reduplicated forms is from H%.

- d. (tù-) *tì-il-í**̀**w-à* ‘we have been feared’
 (tù-) *tì-il-í**̀**w-à-tì-il-ì**̀**w-á*
- e. (tù-) *tì-is-í**̀**w-à* ‘we have been made to fear’
 (tù-) *tì-is-í**̀**w-à-tì-is-ì**̀**w-á*

As seen, the CPH effect is on the first stem in reduplication. This would be expected if the /H/ were from the root—but is different from ISH which is calculated from the end. Thus, compare (14) with (15a) and its reduplication in (15b).

- (15) Soga JE16 (Hyman, personal notes)
- a. *è-bíí-ntú by-é tw-áá-sékwîl-é*
 ‘the things that we pounded’ (general past)
- b. *è-bíí-ntú by-é tw-áá-sékwîl-é-sékwîl-é*
 (-sékul- + -il-e → -sekwiil-e)

In both cases the ISH is assigned to the penultimate mora of the full stem. The fact that the CPH is realised on the first stem in (14c, 14d) suggests that it is behaving similarly to the realisation of the root H in the first stem of the reduplication, or is at least assigned to the first stem.

The second potential non-final effect of CPH is found with reciprocal *-agan* in Ganda and Soga. In this case CPH is realised only when **-i* or **-u* locally ‘interfixes’ between perfect(ive) **-il* and the FV *-a*. In the following examples, the *-il* of *-il-a* imbricates within the second syllable of reciprocal *-agan*, i.e. /-agan-il-e/ → -again-e, as in (16a).

- (16) Ganda JE15 (Hyman, personal notes)
- a. no CPH *tù-lùm-y-àgàin-é* ‘we hurt e.o.’⁹
- b. CPH *tù-lùm-y-àgáin-y-à* (idem)
- c. no CPH *tù-bá-kúb-y-àgàin-é* ‘we made them hit e.o.’
- d. CPH *tù-bá-kúb-àgáin-y-à* (idem)

As seen in (16a) and (16c), if causative *-i* is separated from the perfect(ive) ending, there will be no CPH, and **-il-e* will be realised with final *-e*. On the other hand, in (16b,d), where *-i* is realised right after imbricated *-agan*, there is a CPH and the form ends with *-a*. The same facts are found with passive *-u*. Although it is rare to get the reciprocal co-occurring with the passive, compare the following with *fumb-il-u* ‘marry’ (lit. ‘be cooked for’) in (17).

⁹In (16a) and (16b), ‘hurt’ has the underlying form *lùm-i*, literally ‘cause to bite’ from PB **dɔm* ‘bite’ (BLR 118, Bastin et al. 2002); *kúb-i* ‘make hit’ is transparently derived from PB **kɔb* ‘hit’ (BLR 1984, Bastin et al. 2002).

(17) Ganda JE15 (Hyman, personal notes)

- a. no CPH *bà-fúumb-il-w-àgàin-é* ‘they married e.o.’
- b. CPH *bà-fúumb-il-w-àgáin-w-à* (idem)

A lot of this has to do with an innovative reparsing of *-*agan* as -*a-gan* (Hyman et al. 2017), as seen in (18).

(18) Ganda JE15 (Hyman, personal notes)

- a. original *tù-ty-àgàin-é* < /tù-ti-agan-il-è/ ‘we feared e.o.’
- b. innovated *tù-ti-il-è-gàin-è* < /tù-ti-il-e+gan-il-è/ (idem)

While the inherited situation was one where *-*il-e* followed -*agan*, with which it imbricates, as in (18a), the alternate form in (18b) shows the inflectional ending being spelled out before and after the reciprocal, reparsed as -*gan*.¹⁰ The important thing that the above shows is that for there to be a CPH, the causative or passive extension must locally co-occur with perfect(ive) *-*il*, i.e. be combined, rather than being suffixed in separate positions within the verb stem. Be that as it may, the details found in Ganda, Soga or other languages likely modify and potentially obscure the possible origins of CPH. It is however unlikely that what we see in Soga, Nande, Fuliiru, etc. would have occurred as such in PB. We now turn to consider other cases of H tone extensions in the next section.

3 Other cases of H tone extensions

In the preceding section I enumerated six reasons why I think CPH was probably not a property of PB. A major reason in §2.1 was that CPH is found only in certain interlacustrine languages. Recall the reconstruction by Meeussen (1961; 1967) of all other extensions as *L or toneless. It would certainly be a strong argument in favour of reconstructing CPH in PB, or at least earlier than Proto-Interlacustrine, if CPH could be found outside the JD and JE zones. While most Bantu languages do not contrast tone on any verb extensions, three Bantu languages outside of zone J have been identified with contrastive H tone extensions. While I will argue

¹⁰Note in this context that besides the two forms in (17a) and (17b), *bà-fúumb-íil-w-àgàin-á* is also attested, suggesting that *gan-a* may be coming to be a constituent by itself, i.e. *bà-fúumb-íil-w-à+gàin-á*. In recent work on Ganda I found that *o-ku-láb-àgan-a* ‘to see e.o.’ can not only reduplicate as *o-ku-láb-agan-a+lab-agan-a* with expected full stem reduplication, but also as *o-ku-láb-àgan-a+gan-a*. I have not found any other verb extension that can reduplicate in this way.

that these H tones must have a different origin, I briefly consider each of these in turn.

The best known such case is Chewa N31b which shows the contrasts between toneless and H tone extensions seen in (19).

- (19) Chewa N31b (Mtenje 1986; Kanerva 1989; Hyman & Mtenje 1999a,b; Downing & Mtenje 2017)
- a. Ø tone extensions applicative *mat-il-a* ‘plaster/glue for’
 causative *mat-its-a* ‘cause to plaster/glue’
 reciprocal *mat-an-a* ‘plaster/glue e.o.’
 reversive tr. *mat-ul-a* ‘unplaster/unglue (tr.)’
dial. (Ntcheu variety) passive *mat-idw-a* ‘be plastered/glued’
- b. H tone extensions stative *mat-ik-á* ‘be plasterable/gluable’
 intensive *mat-its-á* ‘plaster/glue a lot/well’
 reversive intr. *mat-uk-á* ‘become unplastered/unglued’
dial. (Nkhotakota variety) passive *mat-idw-á* ‘be plastered/glued’

As seen in (19a) both the verb root *mat* ‘plaster/glue’ and the indicated extensions are toneless. In (19b) the second set of verb extensions assigns a H tone to the verb stem which by general rule links to the FV. Note that intensive *-its-* assigns a H, while segmentally homophonous causative *-its-* does not. The tonal behaviour of the passive extension is different in the Ntcheu and Nkhotakota varieties of Chewa.

The second language is Tonga N15 which, based on personal communications from Lee S. Bickmore and Winifred Mkochi, also contrasts causative *-is* with intensive *-ís*. More intriguing is the segmentally homophonous toneless stative *-ik* vs. passive *-ík* exemplified in (20).¹¹

- (20) Tonga N15 (Lee S. Bickmore & Winifred Mkochi, p.c.)
 medial phrase-final
kù-júl-à kù-júùl-à ‘to open’
 stative *kù-júl-ik-à kù-júl-ìik-à* ‘to be open’
 passive *kù-júl-ik-á kù-júl-ìik-à* ‘to be opened’

¹¹Tonga passive *-ík*, although intriguingly H tone, is clearly not cognate with PB **-ɔ*.

As Lee S. Bickmore and Winifred Mkochi (p.c.) put it: “The H on the stem-initial TBU [Tone-Bearing Unit] in each case is a M[elodic]H, which in stems with toneless roots, lands on V1 [...]. In the Passive you see a second H, from the extension. We analyze the extension H as docking onto the FV, and then shifting one mora to the left when the verb is phrase-final.”

The third language is Herero R30, which is clearly more intricate. Although I had some trouble interpreting the effects in the two sources, I believe that Table 1 summarises the extension tone patterns (omitting the extra L of the FV which the authors cite with the extension tone).

Table 1: Herero Verb Extensions and their Tones

Köhler (1958: 101ff)			Möhlig & Kavari (2008: 146ff)		
causative	<i>-is</i>	L	causative I	<i>-iṣ</i>	L
causative	<i>-ek</i>	L	causative II	<i>-ek, -ik (-e)</i>	L
			causative III	<i>-z(a)</i>	L
applicative	<i>-ir, -er, -in, -en</i>	L	applicative	<i>-er, -ir, -en, -in</i>	L
iterative	<i>-or, -uk</i>	L	reversive-	<i>-or, -ur, -ok,</i>	L
inversive	<i>-orok, -oror</i>	L	intensive (tr. + intr.)	<i>-uk, -oror, -urur, -orok, -uruk</i>	
intransitive	<i>-ák</i>	(H)	intransitive	<i>-ak, -ek, -ok, -uk</i>	L
intransitive- medial	<i>-ík, -ók, -úk</i>	(H)	neutro- passive	<i>-ík, -ik</i>	HL
reciprocal	<i>-asan</i>	(H)	reciprocal	<i>-áṣan</i>	(H)
passive	<i>-u</i>	H	passive	<i>-w, -éw, -íw</i>	H

In Table 1, (H) indicates that the extension will be H if the root is H, otherwise L. Setting these aside, this leaves the passive as consistently H, as shown in (21).

(21) Herero R30 (Köhler 1958: 108)

a.	<i>rw-a</i>	‘fight’	→	<i>ru-w-á</i>	‘be fought’
	<i>hong-a</i>	‘teach’	→	<i>hong-w-á</i>	‘be taught’
	<i>haam-a</i>	‘sit’	→	<i>haám-w-á</i>	‘be sat’
	<i>hakurur-a</i>	‘make hungry’	→	<i>hakúrúr-w-á</i>	‘be made hungry’
b.	<i>nw-á</i>	‘drink’	→	<i>nú-w-á</i>	‘be drunk’
	<i>hír-a</i>	‘water’	→	<i>hír-w-á</i>	‘be watered’
	<i>hínd-á</i>	‘send’	→	<i>hínd-w-á</i>	‘be sent’
	<i>hépek-a</i>	‘make poor’	→	<i>hépék-w-á</i>	‘be made poor’
	<i>húkúr-a</i>	‘undress’	→	<i>húkúr-w-á</i>	‘be undressed’
	<i>hívírik-a</i>	‘praise’	→	<i>hívírik-w-á</i>	‘be praised’
	<i>játátúrur-a</i>	‘unstitch’	→	<i>játátúrúr-w-á</i>	‘be unstitched’

In (21a) the verb root is L, while in (21b) the verb root is H. As seen, the H induced by the passive extension begins with the second syllable much as an ISH would be realised according to Meeussen (1961).

Despite any similarity, I would argue for several reasons that the H tone effects in Chewa, Tonga and Herero are not related to CPH. First, the causative extension is consistently not H. Second, in Herero, the (H) effects might be interpretable as spreading of the root H, which in other cases is sensitive to whether the vowel of the H root was long or short in PB. Concerning the passive, could the H have come from the marker *í* which introduces the prepositional agent phrase following a passive verb? For instance, *etemba má'rí nan-éwá í ókasíno* ‘the car is pulled by the donkey’ (Möhlig & Kavari 2008: 148). Could something similar be behind the H tone passive *-ídw* in certain Chewa varieties and *-ík* in Tonga? Finally, the H tone extensions seem to correlate with intransitivity in all three languages. I doubt that this is an accident. Rather, it suggests that there could have been an earlier H% boundary tone that became associated with intransitive verbs, since they are more likely to occur clause-finally than transitive verbs. Unless passive **-v* could have analogised its H tone to causative **-i*, based on the exceptional V shape of the two extensions and the tendency for both to occur in the last syllable (see below), it is not likely that the above effects are related to CPH. Of course, the similarity between nearby Chewa N31b and Tonga N15 may not be coincidental, or even Herero R30, if Möhlig (2009) is correct in assuming an origin of the language in south-eastern Malawi. The differences and sporadicity convince me, however, that CPH likely had an independent source, to which I now turn in the final section. In any event, even if they were related, it would still be a late development that does not require reconstruction to PB

as these languages belong to a late branch of the Bantu language family tree (i.e. they are East and South-West Bantu, cf. Grollemund et al. 2015).

4 Possible sources of CPH

If CPH did not exist in PB, we are then left with the question of why it exists in the interlacustrine languages enumerated in (3)? Any solution must account first for why CPH is limited to **-i* and **-v* and second why CPH is dependent on there being an inflectional suffixal H (ISH) (Meeussen's **H FV*). There are potentially two types of explanations, one morphological, one phonological: (i) perhaps *-i* and *-v* had a different status or structure from other verb extensions; (ii) perhaps the V shape played a key role, since all other verb extensions have the shape VC. In my past work I have entertained two different morphological explanations: (i) **-i* and **-v* used to be voice suffixes only later acquiring a FV (Hyman 2007a: 161); (ii) **-i* and **-v* used to be enclitic with perfect(ive) *-il* and the FV *-a* (Hyman & Katamba 1990: 153).

According to the first idea, **-i* was originally a FV. The potential relation is often noted to the subject-oriented ("agentive") deverbal nominaliser **-i*, e.g. **dim* 'cultivate' (BLR 968) → **mv-dim-i* 'farmer' (BLR 5491) (Bastin et al. 2002), although this **-i* was toneless (or **L*) in PB. Within the verb system there is a FV *-i* in NW Bantu that often marks stative, but can also be impositive, e.g. in Eton A71 (Van de Velde 2008: 122–123, 132): *búg* 'break (tr.)' → *búg-i* 'break (intr.)' vs. *són-bô* 'squat', *són-i* 'make s.o. squat'.¹² The stative FV *-i* is H tone in Abo A42 and Basaa A43a vs. the L tone verb extensions. I have noted 80 Basaa examples of derived *-i* verbs in Lemb & de Gastines (1973), for instance *sɔp* 'pour' → *sob-i* 'be poured' (cf. Bitjaa Kody 1990: 423–424). Concerning the passive, it is tempting to compare **-v* to **-ú*, which derives stative adjectives from intransitive verbs in certain Bantu languages but has been reconstructed with a close back vowel of the first degree of aperture. I have found 142 Ganda examples in Snoxall (1967), for instance *gum* 'be firm, solid', *gum-ú* 'firm, solid'; *tamiir* 'get drunk', *tamúv-ú* 'drunken'. The only way this hypothesis could be helpful is if the two V-shaped verb extensions were originally **H* tone FVs, which goes against my basic contention that CPH is innovative. Moreover, the above comparisons involve differences in tone or vowel height, not to mention grammatical and semantic differences. Hence, this approach remains at best highly speculative.

¹²Koen Bostoen (p.c.) has suggested that this zone A *-i* suffix may instead derive from Proto-Bantu **-ik-* through loss of the final **k*. While stative and impositive suffixes of this shape do exist elsewhere in Bantu, generally with the degree 2 vowel **ɪ* (which can harmonise to [ɛ]), the *-i* suffix is degree 1 in zone A languages with seven vowels.

6 Causative and passive high tone in Bantu: Spurious or proto?

As a second morphological explanation based on Ganda, Hyman & Katamba (1990: 155) propose that to derive CPH, the perfective *-il* + FV *-a* originally formed an enclitic in the presence of *-i* and *-u*. The basic idea is that the ISH would be assigned twice, once to the base verb, once to the enclitic, much as one finds in West African serial verb tone.¹³ This is illustrated in (22) for both Ganda and Soga, both of which would have the same trace of this alleged earlier structure.

(22) Ganda JE15 and Soga JE16 (Hyman, personal notes)

Input	Ganda	Soga
a. expected: <i>tú- [lek-i- + il-e +H]</i>	<i>*tú-lés-iz-è</i>	<i>*tù-lés-iz-è</i>
b. actual: <i>tú- [[lek-i- + el +H] il-a +H]</i>	<i>tú-lés-ézz-â</i>	<i>tù-lés-éiz-à</i> 'we have caused to leave'

As seen, the incorrect expected output is in (22a), where the spell-out of bi-morphemic **-il-e* surrounds the *-i* of the causativised root: *lek-i* 'make leave' → *les-i* → *les-il-i-e* → *les-iz-y-e* → **les-iz-e*. Instead, the Soga *lés-éiz-à* sequence observed in (22b) requires an even more complex 'cyclic' derivation of the sort discussed in Hyman (2003): *lek-i* → *les-i* → *les-el-i* → *les-ez-i* → *les-eiz-i-a* (imbrication), where *-el* is an extra morph known as a 'stabiliser' elsewhere in Bantu (cf. Cole 1955, Gowlett 1984). As also observed, *l* spirantises to *z*, causative *-i* glides to *y* and is absorbed by the preceding fricative, and the FV is *-a*.¹⁴ As shown, both the stem and the enclitic *-il-i-a* receive an ISH in (22b).

The bipartite structure in (22b) was designed to account not only for the double spell-out of the ISH, but also for the fact that CPH perfectives do not form a tone group (TG) with what follows (Hyman & Katamba 1990: 151). Subject to a number of conditions, a TG in Ganda consists of the verb + the first clitic or phonological word that follows. Within the TG, a sequence consisting of any number of Hs + Ls + Hs plateaus to all H across the two words, as may be seen from (23).

(23) Ganda JE15 (Hyman, personal notes)

- a. i. *y-á-tú-síb-idd-è* 's/he tied for us'
- ii. *y-á-tú-síb-idd-é =kí* 'what did s/he tie for us?'
- b. i. *y-á-tú-kúb-idd-ê* 's/he hit for us'
- ii. *y-á-tú-kúb-idd-é =kí* 'what did s/he hit for us?'

¹³This suggestion would make the most sense if **-il* were originally a verb, following Givón's (1971) general suggestion for Bantu verb suffixes. Although this remains to be confirmed, Voeltz (1977) suggests a pre-PB verb **gid* 'finish'.

¹⁴A full analysis would be more complex than this. It might also be tempting to view the second *e* of *les-eiz-i-a* as a FV, hence *les-e+iz-i-a*. Since a harmonising *-il/-el* 'stabiliser' is found elsewhere in the language, I believe this is the better interpretation.

However, H tone plateauing (HTP) is blocked by CPH as shown in (24).

(24) Ganda JE15 (Hyman, personal notes)

- a. *y-á-síb-ídd-w-â* ‘s/he was tied’
y-á-síb-ídd-w-ââ =*kí* ‘what was s/he tied by’
- b. *y-á-kúb-ídd-w-â* ‘s/he was hit’
y-á-kúb-ídd-w-ââ =*kí* ‘what was s/he hit by’

Note that the input tones are identical to those of the first line of (23b) which undergoes H tone plateauing, so it cannot be the final HL that blocks HTP. The fact that the final syllable is bimoraic in (24) is also irrelevant, since other final CVV syllables undergo HTP, for instance *y-à-ly-â* ‘s/he ate’ vs. *y-à-ly-ââ=kí* ‘what did s/he eat?’.

Importantly, blocking of TG-formation will take place only if the CPH occurs in the last syllable, and this only if *-i* or *-u* combines with **-il* plus the FV *-a*, as in (25a) involving the causative verb *som-es-i* ‘teach, cause to learn’.

(25) Ganda JE15 (Hyman, personal notes)

- a. *tú-sóm-és-ézz-â* ‘we have taught’
tú-sóm-és-ézz-ââ =*kí* ‘what have we taught?’
- b. *tú-sóm-és-è-gàŋ-è* ‘we have taught e.o.’
tú-sóm-és-é-gá-é =*kí* ‘what have we taught e.o.?’

There is however no CPH in (25b), where *-i* appears not to be realised, hence HTP applies. It is likely that the verb stem has been reanalysed as *som-es-i-e+gan-i-e*, where the repeated *-i-e* is from **-il-e*. Note that unlike the above stem + enclitic, this bipartite structure appears to be a compound that does not block TG formation. Since only Ganda and Soga require the perfective *-il* to get CPH, and since they alone require the FV *-a*, I have my doubts about the enclitic explanation. It does however have the merit of attributing the CPH to the ISH, which Jones (2014) implements in a synchronic phonological analysis of Nande JD42: “[...] the Spurious tone is claimed to be nothing more than the second H tone assigned in Complex tone [...]” (Jones 2014: 232).

Turning then to possible phonological accounts, could the CPH derive from the unique phonological properties of the two extensions which are realised late in the verb stem/word, typically occurring in the last syllable? As the only V-shaped extensions, **-i* and **-ɔ* usually form a CVV syllable with the FV. In (26) I consider what would be needed if we assume that CPH is from a single inflectional (‘melodic’) H that has two realisations.

Before concluding, I want to mention another logical source of evidence for the tone of extensions in PB: nominalisation. It would be significant if *-i and *-ɔ were to provide an extra H tone in nominalisations, especially if found in non-CPH languages. Although this has not been exhaustively researched, the data to date are mixed. As Van Otterloo (2011: 260) notes: “[...] the H tones of CS [Causative] and PS [Passive] are not always present in nominal form.” In the following Fuliiru JD63 examples, (27a) shows the transfer of CPH into the noun, while there is no CPH transfer in (27b).

(27) Fuliiru JD63 (Van Otterloo 2011)

- a. *kú-fúú-s-â* ‘to escape’ → *í=kí-fúús-ô* ‘means of escape’
kú-tyábí-rí-z-â ‘to thunder’ → *á=ká-tyábíríz-ô* ‘thunder’
- b. *kú-bút-ír-w-â* ‘to give birth’ → *ú=mú-bútír-wà* ‘native (of a place)’
kú-bùgú-z-â ‘to pay fee’ → *ú=mú-bùgù-zà* ‘tax collector’

Interestingly, of the 70+ nominalisations which Van Otterloo (2011: 288–292) provides, all end in L except for *í-shùvy-ô* ‘answer’, *kí-búúz-ô* ‘question’, and *ká-hùgw-ê* ‘loneliness’, all of which involve an input causative -i or passive -u.¹⁶

In Ganda, class 1 deverbal agentives are generally derived with -i, but with -á after the causative or passive extension. The nominalisations in (28) are from Ashton et al. (1954) and Snoxall (1967), which I cite without the augment.

(28) Ganda JE15 (Ashton et al. 1954; Snoxall 1967)

- a. *ku-lim-a* ‘to cultivate’ → *mu-lim-i* ‘farmer’
ku-kól-à ‘to work’ → *mu-kóz-i* ‘worker’
ku-lagír-a ‘to direct’ → *mu-lagír-i* ‘instructor’
ku-lábir-a ‘to supervise’ → *mu-lábir-i* ‘overseer’
- b. *ku-woóz-a* ‘to exact dues’ → *mu-wóóz-à (wool-i)* ‘tax-collector’
ku-liran-a ‘to be next to’ → *mu-liráàn-w-à (liraan-u)* ‘neighbour’
ku-záàl-is-a ‘to assist at childbirth’ → *mu-záális-à (záal-is-i)* ‘midwife’
ku-yígír-a ‘to teach’ → *mu-yígír-à (yígír-i)* ‘teacher’
- c. *ku-gob-a* ‘to drive’ → *mu-gób-â* ‘driver’
ku-sik-a ‘to inherit’ → *mu-sík-â* ‘heir’
ku-vubuk-a ‘to reach puberty’ → *mu-vúbúk-â* ‘adolescent’

¹⁶I also note that none of them ends with final -a.

As seen most clearly in the first two examples of (28b), the H of *-á* has the same realisation as the ISH, being realised as H on the second mora, followed by all Ls. This is obscured in the third and fourth examples, where the root is also H, hence causing the H of *-á* also to undergo MR. A final HL falling tone is found on the *-a* in (28c), although without a causative or passive morpheme.¹⁷ While the *-i* vs. *-a* nominaliser is attested beyond the CPH languages—and may be PB (Meeussen 1967: 93), it interestingly parallels the Ganda/Soga perfective *-e* vs. *-a* facts. As for the transfer of the CPH to nominalisations, my suspicion is that more transparent derivations or recent nominalisation may be more likely to parallel the tones of the input verb. In any case, a bigger corpus is needed from more languages.

5 Conclusion

To conclude, I repeat the position of Hyman & Katamba (1990) that the CPH was innovated in the interlacustrine area, and that it had to do with the presence of a ‘melodic’ tone, if not also the perfect(ive) suffix. In most of the interlacustrine CPH languages listed in (3), there is a noteworthy prevalence of HTA which ultimately leads to a H > L tone inversion. Since HTA is otherwise rather limited in Bantu (vs. perseverative tone spreading and tone shifting) this correlation should be borne in mind – even though it is not obvious whether or how it might feed into the CPH facts we have seen.¹⁸ In any case, the contrastive H tone extensions attested outside the interlacustrine area are likely a separate development, possibly having to do with marking intransitive verb finality or a H tone agentive preposition following the passive verb. It is also striking how many minimal pairs there are among the extensions.

(29) Chewa N31b and Tonga N15 (Hyman, personal notes), Herero R30
(Wilhelm J. G. Möhlig, p.c.)

- a. Chewa *-its* ‘causative’ vs. *-íts* ‘intensive’;
-ik ‘impositive’ vs. *-ík* ‘stative’
- b. Tonga *-is* ‘causative’ vs. *-ís* ‘intensive’;
-ik ‘stative’ vs. *-ík* ‘passive’

¹⁷The first three nouns in (28c) also have the regular variants *mu-gob-i* and *mu-sisi*, while the third has the variant *mu-vubúf-ù* derived via the deverbal adjectival suffix *-ù* mentioned above.

¹⁸Interestingly, many interlacustrine languages convert final H to HL and ultimately anticipate the H off the FV and onto the penultimate syllable. In Hyman (2007b: 22) I hypothesised that this ‘push’ from the right edge sets more general HTA in motion. Perhaps the proposed change of final H tone C-V-V to HL could be the missing link between HTA and CPH.

- c. Herero *-ak* ‘stative’ vs. *-ák* ‘neutro-passive’;
e.g. *zúv-ak-a* ‘be heard’ vs. *zúv-ák-a* ‘be hearable’

Hopefully we will find more evidence that will lead with certainty to a solution for both groups of H tone extension languages. For now, perhaps the only definitive ‘moral’ we can draw from all of the above comes from the great A.E. Meeussen (1973: 18) himself: “As a general conclusion, one might suggest that future research in comparative Bantu should consist mainly in team work, in which all available evidence, examined critically, is taken into account.” Time to get back to (team) work!

Acknowledgements

I am grateful for the feedback received at the *International Conference on Reconstructing Proto-Bantu Grammar* as well as for the comments from two reviewers.

References

- Ashton, Ethel O., Enoch M. K. Mulira, E. G. M. Ndawula & Archibald N. Tucker. 1954. *A Luganda grammar*. London: Longmans, Green & Co.
- Bastin, Yvonne, André Coupeze, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bickmore, Lee S. 2000. Tones and glides in Namwanga. In Vicky M. Carstens & Frederick Parkinson (eds.), *Advances in African linguistics* (Trends in African Linguistics 4), 135–149. Trenton, NJ; Asmara: Africa World Press.
- Bickmore, Lee S. 2007. *Cilungu phonology* (Stanford Monographs in African Languages). Stanford, CA: Center for the Study of Language & Information.
- Bitjaa Kody, Zachée. 1990. *Le système verbal du basaá*. Yaoundé: University of Yaoundé. (3rd cycle dissertation).
- Cole, Desmond T. 1955. *An introduction to Tswana grammar*. London; New York: Longmans; Green.
- Coupeze, André. 1955. *Esquisse de la langue holoholo* (Annales – Série in-8° – Sciences de l’homme – Linguistique 12). Tervuren: Royal Museum of the Belgian Congo.
- Downing, Laura J. 2003. Compounding and tonal non-transfer in Bantu languages. *Phonology* 20(1). 1–42.

- Downing, Laura J. & Al Mtenje. 2017. *The phonology of Chichewa* (The Phonology of the World's Languages). Oxford: Oxford University Press.
- Ebarb, Kristopher J. 2014. *Tone and variation in Idakho and other Luhya varieties*. Bloomington, IN: Indiana University. (Doctoral dissertation).
- Ebarb, Kristopher J. 2016. An overview of Kabarasi verb tone. *Language* 92(3). e134–e191.
- Ebarb, Kristopher J. & Michael R. Marlo. 2010. Non-spurious H-toned extensions in Nyala-West. Paper presented at the 41st Annual Conference on African Languages, Toronto, 6–8 May 2010. Handout.
- Givón, Talmy. 1971. On the verbal origin of the Bantu verb suffixes. *Studies in African Linguistics* 2(2). 145–163.
- Gowlett, Derek F. 1984. Stabilization in Bantu. *African Studies* 43(2). 187–203.
- Grégoire, Claire. 1979. Les voyelles finales alternantes dans la conjugaison affirmative des langues bantoues centrales. *Journal of African Languages and Linguistics* 1(2). 141–172.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guthrie, Malcolm. 1967–71. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages* (4 volumes). Farnborough: Gregg International.
- Hyman, Larry M. 1990. Boundary tonology and the prosodic hierarchy. In Sharon Inkelas & Draga Zec (eds.), *The phonology-syntax connection*, 109–125. Chicago, IL: University of Chicago Press.
- Hyman, Larry M. 2003. Sound change, misanalysis, and analogy in the Bantu causative. *Journal of African Languages and Linguistics* 24(1). 55–90.
- Hyman, Larry M. 2007a. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville, MA: Cascadilla Press.
- Hyman, Larry M. 2007b. Universals of tone rules: 30 years later. In Tomas Riad & Carlos Gussenhoven (eds.), *Tones and tunes. Volume 1: Typological studies in word and sentence prosody* (Phonology and Phonetics 12.1), 1–34. Berlin: Mouton de Gruyter.
- Hyman, Larry M. 2018. The autosegmental approach to tone in Lusoga. In Diane Brentari & Jackson L. Lee (eds.), *Shaping phonology*, 47–69. Chicago, IL: University of Chicago Press.

- Hyman, Larry M., Sharon Inkelas & Fred Jenga. 2017. Multiple exponence in the Lusoga verb stem. In Claire Bower, Laurence Horn & Raffaella Zanuttini (eds.), *On looking into words (and beyond): Structures, relations, analyses* (Empirically Oriented Theoretical Morphology and Syntax 3), 171–188. Berlin: Language Science Press. DOI: 10.5281/zenodo.495445.
- Hyman, Larry M. & Francis X. Katamba. 1990. Spurious high-tone extensions in Luganda. *South African Journal of African Languages* 10(4). 142–158.
- Hyman, Larry M. & Joyce T. Mathangwane. 1999. Meeussen's rule at the phrase level in Kalanga. In Rosalie Finlayson (ed.), *African mosaic: A festschrift for J.A. Louw on his 75th birthday*, 173–202. Pretoria: University of South Africa Press.
- Hyman, Larry M. & Al Mtenje. 1999a. Non-etymological high tones in the Chichewa verb. *Malilime: The Malawian Journal of Linguistics* 1. 121–156.
- Hyman, Larry M. & Al Mtenje. 1999b. Prosodic morphology and tone: The case of Chichewa. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The prosody-morphology interface*, 90–133. Cambridge: Cambridge University Press.
- Hyman, Larry M. & Armindo S. A. Ngunga. 1994. On the non-universality of tonal association conventions: Evidence from Ciyao. *Phonology* 11(1). 25–68.
- Hyman, Larry M. & Nzama Valinande. 1985. Globality in the Kinande tone system. In Didier L. Goyvaerts (ed.), *African linguistics: Essays in memory of M.W.K. Semikenke* (Studies in the Sciences of Language Series 6), 239–260. Amsterdam: John Benjamins.
- Jones, Patrick J. 2014. *Tonal interaction in Kinande: Cyclicity, opacity, and morphosyntactic structure*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Kanerva, Jonni M. 1989. *Focus and phrasing in Chichewa phonology*. Stanford, CA: Stanford University. (Doctoral dissertation).
- Kisseberth, Charles W. & Mohammad I. Abasheikh. 1975. The perfect stem in Chi-mwi: Ni and global rules. *Studies in African Linguistics* 6(3). 249–266.
- Köhler, Oswin. 1958. Tongestalt und Tonmuster in der Infinitivform des Verbum im Herero. *Afrika und Übersee: Sprachen, Kulturen* 42. 97–110, 59–72.
- Lemb, Pierre & François de Gastines. 1973. *Dictionnaire basaá – français*. Douala: Collège Libermann.
- Liphola, Marcelino M. & David Odden. 1999. The tonal pattern of glides in Shimakonde. In Vicky M. Carstens & Frederick Parkinson (eds.), *Advances in African linguistics: Papers presented at the 28th Annual Conference on African Linguistics, held July 18–22, 1997, Cornell University*, 177–188. Trenton, NJ; Asmara: Africa World Press.

6 Causative and passive high tone in Bantu: Spurious or proto?

- Maho, Jouni F. 2009. *NUGL online: The online version of the new updated Guthrie list, a referential classification of the Bantu languages*. https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf.
- Marlo, Michael R. 2007. *The verbal tonology of Lumarachi and Lunyala-West: Two dialects of Luluyia (Bantu, J.30, Kenya)*. Ann Arbor, MI: University of Michigan. (Doctoral dissertation).
- Marlo, Michael R. 2008. Tura verbal tonology. *Studies in African Linguistics* 37(2). 153–243.
- Marlo, Michael R. 2009. Khayo verbal tonology. *Africana Linguistica* 15. 77–129 + 16 pages of supplementary material online.
- Marlo, Michael R. 2013. Verb tone in Bantu languages: Micro-typological patterns and research methods. *Africana Linguistica* 19. 137–234.
- Meeussen, Achiel Emiel. 1961. Le ton des extensions verbales en bantou. *Orbis, Bulletin international de documentation linguistique* 10(2). 424–427.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeussen, Achiel Emiel. 1971. *Eléments de grammaire lega* (Archives d'ethnographie 15). Tervuren: Royal Museum for Central Africa.
- Meeussen, Achiel Emiel. 1973. Comparative Bantu: Test cases for method. *African Language Studies* 14. 6–18.
- Mkochi, Winifred. 2017. Resolving verbal reduplication paradoxes in Malawian Tonga. *Stellenbosch Papers in Linguistics Plus* 52. 101–126.
- Möhlig, Wilhelm J. G. & Jekura U. Kavari. 2008. *Reference grammar of Herero (Otjiherero): Bantu language of Namibia, with a glossary Otjiherero – English – Otjiherero* (Southern African Languages and Cultures 3). Cologne: Rüdiger Köppe.
- Möhlig, Wilhelm J. G. 2009. Historiography on the basis of contemporary linguistic data: The Herero case. In Wilhelm J. G. Möhlig, Frank Seidel & Marc Seifert (eds.), *Language contact, language change and history based on language sources in Africa* (Sprache und Geschichte in Afrika 20), 159–186. Cologne: Rüdiger Köppe.
- Mtenje, Al. 1986. *Issues in the nonlinear phonology of Chichewa*. London: University College London.
- Mutaka, Ngessimo M. 1994. *The lexical tonology of Kinande* (LINCOM Studies in African Linguistics 1). Munich: Lincom Europa.
- Mutaka, Ngessimo M. & Larry M. Hyman. 1990. Syllables and morpheme integrity in Kinande reduplication. *Phonology* 7(1). 73–119.
- Odden, David & Lee S. Bickmore. 2014. Melodic tone in Bantu: Overview. *Africana Linguistica* 20. 3–13.

- Polak-Bynon, Louise. 1975. *A Shi grammar: Surface structures and generative phonology of a Bantu language* (Annals – Series in-8° – Human Sciences 86). Tervuren: Royal Museum for Central Africa.
- Snoxall, Ronald A. 1967. *Luganda – English dictionary*. Oxford: Clarendon Press.
- Van de Velde, Mark. 2008. *A grammar of Eton* (Mouton Grammar Library 46). Berlin: Mouton de Gruyter.
- Van Otterloo, Karen. 2011. *The Kifuliiru language: Phonology, tone, and morphological derivation*. Vol. 1 (SIL International Publications in Linguistics 146). Dallas, TX: SIL International.
- Van Otterloo, Karen. 2014. Tonal melodies in the Kifuliiru verbal system. *Africana Linguistica* 20. 385–403.
- Voeltz, Erhard F. K. 1977. *Proto Niger-Congo verb extensions*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).

Chapter 7

Reconstructable main clause functions of Proto-Bantu applicative **-id*

Sara Pacchiarotti

Ghent University

This chapter presents evidence in favour of reconstructing at least three main clause-level functions of the Proto-Bantu (PB) applicative **-id*: (i) the introduction of a semantic role which cannot be expressed otherwise with an underived verb root; (ii) the focalisation of a constituent with a Location-related semantic role (most commonly General Location of the event); and (iii) the addition of aspectual and semantic nuances of completeness, iterativity or thoroughness to the meaning of the verb root. With respect to the syntactic function in (i) evidence is provided in favour of the hypothesis that PB **-id* introduced a Spatial/Goal or Location argument and that this function later extended to Human Goals and Beneficiaries. Finally, the chapter establishes possible diachronic relations among the three reconstructed functions of **-id*.

1 Introduction

This chapter deals with the reconstructable main clause functions of the highly polyfunctional and semantically underspecified Proto-Bantu (PB) applicative suffix **-id*.¹ Although virtually all Bantu grammars and other scholarly work consider the applicative synchronically and diachronically first and foremost as a syntactic valence-increasing device, I argue that there are minimally three main

¹Most of this chapter is an updated and revised version of the chapter titled “Historical origin(s) and function(s) of the PB applicative **-id*” in Pacchiarotti (2020). I am much indebted to Thilo C. Schadeberg who made one of his unpublished manuscripts on the PB applicative suffix electronically available to me. His manuscript greatly helped me develop the ideas presented here.



clause functions of **-Id* which can be reconstructed to PB: (i) introducing a semantic role which could not otherwise be expressed with an underived verb root; (ii) narrow-focusing a constituent which usually has a Location-related semantic role; and (iii) adding semantic nuances such as completeness, iterativity or thoroughness to the meaning of the verb root. My evidence comes from the synchronic distribution of these three functions and from attested directions of change within and outside Africa. I propose that one single PB **-Id* suffix carried out these three functions. However, the possibility that two or more originally distinct verbal derivational suffixes ended up as **-Id* in PB, due to phonological mergers (see Hyman 2007), can by no means be ruled out.

Second, with respect to the function in (i), I argue along the lines of Voeltz (1977) and Schadeberg (2003) and contra Trithart (1983) that **-Id* had both a Spatial Goal and Beneficiary function in PB or further back. Specifically, I offer evidence for the fact that **-Id* originally introduced a Spatial Goal argument and was only later extended to Human Goals, though still at the PB stage. Given that current scholarship believes that it is highly likely that minimally some Niger-Congo (NC) node higher than Benue-Congo had a system of so-called “extensions” (i.e. verbal derivational suffixes) (Hyman 2007; 2011; 2014; 2018; Blench (2022 [this volume])), these new historical insights on the original functions of PB **-Id* are immediately relevant not only for the reconstruction of PB but also for higher nodes within NC.

In line with these objectives, this chapter is organised as follows. In §2, I present the synchronic construction types involving **-Id*. In §3, I discuss two attempts at reconstructing the form(s) and function(s) of an applicative derivational suffix in PB and/or further back in NC (Voeltz 1977, Trithart 1983). Both attempts concur in reconstructing the applicative as a valence-increasing syntactic device, but they diverge in the peripheral semantic role that it would have introduced. In §4, I argue in favour of an original Spatial Goal or Location-oriented function of **-Id* and against an original Beneficiary function as proposed by Trithart (1983). My argumentation assumes that one of the functions of PB **-Id* was to introduce a Spatial Goal or a General Location participant to the argument structure of a root. In §5, I assess the synchronic distribution of the selected synchronic construction types laid out in §2 and argue that their corresponding functions should be reconstructed (minimally) to PB. In §5.1, I show the underlying conceptual relation (and hence diachronic link) between functions (i) and (iii) mentioned above. In §5.2, I do the same for functions (i) and (ii) mentioned above. In §5.3, I propose a diachronic path that could account for the conceptual and diachronic relatedness of the three functions. §6 concludes this chapter.

Because an internal classification of Bantu languages based on shared phonological and morphological innovations is lacking, in discussing the synchronic distribution of applicative constructions expressing the functions in (i), (ii), and (iii) in §5, I resort to the latest and most comprehensive lexicon-based Bantu phylogeny, i.e. Grollemund et al. (2015). Thus, by PB, I mean the ancestral language spoken at node 0 or 1 in their phylogeny, and all capitalised uses of cardinal directions refer to their subgroupings (e.g. Eastern, Central-Western, etc.).

2 Functions of widespread main clause construction types involving **-id*

Based on novel data and on previous works on **-id* (see especially Trithart 1983, but also Rugemalira 1993; Kimenyi 1995; Mabugu 2001; Creissels 2004; Jerro 2016), Pacchiarotti (2020) argues that es of **-id* in modern Bantu languages can participate minimally in five structurally and functionally distinct constructions, called, for lack of better names, Types A, B, C, D and E. The structure and function(s) of each construction type are summarised in Appendix A. I take some of these structures and their corresponding functions as a synchronic point of departure for the diachronic considerations in the remainder of this chapter. For a more in-depth discussion of each type, see Pacchiarotti (2020).

In Type A applicative constructions, the applicative morpheme expands the argument structure of the verb root by introducing an obligatorily present applied phrase which could not otherwise be expressed with that root.² This expansion may result in a clear-cut, indisputable increase in the syntactic valence of the derived verb stem, but need not to. Roots participating in this construction type do not “subcategorise” for a particular semantic role and the sole morphological device to express such semantic role is applicative morphology. Within any given language, the semantic roles for which an applicative is required are a lexical property of individual verb roots.³ Thus, the semantic roles that can be mapped onto the applied phrase vary on a root-by-root and language-by-language basis

²I use the term “applied phrase” to refer to the morphosyntactic entity introduced or semantically/pragmatically manipulated by the applicative without any specification of its syntactic category or argumenthood status. This means that, on a language-specific basis, an applied phrase could be an adjunct phrase (infinitive complements and clausal adjuncts; see Hawkins & Hyman 1974; Harford 1993), a prepositional phrase, a noun phrase marked by a locative noun class marker, or an unmarked noun phrase with (some) object properties.

³By “lexical” in this context I mean the properties of a linguistic unit which are memorised or cognitively stored in long-term memory, such as form, meaning, argument structure, and restrictions on morphological, syntactic, and pragmatic use.

but never include Agent or Patient. This mapping depends heavily on the lexical meaning of the verb root, the meaning of other constituents present in the clause, and on the communicative intention of the speaker (Stapleton 1903; Voeltz 1977; Schaefer 1985; Bresnan & Moshi 1990; du Plessis & Visser 1992; Rugemalira 1993; Rapold 1997; Mabugu 2001; Creissels 2004; Thwala 2006; Cann & Mabugu 2007; Seidel 2008; Jerro 2016; Sibanda 2016, among others). For instance, in the Eastern Bantu language Nyambo JE21, the root *gamb* ‘speak’ in (1) requires applicative derivation to co-occur with a locative phrase expressing General Location in a main clause.⁴

- (1) Nyambo JE21 (Rugemalira 1993: 71)
- a. *gamb-ir-á omu-nju*
speak-APPL-FV LOC18-house
‘speak in the house’
 - b. **gamb-á omu-nju*
speak-FV LOC18-house
(intended meaning: ‘speak in the house’)

In Type B constructions, as in Mbuun B87 in (2a), the applicative expands the root’s argument structure by introducing an obligatorily present applied phrase expressing a semantic role which could have been syntactically expressed as an optional oblique, as in (2b).⁵

- (2) Mbuun B87 (Bostoen & Mundeke 2011: 187)
- a. *o-á-kónné máám ó-te*
S3:1-PRS.PROG-plant.APPL mother CL3-tree
‘He is planting a tree for my mother.’
 - b. *o-á-kón ó-te óngírá máám*
S3:1-PRS.PROG-plant CL3-tree for mother
‘He is planting a tree for my mother.’

Unlike in Type A, in Type B, the function of applicative derivation is not purely syntactic. The fact that the free translation of alternations such as (2a) and (2b) is almost always the same in scholarly works might be misleading. Since there

⁴General Location in this chapter means the location where the event takes place.

⁵Bostoen & Mundeke (2011: 182) observe that the Mbuun reflex of the vowel of PB *-*id* is /e/. The reflex of its consonant involves metathesis and assimilation to the root final consonant of the verb root, e.g. *ka-kón* ‘to plant’ > *ka-kónne* ‘to plant-APPL’, *ka-sís* ‘to leave’ > *ka-sísse* ‘to leave-APPL’.

is an alternative way to express a given semantic role, with a given root, in a given construction, in a given language, there might be a semantic or discourse-related difference between the construction with and without the applicative. The applicative construction can imply something about the added semantic role that the construction with the root did not imply (Mabugu 2001), or the optional applicative construction is used when the participant expressed by the applied phrase is a discourse topic (Trithart 1983: 181; Rapold 1997; Peterson 2007). These functions are seldom described or investigated in Bantu literature.

In Type C constructions, the applicative suffix expands the argument structure of the verb root by introducing an obligatorily present applied phrase which could be optionally expressed in the construction with just the root. Unlike in Type B, based on data currently available, the obligatory present applied phrase in Type C usually has a Location-related semantic role, very often General Location, indicating where the event described by the verb root takes place. Besides introducing an obligatorily present applied phrase, the applicative suffix in Type C performs semantic or pragmatic functions on the applied phrase alone or on the whole clausal construction which are different from those observed for Type B. In Type C, the applicative can: (i) place the applied phrase under some kind of narrow focus; (ii) change the “orientation” of the Location applied phrase;⁶ or (iii) indicate that the action described by the root occurs habitually at a certain location. The structure that expresses these three functions is identical (see Appendix A) and this is why I group them together in Pacchiarotti (2020). Crucially, only language-specific Bantu roots which do *not* require an applicative to express General Location within a main clause can participate in Type C applicative constructions. For reasons of space and for the relevance they bear on the present chapter, I only illustrate functions (i) and (ii) for Type C. The narrow-focusing ability of the applicative is illustrated in (3b). In the Fwe sentence in (3a), everything is new information, the locative phrase expressing General Location is not syntactically obligatory, and the clause could be an answer to “What do they do?” In (3b), the presence of applicative derivation on the verb makes the clause

⁶The term “orientation” or “argument orientation” originates in formal semantics (Keenan & Faltz 1985; Nam 1995; Kracht 2002) and refers to the semantic effects of some English locative modifiers in combination with different types of predicate. To give an example, if the English sentence *John saw Mary in the park* is true, then *John saw Mary* and *Mary was in the park* are also true. However, *John was in the park* does not logically or necessarily follow as a true statement from *John saw Mary in the park* – John could have been across the street in a coffee shop when he saw Mary, and the sentence would still be true (see Keenan & Faltz 1985: 158ff). This means that the orientation of the locative modifier *in the park* in the English sentence *John saw Mary in the park* is towards the direct object NP but not necessarily towards the subject NP.

a more felicitous answer to the question “Where do they build?” Note that the presence of the applicative makes the locative phrase obligatory and the target of new information focus.

- (3) Fwe K402 (Hilde Gunnink, p.c.)
- a. *βàzá:kà kùmbári yórwìzì.*
βa-zá:k-a (ku-N-βári í-o-ru-ízi)
 s3:2-build-FV CL17-CL9-near PA9.CONN-CL11-river
 ‘They build (close to the river).’
- b. *βàzá:kìrà kùmbári yórwìzì.*
βa-za:k-ir-á ku-N-βári í-o-ru-ízi
 s3:2-build-APPL-FV CL17-CL9-near PA9-CONN-CL11-river
 ‘They build close to the river.’

In some Bantu languages, applicative morphology can be used to widen the orientation of a locative phrase from involving the object of a transitive verb root to also involving the subject of that transitive verb root. Trithart (1977) and Hyman et al. (1980) report this phenomenon in Haya JE22. Consider (4a) vs. (4b).

- (4) Haya JE22 (Hyman et al. 1980: 578)
- a. *η-ka-bón-a kat’ ómú-nju*
 I-PST-see-FV Kato LOC18-house
 ‘I saw Kato [while he was] in the house.’
- b. *η-ka-bón-el-a kat’ ómú-nju*
 I-PST-see-APPL-FV Kato LOC18-house
 ‘I saw Kato [while I was] in the house.’⁷

Hyman et al. (1980) argue that in (4a) the locative phrase ‘in the house’ is part of the verb complement (i.e. it modifies ‘Kato’), while in (4b) ‘in the house’ is outside of the verb complement and relates to the entire assertion, including the subject’s relationship to the event. In other words, once the applicative is present in the construction, the locative phrase expressing General Location has scope over the entire event (see also Grégoire 1998).

In Type D constructions, the applicative suffix does not introduce an applied phrase. Instead, it indicates that the action described by the root is performed

⁷The apostrophe in (4a) and (4b) means that the final vowel of the word *Kato* elides when a following word starts with a vowel, as *omunyu* does here.

to completion, or that the action is performed continuously, with intensity, persistence, excess, or repetition, among other qualities, as in (5). The number of applicative suffixes used to convey these meanings in North Boma B82 depends on the syllable shape: CVC shapes require one suffix, while C shapes require two.⁸

(5) North Boma B82 (Stappers 1986: 41)

<i>lab-a</i>	‘leave’	<i>lib-il-e</i>	‘leave earnestly’
<i>bɔm-a</i>	‘kill’	<i>bɔm-εη-ε</i>	‘kill everything’
<i>l-é</i>	‘eat’	<i>l-íl-il-é</i>	‘eat up everything’

Finally, Type E pseudo-applicative constructions are irregular, non-productive results of applicative derivation. In this construction type, lexicalised applicativised verb stems (usually displaying one or two applicative derivations) do not introduce an applied phrase to the argument structure of the verb root from which they are synchronically or historically derived. Applicative suffixes present on the verb stem also do not perform semantic or pragmatic functions like those described for Types B, C, or D. As an example, consider the Tswana S31 applicative stem *lalel* [lál-él] ‘have dinner’ which synchronically looks as derived from *lal* [lál] ‘lie down, stay overnight, spend the night’.⁹ As shown in (6a), *lal* is syntactically intransitive; *lalel* in (6b) is also syntactically intransitive, as the thing being eaten is optionally introduced by an instrumental preposition.

(6) Tswana S31 (Creissels 1999: 148–149)

a. *Re tlaa lala mo nageng.*

rì-ttàà-lál-à (mó nàxé-ŋ)

s1PL-FUT-lie_down-FV LOC18 CL9.bush-LOC

‘We will lie down/spend the night/sleep (in the bush).’

b. *Re tlaa lalela ka dikgobe.*

rì-ttàà-lálél-à (ká dí-q^h:bè)

s1PL-FUT-have_dinner-FV INS CL10-beans_and_maize

‘We will have dinner (with beans and maize).’

⁸This is true not only of North Boma, but of Bantu more generally. In some languages up to three applicative suffixes are required, depending on the phonotactics of the verb root, cf. e.g. Sharman (1963: 67–69) for Bemba M42. This suggests that Guthrie’s (1967–71) *-idid ‘persistent’ (Comparative Series 2189) should be amended to *-id ‘persistent’. Very likely, Guthrie’s “persistent” label could be replaced by “applicative” (see §5.1).

⁹The pseudo-applicative *lalel* is the regular reflex of PB **dáadid* ‘have supper, look after, brood’, present in zones J, L, M, and S according to BLR3 (Bastin et al. 2002), and derived from PB **dáad* ‘lie down, sleep, spend the night’, attested in all Bantu zones except P. The applicative verb stem likely already existed in some higher nodes within Bantu.

In the remainder of this chapter I focus on Types A, C and D, because Type B constructions likely developed in languages which came to have other ways of expressing a given semantic role, for instance through prepositions, and Type E are the “death point” of productive applicative morphology. These three construction types express functions (i), (ii), and (iii) introduced in §1, respectively. From a historical perspective, there are at least three possibilities for how they might be related: (a) Types A, C, and D are not diachronically related (or only some are, e.g. Type A and Type D but not Type C), meaning that there would have been two or more morphemes which ended up looking like **-id* through phonological mergers in PB or further back (see Hyman 2007); (b) there is a conceptual link between Types A, C, and D so that these constructions are diachronically related, i.e. **-id* originally expressed one or more of the functions of these constructions from which others evolved by the PB stage or further back; and (c) PB **-id* had only one function (i.e. either purely syntactic as in Type A, semantic as in Type D, or discourse-related as in Type C) and all synchronically attested construction types are independent, parallel innovations. All these possibilities will be assessed in §5. In §3, I briefly discuss the most elaborate attempts (Voeltz 1977; Trithart 1983) at reconstructing a valence-related applicative suffix at some higher node of Niger-Congo (i.e. Type A) and offer new insights on the semantic role it originally introduced.

3 Reconstructions of PB **-id* and Proto-Niger-Congo **-de* as a syntactic device

Despite the difficulties involved in reconstructing verbal derivational suffixes in the NC phylum (Hyman 2007; 2014), current scholarship argues that there are unmistakable cognate suffixes going back to some earlier NC node which would include minimally Benue-Congo (of which Bantu is a prominent member), Kwa, and Gur-Adamawa (Hyman 2007; 2011; 2014; 2018; Blench (2022 [this volume])). Most scholars (see Hyman 2014 for details, Blench (2022 [this volume])) believe that the synchronically richer systems of verb extensions (e.g. in Bantoid and Central Nigerian within Benue-Congo, Gur, and Atlantic) represent the original situation of the proto-language, if they are not retentions from Proto-Niger-Congo (PNC).

Of all main clause affirmative functions of PB **-id* in §2 and illustrated in Appendix A, the syntactic function of Type A constructions is believed to have been

the original function in PB or some higher node within NC.¹⁰ In general, most scholars agree that **-id* in PB (and possibly further back in NC) had a valence-increasing function in that it added a participant to the argument structure of a verb root. Authors differ, however, in what the original semantic role of the applied phrase might have been. Some (e.g. Trithart 1983) argue that it was a Beneficiary, others that it was more likely a Location or a Spatial Goal (e.g. Endemann 1876; van Eeden 1956; Kähler-Meyer 1966; Schadeberg 2003; Hyman 2007). Voeltz (1977), the first who attempted to reconstruct verbal extensions for what used to be known as Niger-Kordofanian,¹¹ believes that the phrase introduced by the applied suffix **-de* could originally be either Beneficiary or Location/Spatial Goal. Specifically, he argues:

APPLIED [...] is a cover term for a variety of semantic relationships also referred to as DIRECTIVE, BENEFACTIVE, APPLICATIVE, RELATIVE, PREPOSITIONAL, and others. Each of these constitutes a correct label for one of a number of semantic forces [the] applied [morpheme] “adds” to a given verb base [...] The extent to which any or all of these notions were present in Niger-Congo-Kordofanian is unclear from the semantic data available on the individual languages outside of the Bantu domain. We feel safe in conjecturing that the APPLIED had minimally the benefactive *do for someone, on someone’s behalf* reading and the directive reading of *move toward, to*. (Voeltz 1977: 59–60, capitalisation and italics in the original)

Like Voeltz (1977), Trithart (1983) attempts to establish Niger-Kordofanian cognates, but only of Bantu applicative **-id* and with different criteria for the establishment of cognacy than those of Voeltz (1977) (see Pacchiarotti 2020: 264 for discussion). Trithart (1983: 155) claims that the indirective function of **-id* appears throughout Bantu and should be reconstructed for PB; by “indirective” she means animate (usually human) NPs with the semantic roles of Beneficiary, Maleficiary, Recipient, Ethical Dative, and (certain instances of) Possessor. After reviewing all other functions that she finds for **-id*, Trithart (1983: 198–199) concludes: “The earliest function of the applied affix was that of a marker for benefactive NPs.

¹⁰ An anonymous reviewer suggests that the emphasis on the syntactic function of **-id* is probably just a reflection of how the field of linguistics developed since the 1960s, i.e. with a predominance of syntax over semantics.

¹¹ At the time Voeltz was writing, NC and Kordofanian were considered sister branches of a higher node called Niger-Kordofanian. Today Kordofanian or Kordofan is at best seen as a geographic group and its affiliation to NC is doubtful, at least for some of its members (see Hammarström et al. 2018).

Throughout Niger-Kordofanian, up to Proto-Bantu, this is the only function consistently exemplified. In Proto-Bantu, the affix began to spread to a variety of additional semantic relations: indirective, motive, locative, and time.¹² Trithart (1983) follows Heine's (1972/73) Bantu internal genetic classification and model of expansion. Before the breakup of the proto-language, Trithart (1983) posits the following steps for the functional development of **-id*: benefactive > recipient > locative > (adverbs of) time.

Trithart (1983) does not provide specific evidence for this proposed directionality of change, except perhaps semantic plausibility. As we will see below, attested directions of change in other languages do not seem to support this directionality. Through different waves of migration the uses of the applicative suffix broaden and the applicative develops discourse functions (see Type C in §2). The path of changes further develops as follows: benefactive > recipient > locative > (adverbs of) time > (adverbs of) manner > instrumental.

Unlike Trithart (1983), other scholars argue that PB **-id* originally added a Spatial Goal or a Location to the argument structure of its verb root (Endemann 1876; van Eeden 1956; Kähler-Meyer 1966; Schadeberg 2003; Hyman 2007). Schadeberg's view stands out among these in that he argues that original function of the applicative "was to *tie* the nonpatient complement *closer* to the verb. The first of such nonpatient complements may well have been locative ones, from which the other roles of the dative object have evolved" (Schadeberg 2003: 74, my emphasis). An original Spatial Goal function finds support in analyses of applicative functions (mostly Types A, B, and C) in individual Bantu languages, such as Shona S10 (Cann & Mabugu 2007) and Luba-Kasai L31a (De Kind & Bostoen 2012).

4 New insights on the original function of **-id* as a syntactic device

In this section, I offer some degree of evidence in support of Schadeberg's (2003) hypothesis, and – to some extent – also Voeltz's (1977), although the latter does not argue for direction of change. Based on the synchronic behaviour of Type A constructions and attested directions of change, I argue in favour of an original Spatial Goal or Location-oriented function of **-id* and against an original Beneficiary function as proposed by Trithart (1983).

¹²Trithart (1983) does acknowledge, however, that not finding a function other than benefactive listed in the sources for a given branch suggests that functions other than benefactive are absent for a given affix, but obviously there is no guarantee that this is in fact the case.

First, in the literature on grammaticalisation and pathways of semantic change (among others Heine et al. 1993; Heine & Kuteva 2002; Givón 2013), no attested paths of change go from benefactive to allative (i.e. Spatial Goal) or from dative (i.e. Animate Goal) to allative; but many go from allative to benefactive. The extension from allative to dative to benefactive is also a major diachronic trend relevant to language evolution, where concrete words become more abstract (Givón 2015: 174; e.g. go to a place > do something for the benefit of someone). According to Heine et al. (1993: 12), allative markers (case marker or adposition) gave rise to purpose and reason markers in Bodic languages (Tibeto-Kanauri), Rama (Chibchan), and To'aba'ita (Austronesian), and eventually to infinitive markers (e.g. German, English, and Indo-European more generally).¹³ The development allative > dative > benefactive (> causative) is reported by Endresen (1994) for Fula (Atlantic). In addition, Heine & Kuteva (2002) report the following grammaticalisations of allative: allative > dative (including benefactives) (Tamil, Lezgian, several Indo-European languages); allative > purpose (Imonda, Albanian, Lezgian, Basque); and allative > temporal (German, Albanian, Lezgian). The development of allative into a benefactive is also reported by Givón (2013) who proposes that ethical dative markers arose (apparently) independently in several languages (Biblical and Modern Hebrew, Aramaic, and other Semitic languages, Spanish, Polish, and perhaps Akkadian, among others) through a grammaticalisation chain such as allative > dative > benefactive > (reflexive-benefactive) > (ethical dative).¹⁴

Second, two facts stand out when looking at the function of Type A constructions, that is, those that introduce an applied phrase with a semantic role that could not be expressed in the construction with only the verb root. First, there is remarkable language-specific, root-specific variation and idiosyncrasy as to

¹³Heine & Kuteva (2002: 37) add complementiser to the end of the chain of grammaticalisation allative > purpose/reason > infinitive > complementiser – with attested cases in Indo-European (Latin, French) and Maori. Perhaps this path of change could also explain why the applicative in Bantu may appear on verbs in *why*, *how* and other subordinate clauses (see Trithart 1983: 148).

¹⁴Heine & Kuteva (2002: 54) also report instances, however, of benefactive markers developing into dative markers. For instance, in Ewe (Kwa, Volta Congo), the verb 'give' developed into a benefactive marker and further into a dative marker (e.g. *He said it to me*). Further, benefactive markers can also develop into purpose markers (Bulgarian, English, Yaqui, Rapa Nui). Heine & Kuteva (2002: 54) observe that in this case, grammaticalisation appears to be achieved by context expansion, where benefactive adpositions are extended from human to inanimate complements. However, they argue that more diachronic data is needed to substantiate this claim of directionality. Heine et al. (1993) and Heine & Kuteva (2002) do not report any cases where a benefactive marker (case marker or adposition) develops into an allative marker, that is, into a marker for Spatial Goals.

whether a verb root requires the applicative to co-occur with a phrase expressing Spatial Goal and other types of Location-related semantic roles such as Specific Location, General Location, Path, etc. (see Trithart 1983; Rugemalira 1993; Pacchiarotti 2020: 124ff). On a language-specific, root-specific basis, the applicative is very often the only morphological means to introduce Spatial Goals and General Locations. According to Gérard Philippon and Denis Creissels (p.c.), the amount of diversification and accretion of complexity found in the Location-related function across Bantu might suggest that this function is older than the Beneficiary-related function (and the Instrument-related function) and thus has had more time to develop complexity and idiosyncrasy.¹⁵ Second, applicative morphology is more often than not required to introduce a Beneficiary argument, especially in Eastern Bantu.

These two facts, along with the attested directions of change laid out in the preceding paragraphs, suggest that PB **-id* initially introduced or, in Schadeberg's (2003) words, "tied" a Spatial Goal closer to a verb root. This use was (occasionally) extended to Human or Animate Goals (i.e. Beneficiaries, Recipients) already in PB. According to this hypothesis, the obligatory introduction of a Beneficiary into a main clause through applicative morphology only happened as a later innovation. In this sense, PB was probably more like Mbuun in (2) where a Beneficiary can be introduced either by a preposition in the construction of the root or alternatively as a core object argument in the construction with the applicative (likely with some differences in meaning and/or discourse function). Positing that both Spatial and Human Goals were introduced by applicative morphology in the proto-language appears to be more economical than positing hundreds of independent parallel innovations in individual Bantu languages where *es* of **-id* originally introducing a Spatial Goal started to be used for Beneficiaries. Positing an original Spatial Goal function also accounts for the Purpose function of **-id* across Bantu, since Purpose is an abstract extension of a Spatial Goal meaning.

The hypothesis that **-id* originally introduced a Spatial Goal is supported by the meaning of some reconstructed stems in the BLR3 database which look as though an applicative suffix was already present at some node of PB – see Table 1.¹⁶ Note that as suggested by their known distribution, not all reconstructions in Table 1 necessarily go back to PB. Several seem to be later innovations,

¹⁵Strikingly, the obligatoriness of applicative morphology with certain verb roots to introduce either a General or a more Specific Location (e.g. wrap something in a leaf vs. wrap something in the house) is not epiphenomenal to Bantu but also found in Wolof (Atlantic) (Sylvie Voisin, p.c.). This might suggest that the syntactic function of introducing Location semantic roles is older than PB.

¹⁶BLR3 is a lexical database with almost "10,000 form-meaning associations of variable time-depth and reliability" (Bostoen & Bastin 2016: 8), drawing on more than a century of research on Bantu languages.

but some certainly do go back to the most recent common ancestor of all Narrow Bantu languages.

Table 1: Reconstructed applicative stems adding a Spatial Goal

BLR3 index	Reconstructed form and synchronic meanings	Reported distribution in BLR3
514	* <i>cêdid</i> ‘come, go down’	E G K N P
1803	* <i>kímbid</i> ‘run’	E G M N
2273	* <i>nìatid</i> ‘tread’	A B C D
2279	* <i>níngid</i> ‘come, go in’	A B
2817	* <i>têdid</i> ‘slip’	A B F G J M N S
3275	* <i>jédid</i> ‘float’	A F G J M S
3486	* <i>jíngid</i> ‘come, go in’	A B C D E F G J K L M N P R S
3397	* <i>jíbid</i> ‘sink’	J L M N
3504	* <i>jítid</i> ‘pour’	B E G H K L M R S
6243	* <i>jikid</i> ‘come, go down’	J L M
7094	* <i>jógid</i> ‘swim’	H L

BLR3 includes about 190 reconstructions of verbal stems with an applicative suffix (see further discussion in §5.2). The synchronic meanings of the es of some of these forms across Bantu suggest that they originally added a Spatial Goal argument – that is, ‘to, into, towards’ – to their verb roots. Other forms such as BLR1277 **gàbid* ‘give away’ (B E G J M S), BLR 6771 **dòngid* ‘speak’ (N, P) and BLR 6986 **còdid* ‘tell’ (D, J) point rather to an original Animate Goal (e.g. Addressee, Recipient, Beneficiary). The problem is that none of these forms (except perhaps **gàbid*) can be reconstructed to PB given their limited distribution across different Bantu zones following BLR3. According to Schadeberg (1978–79), the applicative developed its syntactic function to the fullest when it started to introduce Beneficiaries at the PB stage, because these complements assumed the syntactic role of the direct object as evidenced by rules of pronominalisation and passivisation (see Wald (2022 [this volume])).

The instrumental function of **-id* was probably an innovation limited to particular branches (see also Trithart 1983). According to Trithart (1983: 179), within individual Bantu languages, the instrumental function of es of **-id* looks newer compared to the benefactive and locative functions. There are two arguments in support of this statement. First, unlike for Beneficiaries and Location-related

semantic roles, the use of the applicative to introduce Instruments is never obligatory; that is, a verb root can always co-occur with an instrumental prepositional phrase to express nearly the same meaning expressed by its applicative counterpart. Second, lexicalised applicative forms, which would reflect an early instrumental function, are almost completely absent in Bantu.

An unanswered question at this point is how **-id* became the only means to introduce other Location-related semantic roles such as General Location, Specific Location, Path, etc. In this respect, Schadeberg (1978–79) argues that the primary function of the applicative suffix in PB “was to relate the action expressed by the verb to a place”. This locative use was expanded to Beneficiaries, Recipients, Time, Cause, and Reason. Schadeberg (1978–79; 2003) does not specify whether ‘locative’ is to be understood as General Location, Spatial Goal, or perhaps both. There are at least three logical possibilities. The first possibility is that originally **-id* was the only means to express General Location with certain verb roots, and then extended its usage to cover Spatial Goals and other semantic roles, as in (7). Note that Animate Goal and Purpose/Reason could actually develop out of Spatial Goal simultaneously.

- (7) General Location > Spatial Goal > Animate Goal > Purpose/Reason > Time

The second possibility is that originally **-id* was the only means to express a Spatial Goal with certain verb roots, and then extended its usage to General Location and other semantic roles, as in (8).

- (8) Spatial Goal > General Location > Animate Goal > Purpose/Reason > Time

The higher likelihood of (7) vs. (8) should be tested against attested directions of change. Another question worth answering is whether other Niger-Congo languages require special morphology to express General Location within a clause. For instance, in some Atlantic languages, such as Seereer (Renaudier 2012), cited in Creissels (Forthcoming), certain roots require an applicative suffix to express Location-related semantic roles such as Source. The existence of obligatory applicative constructions in both Atlantic and Bantu does not necessarily provide evidence in favour of reconstructing them to Proto-Niger-Congo, and by extension PB. It does show, however, that obligatoriness is a recurrent feature in Niger-Congo. This is seldom noted in the literature, but see Creissels et al. (2007: 109).

Whatever the case might be, (7) and (8) assume the existence of a single **-id* suffix, originally used for one semantic role (either Spatial Goal or General Location), then broadening its meaning to include others.

Let us now consider this hypothesis within the broader Niger-Congo verbal suffixation system. Hyman (2007) observes that in what he calls “Central Bantu” (as opposed to “NW Bantu and other Niger-Congo languages”), **-id* introduces a multitude of semantic roles (locative, allative, benefactive, instrumental, etc.), but that in Atlantic languages such as Temne and Fula different functions (different semantic roles) are covered by more than one suffix (e.g. in Temne *-r* is used for allative, locative, and recipient meanings, *-a* for benefactive, circumstance, and manner, and *-a/-nɛ* for instrumental). According to Hyman, there are two possible logical scenarios for the development of the polysemy of Bantu **-id*:

1. PB **-id* introduced only one semantic role and then acquired additional meanings or functions, for example through the semantic pathway I described above: Spatial Goal > Human Goal (Recipient/Beneficiaries) > Purpose, and so on. This scenario presupposes relatively few applicative-like extensions in PNC, perhaps only the one reconstructed for PB.
2. PNC had formally distinct verbal suffixes/extensions. Each of these had one function (e.g. one introduced an Instrument, another a Location, another a Beneficiary, etc.). This scenario presupposes a wider inventory of applicative-like extensions in PNC, similar to what is observed in Atlantic languages where there are formally distinct applicatives for different semantic roles (Beneficiary, Instrument, etc.). Some of these merged into a single extension in PB, which, for example, lacks the itive (i.e. allative)/ventive contrast present in the verbal suffix system of several putative Niger-Congo subgroups such as Atlantic (Hyman 2014: 108) and the Nuba Mountain languages (Hyman 2020: 32–33).¹⁷

Hyman (2007: 158) has a preference for this latter scenario, where “Bantu has merged a richer system of applicative-like extensions, but until Atlantic is understood better, the possibility always remains open that some of the extension properties found in that group are actual innovations”. Following Hyman’s (2007) second scenario, a third possibility is that PB **-id* might have been originally an itive marker which merged phonologically with a functionally distinct verbal suffix whose main function was that of introducing General Location or place narrow focus on such constituents (see discussion in §5.2).

¹⁷Koen Bostoën (p.c.) suggests that PB **-id* ‘applicative’ and **-ɔd* ‘separative’ could possibly be an itive/ventive pair. While **-id* can imply movement towards or into, **-ɔd* implies separation or movement away from. It would be worth investigating how many reconstructed verb forms in BLR3 support this hypothesis. The argument orientation function illustrated in (4) could be seen as a relic of an erstwhile productive deictic/motion affix which specified direction and/or the spatial deixis of the speaker.

5 Which functions of **-id* are reconstructable to PB?

In Table 2, I summarise the synchronic presence of Type A, C and D applicative constructions as discussed in §2. This reflects our current state of knowledge. A blank in Table 2 does not necessarily imply absence of a type. It could rather mean absence of data, which are particularly limited for the discourse-oriented Type C. Probably due to the templatic structure of Bantu grammars, where the applicative suffix is virtually always placed under the rubric of (valence-increasing) verbal derivational suffixes (with examples of an added Beneficiary argument as a default), one finds only sporadic mentions of other functions, especially those related to discourse (see also Creissels 2004). In Table 2, the rows represent the major clades in the Bantu phylogeny of Grollemund et al. (2015). The checkmark indicates that a given construction type is present in at least one language in the corresponding subgroup based on available literature.

Table 2: Synchronic distribution of constructions involving **-id*

	Type A	Type C		Type D
		narrow focus	orientation	
North-Western	✓	✓		✓
Central-Western	✓	✓	✓(?)	✓
West-Western	✓	✓		✓
South-Western	✓	✓		✓
Eastern	✓	✓	✓	✓

Type A, where the applicative is the only grammatical means to introduce an applied phrase with a given semantic role, are very common in Niger-Congo, including Bantu (Creissels 2004; Creissels et al. 2007: 109). Koen Bostoen (p.c.) is of the opinion that most Bantuists, certainly those working in the east, would take Type A as the “standard” type. Nevertheless, grammars and other scholarly works on Bantu languages seldom state whether applicative derivation is the *only* morphological means in a given language to introduce a particular semantic role with certain verb roots. For example, Trithart (1983: 148) reports the use of applicative derivation to introduce Beneficiary, Human Goal, or Spatial Goal in all the languages she surveys (from zone A to S), but does not explicitly state if it is obligatory. See in a similar vein Hyman (2003: 275) for Basaa A43a, Mous (2003: 290) for Nen A44, Grégoire (2003: 365) for the languages of the Forest, i.e. zones B and C, Rekanga (2000: 316) for Himba B302, and Bolekia

Boleká (1991: 123) for Bubi A31. Pacchiarotti (2020: 118–134) shows that Type A constructions are the only means to express minimally some semantic roles in Central-Western, South-Western and Eastern Bantu. I assume that the same is true in North-Western and West-Western Bantu (see Table 2). In fact, the historical debate around the original function of *-*id* in PB (see §3) would seem to assume that Type A constructions date back at least to PB.

For Type C, I ticked North-Western and West-Western based on the fact that Trithart (1983: 148) reports the use of applicative morphology in *why* questions in Nen A44 (North-Western) and some variety of Kongo H10 (West-Western). Additionally, Bostoen & Mundeke (2011: 192) report the use of the applicative in Mbuun B87 (West-Western) in *why* questions and answers.¹⁸ The tick for Central-Western is based on Rapold (1997) who reports the use of the applicative in *wh*-questions (e.g. *where*), but not answers, in Lingala C30B. In doing so, I assume that the occurrence of the applicative in typical focus-related discourse environments such as *wh*-questions is related to the focus function on locative phrases described in other languages. South-Western and Eastern were ticked based on the recent survey in Pacchiarotti (2020: 144–157) where multiple languages within these two subgroups are reported as having applicative morphology expressing several distinct types of narrow focus, always on an applied phrase with a Location-related semantic role. The ticks for Type C “orientation” are based on Pacchiarotti (2020: 141–144). The question mark in parenthesis after the checkmark means that Trithart (1983) reports the orientation function of Type C in Mongo C61, but I was unable to find a mention of this function in Hulstaert’s (1966) grammar of this language.

For Type D, where the applicative morpheme conveys repetitiveness, completeness, thoroughness, excess, persistence, intensity, or intentionality, among other concepts, to the action described by the verb root, I ticked all branches based on the survey in Trithart (1983: 153) and Pacchiarotti (2020: 159–166).

If one applies the same principles used in the reconstruction of phonology, morphology and lexicon to the construction types in Table 2 (see for instance Campbell 2004), one would reconstruct in all likelihood Type A, Type C “narrow focus” and Type D to PB, based on majority rule and economy. Type C “orientation” occurs in few branches, but this might simply be the result of lack of data. The synchronic distribution of construction types involving *-*id* in Table 2 makes it unlikely that they would all be parallel independent innovations across the

¹⁸In the answer to the *why* question, the applicative in Mbuun B87 appears on the main verb and co-occurs with a prepositional phrase expressing Reason which is placed under narrow focus according to the authors.

Bantu domain. In the following subsections I argue, based on attested directions of change within and outside of Africa, that Types A, C, and D are diachronically related and that *-*id* had one or more functions from which others had already evolved at the PB stage or further back. However, the possibility that only some of these functions are related and that they might have been two or more morphemes which ended up looking like *-*id* in PB or further back (see Hyman 2007) can by no means be excluded.

5.1 Diachronic link between Types A and D

The diachronic relatedness of Type A, where the applicative obligatorily introduces an applied phrase for which there is no alternative means of expression, and Type D, where the applicative nuances the lexical meaning of the verb root (by adding iterativity, completeness, excess, etc.), is relatively well attested in the literature. Hyman (2014; 2018) argues that over time valence-related extensions develop aspectual-like functions. Hyman (2018: 191) proposes a three-stage process for this shift, reproduced in (9).

(9) <i>Stage I</i>	<i>Stage II</i>	<i>Stage III</i>
valence \supset aspect	aspect \supset valence	aspect

In Stage I, valence suffixes develop aspectual meanings. In Stage II, the aspectual functions take over the valence-related functions which are pushed to residual, lexicalised areas of the grammar; and they are eventually entirely lost in Stage III. Within Benue-Congo, this direction of change is observable in the following subgroups: Bantoid, where most verb extensions are aspectual-like (i.e. Type D) but were formerly more like PB (i.e. more like Type A or perhaps Type B; see Hyman 2014; 2018); Platoid (Gerhardt 1988; 1989), where several languages show mostly aspectual-like extensions cognate with PB valence-like extensions; and Ring (a subgroup of Grassfields Bantu), where two distinct causative suffixes reconstructable to Proto-Grassfields developed intensive and frequentative meanings (Kießling 2004: 171). Reflexes of PB causative *-*ic* (and not *-*ici* as proposed by Bastin (1986), see Bostoen & Guérois (2022 [this volume])) are also used with an intensifying function in many Bantu languages outside the north-western area (Larry M. Hyman, p.c.). Beyond Africa, applicative morphology with aspectual functions such as perfectivity, iterativity, and intensification is reported minimally in several branches of Indo-European (see Kozhanov 2016 and references therein) and Austronesian (Bowden 2001). Additionally, applicative morphology can develop not only aspectual but also modal functions. Epps

(2010) reports applicative morphology developing into a modal marker in Hup, a Nadahup language of Amazonia.

5.2 Diachronic link between Types A and C

Type C are constructions where applicative morphology is used for several discourse-related functions having to do with locative phrases usually expressing General Location: expansion in the orientation of the locative applied phrase, narrow focus on the locative phrase, and habituality of the action at a certain location. More research is needed to fully understand these discourse functions. It is striking, however, that all of them are available only for locative phrases which most usually have a General Location semantic role, even if this could be an artefact of the few examples used across sources to describe these functions. For convenience, the following discussion is centered around the narrow focus function, which at present is the most described discourse function of es of **-id*.

Creissels (2004) observes that knowing how extensive the use of the applicative is as a focalising device within Bantu is crucial to determining whether this use is an innovation or the relic of a usage already present in the proto-language. He suggests that the latter is more probable under the hypothesis that syntactic structures are the result of the fossilisation of discursive devices. His argument builds on Givón's (1979) grammar ontogenesis, whereby pragmatics develops into syntax; for example, topics evolve into subjects and topicalisation gives rise to passivisation. De Kind & Bostoen (2012) have a different take on this issue and posit that the focus function might have developed out of the applicative's syntactic function of introducing an applied phrase. According to De Kind & Bostoen (2012), the focalising function of the applicative in Bantu can only be accounted for by positing that the applicative originally added a Goal meaning in PB. The fact that Goals are usually spatial/locational would explain the extension of the applicative effect of introducing applied phrases in (usually) immediately postverbal focus position to focalising locative phrases, which usually do not occur in this focus position. Apart from the very general tendency whereby discourse develops into syntax, I have no strong arguments at present to claim that De Kind & Bostoen's (2012) hypothesis is less appealing than that of Creissels (2004).

As we saw in §5, the focalising function of applicative morphology with scope over mostly General Location locative phrases is attested at least in Central-Western, South-Western, and Eastern branches. If we assume that there is an intimate relationship between this usage of the applicative in main clause affirmative contexts and its use in some *wh*-questions (why, where, and how), then

the use of PB **-id* in *wh*-questions in North-Western (see Trithart 1983), West-Western (Trithart 1983; Bostoen & Mundeke 2011), and Central-Western (Rapold 1997) languages can be seen as a relic of the focus function (see Table 2). Given that this focus function is present in some way or another in all branches, it is probably most economical to reconstruct it to PB. The argument that the narrow focus function should be reconstructed to PB is further supported by the fact that applicative morphology in distantly related language families is also associated with focalising functions; cf. e.g. Hernández-Green (2016) for Mesoamerican languages, Rose (2019) for Mojeño Trinitario (Arawak), and Nouguié Voisin (2002) for Wolof (Atlantic). For instance, Mora-Marín (2003) reconstructs both a valence-increasing and a focalising function for **b'e* in Proto-Mayan.

In an unpublished manuscript titled *Applicative* written at Leiden University in the late 1970s, Schadeberg (1978–79) also entertains the possibility that at the PB stage **-id* was already used to express assertive focus on a non-object constituent.¹⁹ He thinks along the lines of Creissels (2004) in considering that the focus function is earlier than the function of tying a non-complement closer to its verb root. The fact that this original focus function was specifically dedicated to non-objects would explain why synchronically all pragmatic functions of Type C applicative constructions have to do with locative phrases. Perhaps the original non-object NPs to which this focus function was applied had a Location-related semantic role. Conceivably, once the applicative developed its syntax-related function, the focus function originally available only with locative phrases was extended to full lexical NPs with other semantic roles (e.g. Beneficiary, Recipient, etc.), which gained the focus-sensitive immediately after the verb position thanks to the applicative.

Schadeberg (1978–79) argues that reconstructions of verb forms seemingly containing **-id* at some node of PB are pivotal in tracing back the history of this suffix. He takes reconstructions where verb stems with **-id* and corresponding verb roots without **-id* have the same meaning as evidence for an original assertive focus function of **-id*. In his words: “The frequency of PB verbs in which the presence versus absence of **-il-* does not appear to mark [in the reconstructed glosses] any clear functional or syntactic difference is interpreted as attesting an earlier role of assigning assertive focus to a non-object [...] I interpret this, in conjunction with the observation that many instances of petrified **-il-* occur in

¹⁹Thilo C. Schadeberg informs me that this unpublished manuscript was finalised after the death of Meeussen in 1978 and before 1980 when Schadeberg started to work in Angola. Meeussen encouraged him to work on the applicative and he read and commented on preliminary drafts of this unpublished manuscript.

verbs of motion which are likely to be used with locative complements, as attesting the chronological priority of **-il-* referring to locatives” (Schadeberg 1978–79: 35).

I reproduce in Table 3 some cases cited by Schadeberg (1978–79) where proto-forms with **-id* and their corresponding roots appear to have very similar or identical meanings. I have updated the forms found in Schadeberg (1978–79) against BLR3 and added the last two entries in Table 3. The question mark next to BLR3 1122 means that BLR3 does not report distribution zones for this entry.

In terms of distribution, no generalisations can be drawn on proto-roots and proto-applicative stems in Table 3. In some cases, the proto-root and the proto-applicative stem have almost identical geographical spreads, and both are largely present in the same zones (**dìnd/*dìndid*, **tù/*tùid*, **pòk/*pòkid*). In others, the proto-applicative stem has a slightly more restricted distribution (**díng/*díngidid*, **dèm/*dèmíd*) but still covers almost the entire Bantu area (especially **gàb* and **gàbid*, where the latter covers zone B down to zone S). In yet other instances, root and applicative stem seems to be in complementary distribution (**tám/*támíid*). There are cases where the proto-applicative stem is more widespread than the root (**jòng/*jòngid*) and vice versa (**jím/*jímíid*).

One of the major problems with BLR3 reconstructions is that the glosses of the forms are synchronic attestations of meanings across Bantu zones and not real etymologies (see Bostoen & Bastin 2016 for a detailed discussion). For example, by looking at synchronic meanings of **cèk* and **cèkid*, one wonders whether these two are a case of synonymy or meaning specialisation (either of the root or of the applicative). As Schadeberg (1978–79) observes, there are multiple instances in Table 3 where the reconstructed root and applicative stem seem have exactly the same meaning, see **dìnd/*dìndid*, **jèp/*jèpíid*, **tám/*támíid*, **pòk/*pòkid*, **dòng/*dòngíid*, and **jòng/*jòngíid*. However, there are also cases where the reconstructed applicative stem is reported in BLR3 as having only one of the meanings attributed to its corresponding root: see **díng/*díngíid*, **dèm/*dèmíd*, **gàb/*gàbid*, **tù/*tùid*, and **jím/*jímíid*. There are at least three possible ways to interpret this second trend: (i) the applicative stem has undergone meaning narrowing or specialisation; (ii) the reported synchronic meanings in BLR3 are not always accurate, and it might turn out that these applicative stems have as many meanings as those attributed to their corresponding roots; and (iii) the roots have undergone meaning expansion or broadening with respect to their applicative stems. Given the present state of knowledge in Bantu etymology, it is essentially a matter of subjective interpretation whether one decides to go with option (i), (ii), or (iii).

Schadeberg (1978–79) takes identity or near-identity in meaning as a relic of an original focus function of **-id* on locative phrases. That is, he considers that

Table 3: Proto-roots and putative proto-applicative stems with similar/identical meanings

BLR3 Index	Root	Distribution	BLR3 Index	Applicative stem	Distribution
1062	* <i>díng</i> ‘turn round, wind round, wrap up’	A B C G H J K L N R	1064	* <i>díngɪdɪd</i> ‘wind round’	E G J M R S
987	* <i>dìnd</i> ‘wait, watch, desire’	B C G H J M N P S	988	* <i>dìndɪd</i> ‘wait, watch, desire’	C D F J M N P S
907	* <i>dèm</i> ‘be heavy, be honoured’	A B C E G H J K L M N R S	908	* <i>dèmɪd</i> ‘be heavy’	G J M N P S
3320	* <i>jèp</i> ‘avoid, get out of the way’	G L M R	3321	* <i>jèpɪd</i> ‘avoid, get out of the way’	C E L
2751	* <i>támb</i> ‘take, receive’	B C H	2752	* <i>támbɪd</i> ‘take, receive’	A N S
1274	* <i>gàb</i> ‘divide, give away’	A B C D E F G H J K L M N R S	1277	* <i>gàbɪd</i> ‘give away’	B E G J M S
2597	* <i>pòk</i> ‘take, receive’	E F G L M N	2598	* <i>pòkɪd</i> ‘take, receive’	E G L M P
1122	* <i>dòng</i> ‘pack carefully’	?	1123	* <i>dòngɪd</i> ‘pack carefully’	H M R S
3096	* <i>tú</i> ‘spit, fix the price’	A C D E F G J M S	3097	* <i>túd</i> ‘spit’	A B C D E J L M S
3361	* <i>jímb</i> ‘sing, dance’	A B C D E F G H J K L M N P R S	3363	* <i>jímbɪd</i> ‘sing’	A C H
3563	* <i>jòng</i> ‘add to’	A M	3564	* <i>jòngɪd</i> ‘add to’	A E F G J M R S
522	* <i>cèk</i> ‘laugh, joke’	B C D E F G J K L M N P R S	523	* <i>cèkɪd</i> ‘be pleased, mock, make fun of’	E G L M N

because the function was discourse-oriented there is no semantic difference in meaning between the two proto-forms. One way to (dis)prove this statement is to look more broadly at all reconstructed verb forms in BLR3 seemingly carrying one (or two) applicative suffixes at some PB stage. However, this step does not provide conclusive evidence in favour of or against Schadeberg's tempting argument. I identified 190 reconstructed verb forms including **-id* in BLR3, including those listed in Table 3. Of these: 96 (51%) have no corresponding reconstructed root; 59 (31%) have a meaning which is identical to the meaning of their root; while the remaining 35 forms (18%) have a different meaning compared to that of the corresponding root. Obviously, BLR3 is work-in-progress and as such these percentages (as well as reported "meanings") would likely change if we had access to additional data. Nevertheless, it is noteworthy that in 30% of total cases, the verb form carrying applicative morphology and its corresponding root have the same meaning.

Unfortunately, an argument against Schadeberg's hypothesis (supported by the 30% above) is that one finds reconstructed verb stems with other PB verbal suffixes which also have (nearly) identical meanings to those of their root and/or applicative counterparts. Thus, alongside **jèp* and **jèpid* 'avoid, get out of the way' in Table 3, there is BLR3 3322 **jèpɔk* (E G H K L S) 'avoid, get out of the way'. Similarly, there is **tám̄b*, **tám̄bid* 'take, receive', but also BLR3 2753 **tám̄bɔd* (D H J K L M R) 'take, receive'. Likewise, alongside **gàb* 'divide, give away, make present' and **gàbid* 'give away', there is BLR3 1275 **gàbɔd* (C H J S) 'divide'.

5.3 Diachronic permutation: Are Types A, C, and D all related?

In §5.1 of this chapter we saw that syntactically oriented verbal affixes (Type A) can develop aspectual meanings (Type D). In §5.2, we saw that syntax-oriented verbal affixes might have a focus function (Type C) and that, if so, several scholars believe that the syntactic function might have developed out of the pragmatic function. Drawing an analogy with logic, if A is related to C and A is related to D, are C and D also related? An evolutionary pathway that would link these three construction types is schematically represented in (10).

- (10) *DISCOURSE* *SYNTAX* *SEMANTICS*
 Type C > Type A > Type D
 Focus function Introduction of AP Aspectual nuances

It goes without saying that given the time depth of **-id* and other PB verbal extensions, the evolution laid out in (10) is extremely over-simplistic. There must have been multiple intermediate steps and cycles of evolution happening over

and over again. Nevertheless, this pathway basically links Givón's (1979) claim that syntax arises from discourse with the claim that syntax-related suffixes can over time develop aspectual functions. At present, apart from the discussion in the preceding paragraphs, I am unable to adduce additional evidence for the evolution in (10). This evolution assumes that there was a single suffix **-ɪd* in PB or before which had one original main function related to discourse and that other functions developed diachronically out of this original one. In this scenario, all three functions would have developed at the PB stage (i.e. node 1 in Grollemund et al. (2015), if not before.

However, as observed in §4, it is entirely possible that there were two or more functionally distinct suffixes that ended up being formally identical in PB due to phonological mergers. The hypothesis, originally suggested to me by Koen Bostoen, is that PB might have had an itive/ventive distinction where **-ɪd* was the ventive, i.e. a *come*-type directional towards a deictic centre, while so-called "separative" **-ʊd* was the itive, i.e. a *go*-type directional out of a deictic centre (see Schadeberg & Bostoen 2019: 186). Given the presence of multiple derivational suffixes expressing directional notions in several Bantoid languages (see Blench (2022 [this volume])), this scenario looks all the more plausible. However, it is hard to find synchronic evidence in support of this hypothesis. The poorly understood function of PB **-ɪd* variously called "implicit contrast" (Trithart 1983), "event localiser" (Kimenyi 1995), "event locative" (Rugemalira 1993; 2004), or argument orientation (Pacchiarotti 2020), see (4) in §2, could perhaps be considered as a fossilised remnant of an erstwhile **-ɪd* suffix which was functionally distinct from applicative **-ɪd* and was perhaps part of a directional system within PB verbal derivation. At the same time, it is not clear how an original ventive morpheme might have developed the function of localising subjects of transitive clauses with respect to the position of the objects in the event described by the verb root. More data on the poorly understood function of argument orientation might provide evidence for this hypothesis.

6 Conclusions

Historical linguistics is an exercise in speculation when there are no written records of the older structures or functions that one is attempting to reconstruct. In the case of **-ɪd* this exercise in speculation is complicated by the myriads of functions associated with its es. As initially observed by Dammann (1961) and Kähler-Meyer (1966), multiplicity of meanings is a typical feature of PB verb extensions (see also Voeltz 1977: 12). However, **-ɪd* stands out among other PB ver-

bal derivational suffixes for the number of functions it can perform synchronically. Although other extensions can add semantic nuances to the meaning of their roots, there is to my knowledge no other PB verbal suffix which has dedicated discourse functions such as those described for **-id*. Another remarkable feature of **-id* is its resistance to renewal. The PB causative suffix **-i* has a long history of renewals, possibly due to its exceptional vocalic shape, but also due to the fact that it often develops aspectual-like functions (see Bostoen & Guérois (2022 [this volume])). Even though the applicative is famous for conveying aspectual meanings (see §2 and §5.2), no applicative morphology has been innovated since PB.

In this chapter I have argued that the traditional view of PB applicative **-id* as a purely valence-increasing syntactic device should be revised against new evidence. Reflexes of **-id* minimally perform three main clause functions which are reconstructable to the PB stage: (i) introducing in a main clause a semantic role which could not otherwise be expressed in that main clause with an underived verb root; (ii) narrow-focusing a constituent which usually has a Location-related semantic role; and (iii) adding semantic nuances such as completeness, iterativity or thoroughness to the meaning of the verb root combining with the applicative. The reconstructability of these functions to the PB stage was supported by their synchronic distribution across the Bantu domain, by attested directions of change within and outside of Africa, and by similar functions of applicative morphology in geographically distant and genetically unrelated language families. As for function (i), I provided evidence *contra* Trithart (1983) and in favour of Schadeberg (2003) that PB applicative **-id* originally added a Spatial Goal into a main clause and extended its usage to Human Goals at some point in PB before Bantu languages started to drift away from the homeland.

I also presented some evidence that functions (i), (ii), and (iii) might be diachronically related and might have developed out of a single **-id* form. At the same time, I entertained the possibility of there being at least two functionally distinct **-id* morphemes at some point in PB which might have gained the same form due to phonological mergers. One of these two **-id* forms was an applicative suffix semantically specified to introduce Spatial Goals/Locations. The other **-id* was a ventive directional and possibly came in a pair with an itive suffix **-vd*, usually called separative in Bantu studies (see Schadeberg & Bostoen 2019). This hypothesis, for which there is currently little to no synchronic evidence, is nevertheless appealing in that it could explain the complexities and idiosyncrasies observed in the uses of **-id* employed in Location-related functions, i.e. they would be fossilised uses of two suffixes with distinct spatial-related functions.

Future research in this domain is hindered by the difficulty in saying anything reliable about the forms and functions of NC verbal suffixes (for a detailed discussion, see Hyman 2007), and by the huge time depth of these eroded morphemes. Nevertheless, promising directions for future research aimed at understanding whether the evolutionary pathway in (10) is possible include: gathering more data on the less described functions of es of *-*rd* in Jarawan Bantu and Bantu zone A (if any); determining whether other Benue-Congo or Niger-Congo languages require dedicated derivational morphology in order for a verb root to combine with a General Location or a Spatial Goal element in the clause; finding etymologies either in Bantu or elsewhere in Niger-Congo for *-*rd* that would support or disprove the pathway in (10); and gaining a deeper understanding of the exact relationship between focus and valence-increasing morphology within and outside of Africa.

Acknowledgements

I am grateful to Thilo C. Schadeberg, Koen Bostoen and two anonymous reviewers for useful comments and feedback on earlier versions of this chapter. The usual disclaimers apply.

Abbreviations

1 (followed by PL)	first person	LOC	locative affix or preposition
APPL	applicative affix		
CAUS	causative affix	PA	pronominal affix
CLX	noun class prefix of class x (where x is a number)	PFV	perfective
		PL	plural
		PROG	progressive
CONN	connective	PRS	present
FUT	future	PST	past
FV	final vowel	S	subject
IMP	imperative	S3:x	third person subject of class x
INS	instrumental preposition		

Appendix A Bantu applicative construction types

		Structural and functional features			
		(i)	(ii)	(iii)	(iv)
Type A (syntax)		yes	no	yes	yes
Root	NP V (NP) (NP) (PP θ_Y)				
cxn:					
Appl	NP V _{APPL} AP θ_X (NP) (NP) (PP θ_Y)				
cxn:					
Type B (syntax, discourse, semantics?)		yes	yes?	yes	no
Root	NP V (NP) (NP) (PP θ_X)				
cxn:					
Appl	NP V _{APPL} AP θ_X + F? (NP) (NP) (PP θ_Y)				
cxn:					
Type C (syntax, discourse, semantics limited to LOCP)		yes	yes	yes	no
Root	NP V (NP) (NP) (LOCP) (PP θ_Y)				
cxn:					
Appl	NP V _{APPL} (NP) (NP) AP(=LOCP) + F α				
cxn:	(PP θ_Y)				
Type D (semantics/aspect)		no	yes	restricted	yes
Root	NP V (NP) (NP) (PP)				
cxn:					
Appl	NP V _{APPL(APPL)(APPL)} + F β (NP) (NP)				
cxn:	(PP)				
Type E pseudo-applicatives (frozen forms)		no	no	no	yes
Root	NP V (NP) (NP) (PP)				
cxn:					
Appl	NP [V _{APPL(APPL)}] LEXICALISED (NP)				
cxn:	(NP) (PP)				

(Adapted from Pacchiarotti 2020: 111)

Column headings are:

(i) introduces an obligatorily present applied phrase; (ii) semantic or pragmatic functions of the applicative construction; (iii) productive; (iv) subject to lexicalisation. A question mark indicates uncertainty/lack of data.

Abbreviations used in this table:

AP = applied phrase

F = function(s) (F α is different from F β)

LOCP = locative phrase

NP = noun phrase

PP = prepositional phrase (for convenience)

θ = theta-role/semantic role (θ_X indicates a semantic role different from θ_Y)

V = verb

APPL = applicative suffix.

References

- Bastin, Yvonne. 1986. Les suffixes causatifs dans les langues bantoues. *Africana Linguistica* 10. 55–145.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Blench, Roger M. 2022. The relevance of Bantoid for the reconstruction of Proto-Bantu verbal extensions. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 235–280. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575823.
- Bolekia Boleká, Justo. 1991. *Curso de lengua bubi* (Colección Ensayos 8). Malabo: Centro Cultural Hispano-Guineano.
- Bostoen, Koen & Yvonne Bastin. 2016. Bantu lexical reconstruction. In *Oxford handbooks online*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199935345.013.36.
- Bostoen, Koen & Rozenn Guérois. 2022. Reconstructing suffixal phrasemes in Bantu verbal derivation. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 343–383. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575829.
- Bostoen, Koen & Léon Mundeke. 2011. The causative/applicative syncretism in Mbuun (Bantu B87, DRC): Semantic split or phonemic merger? *Journal of African Languages and Linguistics* 32(2). 179–218.
- Bowden, John. 2001. *Taba: Description of a South Halmahera language* (Pacific Linguistics 521). Canberra, ACT: Pacific Linguistics, Research School of Pacific & Asian Studies, The Australian National University.
- Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2). 147–185.
- Campbell, Lyle. 2004. *Historical linguistics: An introduction*. 2nd edn. Edinburgh: Edinburgh University Press.
- Cann, Ronnie & Patricia Ruramisai Mabugu. 2007. Constructional polysemy: The applicative construction in chiShona. In Marina Rakova, Gergely Pethö & Csilla Rákosi (eds.), *The cognitive basis of polysemy. New sources of evidence for theories of word meaning* (Metalinguistica 19), 221–45. Bern: Peter Lang.
- Creissels, Denis. 1999. Dictionnaire tswana – français. Unpublished manuscript.
- Creissels, Denis. 2004. Non-canonical applicatives and focalization in Tswana. Paper presented at The First Syntax of World's Languages Conference, Leipzig.

- Creissels, Denis. Forthcoming. *Transitivity, valence and voice*. Oxford: Oxford University Press.
- Creissels, Denis, Gerrit J. Dimmendaal, Zygmunt Frajzyngier & Christa König. 2007. Africa as a morphosyntactic area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa* (Cambridge Approaches to Language Contact), 86–150. Cambridge: Cambridge University Press.
- Dammann, Ernst. 1961. Das Applikativum in den Bantusprachen. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 111(1). 160–169.
- De Kind, Jasper & Koen Bostoen. 2012. The applicative in ciLubà grammar and discourse: A semantic goal analysis. *Southern African Linguistics and Applied Language Studies* 30(1). 101–124.
- du Plessis, Jacobus A. & Marianna Visser. 1992. *Xhosa syntax*. Pretoria: Via Afrika.
- Endemann, Karl. 1876. *Versuch einer Grammatik des Sotho*. Berlin: W. Hertz.
- Endresen, Rolf Theil. 1994. *The etymological relationship between the benefactive and causative suffixes in Fula* (University of Trondheim Working Papers in Linguistics 22). Trondheim: University of Trondheim.
- Epps, Patience. 2010. Linking valence change and modality: Diachronic evidence from Hup (Amazonia). *International Journal of American Linguistics* 76(3). 335–356.
- Gerhardt, Ludwig. 1988. A note on verbal extensions in Jarawan Bantu. *Journal of West African Languages* 18(2). 3–8.
- Gerhardt, Ludwig. 1989. Kainji and Platoid. In John Bendor-Samuel (ed.), *The Niger-Congo languages: A classification and description of Africa's largest language family*, 359–376. Lanham, MD: University Press of America, by arrangement with SIL.
- Givón, Talmy. 1979. *On understanding grammar*. New York, NY: Academic Press.
- Givón, Talmy. 2013. The diachrony of the so-called 'ethical dative'. In Tim Thornes, Erik Andvik, Gwendolyn Hyslop & Joana Jansen (eds.), *Functional-historical approaches to explanation: In honor of Scott DeLancey* (Typological Studies in Language 103), 43–66. Amsterdam: John Benjamins.
- Givón, Talmy. 2015. Diachrony, ontogeny, and evolution. In Talmy Givón (ed.), *The diachrony of grammar*, vol. 2, 695–730. Amsterdam: John Benjamins.
- Grégoire, Claire. 1998. L'expression du lieu dans les langues africaines. *Faits de langues* 11-12. 285–303.
- Grégoire, Claire. 2003. The Bantu languages of the forest. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 349–370. London: Routledge.

- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guthrie, Malcolm. 1967–71. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages (4 volumes)*. Farnborough: Gregg International.
- Hammarström, Harald, Robert Forkel & Martin Haspelmath (eds.). 2018. *Glottolog 3.3*. Jena: Max Planck Institute for the Science of Human History. <http://glottolog.org>.
- Harford, Carolyn. 1993. The applicative in Chishona and lexical mapping theory. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar 1* (CSLI Lecture Notes 38), 93–112. Stanford, CA: Center for the Study of Language & Information.
- Hawkinson, Ann Katherine & Larry M. Hyman. 1974. Hierarchies of natural topic in Shona. *Studies in African Linguistics* 5(2). 147–170.
- Heine, Bernd. 1972/73. Zur genetischen Gliederung der Bantu-Sprachen. *Afrika und Übersee* 56(3). 164–185.
- Heine, Bernd, Tom Güldemann, Christa Kilian-Hatz, Donald A. Lessau, Heinz Roberg, Mathias Schladt & Thomas Stolz. 1993. *Conceptual shift: A lexicon of grammaticalization processes in African languages* (Afrikanistische Arbeitspapiere 34-35). Cologne: Institut für Afrikanistik.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Hernández-Green, Néstor. 2016. Registration versus applicative constructions in acazolco otomí. *International Journal of American Linguistics* 82(3). 353–383.
- Hulstaert, Gustaaf. 1966. *Grammaire du lomóngó. Troisième partie : syntaxe* (Annales – Série in-8° – Sciences humaines 58). Tervuren: Royal Museum for Central Africa.
- Hyman, Larry M. 2003. Basaá (A43). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 257–282. London: Routledge.
- Hyman, Larry M. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville, MA: Cascadilla Press.
- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.

- Hyman, Larry M. 2014. Reconstructing the Niger-Congo verb extension paradigm: What's cognate, copied or renewed? In Martine Robbeets & Walter Bisang (eds.), *Paradigm change: In the Transeurasian languages and beyond* (Studies in Language Companion Series 161), 103–126. Amsterdam: John Benjamins.
- Hyman, Larry M. 2018. Common Bantoid verb extensions. In John R. Watters (ed.), *East Benue-Congo: Nouns, pronouns, and verbs* (Niger-Congo Comparative Studies 1), 173–198. Berlin: Language Science Press. DOI: 10.5281/zenodo.1314327.
- Hyman, Larry M. 2020. A note on Nuba Mountain verb extensions. *Faits de langues* 51(1). 29–36.
- Hyman, Larry M., Alessandro Duranti & Malillo Morolong. 1980. Towards a typology of the direct object in Bantu. In Luc Bouquiaux (ed.), *L'expansion bantoue : actes du colloque international du CNRS, Viviers (France) – 4–16 avril 1977*, vol. 2 (Société d'études linguistiques et anthropologiques de France, Numéro spécial 9), 563–582. Paris: SELAF.
- Jerro, Kyle J. 2016. *The syntax and semantics of applicative morphology in Bantu*. Austin, TX: University of Texas at Austin. (Doctoral dissertation).
- Kähler-Meyer, Emmi. 1966. Die örtliche Funktion der Applikativendung in Bantusprachen. In Johannes Lukas (ed.), *Neue afrikanistische Studien: Festschrift für A. Klingenheben*, 126–136. Hamburg: Deutsches Institut für Afrika Forschung.
- Keenan, Edward L. & Leonard M. Faltz. 1985. *Boolean semantics for natural language*. Dordrecht: D. Reidel.
- Kießling, Roland. 2004. Kausation, Wille und Wiederholung in der verbalen Derivation der westlichen Ring-Sprachen (Weh, Isu). In Raimund Kastenholtz & Anne Storch (eds.), *Sprache und Wissen in Afrika: Beiträge zum 15. Afrikanistentag, Frankfurt am Main und Mainz: 30. September – 2. Oktober 2002*, 159–181. Cologne: Rüdiger Köppe.
- Kimenyi, Alexandre. 1995. Kinyarwanda applicatives revisited. Paper presented at The 8th Niger-Congo Syntax-Semantics Workshop, Boston University.
- Kozhanov, Kirill. 2016. Verbal prefixation and argument structure in Lithuanian. In Axel Holvoet & Nicole Nau (eds.), *Argument realization in Baltic*, 363–402. Amsterdam: John Benjamins.
- Kracht, Marcus. 2002. On the semantics of locatives. *Linguistics and Philosophy* 25(2). 157–232.
- Mabugu, Patricia Ruramisai. 2001. *Polysemy and the applicative verb construction in Chishona*. Edinburg: University of Edinburg. (Doctoral dissertation).

- Mora-Marín, David F. 2003. Historical reconstruction of Mayan applicative and antidative constructions. *International Journal of American Linguistics* 69(2), 186–228.
- Mous, Maarten. 2003. Nen (A44). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 283–306. London: Routledge.
- Nam, Seungho. 1995. *Semantics of locative prepositional phrases in English*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Nouguier Voisin, Sylvie. 2002. *Relations entre fonctions syntaxiques et fonctions sémantique en wolof*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Pacchiarotti, Sara. 2020. *Bantu applicative constructions* (Stanford Monographs in African Languages). Stanford, CA: Center for the Study of Language & Information.
- Peterson, David A. 2007. *Applicative constructions*. Oxford: Oxford University Press.
- Rapold, Christian. 1997. *The applicative construction in Lingala*. Leiden: Leiden University. (MA thesis).
- Rekanga, Jean-Paul. 2000. *Essai de grammaire himba (langue bantoue du Gabon, B36), Tome II : morphologie*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- Renaudier, Marie. 2012. *Dérivation et valence en Seereer*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Rose, Françoise. 2019. From classifiers to applicatives in Mojeño Trinitario: A new source for applicative markers. *Linguistic Typology* 23(3), 435–466.
- Rugemalira, Josephat M. 1993. *Runyambo verb extensions and constraints on predicate structure*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).
- Rugemalira, Josephat M. 2004. Locative arguments in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 285–295. Cologne: Rüdiger Köppe.
- Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 71–89. London: Routledge.
- Schadeberg, Thilo C. 1978–79. *Applicative*. Unpublished manuscript. Leiden: Leiden University.
- Schadeberg, Thilo C. & Koen Bostoen. 2019. Word formation. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 2nd edn. (Routledge Language Family Series), 172–203. Milton Park, Abingdon: Routledge.

- Schaefer, Ronald P. 1985. Motion in Tswana and its characteristic lexicalization. *Studies in African Linguistics* 16(1). 57–87.
- Seidel, Frank. 2008. *A grammar of Yeyi: A Bantu language of southern Africa* (Grammatical Analyses of African Languages 33). Cologne: Rüdiger Köppe.
- Sharman, John Campton. 1963. *Morphology, morphophonology and meaning in the single-word verb-forms in Bemba*. Pretoria: University of South Africa. (Doctoral dissertation).
- Sibanda, Galen. 2016. The Ndebele applicative construction. In Doris L. Payne, Sara Pacchiarotti & Mokaya Bosire (eds.), *Diversity in African languages: Selected papers from the 46th Annual Conference on African Linguistics* (Contemporary African Linguistics 1), 309–333. Berlin: Language Science Press. DOI: 10.17169/langsci.b121.488.
- Stapleton, Walter Henry. 1903. *Comparative handbook of Congo languages: Being a comparative grammar of the eight principal languages spoken along the banks of the Congo river from the west coast of Africa to Stanley Falls, and of Swahili, the “lingua franca” of the country stretching thence to the east coast, with a comparative vocabulary giving 800 selected words from these languages, with their English equivalents, followed by appendices on six other dialects*. Bolobo: Hannah Wade Printing Press of the Baptist Missionary Society.
- Stappers, Leo. 1986. *Boma: Eine Sprachskizze (aus dem Nachlass übersetzt, überarbeitet und herausgegeben)*. Hamburg: Helmut Buske.
- Thwala, Nhlanhla. 2006. Parameters of variation & complement licensing in Bantu. *ZAS Papers in Linguistics* 43. 209–232.
- Trithart, Mary Lee. 1977. Locatives. In Ernest Rugwa Byarushengo, Alessandro Duranti & Larry M. Hyman (eds.), *Haya grammatical structure: Phonology, grammar, discourse* (Southern California Occasional Papers in Linguistics 6), 89–98. Los Angeles, CA: Linguistics Department, University of Southern California.
- Trithart, Mary Lee. 1983. *The applied suffix and transitivity: A historical study in Bantu*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- van Eeden, Bernardus Izak Christiaan. 1956. *Zoeloe-grammatika*. Stellenbosch: Die Universiteitsuitgewers en -boekhandelaars.
- Voeltz, Erhard F. K. 1977. *Proto Niger-Congo verb extensions*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Wald, Benji. 2022. On reconstructing the Proto-Bantu object marking system. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 423–463. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575833.

Chapter 8

Reconstructing suffixal phrasemes in Bantu verbal derivation

Koen Bostoen^a & Rozenn Guérois^b

^aGhent University ^bLLACAN - Langage, Langues et Cultures d'Afrique (CNRS, INALCO, EPHE) and University of KwaZulu-Natal

This chapter introduces the notion of suffixal phrasemes to designate the semantically non-compositional complexes of suffixes which emerged across time and space in Bantu to renew morphology in several verbal derivation categories. It is shown that such verb derivational phrasemes can be reconstructed to different ancestral stages as far back as Proto-Bantu (PB) and possibly beyond. The oldest instance of such a suffixal phraseme in Bantu is the causative **-idi*, which is reconstructed to PB as the phraseologisation of applicative **-id* and the short causative **-i*, in addition to the previously reconstructed simplex PB causative suffixes **-i* and **-ic*. The Bantu ancestral language that emerged after the North-Western Bantu branches had split off created a new causative marker, i.e. **-iki*, through the non-compositional reanalysis of neuter **-ik* and short causative **-i*. Around the same stage, the long passive suffix **-ibv* rose as an aggregation of the middle suffix **-Vb*, well-attested in North-Western Bantu, and the short PB passive suffix **-v*. Much younger but still of considerable time-depth are reciprocal phrasemes produced out of a complex of PB associative/reciprocal **-an* preceded by either causative **-idi* (i.e. **-izyan*) or intensive **-ang/*-ag/*-ak* (most often **-angan*). These causative, passive and reciprocal suffixes are all built on a final element that goes back to at least PB and whose semantics and syntax it copied. Other suffixal phrasemes rather adopted the role of their initial element, while still others developed idiosyncratic functions in which the input of their historical components can only be inferred.

1 Introduction

The propensity of Bantu verbal derivation suffixes to fuse or combine into a new suffix conveying a meaning that is not simply the direct sum of the meanings of



its historical components has been recognised by numerous scholars (see among others Meinhof 1910; Dammann 1954; Stappers 1967; Guthrie 1970; Meeussen 1973; Bastin 1986; Hyman 2007; 2018). In this chapter, we adopt the proposal of Beck & Mel'čuk (2011) to consider such semantically non-compositional suffixal complexes as “morphological phrasemes”, more specifically “suffixal phrasemes”, and (re)assess whether such complexes can be reconstructed to Proto-Bantu (PB).

Meeussen (1967: 92) did not reconstruct complex derivation suffixes to PB. His nine reconstructed verbal derivation suffixes, also known as extensions, are all considered to be simplex: **-i* “causative”, **-id* “applicative”, **-ik* “impositive”, **-ik* “neuter”, **-am* “stative”, **-an* “reciprocal”, **-at* “contactive”, **-ú* “passive”, **-ud* “tr. reversive” and **-uk* “intr. reversive”. Among these suffixes, causative **-i* and passive **-v* (as they are usually noted today) stand out in three regards, i.e. they bear a high tone, they consist of only a vowel segment and they occupy a far-right position in the morphological template of the verb stem. Although Hyman (2022 [this volume]) shows that their exceptional high tone is a later innovation, their V shape still contrasts with the more common VC shape of other PB extensions. Moreover, their morphotactic behaviour is particular in that their templatic position in the verb stem's derivational suffix slot is the one furthest removed from the root. They tend to be stacked after all other derivational suffixes, i.e. just before the final vowel (Hyman 2003c; Good 2005). These two remarkable features, i.e. their vocalic form and their specific position in the verb template, have been taken as possible evidence for them being old Niger-Congo voice suffixes, which were possibly integrated in different later derivational suffixes (see Hyman 2007: 161).

Another special feature that causative **-i* and passive **-v* share is that after Meeussen (1967) they have both been reconstructed as having a phonologically conditioned allomorphy. As for the passive, following Stappers (1967), Schadeberg (2003: 78) posits **-v* occurring after C and **-ibv* after V (repeated in Schadeberg & Bostoen 2019: 186). As for the causative, Meeussen (1967: 92) already posits a possible allomorph **-íc* (**-ic-* ? in his writing), but without specifying any conditioning. Following Bastin (1986: 130) and in line with the conditioning of the passive allomorphy, Schadeberg (2003: 78) reconstructs an original complementary distribution in PB: **-i* after C and **-ici* after V (repeated in Schadeberg & Bostoen 2019: 174). Bastin (1986: 130) furthermore reconstructs a second long (“polyphonic”) causative suffix **-idi*, which she considers to be a later innovation resulting from the “fixing” (“*figement*” in her words) of PB applicative **-id* and PB causative **-i*. Strikingly, **-idi* ends in the same vowel as that of short causative **-i*, just like the other long causative **-ici*, and just like the long passive **-ibv*, which also ends in the same vowel as that of the short passive **-v*.

In this chapter, we analyse these semantically non-compositional complex causative and passive suffixes as “morphological phrasemes”, in line with Beck & Mel’čuk (2011). We also critically reassess their actual time depth with regard to the Bantu family tree. We claim that, contrary to common acceptance, causative **-ɪdi* should be reconstructed to PB, while passive **-ɪbʊ* only emerged at a later ancestral stage. We also argue against the reconstruction of VCV shape for the long causative suffix **-ɪci*. It should be reconstructed as Meeussen (1967: 92) proposed, i.e. **-ɪc* without a final vowel. This latter suffix is not a Bantu-internally created morphological phraseme, but a Niger-Congo retention.

In §2, we introduce the concept of “morphological phraseme” and show that semantically non-compositional sequences of verb derivational suffixes are widespread in Bantu. In §3, we demonstrate that reciprocal suffixes ending in PB **-an* are among the most common morphological phrasemes in Bantu verbal derivation and that they can be reconstructed to ancestral nodes with considerable time depth in the Bantu family, but not to PB (see Dammann 1954; Bostoen et al. 2015; Bostoen Forthcoming; Dom et al. Forthcoming). In §4, we claim that passive **-ɪbʊ* is a morphological phraseme that emerged through the non-compositional reanalysis of a suffixal aggregation consisting of middle **-Vb* and passive **-ʊ*. We furthermore argue that the long passive suffix should be reconstructed as **-ɪbʊ*, with an initial half-close front vowel instead of a close one, and not to PB, but to a later stage. In §5, we analyse causative **-ɪci* and **-ɪdi* along the same lines before reconsidering their distribution within and outside of Bantu. Conclusions follow in §6.

2 Suffixal phrasemes in Bantu verbal derivation

A well-known feature of Bantu languages is that they can stick two or more derivational verb suffixes to the verb root. Reconstructing such combinations of extensions to PB is challenging, as Meeussen (1967: 92) already admitted: “A verbal base can have more than one suffix, but such suffix sequences are difficult to illustrate with reconstructed bases, since these forms are productive and highly unstable”. He does recognise, nonetheless, that the combination of suffixes in Bantu languages is governed by certain principles: “Some characteristics of suffix sequences can, however, be given: -ik-, -am-, (-ad-), -at- would occupy first position; -í- and -ú- have last position (even after pre-final and after C of -ɪde), and -ú- absolute last (even after -í-); a tentative and probably too strict order of possible succession is the following: (ad) at am/ik, ud/uk an id í ú.” Considering extensive comparative data, both Hyman (2003c) and Good (2005) confirm that the ordering of Bantu derivational suffixes indeed does not happen haphazardly, but is

ruled by a historical template. The recurrent templatic suffix order they identify in present-day Bantu languages only partially corresponds to the one proposed by Meeussen (1967: 92), in part because they do not consider all reconstructed extensions. Hyman (2003c: 261–262) proposes a pan-Bantu “CARP” template, actually “CARCP”, i.e. CAUS-APPL-RECP-CAUS-PASS or **-ici-id-an-i-ɔ* (in our notation), in which the long and short PB causative suffixes occupy distinct positions. Hyman (2003c) postulates that this template goes back to PB. Good (2005) provides evidence to reconstruct part of it, i.e. the “CAT” **-ici-id-i* or “causative-applicative-transitive” sequence. He uses “causative” to refer to the so-called “long causative” **-ici* and “transitive” to refer to the so-called “short causative” **-i*. The fact that the ordering of productive Bantu derivational suffixes obeys to such a template does not mean that suffixes are always ordered in that way. The default order can be overruled by other constraints, such as the so-called “Mirror Principle” (MP) (Baker 1985), according to which affix order mirrors the order of syntactic operations. As for the sequencing of verbal derivation suffixes in Bantu, this implies that the suffix furthest removed from the root has syntactic scope over the one closest to the root, as illustrated in (1) for Swahili G42d. While *pigiana* in (1a) is a reciprocalised applicative (lit. [[beat an eyelid to] each other]), *pigania* in (1b) is an applicativised reciprocal (lit. [[beat each other] for that salt]). While (1b) respects both the CARCP template and the MP, MP overrules CARCP in (1b) in that the reciprocal suffix occurs before the applicative.

(1) Swahili G42d

- a. *Yule mtu na Luteni Pinju walipigiana kope.* (Mwenegoha 1975: 87)
yu-le m-tu na L.P. wa-li-pig-i-an-a kope
 PP₁-DIST.DEM 1-person and L.P. SP₂-PST-beat-APPL-RECP-FV 9.eyelid
 ‘That person and Luteni Pinju winked at each other.’
- b. *[W]akiona chumvi hupigania ile chumvi.* (Velten 1901: 69)
wa-ki-on-a chumvi hu-pig-an-i-a i-le
 SP₂-COND-see-FV 9.salt HAB-beat-RECP-APPL-FV PP₉-DIST.DEM
chumvi
 9.salt
 ‘If they see salt, they usually fight with each other for that salt.’

However, the MP can also be overruled by CARCP as shown in (2) with data from Chewa N31b. Both in (2a) and in (2b) the templatic CARCP is followed. In terms of syntactic operations, however, (2a) is an applicativised causative ([[make cry] with sticks]), while (2b) is a causativised applicative ([make [stir with spoon]]). The MP is violated in (2b), because the applicative suffix occurs after the causative.

(2) Chewa N31b (Hyman 2003c: 248)

- a. *a-lenjé a-ku-lil-íts-il-a mw-aná n-dodo*
 2-hunter SP₂-PROG-cry-CAUS-APPL-FV 1-child 9-stick
 ‘The hunters are making the child cry with sticks.’
- b. *a-lenjé a-ku-tákás-íts-il-a m-kázi m-thíko*
 2-hunter SP₂-PROG-stir-CAUS-APPL-FV 1-woman 3-spoon
 ‘The hunters are making the woman stir with a spoon.’

While templatic suffix orders can be both mirroring and non-mirroring, as in (2), non-templatic orders, as in (1b), can only be mirroring. According to Hyman (2003c) and Good (2005), there are no cases in Bantu of non-templatic suffix sequences that are not mirroring. Additionally, every language which allows non-templatic orders also has the templatic equivalent. Given that from a synchronic point of view non-mirroring templatic orders can be accounted for neither syntactically nor semantically, they are best considered as the product of history and, as such, they challenge the assumedly non-arbitrary relation between morphology and syntax/semantics.

Even more challenging for the correlation between verbal derivation morphology and syntax/semantics are those suffix sequences in which the syntactic role and/or the semantic import of each separate suffix are no longer clearly identifiable. Unlike the suffix orders dealt with by Hyman (2003c) and Good (2005), such complex suffixes are semantically and/or syntactically non-compositional. Take for example the suffix *-anil* in Mozambican Ngoni N122 (Kröger 2016). It is a disyllabic extension in which one can clearly identify the reflexes of RECP **-an* and APPL **-id*. Synchronically, however, this extension is one and indivisible and functions as a “pluractional” marker. It signals that the action expressed by the verb is done by many subjects simultaneously or successively, in contrast to *-ang* which rather marks that the action affects several objects, as shown in (3).

(3) Ngoni N122 (Kröger 2016)

- Xi-pexa a-pêt-a kw-a-kem-ang-a aka-ganja-mundu.*
 7-hare SP₁-PASS-FV INF-OP₂-call-PL-FV 2a-friend-his
A-hik-anil-a v-oha.
 SP₂-COME-PL-FV 2-all
 ‘Hare went to call his friends [one by one, like going from door to door].
 They all came [one by one].’

Semantically, *-anil* evokes “plurality of participants” (see Lichtenberk 1985), which Bostoen et al. (2015) propose as the underlying semantic notion accounting

for the semantic shifts that **-an* underwent across Bantu. It also evokes the notions of “intensity”, “iterativity”, “persistence”, “duration”, “continuation”, which reflexes of applicative **-id* may express across Bantu, often in reduplicated or triplicated form depending on the language and the phonotactics of the root with which it combines (see Trithart 1983: 153; Pacchiarotti 2020: 159–166). Nevertheless, *-anil* conveys neither reciprocity, the productive grammatical meaning of the reflexes of **-an* in Ngoni, nor any of the productive uses of **-id*, such as licensing a supplementary object which can be a beneficiary, an instrument or a location (Heidrun Kröger, p.c.). What is more, it is definitely not a combination of the productive meanings of its two components. Given that the suffix ordering in *-anil* is at odds with the CARCP template, its original compositional meaning must have obeyed the MP with the applicative having syntactic scope over the reciprocal, i.e. [[do each other X] for Y]), just like Swahili *pigania* in (1b), which is synchronically still compositional. The present-day *-anil* suffix does not reflect this configuration at all.

Syntactically too, it no longer reflects its historical components, as it is neither valence-decreasing as **-an* tends to be, nor valence-increasing as **-id* often is. Synchronically, *-anil* is valence-neutral.

A suffix like Ngoni N122 *-anil*, which is historically aggregated but synchronically non-compositional, is an instance of what Beck & Mel'čuk (2011) call a “morphological phraseme”. Phrasemes are best known in the domain of multiword expressions, such as clichés, collocations, and idioms, but Beck & Mel'čuk (2011) show that restricted or phraseologised complex expressions not only exist at the level of the phrase. They equally occur at other language levels, especially in morphology. Sequences of bound morphemes may manifest the same features as lexical-syntactic phrasemes, i.e. paradigmatic restrictedness and syntagmatic non-compositionality.

Let us illustrate these two features with the Swahili proverb *Heri kufa macho kuliko kufa moyo* ‘It’s better to go blind than to despair’. This proverb is in itself a conventionalised saying containing two phrasal idioms built on the verb *kufa* ‘to die’, i.e. one with *macho* ‘eyes’ and another with *moyo* ‘heart’. The same verb serves as the matrix of a number of other Swahili idioms, e.g. *kufa masikio* ‘to go deaf’ [lit. ‘to die’ + ‘ears’] and *kufa sauti* ‘to lose voice, be hoarse’ [lit. ‘to die’ + ‘voice’]. All of these sayings are paradigmatically restricted in that *kufa* cannot be replaced by any other verb commonly used to express loss or disappearance, such as *kupotea* ‘to get lost, be lost, disappear’ or *kukata* ‘to cut’. The same holds for the nouns combining with *kufa* ‘to die’. The paradigmatic restrictedness of these sayings is nicely illustrated by comparing them to the idiom *kukata tamaa* ‘to despair, lose hope’ [lit. ‘to cut’ + ‘desire, greed, lust, passion’]. It is a synonym

of *kufa moyo*, but neither *kufa tamaa* nor *kukata moyo* are appropriate sayings in Swahili. All of these phrasal idioms are also syntagmatically non-compositional in that their meaning is not simply the sum of the semantic values of its components. Beck & Mel'čuk (2011) would consider the Swahili sayings with *kufa* as non-compositional phrasemes or idioms, because unlike in collocations, none of the components serves as “semantic pivot” of the complex expression. In a common Swahili collocation like *kufa ajali* ‘to die in/from an accident’, *kufa* is the semantic pivot since the complex expression is about dying and *ajali* ‘accident’ simply determines the cause of death. Similarly, *macho ya kuangaza* ‘bright eyes’ is a collocation in which *macho* ‘eyes’ is the semantic pivot and *ya kuangaza* ‘of shine’ the modifier. In an idiom like *kufa macho* ‘to go blind’, however, neither *kufa* nor *macho* is the semantic pivot, even if their respective semantic contribution is transparent.

In the same way as *kufa macho* is a lexical-syntactic phraseme, the above-cited Ngoni N122 pluractional suffix *-anil* is a morphological phraseme. The sequence of suffixes *-an* and *-il* is paradigmatically restricted in that none of them can be replaced by another suffix to generate the same meaning. It is also syntagmatically non-compositional as none of the historical components serves as the semantic pivot, even if the possible semantic contribution of its two components has not become entirely opaque.

Just like non-compositional lexical phrasemes at the syntactic level, morphological phrasemes can also manifest variable degrees of semantic transparency. A good case in point in comparison with Ngoni *-anil* is Swahili *-ikan*, which is a lexically conditioned allomorph of the so-called “neuter” or “stative” *-ik* (see Ashton 1944: 226–229). Most verb roots select the simplex suffix *-ik*, whose vowel displays harmony with mid root vowels, e.g. *vunj-a* ‘break (tr.)’ > *vunj-ik-a* ‘get/be broken, be breakable’, *som-a* ‘read’ > *som-ek-a* ‘be read(able)’. However, a restricted set of roots only occur with the complex allomorph *-ikan*, e.g. *pat-a* ‘get’ > *pat-ikan-a* ‘be available’, *wez-a* ‘can’ > *wez-ekan-a* ‘be possible, feasible’. Other roots can take both, e.g. *on-a* ‘see’ > *on-ekan-a* ‘be visible’ (but *on-ek-a* ‘appear, be visible, perceptible’ is attested), *changany-a* ‘mix’ > *changany-ik-a* ‘be mixed’ (but *changany-ikan-a* ‘be mixed’ may also be heard). In other words, Swahili stative verbs with *-ik* and *-ikan* do not fall into neat categories allowing either one or the other or both, but manifest a cline with marked preferences at each end (Schadeberg 2004). The fact that certain verb stems may take *-ik* and *-ikan* suggests that the addition of *-an* must have been semantically motivated at some point in time, most likely conveying that the stative event involved multiple participants. Synchronically, however, this semantic motivation has become opaque. Therefore, the neuter suffix *-ikan* is to be considered semantically

non-compositional. Compared to Ngoni pluractional *-anil*, both are combinatorily constrained but manifest variable degrees of semantic non-compositionality. In Swahili, *-ikan* conveys the same neuter meaning as *-ik*. Hence, only the semantic contribution of *-an* has become opaque. In Ngoni, however, the pluractional meaning of *-anil* is reducible to the productive meaning of neither *-an* nor *-il*. Thus, in the case of Swahili, given that *-ikan* conveys the same meaning as the simplex allomorph *-ik*, one could analyse *-ik* as the semantic pivot of the morphological aggregation and thus question whether it is not rather a collocation than an idiom. Beck & Mel'čuk (2011: 192) use the term “derivational affixal collocations” to refer to such “combinations of derivational affixes, one of which is chosen freely based on its meaning and the other of which is added automatically as its collocate”.

Another feature that Ngoni *-anil* and Swahili *-ikan* have in common is that they are quite language-specific. They do not seem to have a very large geographic spread within the Bantu family and can thus be assumed to be of recent origin.¹ However, there are several morphological phrasemes which do have a wide distribution across Bantu and are suitable for reconstruction at some ancestral Bantu stage. Before we consider the reconstruction of the reciprocal, passive and causative suffixal phrasemes, which are at the core of this chapter, we briefly deal with frequentative/iterative/intensive *-agvɔd* (transitive) and *-agvɔk* (intransitive). In some languages, such as Mbukushu K333 in (4), both the transitive and intransitive equivalents are reported; in others, such as Nyamwezi F22 in (5), only one of the two. As the Mbukushu data in (4b) show, the simplex underived root is not always attested in the language, a fact that points towards a certain degree of lexicalisation.

(4) Mbukushu K333 (Wynne 1980; Fisch 1998: 126)

- a. *ghamb-a* ‘speak’ > *ghamb-aghur-a* ‘talk a lot’
yend-a ‘go’ > *yend-aghur-a* ‘walk around (aimlessly)’
nw-a ‘drink’ > *nw-aghur-a* ‘be addicted to alcohol’
- b. *dham-a* > *dham-aghuk-a* ‘roll and swing of ship by waves’
‘sink to bottom’ *tjoth-aghuk-a* ‘be very much ashamed’
túk-aghuk-a ‘make slight cracking noise, as boiling fat’

¹Similar complex derivational suffixes have been observed though in other Bantu languages. For instance, Maganga & Schadeberg (1992: 164) report some lexicalised instances of *-anil* in Nyamwezi F22. However, these do not have the same pluractional meaning as in Ngoni. This suggests that Nyamwezi *-anil* is probably an independent development.

- (5) Nyamwezi F22 (Maganga & Schadeberg 1992: 167: 167)
but-á ‘cut (sth. big)’ > *but-ágvɔl-a* ‘cut into small pieces’
lum-á ‘bite’ > *lum-ágvɔl-a* ‘bite many times’
ol-á ‘drink’ > *ol-ágvɔl-a* ‘draw many lines’

In Mbukushu, the simplex suffix *-ag* is not reported, while both *-ul* and *-uk* are labelled “inversive” (Fisch 1998: 127–129), also known elsewhere in Bantu as “separative” or “reversive” (see Dammann 1959; Schadeberg 1982). In Nyamwezi, *-ag* is an inflectional marker carrying a habitual meaning, among other things, while simplex *-ul* is a transitive “separative” as in Mbukushu (Maganga & Schadeberg 1992: 167). Maganga & Schadeberg (1992: 167) consider the iterative or pluractional meaning of *-agvɔl* as the sum of the meanings of its components, but this seems hard to sustain. It is true that the meaning of the complex *-agvɔd/-agvɔk* suffix in Mbukushu and Nyamwezi is close to the one reconstructed for its first element, see Sebasoni (1967: 134): “*La préfinale du verbe bantou a dû être -ag-, avec le sens de durée, de répétition, de continuité*” [“The pre-final of the Bantu verb must have been *-ag-*, with the meaning of duration, repetition, continuity”].² However, the contribution of the second element has become strictly syntactic, i.e. signalling the difference between transitive and intransitive. Neither *-vɔd* nor *-vɔk* has retained the “reversive” (Dammann 1959) or “separative” (Schadeberg 1982) semantics reconstructed as their original meaning, but only their transitivity and intransitivity respectively. In this regard, *-agvɔd/-agvɔk* differ from Swahili *-ikan* and Ngoni *-anil*, in that in the latter two morphological phrasemes the syntactic impact of the second element is less apparent: *-an* is valence-decreasing just like *-ik*, while the usual valence-increasing role of applicative *-il* is lost. Semantically, however, *-agvɔd/-agvɔk* is non-compositional, just like *-ikan* and *-anil*. Moreover, as is the case for *-ikan*, and to a lesser extent for *-anil*, the meaning of *-agvɔd/-agvɔk* is also closely related to the historical meaning of its first element, while the second element seems to have become semantically opaque.

In terms of geographical distribution, *-agvɔd/-agvɔk* are attested in a cluster of more or less adjacent languages belonging to zones F, J, K, L, and M, and to group S10, as far as we can tell from a preliminary, non-exhaustive survey.

²The *-ag* suffix is both functionally and positionally distinct from Bantu derivational suffixes and therefore called “pre-final” instead of “extension”. Due to this peculiar status it has not been examined with regard to the CARCP template. Sebasoni (1967: 131) considers this “pre-final” to have three distinct forms which are historically related but synchronically largely in complementary distribution: “[...] *-ag- prédomine au nord-est et à l’est du domaine bantou, -ak- au nord, -anga- à l’ouest et au sud*” [“... *-ag-* prevails in the north-east and east of the Bantu domain, *-ak-* in the north, *-anga-* in the west and south”].

These are some of the westernmost Eastern Bantu (EB) languages and easternmost South-Western Bantu (SWB) languages. Although SWB and EB are actually not discrete clades in the phylogeny of Grollemund et al. (2015), the contiguous spread of *-agvɔd/-agvɔk* does crosscut several subclades. Hence, these morphological phrasemes can hardly be posited as an innovation reconstructable to a specific ancestral node in the Bantu family tree. Their geographic pattern rather suggests that they are an areal feature. Morphology is commonly seen as more resistant to borrowing in contact situations than other aspects of language. Nonetheless, morphological copying has been shown to happen, especially between related languages that are typologically similar, in which case its effects are hard to distinguish from both common inheritance and drift or parallel innovation within a language family (see Dimmendaal 1987; Mithun 2013). At the same time, even if morphological copying did underlie the current distribution of *-agvɔd/-agvɔk* within Bantu, more in-depth research would be needed to explain how such a specific morphological innovation could have spread over such large distances.

In any event, what we retain for our current purposes from all that precedes in this section are the following three observations:

1. Morphological phrasemes do exist in Bantu verbal derivation and commonly consist of a sequence of two suffixes that go back to at least PB (as we discuss in the next section, sequences of three such suffixes also occur);
2. They commonly convey a meaning that is identical or closely related to that of the first element in the sequence, while this second element tends to become semantically opaque and has at most a syntactic role if any;
3. Some of these verb derivational phrasemes are language-specific and thus of recent origin, while others have a wider cross-linguistic distribution and must have originated in earlier ancestral times.

These insights are important for our historical analysis of reciprocal, passive and causative suffixal phrasemes that follows. Unlike frequentative *-agvɔd/-agvɔk*, for each of these derivations, non-compositional complex suffixes can be reconstructed to different ancestral nodes in the Bantu family tree. Moreover, unlike for *-anil*, *-ikan* and *-agvɔd/-agvɔk*, reciprocal, passive and causative phrasemes rather adopt the original meaning of their last element than that of their first element.

3 Reciprocal suffixal phrasemes

Reflexes of PB **-an* are known to be extremely polysemous (Dammann 1954; Mugane 1999; Maslova 2007). As surveyed in Bostoen et al. (2015), they convey, across Bantu, meanings as diverse as sociative/collective, reciprocal, natural collective, natural reciprocal, chaining, antipassive, intensive/extensive, iterative, comitative/instrumental, body action middle, cognition middle, spontaneous event middle, potential, etc. Verb stems incorporating *-an* tend to be highly lexicalised and to cover meanings which are associated with the agent-oriented part of the semantic middle domain (Dom et al. 2016), especially – but not exclusively – in languages having a long productive reciprocal marker. Dammann (1954) already noticed that several Bantu languages have at least two reciprocal markers, i.e. the direct reflex of **-an* and a longer suffix in which *-an* is preceded by another element. He also observed that the simplex marker tends to be “frozen” (“*erstarrt*”) in those languages, while the complex one is productively used in new derivations (“*Neubildungen*”). Dammann (1954) furthermore discerned that historically speaking the first element is very often either a causative suffix (commonly a reflex of **-ici* or **-idi*) or an intensive suffix (commonly a reflex of **-ang*, **-ag* or **-ak*), whose original meaning got bleached, given that the productive non-compositional meaning of the complex suffix is simply reciprocal. Each type of complex reciprocal suffix identified by Dammann (1954) is illustrated in (6a) and (7a) respectively. In both Woyo H16dK and Kwezo L13, these complex suffixes are productively used to express reciprocity. As shown in (6b) and (7b), the two languages also still have verb stems with *-an* in their lexicon. These verbs very often refer to natural reciprocal situations, i.e. symmetrical events that inherently involve two or more participants (Dom et al. Forthcoming).

(6) Woyo H16dK (Dom et al. Forthcoming)

- a. *Bôbá ba bacyentó kunizyana betikunizyana mpyanza.*
boba ba ba-cyento kun-izyan-a ba-iti-kun-izyan-a
 old_person CONN₂ NP₂-woman plant-RECP-FV SP₂-HAB-plant-RECP-FV
N-pyanza
 NP₉-cassava
 ‘The old women often plant cassava for each other.’
- b. *kwel-án-a* ‘marry’
mon-án-a ‘meet’
sak-án-a ‘play, have fun’

(7) Kwezo L13 (Forges 1983: 285–286)

- a. *Muwáya nēnzi mugúdàlangăna îfu.*
mu-way-a ne-nzi mu-gu-dal-angan-a i-fu
 SP_{2PL}-leave-FV with-her LOC₁₈-INF-observe-RECP-FV NP₈-habit
 ‘You leave with her to observe each other’s habits.’
- b. *gu-z-ăn-a* ‘to bump into each other’
gú-fw-ăn-a ‘to resemble’
gú-m-ăn-a ‘to discuss, argue with’

While a systematic comparative study of the geographic distribution and various functions of complex reciprocal markers ending in *-an* is pending, we show in this chapter that certain derivational phrasemes involving reciprocals such as *-izyan* in Woyo and *-angan* in Kwezo have a greater time depth than others (cf. e.g. *-anil* in §2) and can be reconstructed to given nodes in the Bantu family tree. As argued in great detail in Dom et al. (Forthcoming), this is certainly the case for Woyo *-izyan*, the most conservative reflex of the reciprocal phraseme **-izyan*, reconstructable to Proto-Kikongo, the most recent common ancestor of the Kikongo Language Cluster (KLC), a discrete sub-branch of the West-Coastal Bantu (WCB) branch (de Schryver et al. 2015; Pacchiarotti et al. 2019). **-izyan* is a non-compositional complex of causative **-idi* (see infra) and reciprocal **-an*. Dom et al. (Forthcoming) argue that **-izyan* rose as a productive reciprocal marker through generalisation of its original compositional meaning ‘reciprocity of causation’, i.e. ‘cause each other to do X’ (satisfying both the CARCP template and the MP), to “reciprocity” more generally. This generalisation was followed by a usage expansion from primarily intransitive verb types to other verb types. The initial causative **-idi* must have already become semantically bleached in Proto-Kikongo as the reflex of **-izyan* is attested as a productive reciprocal marker in all KLC subgroups. Given that little derivational verb morphology has survived in the remainder of WCB, it is hard to say whether **-izyan* possibly goes back to the most recent common ancestor of the entire branch.

However, as discussed in Bostoen (Forthcoming) and summarised in (8), several SWB languages have a very similar reciprocal phraseme.

- (8) Mbundu H21 *-ažan* Salampasu L51 *-asyan*
 Lucazi K13 *-asian* Ruund L53 *-ijaan*
 Luvale K14 *-asan* Kanincin L53 *-azyaan*
 Lwalwa L221 *-asyan* Kwanyama R21 *-afan*
 Songye L23 *-ijeen* Ndonga R22 *-athan*
 Luba-Hemba L34 *-izyen* Herero R30 *-asan*

The question is whether the forms in (8) could go back to the same proto-form **-izyan*. Attributing to them a certain time depth as reciprocal markers is definitely plausible if one reckons that they are no longer productive. Synchronically, most languages in (8) use their inherited reflexive prefix to refer to reciprocal situations, whether or not in combination with the long reciprocal suffix. As argued in Bostoen (Forthcoming), compared to the KLC, the SWB languages have initiated a further cycle of innovation in reciprocal marking. In the KLC, **-izyan* replaced **-an* as a productive reciprocal marker in Proto-Kikongo and the simplex suffix became a highly lexicalised middle marker. In SWB, the complex marker met the same fate as **-an* in the KLC, after the reflexive prefix had elbowed it out as a productive marker of reciprocity which developed reflexive-reciprocal polysemy.

Tracing back the suffixes in (8) to a single proto-form **-izyan* is also a likely hypothesis from a formal point of view, as their shapes vary roughly along the same lines as those in the KLC. The only feature not attested in the KLC is the final front mid vowel observed in Songye and Luba-Hemba. Nevertheless, the mid vowel in Songye and Luba-Hemba could be easily explained as a coalescence of the final vowel of **-idi* and the vowel of **-an*. As for the first vowel of the suffixes in (8), the front vowel of the causative suffix was maintained in a few languages, while the low vowel of **-an* was copied to the first syllable in most other languages. The second front vowel of the causative suffix was retained, as in Lucazi K13 *-asian*, underwent gliding, as in Lwalwa L221 *-asyan*, or was absorbed in the preceding fricative, as in Luvale K14 *-asan*, a common phonological process in Bantu known as “Y-absorption” (Bastin 1986; Hyman 2003b; Bostoen 2008). As for the fricative, it is voiced in a minority of languages, while elsewhere voiceless. Dom et al. (Forthcoming) argue that the voiceless reflexes in the KLC are the outcome of “spirant devoicing”, a phonological process common not only in the KLC (Bostoen & Goes 2019), but also elsewhere in Bantu (Nurse & Hinnebusch 1993; Nurse 1999; Labrousse 2000; Bostoen 2009: 206). That is exactly where the shoe pinches for SWB. Several SWB languages in (8) which have a reciprocal marker with a voiceless fricative, such as Lucazi (*-asian*), Luvale (*-asan*), Kwanyama (*-afan*), Ndonga (*-athan*) and Herero (*-asan*), do not undergo spirant devoicing according to the surveys of Janson (2007: 111–115) and Fehn (2019: 249). For those languages one would need to assume a first phraseme component that started out voiceless, such as causative **-ici* (instead of causative **-idi*). This would imply that not all forms in (8) go back to a putative **-izyan* at the level of Proto-SWB. On the other hand, the fricatives /f/, /th/ (= [θ]) and /s/ of the suffixes in Kwanyama, Ndonga and Herero respectively cannot be reflexes of the **c* in **-ici*. The regular reflex of PB **c* in those languages is /h/ (and /x/ in

Kwanyama) (Fehn 2019: 246). Both Janson (2007) and Fehn (2019) only consider spirantisation within the root. It is well-known that sounds in (grammatical) affixes do not necessarily undergo the same regular changes as those in the root (see for instance Nurse 2008: 112 with regard to Bantu TAM affixes). Therefore, it might well be that all suffixes in (8) do go back to **-izyan*.³ If so, this form could be reconstructed to Proto-SWB and, by extension, to an ancestral node overarching both Proto-SWB and Proto-Kikongo.

Let's take a look at whether possible reflexes of **-izyan* are found elsewhere in major Bantu subgroups. In this respect, it is interesting to observe that Bangubangu D27 attests a suffix *-izeen* which marks reciprocity in conjunction with the reflexive prefix *yi-* (Meeussen 1954a: 28), as shown in (9).⁴ This suffix could easily be a regular reflex of **-izyan*, its final mid vowel resulting from a coalescence of the final vowel of **-idi* and the vowel of **-an*, just like in the SWB languages Songye (*-ijeen*) and Luba-Hemba (*-izyen*) discussed above. The genealogical status of this language spoken in the Maniema region of eastern DRC is not straightforward.⁵

- (9) Bangubangu D27 (Meeussen 1954a: 28)
- | | | | |
|--------------------------|------------------------|--------------------------|------------|
| <i>u-yi-móy-ézáén-a</i> | 'to see one another' | cf. <i>u-mon-á</i> | 'to see' |
| <i>u-yi-húmb-izéén-a</i> | 'to punch one another' | cf. <i>u-humb-án-a</i> ; | 'to punch' |
| | | <i>u-humb-á</i> | |
| <i>u-yi-tág-ézáén-a</i> | 'to call one another' | cf. <i>u-tag-án-a</i> | 'to call' |

There are also Central-Western Bantu (CWB) languages which have a non-compositional suffix of the type "causative + reciprocal". One of them is Mongo

³One could also assume that the potential reflexes of **-izyan* attesting irregular spirant devoicing are instances of morphological copying (cf. supra). However, certainly Kwanyama (*-afan*) and Ndonga (*-athan*) manifest rather language-specific outcomes of spirantisation, i.e. /f/ and /th/ respectively. Also the suffix' retention of the front vowel following the fricative in Lucazi (*-asian*) is unique. These idiosyncrasies make scenario of suffix borrowing less likely. Luvale (*-asan*) and Herero (*-asan*) have a more commonly attested potential reflex of **-izyan*, but no languages in the neighbourhood from which they could have borrowed it.

⁴Bangubangu D27 has a second complex reciprocal marker, which is not productive, i.e. *-agan* (Meeussen 1954a: 28).

⁵Bangubangu D27 is not included in the phylogeny of Grollemund et al. (2015), but several close relatives, such as Lega D25 and Holoholo D28, are. They are considered to be part of Eastern Bantu (EB), as they were in the earlier lexicostatistical study of Bastin et al. (1999) (see also Vansina 1995). However, the support values in the Grollemund et al. (2015), which separate the D20 cluster from the Luba cluster L30, which is considered to be SWB, are quite low. The dividing line between SWB and EB is thus not sharp. As a consequence, the D20 cluster could have well been labelled SWB, just like the L30 cluster could have been considered EB instead of SWB.

C61 in (10). Along with several other complex suffixes ending in *-an*, i.e. *-Van* (< **-ikan*), *-Vngan* (< **-angan*), *-Vtan* (< **-atan*), Hulstaert (1965: 241–243) also identifies *-Vsan*. The first vowel of these complex suffixes is always a copy of the root vowel. All of these phrasemes built on *-an*, which Hulstaert (1965) considers to be “unproductive extensions”, occur on lexicalised derived verb stems. As illustrated with *-Vsan* in (10), their middle meanings cannot be directly derived from the extant underived base verb, if any. The fact that none of these suffixes is still productive and that all of them express lexicalised middle meanings rather than productive reciprocity suggests that their phraseologisation is not of recent origin.

(10) Mongo C61 (Hulstaert 1965: 242)

<i>kák-asan</i> ‘be nervous’	cf. <i>kák</i> ‘extract’
<i>kak-asan</i> ‘invade everything’	cf. <i>kak</i> ‘be violent’
<i>kék-esan</i> ‘be crossed’	cf. <i>kék</i> ‘block’
<i>kek-esan</i> ‘scowl, frown’	
<i>líng-isan</i> ‘hide’	cf. <i>líng</i> ‘wrap, roll up’

Nonetheless, it is rather unlikely that Mongo *-Vsan* is a reflex of **-izyan* (i.e. PB **-Id-i-an*), as the language has a direct reflex of **-Id-i*, i.e. *-ej* (Hulstaert 1965: 255–257, 289), which is in itself unproductive and quite rare. Verbs marked with *-ej* are always transitive and convey a notion of intensity, which is a common functional reassignment of the causative across Niger-Congo (Hyman 2007: 161). As shown in (11), a limited set of them combines with *-an* to convey reciprocity (Hulstaert 1965: 286).

(11) Mongo C61 (Hulstaert 1965: 256–257, 286)

<i>bók</i> ‘throw’	> <i>bók-ej</i> ‘throw in’	> <i>bók-ej-an</i> ‘throw e.o. in’
<i>im</i> ‘murmur’	> <i>im-ej</i> ‘express agreement’	> <i>im-ej-an</i> ‘believe e.o.’
<i>kɔt</i> ‘cut’	> <i>kɔt-ej</i> ‘make scarifications’	> <i>kɔt-ej-an</i> ‘scarify e.o.’
<i>lend</i> ‘watch’	> <i>lend-ej</i> ‘watch with impatience’	> <i>lend-ej-an</i> ‘watch e.o.’
<i>táng</i> ‘name’	> <i>táng-ej</i> ‘promise’	> <i>táng-ej-an</i> ‘promise e.o.’

Formally speaking, the *-ej-an* sequence in Mongo could be a regular reflex of **-izyan*. However, semantically speaking, unlike **-izyan*, it is compositional. Except maybe in the example *lend-ej-an* ‘watch each other’, the meanings of verbs ending in *-ej-an* in (11) convey both the intensive semantics of *-ej* and the

reciprocity of *-an*. So, *-ej-an* is not a suffixal phraseme in Mongo. Nonetheless, the synchronic situation in Mongo is still relevant to the development of the phraseme **-izyan*, as it could reflect the stage immediately preceding the phraseologisation of a sequence of two distinctive suffixes into one non-compositional suffix. The fact that *-ej* is unproductive in Mongo and quite rare makes it the perfect candidate to become the first and semantically void component of a morphological phraseme signalling reciprocity.

In North-Western Bantu, we could not retrieve any reciprocal phrasemes ending in *-an* and having a causative suffix as the semantically empty first component. We did not discover any formally matching but semantically compositional equivalents of **-izyan* either, as we did with *-ej-an* in Mongo. One does find, however, sequences of causative and reciprocal suffixes, which are not entirely compositional and do not express a reciprocal meaning. Their causative suffix looks like a reflex of **-ici*. In Kundu A122, for instance, Ittmann (1971: 297) reports that the combination of causative *-ise* with *-ana* expresses a “causal state”, i.e. a middle situation type as illustrated in (12). The same sequence, also expressing a (causal) state, occurs in Duala A24, as shown in (13).

- (12) Kundu A122 (Ittmann 1971: 297)
kɛlɛ ‘become sick’ > *kélisanɛ* ‘be sick-making’
tángá ‘quarrel’ > *tángisanɛ* ‘be quarrelsome’

- (13) Duala A24 (Ittmann 1939: 147)
bɔbisanɛ ‘be incapable of resistance’
tongwisánɛ ‘be conductive, get along’
bwésànɛ ‘be deadly’
bólisanɛ ‘be curative’

In sum, a reciprocal phraseme **-izyan*, which developed from the sequence of causative **-idi* and reciprocal **-an*, seems to be reconstructable to an ancestral stage from which both the WCB and SWB subgroups emerged. This ancestor could correspond to node 6 in the phylogenetic tree of Grollemund et al. (2015). However, one should then suppose that it got lost in EB, at least as far as we can tell from our admittedly incomplete assessment of its geographic distribution. According to this same survey, **-izyan* is not attested as a reciprocal phraseme in languages descending from any of the branches higher up in the tree, although we do find similar but compositional sequences in CWB. In this branch, we find phrasemes built on the sequence of causative **-ici* and reciprocal **-an*, suggesting that this specific suffix order has also been subject to phraseologisation into **-izyan*. A systematic comparative study of these causative-reciprocal sequences

across Bantu would be beneficial to tease apart reflexes of **-izyan* from those of **-isyan* and to gain a better understanding of their time depth within Bantu.

The same holds for reciprocal phrasemes ending in *-an* and taking as first element the intensive suffixes **-ang*, **-ag* or **-ak*, which Sebasoni (1967: 131) considers to be largely in complementary geographic distribution. Unlike reflexes of **-izyan* and **-isyan*, this kind of reciprocal phrasemes is scattered across EB. In the West Nyanza subgroup of Great Lakes Bantu, for instance, *-angan/-agan* is the productive reciprocal marker in Talinga JE102 (Paluku 1998: 229), Nyoro JE11 (Maddox 1938: 37), Tooro JE12 (Rubongoya 1999: 202), Ganda JE15 (Livinhac et al. 1921: 116; Hyman (2022 [this volume])),⁶ Soga JE16 (Nabirye 2016: 326), Nyambo JE21 (Rugemalira 1993: 148), and Haya JE22 (Kuijpers 1922: 98). It is also found further south in Ndengeleko P11 (Ström 2013: 210–211) and Yao P21 (Mchombo & Ngunga 1994). In Lamba M54, *-akan/-añkan* is an associative marker which “indicates that two or more subjects are associated together in the action of the verb” (Doke 1938: 198).

In SWB, *-angan* is or once was a productive reciprocal marker in several zone L languages (Bostoen Forthcoming), such as Kwezo L13 (Forges 1983: 261, 285), Kete L21 (Kamba Muzenga 1980: 132, 137), Luba-Kasai L31a (Kabuta & Schiffer 2009: 102), Kanyok L32 (Mukash Kalel 1982: 156; Stappers 1986: 14), and Luba-Katanga L33 (Nkiko 1975: 39).

The suffixal phraseme *-angan* is also attested in WCB, especially in the KLC. In Manyanga H16b, for example, Laman (1936: 199) describes how “semireciprocal verbs are formed by adding the suffix *-angana* to the primary stem of the verb” and “express that one of the parts in the action is active while the other is indifferent”, e.g. *fin-angan-a* ‘approach’, *nam-angan-a* ‘follow something, attach oneself to’. In Ntandu H16g, Daeleman (1966: 185) labels the suffix as “alterative”. Its semantics are close to those of its cognate in Manyanga: “The bases with *-angan-* appear to indicate a reciprocal event in which the effective contribution comes from one side, i.e. an event that is directed towards others or elsewhere (and therefore can also be called extensive)”, e.g. *bul-angan-a* ‘bump into someone else, encounter, meet, debouch into’, *fil-angan-a* ‘approach, be near, be right behind’. Remarkably, traces of *-angan* are even found in WCB languages outside of the KLC, where the verbal derivation system has usually become severely eroded. In Tiene B81, for example, Hyman (2010: 31) considers the *-neŋa* extension occurring in some rare relic reciprocal verbs, such as *lé-neŋa* ‘eat with each

⁶Hyman (2022 [this volume]) argues that **-agan* has been phonologically reparsed in Ganda JE15 as *-a-gan*, which can be taken as synchronic evidence for the fact that the historical complex of *-ag* and *-an* suffix became monomorphemic.

other', *nú-neya* 'drink each other', *pé-neya* 'give each other', *té-neya* 'injure each other', as a reflex of **-angan*.⁷

As discussed above, reflexes of **-angan* also occur in CWB languages such as Mongo. According to our current sketchy documentation, **-a(n)gan/*-akan* phrasemes, which express reciprocity or a closely related meaning, are scattered across languages of the CWB, WCB, SWB and EB branches, but have not been observed in NWB. In other words, they could go back as early as node 5 in the phylogenetic tree of Grollemund et al. (2015). A dedicated study would be needed, however, to corroborate this preliminary assessment. Not only the geographic distribution of phrasemes ending in *-an* and having one of the allomorphs of the PB intensive suffix (**-ang*, **-ag*, **-ak*) as first element should be studied more systematically, but also the question of whether all current-day attestations really result from one single phraseologisation at a given ancestral node or should rather be seen as parallel innovations. Further research is also needed on whether the complementary geographic distribution between **-ang*, **-ag* and **-ak* observed for the simplex intensive suffix also persists in the phraseme. This would help discern whether **-angan*, **-agan* and **-akan* are allomorphs of the same underlying morpheme or whether they should be taken as independent morphological phrasemes.

4 Passive suffixal phrasemes

Following Stappers (1967), Schadeberg (2003: 78) reconstructs a phonologically conditioned allomorphy for the passive suffix, i.e. **-v* occurring after C and **-ibv* after V (repeated in Schadeberg & Bostoen 2019: 186).⁸ Hyman (2003c) resumes both allomorphs under “P” in the CARCP template, unlike the causative suffixes which are assigned distinct positions. Neither Stappers (1967) nor Schadeberg (2003: 78) are explicit on the ancestral stage to which this allomorphy should be

⁷“The above four C(V)- roots occur with traces of the reciprocal extension *-ney-* inherited from the PB plural + reciprocal sequence **-a(n)g-an-* found in a number of daughter languages (cf. Haya *-angan-*, Ganda *-agan-*). In the Tiene reflex, the velar + coronal sequence is metathesised to coronal + velar, in conformity with the place restrictions on prosodic stems. Significantly, there are no vestiges of the reciprocal with CVC- or CVCVC- verb bases, precisely because *-ney-* would require a fourth syllable. It is again clear that derived stems are maximally trisyllabic in Tiene.” (Hyman 2010: 31)

⁸If this was indeed the original conditioning, it was not conserved as such in many present-day Bantu languages. In some languages, such as Swahili G42d (Mpiranya 2015: 110–115; Racine 2015: 56–58) and Soga JE16 (Nabirye 2016: 330), the functional distribution between the reflexes of the short and long allomorph is different. In others, the allomorphy has been given up entirely in favour of one form, for instance *-iibw* in Luba-Kasai L31a (Meeussen 1962: 10).

reconstructed, but one could implicitly assume that it is PB. We argue here that it should not be reconstructed to PB, i.e. node 1 in Grollemund et al. (2015), but that it only emerged after NWB had branched off.

Before we elaborate on this new hypothesis, we note that the reconstruction of the short passive suffix $*-v$ to PB is well established (see Meeussen 1967: 92; Stappers 1967; Guthrie 1971: 9; Heine 1972/73: 177; Schadeberg 2003: 78). Ever since Torrend (1891: 272–273), the wide distribution of $*-v$ across Bantu has been acknowledged (see also Werner 1919: 147). Its reflexes are attested in all major branches of Narrow Bantu, including NWB, where it is quite rare. We have retrieved reflexes of $*-v$ in Bubi A31, viz. $-v$ (Bolekia Boleká 1991: 151), Mpongwe B11a, viz. $-o$ (Gautier 1912: 116–119), Orungu B11b, viz. $-o$ (Ambouroué 2007: 205), and in Tsogo B31, viz. $-u$ (Raponda-Walker 1937: 47). In all of these languages the passive is realised as the final vowel of the verb form, unlike in Benga A34 where it is reported as $-w$ in front of the final inflectional vowel (Mackey 1855: 34, 44), as is usually the case in Bantu. Decisive for reconstructing passive $*-v$ to PB is the existence of Niger-Congo cognates outside of Bantu, in Atlantic languages among others, as reflected in the reconstruction of neutro-passive $*-V_{[+back]}$ to Proto-Atlantic by Doneux (1975: 107) (see Hyman 2007: 151). The occurrence of the short passive suffix at the two extremes of the Niger-Congo area led Voeltz (1977: 64) to reconstruct passive $*O$ to Proto-Niger-Congo. PB passive $*-v$ is therefore to be considered as a Niger-Congo retention.

In contrast with $*-v$, passive $*-ibv$ does not have reported cognates outside of Bantu. Nonetheless, long passive suffixes have a wide distribution within Bantu, as evidenced by the first PB passive reconstruction ever, i.e. $*-igwa$ by Meinhof (1906: 76), who reckons that it is often shortened to $-wa$ on the surface. Apart from the same short suffix $-wa$, Werner (1919: 147) also identifies $-igwa$ along with a series of other long forms, i.e. $-iwa$, $-edwa \sim -idwa$, $-ebwa \sim -ibwa$. The consonantal variation observed in long passive suffixes is one of the arguments which led Stappers (1967) to propose $*-i-v$ as reconstruction for the long form and to posit, for the first time, a complementary distribution between short $*-v$ after C and long $*-i-v$ after V. According to Stappers (1967), the appearance of intervocalic consonants would be a later development restricted to EB and SWB languages. He considers intervocalic $/b/$ as the most widespread, i.e. occurring in a contiguous area comprising most of zones L, D and E (including J). Attestations of intervocalic $/d/$ and $/g/$ are relatively rare and scattered across EB. Stappers (1967) retrieves instances of $/g/$ in Gusii JE42, Shambaa G23, Gogo G11, Bena G63, Yao P21, Tonga M64, and possibly also in Pokomo E71 and Nilamba F31, while he reports occurrences of $/d/$ in Mambwe M15, Nyiha M23, Nyanja N31a, Nyungwe N43, Tonga N15, and Ronga (not clear which one). Stappers (1967: 145) conjectures

that the *-idv* type could have applicative **-id* as first element. Stappers (1967) was also the first one to analyse the long allomorph as a morphological phraseme, which has the short form **-v* as its last element. He believes the preceding front vowel to be a reflex of the short causative **-i*. This hypothesis is implausible given that the long passive allomorph never triggers spirantisation, while causative **-i* commonly does across Bantu (Bastin 1986; Hyman 2003b; Bostoen 2008).

In order to understand why Stappers (1967) proposes **-i-v* as basic form for the long passive allomorph, it is important to see that he does not factor in the effect of diachronic sound change. He does not consider the possibility that *-iw*, its most widespread current-day reflex, could go back to a **-iCv* proto-form whose intervocalic consonant went lost. On the contrary, Schadeberg (2003: 78) does consider diachronic phonology and proposes **-ibv* as reconstruction for the long form. Intervocalic **b* is indeed the most plausible reconstruction here, not only because it is the consonant that occurs most often in those present-day languages having a long passive allomorph with intervocalic consonant, but also because intervocalic **b* lenition and loss is quite common in EB; see for instance Guthrie (1967: 71) for the reflexes of **ba* in root-initial position. In front of a back vowel, **b* elides even more easily than before other vowels; see for instance Nurse (1999: 6) who posits the weakening of **b* before “labial vowels” as a shared innovation of the North-East Coast Bantu subgroup. As for the two other stops observed in the long passive suffix of certain EB languages, /g/ could certainly result from a fortition subsequent to the loss of **b*. Yao P21, for instance, which has an *-igw* passive extension, does sometimes have /g/ where **b* went lost, e.g. **bvmb* ‘mold in clay’ > *ku-gumb-a*, **bvɔŋg* ‘be round’ > *ku-gulung-a* (Viana 1961). The emergence of intervocalic *d/l* is more difficult to account for. An epenthetic *l* seems more plausible than positing it as a reflex of applicative **-id*, but this would need more historical-comparative phonological research. In any event, as these long passive suffixes with *d/l* represent a very local development, their status is insignificant in terms of deep-time reconstruction.

Simply put, we do agree with Schadeberg (2003: 78) that reconstructing **b* as the consonant of the long passive allomorph is the most plausible hypothesis, especially if one reckons that simplex middle suffixes ending in /b/ occur in NWB (see also Schadeberg 2003: 78; Bostoen & Nzang-Bie 2010). As discussed below, this middle suffix ending in /b/ is the one we consider to be the historical first component of passive **-ibv*. However, first, we would like to propose a revision to the reconstruction for the initial vowel proposed by Schadeberg (2003: 78) and copied by Schadeberg & Bostoen (2019: 186). Schadeberg (2003: 78) does not reconstruct the long passive form with a near-close front vowel, i.e. [ɪ], as Stappers (1967) does for the forms with an intervocalic consonant, but with a

close front vowel, i.e. [i]. This seems unjustified, as the long passive allomorph never triggers spirantisation,⁹ which would be expected (at least in some languages) if it were a close vowel. Moreover, it often undergoes vowel harmony with root mid vowels (e.g. Swahili *ib-iw-a* ‘be stolen’ vs. *ol-ew-a* ‘be married’), as PB second-degree front vowels often do (e.g. Swahili *pik-i-a* ‘cook for’ APPL vs. *som-e-a* ‘read for’ APPL, *saf-ish-a* ‘(make) clean’ CAUS vs. *wez-esh-a* ‘enable’ CAUS). Based on this evidence, the long passive allomorph should be reconstructed as **-ibʊ* instead of **-ibʊ*.¹⁰

The key question to be answered here is to which ancestral Bantu stage **-ibʊ* should be reconstructed, and by extension the allomorphy with **-ʊ*. As mentioned above, no cognates have been reported outside of Narrow Bantu. As for its distribution within Bantu, Stappers (1967: 141–142) does not report any attestations of the long allomorph in NWB and CWB languages. Our review of available NWB and CWB sources slightly changes this picture. In both subgroups, we could only identify relics of the **-ʊ*, but none of **-ibʊ*, except in one language that Grollemund et al. (2015) classify as part of NWB, i.e. Kota B25, as shown in (14).¹¹

- (14) Reflexes of passive **-ibʊ* in Kota B25 (Piron 1990: 124)

Édíbwèkè.

à-é-dí-ibù-àk-à

SP₁-NEAR_FUT-eat-PASS-IPFV-FV

‘He will be eaten.’

No attestations of **-ibʊ* have been found in Guthrie’s zone A. What several NWB languages of zone A do have, however, as already pointed out by Schadeberg (2003: 78), is a suffix “of the general shape **(a)b(e)* (the vowels differ from

⁹Spirantisation is not to be confused here with the palatalisation of bilabials which the short passive allomorph *-w* triggers in several zone S languages (see Ohala 1978), unlike the long passive allomorph *-iw* which never has this palatalising effect, e.g. Zulu S42 *lob-a* ‘write’ > *lob-w-a* ‘be written’ > *lotsh-w-a* vs. *ab-a* ‘divide’ > *ab-iw-a* ‘be divided’ (van der Spuy 2014).

¹⁰Note that Hyman (2007: 151, 2018: 177) does write **-ibʊ* for the long passive allomorph, i.e. with a near-close front vowel and as a combination of two suffixes, even if he refers to Schadeberg (2003) as his source.

¹¹The genealogical status of Kota and other languages of Guthrie’s B20 group is problematic. As Bastin & Piron (1999: 156–159) point out, not only does B20 split into two separate genealogical subgroups, but the one including Kota also shifts affiliations among WCB, CWB and NWB depending on the lexicostatistical method applied. This is a typical instance of what they call a “floating group” (“*groupe flottant*”). It is likely that language contact played an important role in the genesis of Kota and its closest relatives.

language to language), with a meaning described as passive(-like), neuter or middle voice". As the list in (15) shows, this middle affix is indeed quite widespread in zone A languages.¹² Its degree of productivity varies from language to language, and in most of them, it may also serve as a grammatical marker of passive voice.

(15) Reflexes of middle *-VbV in NWB

Lundu A11	- <i>âb</i>	(Kuperus 1985)
Kpe A22	- <i>av(-ε)</i>	(Hyman 2007: 160)
Wovia A222	- <i>vê</i>	(Richter 2013)
Duala A24	- <i>Vbê</i>	(Bilola 1994)
Noho A32a	- <i>abe</i>	(Schadeberg 1980)
Basaa A43a	̣̣- <i>b-a</i>	(Hyman 2003a) ¹³
Bakoko A43b	- <i>bê</i>	(Kenmogne 2000)
Nen A44	<i>bí-</i> ~ <i>bé-</i>	(Mous 2003)
Maande A46	<i>pí-</i> ~ <i>pé-</i>	(Taylor 1986)
Gunu A622	<i>bá-</i>	(Orwig 1989)
Ewondo-Fang A70	- <i>VbV</i>	(Alexandre 1966; Essono 2000; Van de Velde 2008; Bostoen & Nzang-Bie 2010) ¹⁴

Formally speaking, the middle suffixes in (15) occur in different shapes, i.e. VC, VCV and CV, mostly as a suffix. In this case it is not clear to what extent their final vowel is distinct from the common Bantu inflectional final vowel. In the A44, A46 and A60 languages, the earliest NWB offshoots (Bastin et al. 1999; Bastin & Piron 1999; Grollemund et al. 2015),¹⁵ for reasons unknown, it is a prefix. Regardless of their morphological status, all shapes in (15) have a non-back final vowel and as such they could never be reflexes of *-*rbv*. As for the first vowel, there is quite some variation, but it is striking that most often it is either /a/ or a copy of the root vowel (hence -*Vb*) in Duala A24 and the A70 languages. The same holds true for all CVCVC verb stems ending in **b* in BLR3 (Bastin et al. 2002), as shown in (16).

¹²So far, we could not retrieve any attestations of middle *-*Vb* in the B10-30 languages, which are also commonly seen as genealogically part of NWB (see Bastin et al. 1999; Bastin & Piron 1999; Grollemund et al. 2015), only relics of the short passive *-*v* (cf. supra).

¹³The symbol ̣̣ indicates a height umlaut that occurs with certain suffixes (Hyman 2003a: 274).

¹⁴As discussed in Bostoen & Nzang-Bie (2010), the most recent common ancestor of the Bantu A70 languages developed a productive passive suffix *-*Vban*, which is a suffixal phraseme combining middle -*VbV* and reciprocal -*an* in a semantically non-compositional way.

¹⁵There is general agreement to classify A44 and A46 languages together with A60 languages, mostly because of the close relatedness of their lexicon (Dieu & Renaud 1983; Mous & Breedveld 1986). Together, these languages from Central Cameroon are known as the "Mbam" subgroup and considered to be an important link between Narrow Bantu and Wide Bantu, also known as Bantoid (Bastin & Piron 1999: 155; Bostoen & Grégoire 2007: 76).

(16)	Bantu Lexical Reconstructions with <i>*-a/Vb</i> extension (Bastin et al. 2002)		
	<i>*jódob</i>	‘become soft’	(attested in Guthrie’s zones C J N M P S)
	<i>*pùdɛsb</i>	‘be seized with convulsions’	(L M)
	<i>*cìdɪb</i>	‘shake one’s feet’	(J L)
	<i>*kɪdɪb</i>	‘walk sp.’	(H L)
	<i>*tùtɪb</i>	‘walk with a slight stoop, walk with difficulty’	(J)
	<i>*jíjab/jíjɪb</i>	‘know’	(B H/B C G H J K M N R S)
	<i>*kokob/kakab</i>	‘walk with a slight stoop’	(R S/L M)
	<i>*kádab</i>	‘wash one’s hands’	(J R)
	<i>*kàdab</i>	‘crawl on all fours’	(L)
	<i>*cadab</i>	‘struggle’	(L M)
	<i>*jikab</i>	‘perforate’	(K S)
	<i>*játab/jítáb/jítáb</i>	‘answer call’	(A S/J K L M N/D R S)

The reconstructions in (16) not only share this formal feature, but nearly all also have in common that their meaning belongs to a subcategory of the semantic domain of the middle (see Kemmer 1993), such as body action, emotion, cognition, (change of) state. Only the last two forms in (16) have meanings that do not really fit into that pattern, but it is well-known that verb stems including non-productive derivational suffixes easily develop idiosyncratic meanings and syntactic features that are at odds with those of the once productive suffix (see Bastin 1985; Good 2007; Pacchiarotti 2020: 167–260). Because the reconstructions in (16) have reflexes well outside NWB (including EB as can be seen from the Guthrie zones included), this probably means that some lexicalised middle verb stems ending in **b* are quite old and represent relics of a derivational *-Vb* suffix that once used to be more productive. The fact that this morpheme is still described as a distinct affix in several NWB languages probably indicates that it was longer productive there than elsewhere in Bantu. Outside of NWB, it is rarely identified as a separate extension, although this might merit more systematic investigation. It could well be mentioned as an unproductive suffix in languages whose morphology was described in quite some detail. A comprehensive perusal of big dictionaries might also prove useful in this regard.

In brief, we wish to propose that the long passive suffix **-ɪbʊ* is a suffixal phraseme that developed out of a sequence of the “middle” **-Vb* suffix and the short passive **-ɔ*. The question that needs to be answered to substantiate this claim is how the long passive allomorph ended up with the near-close vowel **ɪ*

and not with either **a* or a copy of the root vowel. Variations like in **jijab* and **jijb* ‘to know’ in (16) and the fact that certain NWB in (15) have *-ab* instead of *-Vb* suggest that the original suffix had **a* and that the copy of the root vowel is a later innovation. If such is the case and the first element of **-ibv* has indeed its origin in this middle suffix, the long passive allomorph can only have emerged at a stage where the change towards **-ab > *-Vb* had already happened. The stabilisation of the near-close front vowel in **-ibv* could then be seen as a further innovation. The productivity of **-Vbv* as a passive allomorph may have induced paradigm levelling, i.e. the suppression of variation at a morpheme boundary in favour of one vowel. Why this uniformisation privileged **i* is hard to say. Is it because it was the vowel most common in roots taking the **-Vb* suffix? Or by analogy with several other derivational suffixes (i.e. applicative, neuter, impositive) starting with *-i*? Was this the result of a harmony process triggered by the short passive suffix **-v*? More in-depth comparative research is needed to answer these questions.

As to the ancestral stage to which the long passive allomorph **-ibv* should be reconstructed, it can definitely be posited at node 6 in the phylogeny of Grollemund et al. (2015), i.e. the most recent common ancestor of WCB, SWB and EB. The presence of **-ibv* in Kota B25 could indicate that the suffix actually goes back as far as node 3. However, as discussed above, the genealogical status of Kota and its closest relatives is tricky. It straddles NWB, CWB and WCB probably due to the fact that contact between languages from these different branches contributed to Kota as we know it today. For the time being, the sole occurrence of **-ibv* in Kota cannot be taken as solid evidence for its reconstruction above node 6 in the tree of Grollemund et al. (2015). More attestations elsewhere in NWB would be needed, for instance in Kota’s close relatives from Guthrie’s B10-30 groups, once these are better described. If **-ibv* were reconstructed back to node 3, one would also need to explain why it is absent from the B10 and B30 languages and also from the CWB languages of zone C, the two branches that split off after node 3 and before node 6. However, it is well-known that passive morphology underwent quite some innovation in zone C (see Meeussen 1954b; Schadeberg 2003). A more in-depth study might therefore be needed to exclude that no remnants of **-ibv* can be identified in CWB and the B10-30 languages. If no new attestations are identified in these languages, **-ibv* could be seen as a shared innovation indicating that WCB, SWB and EB are more closely related among each other than with NWB and CWB, which would corroborate the internal Bantu classification proposed by Grollemund et al. (2015).

5 Causative suffixal phrasemes

Bastin (1986: 130) reconstructs three distinct causative suffixes: **-ici*, **-i* and **-idi* (or in her orthography of the day: **içj*, **j* and **idi*). She considers the first two to be PB, while the last one would be of more recent origin. In this section, we mainly reassess the abundant data and analyses already present in her in-depth historical-comparative study of Bantu causative morphology to draw some different conclusions.

Bastin (1986: 130) considers the reconstruction of **-ici* and **-i* to PB as beyond any doubt, first and foremost due to their general distribution within Bantu. In the case of **-ici*, Bantu-internal evidence is corroborated by comparative Niger-Congo data. Bastin (1986: 101) links PB **-ici* with Proto-Niger-Congo **ti* and **ci* as proposed by Voeltz (1977: 60–63). These two Niger-Congo suffixes would have merged in Proto-Benue-Congo and resulted in a single reflex **-ici* in PB (see Voeltz 1977: 61; Bastin 1986: 92). More systematic comparative research within Niger-Congo would be needed to either substantiate or discard Voeltz' merger hypothesis, but it is crystal clear that Bantu causative *-is* suffixes, as the reflexes of **-ici* most commonly look like, have cognates across Niger-Congo, as far as Atlantic and Gur (see Hyman 2007). Unlike PB **-ici*, Bastin (1986: 101) considers the PB short causative **-i* to be a Bantu-specific innovation.

In the light of the preceding sections, especially the one on the passive, considering the PB short causative suffix as more recent than the PB long causative suffix sounds counterintuitive, especially since PB **-ici* seems to end in PB **-i*, much like passive **-ibv* ends in PB **-v*. This alleged innovation is also at odds with the conjecture of Hyman (2007: 161) that PB “causative **-i* and passive **-v* are old voice suffixes”. We therefore believe that two assumptions of Bastin (1986) might need revision: (1) that causative **-i* is not attested beyond Bantu; (2) that the long causative suffix **-ici* really ends in a vowel.

As for the occurrence of causative **-i* elsewhere in Niger-Congo, identifying cognates of a vocalic suffix is obviously not an easy job. It is always hard to tell whether similar vowel-only suffixes in other branches of Niger-Congo do not result from the loss of a consonant. Nonetheless, Atlantic languages such as Bijogo (Segerer 2002) and Kisi (Childs 1995), for example, do have a causative suffix *-i* (see Hyman 2007: 154), which could well be a cognate of PB **-i*. In other words, both the short and long PB causative suffixes seem to go a long way in Niger-Congo.

Concerning the VCV shape of the PB long causative suffix, it is important to realise that Bastin (1986: 66) starts out from the question whether the long causative

suffix, which commonly has a voiceless fricative consonant in Bantu, should be reconstructed as **-ɪc*, **-ic*, **-ɪci*, **-ici*, or still as **-ɪki*. She does consider the possibility of a PB long causative suffix **-ic* without final vowel, as actually proposed by Meeussen (1967: 92). Her consideration of reconstructions with final vowel was prompted by earlier proposals that all or part of the present-day causative suffixes with a voiceless fricative (mainly /s/ or /f/) should be seen as the reflexes of a causative phraseme **-ɪki* (see Meinhof 1910: 43), consisting of impositive **-ɪk* and short causative **-i* (Guthrie 1970: 219). Bastin (1986: 100) herself admits that in very few present-day languages the reflex of **-ici* displays a final vowel, neither on the surface nor underlyingly. She also recognises that in numerous languages the reflex of **c* in front of **i* is not different than before any other vowel. Furthermore, she acknowledges that it is impossible in many languages to tell apart the reflexes of **ki* and **ci* (and even **ci* of less relevance here). Finally, and most importantly, she concedes that there are languages where the voiceless fricative cannot be a reflex of **k* followed by **i*, while there are others where it can only be a reflex of **k* followed by **i* (Swahili *-ish* for example), and not of **c(i)* (Bastin 1986: 92–100). In other words, Bastin (1986) provides all evidence to argue against a unified account of all long Bantu causative suffixes having a voiceless fricative, but she still comes up with a single PB **-ici* reconstruction.

Critically reassessing her evidence, we deem it necessary to distinguish between two distinct causative suffixes that gave rise to present-day reflexes with a voiceless fricative or affricate: (1) **-ic* as proposed by Meeussen (1967: 92), which goes back to PB, and (2) **-ɪki* of later origin. The fact that certain current-day languages have two distinct causative suffixes ending in a voiceless fricative/affricate is strong evidence in favour of this hypothesis. Cuwabo P34 is one such language. Its reflex of PB causative **-ic* is *-iʔ*. Its causative *-ec*, realised in free variation as either [ec] or [etʃ], is a regular reflex of **-ɪki* and regularly corresponds to Swahili *-ish*. Similarly, Cuwabo causative *-uc/-oc* is reflex of **-ɪki* (and corresponds to Swahili *-ush*, as in *anguka* ‘fall’ > *angusha* ‘make fall’) (see Guérois & Bostoen 2016). While Cuwabo causative *-uc/-oc* unmistakably results from the unification of separative **-ɪk* and causative **-i*, more research is needed to determine whether causative **-ɪki* results from the phraseologisation of neuter **-ɪk* and causative **-i*, or rather from impositive **-ɪk* and causative **-i* as proposed by Guthrie (1970: 219). Determining the time depth of the causative phraseme **-ɪki* is greatly complicated by the fact that its reflexes are so difficult to distinguish from those of **-ic* and would thus require a new dedicated study.

Once one recognises the need to posit at some ancestral stage the emergence of a causative phraseme **-iki*,¹⁶ then the PB causative suffix **-ic* can perfectly be reconstructed without final vowel, all the more because /s/ or /ʃ/ are the commonest reflexes of **c* across Bantu anyway, also in the absence of a following close front vowel (see Guthrie 1967: 76). This also perfectly ties in with the Bantu-external evidence. Causative suffixes having /s/ or /ʃ/ are widespread throughout Niger-Congo (see Voeltz 1977; Hyman 2007), and beyond (Hyman 2014). Hence, simply reconstructing a VC shape ending in PB **c* seems to do the job. Considering both Bantu-internal and Bantu-external evidence, reconstructing **-ic* to PB, as proposed by Meeussen (1967: 92), is thus more plausible than **-ici* as advanced by Bastin (1986).¹⁷

We are then left with the third widespread Bantu causative suffix, i.e. **-idi*, which Bastin (1986: 130) analyses as a historical aggregation of PB applicative **-Id* and PB causative **-i*, an idea put forth already by Meinhof (1910: 43). Due to spirantisation commonly triggered by causative **-i*, the **d* of **-idi* typically has a voiced fricative reflex, unlike the fricative reflex of **k* in **-iki*, e.g. Swahili G42d *-iz* as in *fany-iz-a* ‘make do’ vs. *-ish* as in *anz-ish-a* ‘make start’ (see Miehe 1989), or **-ic*, e.g. Cuwabo P34 *-eð* as in *weénjêð-a* ‘add, increase (tr.)’ vs. *-i?* as in *téy-i?-a* ‘make laugh’ (Guérois & Bostoen 2016). In contrast to the two other Bantu causative suffixes, i.e. **-i* and **-ic*, Bastin (1986: 130) questions the PB status of **-idi*. Although she acknowledges its wide distribution, she believes it to be of more recent origin and sees its emergence as potentially correlated with the regression of **-i* as a productive causative suffix. She furthermore allows the possibility that the unification of **-Id* and **-i* into causative **-idi* recurrently took place as a parallel innovation.

It seems unlikely that the morphological phraseme **-idi* was innovated multiple times and would thus be a relatively recent creation. The two main reasons to think so are (1) its general distribution in the Bantu domain and (2) its highly lexicalised status. With regard to its spread across Bantu, Bastin (1986: 101–105)

¹⁶Positing **-iki* also accounts for the lengthening of the final inflectional vowel *-a*, which is observed after the long causative suffix in certain Great Lakes Bantu languages, e.g. Shi JD53 *àasunisaa* ‘he made grow’ (Bastin 1986: 100, see also Trithart 1977: 78–79 for the same phenomenon in Haya JE22). The final *i* of the causative is fully assimilated to the final vowel but with conservation of its quantity, which results in a long *aa*.

¹⁷As for the first vowel of **-ic(i)*, Bastin (1986: 73–91) concludes after a systematic review of the comparative Bantu-internal evidence that both the close and half-close front vowel could be reconstructed as the original one. She eventually opts for the first-degree **i*, because Voeltz (1977: 60–63) proposed the same for Proto-Niger-Congo. To put it differently, to possibly revise the first vowel of the PB long causative suffix, one would need to reassess comparative data from elsewhere in Niger-Congo, which goes beyond the scope of this chapter.

herself identifies instances of **-idi*, which commonly have a spirantised reflex of **d* (either a voiced fricative or affricate), in all major Bantu branches except NWB. However, she also reports a causative suffix with the shape *-Vl(V)* in several NWB languages of zone A, i.e. Kpe A22, Su A23, Duala A24, Benga A34, Ewondo A72a and Bulu A74a, which she considers to be a possible reflex of **-idi* (Bastin 1986: 127–129). Our systematic survey of available sources for NWB languages led us to identify several other reflexes, listed in (17). Most of them do have a fricative or affricate consonant. Reflexes of the causative suffix **-idi* are thus also well attested in NWB.

- (17) Reflexes of causative **-idi* in NWB
- | | | |
|--------------|------------------|----------------------|
| Bafo A141 | <i>-dʒi</i> | (Apuge & Neba 2011) |
| Bakoko A43b | <i>-jè</i> | (Kenmogne 2000) |
| Kpa A53 | <i>-zì</i> | (Guarisma 2000) |
| Tuki A601 | <i>-ij</i> | (Kongne Welaze 2004) |
| Kol A832 | <i>-àzà</i> | (Henson 2007) |
| Kako A93 | <i>-idy</i> | (Ernst 1998) |
| Mpongwe B11a | <i>-iz ~ -ez</i> | (Gautier 1912) |

Another argument against the recent origin of **-idi* is the observation that it rarely acts as a productive causative suffix. In most languages, it is attested with a variable number of lexicalised verbs but cannot be used productively to derive causative verbs. As Bastin (1986: 119–120) nicely summarises, this is especially so in WCB, SWB and EB languages, where the reflex of **-ic* or **-iki* is often the most productive causative suffix.¹⁸ The fact that **-idi* manifests such high degree of lexicalisation in the latest offshoots of the Bantu family runs against the hypothesis that it is a late and parallel innovation. Even more decisive in this regard is the fact that **-idi* itself has become one of the constituents of a new phraseme, i.e. reciprocal **-izyan* (see §3), which could be reconstructed as far as node 6 in the phylogenetic tree of Grollemund et al. (2015). To be involved in the creation of a new suffixal phraseme at such a deep ancestral stage, **-idi* must have become non-compositional well before.

After having carefully reconsidered the available evidence, it seems safe to postulate that **-idi* is a third causative that can be reconstructed to PB, i.e. node 1 in Grollemund et al. (2015). While **-i* and **-ic* were inherited from older Niger-Congo ancestral stages, **-idi* seems to be a PB innovation that emerged through the phraseologisation of applicative **-id* and causative **-i*.¹⁹

¹⁸Many CWB languages of zone C only have the reflex of **-idi* as a long causative suffix.

¹⁹Bastin (1986) did not consider the possible distribution of **-idi* beyond Narrow Bantu and, as

In sum, causative morphology turns out to be the most diverse and innovative within Bantu. This is definitely so if one reckons that we have not considered here causative(-like) suffixes, such as impositive **-ɪk* and transitive separative **-ɔd*, and the morphological phrasemes in which these and other suffixes are involved, e.g. **-ɪki*, **-ɔki* and **-ɔdi*. These merit a systematic and comprehensive study. Unlike other verbal derivational categories, PB not only retained two distinct Niger-Congo causative suffixes, i.e. **-i* and **-ic*, but also created a new causative phraseme **-ɪdi*. As we discussed in §2–4, the creation of such phrasemes for the passive and the reciprocal only happened at later ancestral stages within Bantu language history. Similarly, causative morphology, phrasemic innovation for causative morphology happened after PB, as can be seen in the reflexes of **-ɪki* in languages such as Swahili and Cwabo.

6 Conclusions

In this chapter, we have shown that the creation of suffixal phrasemes is a common strategy to innovate Bantu verbal derivation morphology. We have identified semantically non-compositional aggregations of existing suffixes in derivational categories as diverse as the pluractional, neuter, intensive, reciprocal, passive and causative. Some of these phrasemes adopt the semantics and syntax of one of their constituents, either the first or the last element, while others develop idiosyncratic functions in which the original contribution of their historical components can at best be surmised. A more comprehensive typology of morphological phrasemes in Bantu derivational morphology would be most welcome. Interestingly, just like certain verbal derivational categories innovate their morphology by stacking a new suffix to their inherited suffix, interrogatives in Bantu (and elsewhere in the world) also manifest a very strong tendency for continuity in their evolution. As Idiatov (2022 [this volume]) shows, a new interrogative is almost always based on another pre-existing one.

We furthermore demonstrated that verb derivational phrasemes can be reconstructed to different ancestral stages in Bantu history, up to PB. Innovation through the coinage of suffixal phrasemes is most advanced in causative morphology, as the oldest phraseme we reconstruct is PB **-ɪdi*, which emerged out of the concatenation of applicative **-ɪd* and short causative **-i*. Hence, PB did not only have causative **-i* and **-ic* (and not **-ici* as proposed by Bastin 1986, though maybe **-ɪc* instead of **-ic*, see footnote 17), inherited from ancestral Niger-Congo

far as we can judge, possible Niger-Congo cognates of PB **-ɪdi* have also not been reported elsewhere. Admittedly, we did not carry out a systematic perusal of the relevant literature.

stages, but also **-idi*. Two of the reasons why causative morphology started to renew so early are probably the exceptional vocalic shape of **-i* and the specific morphophonological processes it triggers, as well as the fact that it already had a highly lexicalised status in PB. This is to be expected given that it is a Niger-Congo inheritance. The functional distribution of the three PB causative suffixes along the lines of categories such as direct and indirect causation and intensity merits further study. Innovation in causative morphology did not stop in PB as younger causative phrasemes such as **-iki* also occur across Bantu, especially outside of NWB. As suggested by one of the reviewers of this chapter, the fact that causative suffixes are often functionally reassigned to the excessive/intensive marking may also have contributed to their frequent innovation in form.

That more innovation happened at ancestral nodes posterior to the split-off of NWB is clear from the passive and reciprocal phrasemes we propose in this chapter. First of all, we argued that the long passive suffix should be reconstructed with an initial near-close front vowel, i.e. **-ibv* instead of **-ibʊ*, and that it does not go back to PB. The phraseologisation of the middle suffix **-Vb*, well-attested in NWB and possibly going back as far as PB (node 1), and the PB passive suffix **-v* did not happen before node 3 and probably not even before node 6 in the phylogeny of Grollemund et al. (2015), i.e. the most recent common ancestor of WCB, SWB and EB. This morphological phraseme could be a shared morphological innovation suggesting that these subgroups are indeed more closely related to each other than to the rest. The exceptional short vocalic shape of the PB passive suffix **-v* was a good structural motivation to innovate passive morphology.

The reciprocal phrasemes ending in **-an* and having either causative **-idi* (i.e. **-izyan*) or intensive **-ang/*-ag/*-ak* (most often **-angan*) as a first element also have a relatively deep ancestry. Although more dedicated studies are required to better define their exact time depth, we claim that they could have emerged at nodes 5 or 6 in the phylogeny of Grollemund et al. (2015). Just like passive **-ibv*, these reciprocal phrasemes could thus also be diagnostic for Bantu internal classification. Unlike with the causative and passive suffixes, the main motivation for innovation in reciprocal morphology was not the shape of PB **-an*, but the fact that it tends to become lexicalised and undergo semantic shift within the middle domain.

To conclude, we would like to point out that phraseologisation in verbal derivation morphology probably already happened well before PB. As Hyman (2018: 193) suggests, morphological phrasemes also occur in Bantoid languages outside of Narrow Bantu, where CVC-shaped extensions in languages such as Noni and Lamnso? probably result from the fusion of two suffixes. Hyman (2007: 161) also

identifies fusion via prosodic restriction and phonological erosion as a common process of derivational suffix innovation in Niger-Congo. Some PB derivational suffixes, which tend to be seen as simplex, could therefore also be morphological phrasemes in origin. A diachronic reassessment of the separative pair **-ʊk/ *-ʊd* from a wider Benue-Congo/Niger-Congo perspective might be beneficial in this regard. The formal and functional commonalities of neuter **-ɪk* and intransitive separative **-ʊk* on the one hand, and applicative **-ɪd* and transitive separative **-ʊd* on the other, suggest a historical link and the possibility that **-ɪk* and **-ɪd* might have been a diachronic component of **-ʊk* and **-ʊd*, respectively, or the other way around. If some of them are indeed morphological phrasemes, their creation must have happened at the stage of PB or before, as all of them go back to at least the most recent common ancestor of all (Narrow) Bantu languages.

Acknowledgements

We wish to thank Sara Pacchiarotti, Larry M. Hyman, Sebastian Dom, and two anonymous reviewers for their very relevant and helpful feedback. The usual disclaimers apply. Research of both authors for this chapter was funded by the Special Research Fund (BOF) of Ghent University.

Abbreviations

APPL	applicative
C	consonant
CARCP	causative-applicative-reciprocal-causative-passive
CARP	causative-applicative-reciprocal-passive
CAT	causative-applicative-transitive
CAUS	causative
COND	conditional
CONN	connective
CWB	Central-Western Bantu
DEM	demonstrative
DIST	distal
EB	Eastern Bantu
FUT	future
FV	final vowel
HAB	habitual
INF	infinitive

IPFV	imperfective
LOC _x	locative prefix of class x
N	nasal
NWB	North-Western Bantu
OP _x	object prefix of class/person x
PASS	passive
PB	Proto-Bantu
PL	pluractional
PP _x	pronominal prefix of class x
PROG	progressive
PST	past
RECP	reciprocal
SP _x	subject prefix of class/person x
SWB	South-Western Bantu
V	vowel
WCB	West-Coastal Bantu (aka West-Western Bantu)

References

- Alexandre, Pierre. 1966. *Système verbal et prédicatif du bulu* (Langues et littératures de l'Afrique noire 1). Paris: Librairie C. Klincksieck.
- Ambouroué, Odette. 2007. *Éléments de description de l'orungu : langue bantu du Gabon (B11b)*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- Apuge, Michael E. & Ayu'nwi N. Neba. 2011. The verb morphology of Bafaw. In Emmanuel Chia, Vincent A. Tanda & Ayu'nwi N. Neba (eds.), *The Bafaw language*, 65–123. Oxford: African Books Collective.
- Ashton, Ethel O. 1944. *Swahili grammar (including intonation)*. London: Longmans, Green & Co.
- Baker, Mark C. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3). 373–415.
- Bastin, Yvonne. 1985. *Les relations sémantiques dans les langues bantoues* (Classe des Sciences Morales et Politiques, Mémoires in-8°, Nouvelle Série 48(4)). Brussels: Académie royale des sciences d'outre-mer.
- Bastin, Yvonne. 1986. Les suffixes causatifs dans les langues bantoues. *Africana Linguistica* 10. 55–145.

- Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and divergence in the Bantu languages: Perspectives from a lexicostatistic study* (Annals – Human Sciences 162). Tervuren: Royal Museum for Central Africa. https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/lexicostatistic-study-bantu-languages.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bastin, Yvonne & Pascale Piron. 1999. Classifications lexicostatistiques : bantou, bantou et bantoïde (de l'intérêt des « groupes flottants »). In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 149–163. Stanford, CA: Center for the Study of Language & Information.
- Beck, David & Igor Mel'čuk. 2011. Morphological phrasemes and Totonacan verbal morphology. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 49(1). 175–228.
- Bilola, Edmond. 1994. Passive incorporation and clause structure. *Journal of West African Languages* 24(1). 101–107.
- Bolekia Boleká, Justo. 1991. *Curso de lengua bubi* (Colección Ensayos 8). Malabo: Centro Cultural Hispano-Guineano.
- Bostoen, Koen. 2008. Bantu spirantization: Morphologization, lexicalization and historical classification. *Diachronica* 25(3). 299–356.
- Bostoen, Koen. 2009. Shanjo and Fwe as part of Bantu Botatwe: A diachronic phonological approach. In Akinloye Ojo & Lioba Moshi (eds.), *Selected Proceedings of the 39th Annual Conference on African Linguistics*, 110–130. Somerville, MA: Cascadilla Proceedings Project.
- Bostoen, Koen. Forthcoming. Reflexive-reciprocal polysemy in south-western Bantu: Distribution, typology and origins. In Eva-Marie Bloom Ström, Hannah C. Gibson, Rozenn Guérois & Lutz Marten (eds.), *Current approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.
- Bostoen, Koen, Sebastian Dom & Guillaume Segerer. 2015. The antipassive in Bantu. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 53(4). 731–772.
- Bostoen, Koen & Heidi Goes. 2019. Was Proto-Kikongo a 5 or 7-vowel language? Bantu spirantization and vowel merger in the Kikongo Language Cluster. *Linguistique et Langues Africaines* 5. 25–68.
- Bostoen, Koen & Claire Grégoire. 2007. La question bantoue : bilan et perspectives. *Mémoires de la Société de Linguistique de Paris, Nouvelle Série* 15. 73–91.

- Bostoen, Koen & Yolande Nzang-Bie. 2010. On how 'middle' plus 'associative/reciprocal' became 'passive' in the Bantu A70 languages. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 48(6). 1255–1307.
- Childs, Tucker G. 1995. *A grammar of Kisi*. Berlin: Mouton de Gruyter.
- Daeleman, Jan. 1966. *Morfologie van naamwoord en werkwoord in het Kongo (Ntandu) met ontleding van het foneemsysteem*. Leuven: Katholieke Universiteit te Leuven. (Doctoral dissertation).
- Dammann, Ernst. 1954. Reziprok und Assoziativ in Bantusprachen. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 104(1). 163–174.
- Dammann, Ernst. 1959. Inversiva und Repetitiva in Bantusprachen. *Afrika und Übersee: Sprachen, Kulturen* 43(2). 116–127.
- de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87–162 + supplementary material online.
- Dieu, Michel & Patrick Renaud. 1983. *Situation linguistique en Afrique centrale : inventaire préliminaire : le Cameroun* (Atlas linguistique de l'Afrique centrale; Atlas linguistique du Cameroun). Paris; Yaoundé: ACCT; CERDOTOLA & DGRST.
- Dimmendaal, Gerrit J. 1987. Drift and selective mechanisms in morphological change: The eastern Nilotic case. In Anna Giacalone-Ramat, Onofrio Carruba & Giuliano Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*, 193–210. Amsterdam: John Benjamins.
- Doke, Clement M. 1938. *Text book of Lamba grammar*. Johannesburg: Witwatersrand University Press.
- Dom, Sebastian, Heidi Goes & Koen Bostoen. Forthcoming. Multiple reciprocity marking in the Kikongo Language Cluster: Functional distribution and origins. In Eva-Marie Bloom Ström, Hannah C. Gibson, Rozenn Guérois & Lutz Marten (eds.), *Current approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.
- Dom, Sebastian, Leonid Kulikov & Koen Bostoen. 2016. The middle as a voice category in Bantu: Setting the stage for further research. *Lingua Posnaniensis* 59(2). 129–149.
- Doneux, Jean Léonce. 1975. Hypothèses pour la comparative des langues atlantiques. *Africana Linguistica* 6. 41–129.
- Ernst, Urs. 1998. *Le groupe verbal et ses éléments en kakɔ (1ère partie)*. Yaoundé: Société Internationale de Linguistique.

- Essono, Jean-J. Marie. 2000. *L'ewondo : langue bantu du Cameroun : phonologie – morphologie – syntaxe*. Yaoundé: Presses de l'Université catholique d'Afrique centrale.
- Fehn, Anne-Maria. 2019. Kuvale: A Bantu language of southwestern Angola. *Journal of African Languages and Linguistics* 40(2). 235–270.
- Fisch, Maria. 1998. *Thimbukushu grammar*. Windhoek: Out of Africa Publishers.
- Forges, Germaine. 1983. *Phonologie et morphologie du kwezo* (Annales – Série in-8° – Sciences humaines 113). Tervuren: Royal Museum for Central Africa.
- Gautier, Jean-Marie. 1912. *Grammaire de la langue mpongwée*. Paris: Procure des PP. du Saint-Esprit.
- Good, Jeff. 2005. Reconstructing morpheme order in Bantu: The case of causativization and applicativization. *Diachronica* 22(1). 3–57.
- Good, Jeff. 2007. Slouching towards deponency: A family of mismatches in the Bantu verb stem. In Matthew Baerman, Greville G. Corbett, Dunstan Brown & Andrew Hippisley (eds.), *Deponency and morphological mismatches* (Proceedings of the British Academy 145), 203–230. Oxford: Oxford University Press.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guarisma, Gladys. 2000. *Complexité morphologique – simplicité syntaxique. Le cas du bafia, langue bantoue périphérique (A 50) du Cameroun* (Société d'études linguistiques et anthropologiques de France 379). Paris: Peeters.
- Guérois, Rozenn & Koen Bostoen. 2016. Syncretism in Cuwabo valency-increasing verbal extensions. Paper presented at the 6th International Conference on Bantu Languages, Helsinki.
- Guthrie, Malcolm. 1967. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 1: The comparative linguistics of the Bantu languages*. Farnborough: Gregg Press.
- Guthrie, Malcolm. 1970. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part II, Volumes 3 & 4: A catalogue of Common Bantu with commentary*. Farnborough: Gregg International.
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Heine, Bernd. 1972/73. Zur genetischen Gliederung der Bantu-Sprachen. *Afrika und Übersee* 56(3). 164–185.
- Henson, Bonnie J. 2007. *The phonology and morphosyntax of Kol*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).

- Hulstaert, Gustaaf. 1965. *Grammaire du lomóngó. Deuxième partie : morphologie* (Annales – Série in-8° – Sciences humaines 57). Tervuren: Royal Museum for Central Africa.
- Hyman, Larry M. 2003a. Basaá (A43). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 257–282. London: Routledge.
- Hyman, Larry M. 2003b. Sound change, misanalysis, and analogy in the Bantu causative. *Journal of African Languages and Linguistics* 24(1). 55–90.
- Hyman, Larry M. 2003c. Suffix ordering in Bantu: A morphocentric approach. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2002*, 245–281. Dordrecht: Kluwer Academic Publishers.
- Hyman, Larry M. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville, MA: Cascadilla Press.
- Hyman, Larry M. 2010. Affixation by place of articulation: The case of Tiene. In Jan Wohlgemuth & Michael Cysouw (eds.), *Rara & Rarissima: Documenting the fringes of linguistic diversity* (Empirical Approaches to Language Typology 46), 145–184. Berlin: Mouton de Gruyter.
- Hyman, Larry M. 2014. Reconstructing the Niger-Congo verb extension paradigm: What's cognate, copied or renewed? In Martine Robbeets & Walter Bisang (eds.), *Paradigm change: In the Transeurasian languages and beyond* (Studies in Language Companion Series 161), 103–126. Amsterdam: John Benjamins.
- Hyman, Larry M. 2018. Common Bantoid verb extensions. In John R. Watters (ed.), *East Benue-Congo: Nouns, pronouns, and verbs* (Niger-Congo Comparative Studies 1), 173–198. Berlin: Language Science Press. DOI: 10.5281/zenodo.1314327.
- Hyman, Larry M. 2022. Causative and passive high tone in Bantu: Spurious or proto? In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 281–308. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575825.
- Idiatov, Dmitry. 2022. Diachronic typology and the reconstruction of non-selective interrogative pronominals in Proto-Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 667–737. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575843.
- Ittmann, Johannes. 1939. *Grammatik des Duala (Kamerun) unter Mitarbeit von Carl Meinhof* (Beihefte zur Zeitschrift für Eingeborenen-Sprachen 20). Berlin: Dietrich Reimer.

- Ittmann, Johannes. 1971. *Sprichwörter der Kuntu (Kamerun)* (Deutsche Akademie der Wissenschaften, Institut für Orientforschung 75). Berlin: Akademie.
- Janson, Tore. 2007. Bantu spirantisation as an areal change. *Africana Linguistica* 13. 79–116.
- Kabuta, Ngo Semzara & Maya Schiffer. 2009. *Mufòlòji wà Cilubà: Algemene inleiding tot de structuur van het Cilubà*. Ghent: Recall.
- Kamba Muzenga, Jean-Georges. 1980. *Esquisse de grammaire kete* (Annales – Série in-8° – Sciences humaines 104). Tervuren: Royal Museum for Central Africa.
- Kemmer, Suzanne. 1993. *The middle voice* (Typological Studies in Language 23). Amsterdam: John Benjamins.
- Kenmogne, Michel. 2000. *The lexical phonology of Bakoko*. Buea: University of Buea. (Doctoral dissertation).
- Kongne Welaze, Jacques. 2004. *Morphologie verbale du tuki*. Yaoundé: Université of Yaoundé I. (MA thesis).
- Kröger, Heidrun. 2016. Verbal plurality – not so common extensions in Mozambican Ngoni. Paper presented at the 6th International Conference on Bantu Languages, Helsinki.
- Kuijpers, Em. 1922. *Grammaire de la langue haya*. Boxtel: Prokuur van de Witte Paters.
- Kuperus, Juliana. 1985. *The Londo word: Its phonological and morphological structure* (Annals – Human Sciences 119). Tervuren: Royal Museum for Central Africa.
- Labroussi, Catherine. 2000. Spirant-devoicing and genetic classification in eastern Bantu. In H. Ekkehard Wolff & Orin D. Gensler (eds.), *Proceedings of the 2nd World Congress of African Linguistics, Leipzig 1997*, 323–358. Cologne: Rüdiger Köppe.
- Laman, Karl E. 1936. *Dictionnaire kikongo – français, avec une étude phonétique décrivant les dialectes les plus importants de la langue dite kikongo* (Institut royal colonial belge, Section des sciences morales et politiques, Mémoires – Collection in-8°, Tome II). Brussels: Georges van Campenhout.
- Lichtenberk, Frantisek. 1985. Multiple uses of reciprocal constructions. *Australian Journal of Linguistics* 5(1). 19–41.
- Livinhac, Léon, Camille Denoit & Albert Wolters. 1921. *Grammaire luganda*. 4th revised edition. Maison-Carrée (Alger): Imprimerie des Missionnaires d’Afrique (Pères Blancs).
- Mackey, James L. 1855. *A grammar of the Benga language*. New York, NY: Mission House.

- Maddox, Harry E. 1938. *An elementary Lunyoro grammar*. London: Society for Promoting Christian Knowledge.
- Maganga, Clement & Thilo C. Schadeberg. 1992. *Kinyamwezi: Grammar, texts, vocabulary* (East African Languages and Dialects 1). Cologne: Rüdiger Köppe.
- Maslova, Elena S. 2007. Reciprocal and polyadic (remarkable reciprocals in Bantu). In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions* (Typological Studies in Language 71), 335–352. Amsterdam: John Benjamins.
- Mchombo, Sam A. & Armindo S. A. Ngunga. 1994. The syntax and semantics of the reciprocal construction in Ciyao. *Linguistic Analysis* 24(1-2). 3–31.
- Meeussen, Achiel Emiel. 1954a. *Linguistische schets van het Bangubangu* (Annalen - Reeks in-8° - Wetenschappen van de Mens - Linguïstiek 5). Tervuren: Royal Museum of the Belgian Congo.
- Meeussen, Achiel Emiel. 1954b. Werkwoordafleiding in Mongo en Oerbantoe. *Æquatoria* 17. 81–86.
- Meeussen, Achiel Emiel. 1962. *Notes de grammaire luba-kasayi*. Unpublished ms., 14 pp. Tervuren: Royal Museum for Central Africa.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeussen, Achiel Emiel. 1973. Comparative Bantu: Test cases for method. *African Language Studies* 14. 6–18.
- Meinhof, Carl. 1906. *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. Berlin: Dietrich Reimer.
- Meinhof, Carl. 1910. *Grundriß einer Lautlehre der Bantusprachen, nebst Anleitung zur Aufnahme von Bantusprachen. Anhang: Verzeichnis von Bantuwortstämmen*. 2nd revised edition. Berlin: Dietrich Reimer (Ernst Vohsen).
- Miehe, Gudrun. 1989. Les extensions verbales en swahili et dans les langues voisines. In Marie-Françoise Rombi (ed.), *Le swahili et ses limites : ambiguïtés des notions reçues*, 107–120. Paris: Editions Recherches sur les Civilisations.
- Mithun, Marianne. 2013. Challenges and benefits of contact among relatives: Morphological copying. *Journal of Language Contact* 6(2). 243–270.
- Mous, Maarten. 2003. Nen (A44). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 283–306. London: Routledge.
- Mous, Maarten & Anneke Breedveld. 1986. A dialectometrical study of some Bantu languages (A. 40 – A. 60) of Cameroon. In Gladys Guarisma & Wilhelm J. G. Möhlig (eds.), *La méthode dialectométrique appliquée aux langues africaines*, 177–241. Berlin: Dietrich Reimer.
- Mpiranya, Fidèle. 2015. *Swahili grammar and workbook*. Milton Park, Abingdon: Routledge.

- Mugane, John M. 1999. The recalcitrant nature of the Bantu verbal morpheme -an-. *Linguistic Analysis* 29(1-2). 160–181.
- Mukash Kalel, Timothée. 1982. *Le kanyok, langue bantoue du Zaïre : phonologie, morphologie, syntagmatique*. Paris: Université Sorbonne Nouvelle, Paris 3. (Doctoral dissertation).
- Mwenegoha, Hamza A. K. 1975. *Kifo cha furaha*. Nairobi: Transafrica.
- Nabirye, Minah. 2016. *A corpus-based grammar of Lusoga*. Ghent: Ghent University. (Doctoral dissertation).
- Nkiko, Munya Rugero. 1975. *Esquisse grammaticale de la langue luba-shaba : parler de Kasongo Nyembo* (Travaux et Recherches du CELTA). Lubumbashi: Centre de linguistique théorique et appliquée, Université nationale du Zaïre.
- Nurse, Derek. 1999. Towards a historical classification of East African Bantu languages. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 1–41. Stanford, CA: Center for the Study of Language & Information.
- Nurse, Derek. 2008. *Tense and aspect in Bantu* (Oxford Linguistics). New York, NY: Oxford University Press.
- Nurse, Derek & Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history* (University of California Publications in Linguistics 121). Berkeley, CA: University of California Press.
- Ohala, John J. 1978. Southern Bantu vs. the world: The case of palatalization of labials. In Jeri J. Jaeger, Anthony C. Woodbury, Farrell Ackerman, Christine Chiarello, Orin D. Gensler, John Kingston, Eve E. Sweetser, Henry Thompson & Kenneth W. Whistler (eds.), *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, 370–386. Berkeley, CA: Berkeley Linguistics Society, University of California.
- Orwig, Carol. 1989. Les extensions verbales en nugu. In Daniel Barreteau & Robert Hedinger (eds.), *Descriptions de langues camerounaises* (DELAN: Description systématique des langues nationales), 283–314. Paris: SIL & ACCT & ORSTOM.
- Pacchiarotti, Sara. 2020. *Bantu applicative constructions* (Stanford Monographs in African Languages). Stanford, CA: Center for the Study of Language & Information.
- Pacchiarotti, Sara, Natalia Chousou-Polydouri & Koen Bostoen. 2019. Untangling the West-Coastal Bantu mess: Identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica* 25. 155–229.
- Paluku, André Mbula. 1998. *Description grammaticale du kitalinga : langue bantu du Nord-Est du Zaïre* (LINCOM Studies in African Linguistics 22). München: Lincom Europa.

- Piron, Pascale. 1990. *Éléments de description du kota, langue bantoue du Gabon*, B.25. Brussels: Université libre de Bruxelles. (MA thesis).
- Racine, Odile. 2015. *Les extensions verbales en swahili standard* (East African Languages and Dialects 24). Cologne: Rüdiger Köppe.
- Raponda-Walker, André. 1937. *Essai de grammaire tsogo*. Brazzaville: Government of Brazzaville.
- Richter, Doris. 2013. *Wovia dictionary*. Buea: University of Buea.
- Rubongoya, L. T. 1999. *A modern Runyoro-Rutooro grammar* (East African Languages and Dialects 9). Cologne: Rüdiger Köppe.
- Rugemalira, Josephat M. 1993. *Runyambo verb extensions and constraints on predicate structure*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).
- Schadeberg, Thilo C. 1980. La morphologie verbale du bantu commun et les langues bantoues du Cameroun. In Luc Bouquiaux (ed.), *L'expansion bantoue : actes du colloque international du CNRS, Viviers (France) – 4–16 avril 1977*, vol. 2 (Société d'études linguistiques et anthropologiques de France, Numéro spécial 9), 503–509. Paris: SELAF.
- Schadeberg, Thilo C. 1982. Les suffixes verbaux séparatifs en bantou. *Sprache und Geschichte in Afrika* 4. 55–66.
- Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 71–89. London: Routledge.
- Schadeberg, Thilo C. 2004. Kutendekana: Socio-stative verbs in Swahili. Paper presented at the 17th Swahili Colloquium, Bayreuth.
- Schadeberg, Thilo C. & Koen Bostoen. 2019. Word formation. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 2nd edn. (Routledge Language Family Series), 172–203. Milton Park, Abingdon: Routledge.
- Sebasoni, Servilien. 1967. La préfinale du verbe bantou. *Africana Linguistica* 3. 122–135.
- Segerer, Guillaume. 2002. *La langue bijogo de Bubaque (Guinée Bissau)* (Afrique et Langage 3). Louvain: Peeters.
- Stappers, Leo. 1967. Het passief suffix -u- in de Bantoe-talen. *Africana Linguistica* 3. 137–145.
- Stappers, Leo. 1986. *Kanyok: Eine Sprachskizze*. Cologne: Institut für Afrikanistik, Universität zu Köln.
- Ström, Eva-Marie. 2013. *The Ndengeleko language of Tanzania*. Gothenburg: University of Gothenburg. (Doctoral dissertation).
- Taylor, Carrie. 1986. Extensions du radical en nɔmaante : la morphologie dérivationnelle. *Journal of West African Languages* 16(1). 53–62.

- Torrend, Julius. 1891. *A comparative grammar of the South-African Bantu languages, comprising those of Zanzibar, Mozambique, the Zambezi, Kafirland, Benguela, Angola, the Congo, the Ogowe, the Cameroons, the Lake Regions*. London: Kegan Paul, Trench, Trübner & Co.
- Trithart, Mary Lee. 1977. Causatives and instrumentals. In Ernest Rugwa Byarushengo, Alessandro Duranti & Larry M. Hyman (eds.), *Haya grammatical structure: Phonology, grammar, discourse* (Southern California Occasional Papers in Linguistics 6), 73–88. Los Angeles, CA: Linguistics Department, University of Southern California.
- Trithart, Mary Lee. 1983. *The applied suffix and transitivity: A historical study in Bantu*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Van de Velde, Mark. 2008. *A grammar of Eton* (Mouton Grammar Library 46). Berlin: Mouton de Gruyter.
- van der Spuy, Andrew. 2014. Bilabial palatalisation in Zulu: A morphologically conditioned phenomenon. *Stellenbosch Papers in Linguistics Plus* 44. 71–87.
- Vansina, Jan. 1995. New linguistic evidence and ‘the Bantu expansion’. *Journal of African History* 36(2). 173–195.
- Velten, Carl. 1901. *Safari za wasuaheli*. Göttingen: Vandenhoe & Ruprecht.
- Viana, Miguel José. 1961. *Dicionário de português-chi-yao e chi-yao-português; elementos de gramática*. Lourenço Marques: Imprensa Nacional de Moçambique.
- Voeltz, Erhard F. K. 1977. *Proto Niger-Congo verb extensions*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Werner, Alice. 1919. *Introductory sketch of the Bantu languages*. London: Kegan Paul, Trench, Trübner & Co.
- Wynne, R. C. 1980. *English – Mbukushu dictionary*. Amersham: Avebury.

Part III

**Proto-Bantu clausal
morphosyntax and information
structure**

Chapter 9

Predicate structure and argument indexing in early Bantu

Tom Güldemann

Humboldt University of Berlin and Max Planck Institute for Evolutionary Anthropology in Leipzig

Meeussen's (1967: 108–111) Proto-Bantu reconstruction involves a morphologically compact predicate with bound cross-reference on the verb for core arguments, which indeed characterises the majority of modern languages in the Bantu spread zone. In the north-west, however, numerous Bantu languages possess a split predicate structure with free pronouns or person-inflected portmanteau morphemes that also encode tense, aspect, modality, and polarity. This feature is also found in many languages of the Macro-Sudan Belt, a large convergence area neighbouring the Bantu spread zone and hosting its homeland and Bantu's closest relatives in Benue-Kwa (Güldemann 2008; 2018). Moreover, several Proto-Bantu subject and object prefixes reconstructed by Meeussen (1967) and other researchers deviate considerably from pronoun forms that can be assumed for early Benue-Kwa and Niger-Congo in general (Güldemann 2017). Against this background, the present chapter proposes a revised conceptualisation of pronominal participant marking in early Bantu that can reconcile the modern empirical data in this group with the typological profile of the area where Proto-Bantu originates. It implies that Meeussen's verbal argument cross-reference reconstructions are themselves valid, both in terms of morphosyntactic status and segmental form, but should not be projected back to the proto-stage that gave rise to the entire Narrow Bantu family as traditionally defined. Since these reconstructions differ from argument cross-reference in predicates elsewhere in Benue-Kwa, they should be seen as innovations in later ancestral stages of Bantu.



The diversity of predicate structures in modern Bantu is far greater, however, than the template in Table 1 would suggest. In the north-west, one finds many languages with verb structures such as the one in (2) from Ewondo A72a.

- (2) Ewondo A72a (Redden 1979: 56)
a-kad m̩ə soób bī-yé
 3SG-HAB 1SG wash 8-cloth
 ‘He washes clothes for me.’

Similar patterns are also widespread in the closest relatives of Narrow Bantu in the Macro-Sudan Belt, as illustrated in (3) with an example from Aghem.

- (3) Aghem [Grassfields Bantu, Bantoid, Benue-Kwa, Niger-Congo] (Hyman 2010: 101–102)
ò m̩ó zì kí-bé ‘né
 3SG PROX.PST eat 7-fufu today
 ‘He ate fufu today.’

The examples in (2) and (3) show that both Narrow Bantu languages from the north-west and their closest relatives outside of Narrow Bantu feature independent subject and object pronouns and/or so-called STAMP morphemes (see Anderson 2011; 2012; 2015; 2016), combining Subject cross-reference with the marking of Tense, Aspect, Modality and/or Polarity, such as *a-kad* (3SG-HAB) in (2).

There is an important caveat to make. On this scale of observation, the assessment of elements referring pronominally to subject and object as more independent from the verb lexeme has to rely to a large extent on the orthographic conventions applied in the hundreds of languages concerned. It has been claimed that in West African languages argument cross-reference on the verb is largely prefixal/bound (cf. e.g. Creissels 2000: 235). Unfortunately, this claim has not yet been supported by conclusive evidence. Until proven otherwise, I cannot help identifying a consistent areal pattern in the fact that clausal argument indexation in the north-west of the Bantu spread zone and in the adjacent Macro-Sudan Belt is so often written separately from the verb lexeme, as opposed to the consistent conjunctive writing in most languages of the core Bantu area. Doing otherwise would imply that large parts of all previous assessment of languages in Africa and beyond regarding their morphological typology are spurious.

The modern and geographically structured diversity sketched above begs the question of what Meeussen’s reconstructed template in Table 1 represents. Nurse (2008: §6) provides a thorough discussion of issues revolving around the reconstruction of Bantu verb structure, which is viewed to involve the lexical verb,

verbal argument cross-reference for subject and object, and various types of predicate operators expressed by auxiliaries, particles, and affixes of variable position and host. While a complete treatment needs to account for tone (e.g. Kisseberth & Odden 2003: 61–62; Downing 2011; Marlo 2013; Odden & Bickmore 2014), I focus here on the segmental aspects of Bantu predicates.

So far, there is no consensus on the historical interpretation of the diversity illustrated with (1), (2) and (3) above. Three major proposals have been made to derive the different verbal structures in modern languages from an early Bantu predicate structure. These are given in (4). Capital letters stand for individual morphemes as meaning-bearing units with C representing the verb root, the lexical core of the predicate.

(4) Different proposals for an early Bantu predicate structure

- a. I [A] [B] [C] [D] [E] [F] e.g. Meinhof (1938)
- b. II [A-B-C-D-E-F] e.g. Meeussen (1967: §2, §6–7)
- c. III [A-B] [C] [D-E-F] [A-B] [C-[D-E-F]] + other patterns
e.g. Güldemann (2003)

Meinhof’s proposal I in (4a), which derives all agglutinative structures in Bantu from the isolating language type found recurrently in West and Central Africa, is not dealt with further here, as I consider it completely discarded today by African linguists. The pattern II in (4b), which I label the “compact predicate hypothesis”, represents the general consensus since Meeussen’s (1967) work. It derives the present-day structures in (2) and (3) by means of erosion (cf. e.g. Schadeberg 2003: 156) or erosion and partial dismantling (Hyman 2007; 2011) of the assumed inherited agglutinative structure. Profile III in (4c), which is intermediate between the extremes of I and II and involves various patterns, is referred to here as the “split predicate hypothesis”, where multi-word predicates separating subject marking and verb stem were typical despite the existence of a certain amount of bound morphology. This pattern has been proposed more recently, notably by Güldemann (2003; 2007; 2011b,a; 2013), Good & Güldemann (2006), and Nurse (2007; 2008: 62–72). It considers the highly agglutinative predicates characteristic of many modern-day Bantu languages as a later innovation through phonological fusion of the verb stem domain with preceding material.

Recent macro-areal research (cf. Güldemann 2010; 2011a; 2018) argues that the Bantu family forms its own large spread zone and differs strongly from the typological profile of the Macro-Sudan Belt from where Bantu originally spread out. One of the most striking differences is the degree of morphological synthesis in the verb. Against this background, at least two opposite scenarios can account

for the emergence of the modern geographical gradient between split predicates in the north(west) and increasingly compact predicates in the south(east). These are schematised in Figure 1. The panel on the left side of the figure represents the traditional “compact predicate hypothesis” (see II in (4) above), while that on the right side illustrates the ‘split predicate hypothesis’ (see III in (4) above). The upper and lower boxes in each panel represent the two geographical areas Macro-Sudan Belt and Bantu spread zone, respectively. The arrows symbolise the major typological shift from Proto-Bantu to the modern situation, as implied by each scenario, i.e. from more to less agglutinative in scenario II (left panel) and from less to more agglutinative in scenario III (right panel). According to the last hypothesis, Meeussen’s (1967) reconstruction in Table 1 would be a later stage in Bantu. The situation in divergent north-western Bantu languages is not ascribed to erosion let alone morphological dismantling, as is assumed in commonly held positions. Rather, it reflects an earlier stage out of which compact predicates developed during the southward expansion of Bantu.

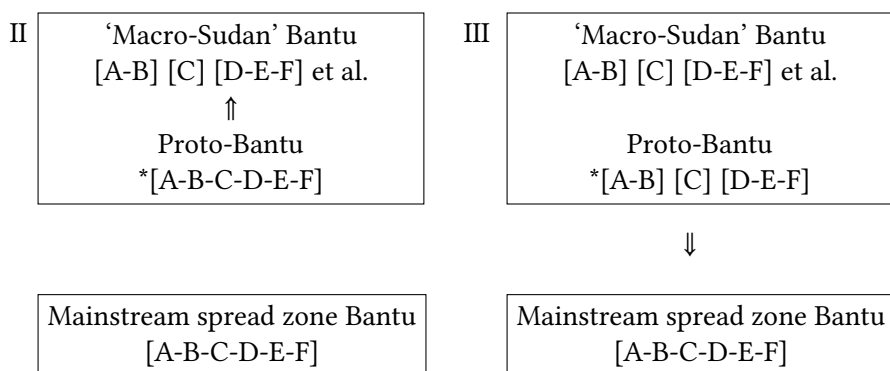


Figure 1: Two areal-historical models for the modern verb-synthesis Bantu profile

Güldemann (2011a: 126) writes on this stage:

Pre- or even Proto-Bantu possessed a split predicate distributed over more than one phonological word. Its basic constituents would have been the preverbal complex of predicate markers for the subject and predication operators, and secondly the verb stem involving (possibly multiple) extension suffixes but with some degree of size restriction. Non-subject pronouns occurred alternatively before or after the verb stem. If preceding it, object pronouns could enter with the verb into a tighter prosodic constituent known in Bantu linguistics as the ‘macrostem’. It should also be considered that subject pronouns or other class-indexing markers that immediately preceded

a verb stem (like in some simple verb forms or verbal nouns) also entered the macrostem domain and thus fused here earlier than in more complex predicate types.

As partly sketched in (5a) for simplex and (5b) for complex predicates, my envisaged profile allows for a diverse range of morphological patterns of predicates and narrow verb forms. The proposal even involves cases of simple phonological words with pronominal marking prefixed to a verb stem or auxiliary. The interpretation that my hypothesis implies that “Proto-Bantu and Proto-Niger-Congo had no inflectional verb prefixes” – cf. Bostoen (2019: 324); similarly Hyman (2011: 3, 5, 31) – is thus inadequate. It does imply, however, that Proto-Bantu did not have the morphologically complex compact predicate structure in Table 1.

(5) Range of morphological patterns of predicates and narrow verb forms

- a. i. [SBJ-STEM]
- ii. [OBJ-STEM]
- iii. [INF-STEM]
- b. i. [SBJ-AUX] [∅ STEM]
- ii. [SBJ-AUX] [SBJ-STEM]
- iii. [SBJ-AUX] [OBJ-STEM]
- iv. [SBJ-AUX] [INF-STEM]

Some amount of the diverse structural profile implied by hypothesis III, notably split patterns as in (5b), still exists widely in the Bantu spread zone where the compact predicate of Table 1 clearly predominates today. The patterns not only persist there but can also be observed to transform to the standard compact type. Example (6) from Shona illustrates the origin of a compact predicate from the split pattern in (5b-iv), and (7) from Zulu shows a case of a compact predicate emerging from a structure close to that in (5b-ii).

(6) Shona S10 (Fortune 1955: 271)

ndi-ri ku-tora > ndi-riku-tora
1SG-COP 15INF-take 1SG-PROG-take
'I am taking.'

(7) Zulu S42 (Doke 1927: 169)

ngi-be ngi-thanda > bengi-thanda
1SG-be:PST 1SG-love PROX.PST.IPFV:1SG-love
'I was loving.'

In the following, I try to substantiate scenario III by looking at the cross-Bantu diversity regarding the form of speech-act participant cross-reference (1SG/PL and 2SG/PL) before the verb stem and comparing it with the relevant earlier language states of the larger family. I start with outlining my recent findings about the pronoun system of early Niger-Congo in §2.1 and contrast them with the current state of reconstruction within Bantu in §2.2. In §2.3, I re-examine the available Proto-Bantu reconstructions regarding two central aspects of pronominal indexation of clausal arguments, namely their fusion with the lexical verb in §2.3.1 and their segmental forms in §2.3.2. In §3, I summarise the results.

While I give further details in §2.3 on the scope and methodology of my investigation, it should be clear already that I do not intend here to provide a full-scale reconstruction of pronominal argument indexation in Proto-Bantu. In view of the scale of such a task, this would be a major project in its own right. This contribution is primarily an arguably viable exercise in diachronic (and partly areal) typology, which I think is needed in the current state of Bantu historical linguistics, including a plea for rethinking the general historical approach to the emergence of modern Bantu diversity.

2 Syntax and form of preverbal participant cross-reference

In the main body of §2, I assess central diachronic issues of pronominal forms in Bantu and its ancestors. I first look at historical stages prior to Bantu, namely Niger-Congo and Benue-Kwa (§2.1). I then discuss Proto-Bantu as currently reconstructed but differing considerably from the former (§2.2). Finally, I undertake an evaluation of the full array of argument cross-reference in modern Bantu languages regarding morphological status as independent or bound (§2.3.1) and segmental form (§2.3.2) in order to compare it with that in earlier states with a view to reconstruction.

2.1 Pronouns in early Niger-Congo

It may appear strange to try to approach the reconstruction of pronominal marking in a relatively young and still tightly knit family like Bantu from the perspective of Niger-Congo as it is old and highly diverse. Nevertheless, the historical assessment of its pronouns has both general and specific advantages in the present context. Pronouns are historically relatively stable and form paradigms that are not only historically more diagnostic but also quite restricted, as opposed, for instance, to the multiplicity of TAMP operators as attested in certain Niger-Congo

Table 2: Pronoun paradigms in Early Niger-Congo and some conservative subgroups (Güldemann 2017: 114)

Genealogical pool	Lineage ⁺	1SG	2SG	1PL	2PL
Early Niger-Congo		<i>*mV^{front}</i>	<i>*mV^{back}</i>	<i>*TV^{close}</i>	<i>*NV^{close}</i>
Ubangi	Gbayaic	<i>*mí</i>	<i>*mɛ</i>	<i>*-lɛ́</i>	<i>*-nɛ́</i>
Adamawa	Mumuyic	<i>*mE/ *N</i>	<i>*mo</i>	<i>*rO</i>	<i>*noO</i>
Adamawa	Kwa-Baa		<i>ĪyŌ -mù</i>	<i>-(t)</i>	<i>-n</i>
Adamawa	Fali	<i>(-)mì</i>	<i>*mu</i>	<i>*-to</i>	<i>*-no</i>
Gur	Central: Oti-Volta	<i>*mV</i>	<i>*bV/(f)V</i>	<i>*tV</i>	<i>*(n)yV</i>
Atlantic	Mel: Temnic	<i>*mi</i>	<i>*mO</i>	<i>*sV</i>	<i>*nV</i>
Atlantic	Sua	<i>meN-</i>	<i>mɔɔ</i>	<i>nrɔ</i>	<i>nɔɔ</i>
Early Benue-Kwa		<i>*mV^{front}</i>	<i>*(B)V^{back}</i>	<i>*TV^{close}</i>	<i>*NV^{close}</i>
Benue-Kwa	Oko	<i>-mɛ</i>		<i>-wɔ</i> <i>-tɔ</i>	<i>-nɔ</i>
Benue-Kwa	Lagoon: Abé	<i>mə</i>		<i>fə</i> <i>-lə</i>	<i>-ɲə</i>

lineages, notably Bantu (cf. Nurse & Philippson 2006; Nurse 2008; Nurse et al. 2016). Accordingly, one can observe considerable recent advances in pronoun reconstruction of Niger-Congo and Benue-Kwa.²

In Güldemann (2017), I propose an approximate proto-paradigm for speech-act participant pronouns given in the second line of Table 2.³ While these are not proto-forms in the canonical sense of the Comparative Method, as explained in the article and marked accordingly by subscript *, the paradigms of selected Niger-Congo cases in Table 2 represent evidence for their plausibility (close cognates are left-aligned). Given the amount of data involved, the hypothetical exponents, i.e. the phonological expressions of the relevant morphosyntactic categories, are necessarily abstract. However, they are still concrete enough for an informative comparison with forms attested across modern Bantu.

²Benue-Kwa is a major Niger-Congo branch, also known as East Volta-Congo, and includes Kwa and Benue-Congo (cf. Williamson & Blench 2000: 18). Bantu is one of its lower-level offshoots.

³The reconstruction does not necessarily represent Proto-Niger-Congo but may well reflect a later stage. For example, the eastern Ubangi lineages do not give evidence for the full pronoun set and could well be outside the clade whose ancestor possessed the proto-paradigm in Table 2.

A couple of other points need to be made in the present context. First, the paradigms from outside Bantu hardly ever involve cross-reference that is bound to the verb stem. Subject pronouns are either independent or enter so-called STAMP morphemes within the above-mentioned split predicate structure. The latter is an areal feature of the Macro-Sudan Belt (Güldemann 2011a; 2013; 2018) and is even reconstructed by Anderson (2011; 2012; 2015; 2016) for various lineages of this area, including some of Niger-Congo. In other words, argument indexation bound to the verb neither appears to be deeply entrenched in Niger-Congo nor is it a trait that characterises the areal context of the Bantu homeland. Finally, in Güldemann (2017), I discuss evidence for the narrower context of Benue-Kwa (and independently in a few other cases) that the denasalised 2SG form $*(B)V^{back}$ is a later innovation, which is particularly relevant for its possible reflex in Proto-Bantu at issue here.

2.2 Pronouns and bound verbal argument cross-reference in Bantu

Previous reconstructions of Proto-Bantu bound argument cross-reference on the verb show considerable agreement, not only in assuming all markers to be affixes but also regarding their specific forms. This is apparent from the various proto-paradigms in Table 3, even if the later ones may well build to some extent on Meeussen's (1967) first reconstruction.

Table 3: Various versions of the Proto-Bantu bound verbal cross-reference paradigm

Reconstruction	Subject				Object			
	1SG	2SG	1PL	2PL	1SG	2SG	1PL	2PL
Meeussen (1967)	*n-/ɲ-	*u-	*tu-	*mu-	*n-	*ku-	*tu-	*mu-
Guthrie (1971: 10)	*NI-/NY-	*Ū-	*TŪ-	*MŪ-	–	*KŪ-	*TŪ-	*MŪ-
Schadeberg (2003: 151)	*N-	*Ū-	*tŪ-	*mŪ-	*-N-	*kŪ-	*tŪ-	*mŪ-
Nurse (2007: 250)	*n-	*Ū-	*tŪ-	*mŪ-	–	–	–	–
Babaev (2008: 148)	*ɲi-	*Ū-	*tŪ-	*mŪ-	–	*kŪ-	–	–

Note: <Ū> renders the vowel commonly represented as <u> in Bantu historical studies

When comparing the Bantu reconstructions in Table 3 with those for higher genealogical levels, one can observe a considerable amount of cognacy. Table 4 provides a comparison between the three historical stages of Benue-Kwa (cf. Güldemann 2017, Table 2 above), Bantoid (and, as will be shown, parts of north-western Bantu; cf. Babaev 2008), and Narrow Bantu. Importantly, the similarity

Table 4: The reconstruction of pronominal marking in Bantu and beyond

Reconstructions	1SG	2SG	1PL	2PL	3SG. HUMAN	3PL. HUMAN
Benue-Kwa	* <i>mV^{front}</i>	* <i>(B)V^{back}</i>	* <i>TV^{close}</i>	* <i>NV^{close}</i>	* <i>(V^{back})</i>	* <i>(ba)</i>
Bantoid	* <i>mi</i> , * <i>N-</i>	* <i>ɔ</i>	* <i>tɔ~tɪ</i>	* <i>nɔ~nɪ</i>	* <i>(j)ɔ</i> , * <i>a</i>	* <i>ba</i>
Bantu ⁺						
Non-verbal ^o	* <i>-mi-</i>	* <i>-w-</i>	* <i>-cu-</i>	* <i>-nu-</i>	* <i>-w-</i> , * <i>-j-</i>	* <i>-ba-</i>
Verbal subject	* <i>ɲi-</i>	* <i>ɔ-</i>	* <i>tɔ-</i>	* <i>mɔ-</i>	* <i>ɔ-</i> , * <i>a-</i>	* <i>ba-</i>
Verbal object	* <i>ɲi-</i>	* <i>kɔ-</i>	* <i>tɔ-</i>	* <i>mɔ-</i>	* <i>mɔ-</i>	* <i>ba-</i>

Notes: after Schadeberg (2003: 149, 51), Kamba Muzenga (2003), Babaev (2008); hyphens do not indicate the status as infixes but that morphemes can be prefixed and/or suffixed to these pronominal roots

exists between largely *free* forms in the first two units and *bound* forms in currently conceived Proto-Bantu, particularly the exponents in the line “non-verbal” of Table 4 for non-verbal morphosyntactic contexts like independent and possessive pronouns (cf. Stappers 1986; Kamba Muzenga 2003). This picture already suggests that cognate forms are unlikely to have been involved in the early past in a compact predicate with participant cross-reference. With the enormous time depth assumed for Benue-Kwa (or Niger-Congo), any bound exponents of such early stages would be expected today to show signs of erosion in this context rather than being largely identical to their free counterparts.

The differences in Table 4 are equally revealing. First, Benue-Kwa and Bantoid pronouns do not give systematic evidence for the functional differentiation in Bantu in the form of distinct paradigms. Second, the Proto-Bantu bound argument cross-reference on the verb deviates significantly from the Benue-Kwa and Bantoid pronoun canon in four person-number positions.⁴ These are 1SG subject and object **Ni* vs. **mV^{front}*, 2SG object **kU* vs. **(B)V^{back}* (from earlier **mV^{back}*), 2PL subject and object **mU* vs. **NV^{close}*, and 3SG human object (= noun class 1) **mU* vs. **V^{back}*.

The received Bantu reconstructions as such are not assumed to be invalid, as they are supported by extensive empirical evidence within this group. The

⁴Certain details in the different available reconstructions vary and thus remain indeterminate but at the same time are largely irrelevant for the present topic. While the morphological status and the consonants of the markers are important, the exact quality and tone of the vowel part are secondary. Hence, the latter are represented from now on by means of abstract capitalised segments. In a similar vein, capital *N* in 1SG forms stands for a non-bilabial nasal, which is non-committal to the exact place of articulation.

question is whether they really pertain to Proto-Bantu in terms of Guthrie's (1948; 1971) delimitation of the family or whether they are innovations in lower clades of the phylogeny (cf. Henrici 1973; Stewart 1976: 4, for a similar approach to certain lexical reconstructions).

2.3 Present methodology and data survey

Since Bantu is comprised of several hundred languages, any attempt at reconstructing a Proto-Bantu feature is an enormous task. This is even more relevant for the complex domain of pronoun paradigms as they are central to Bantu morphosyntax and thus tend to enter in construction with lexemes and other form paradigms, making them highly prone to change and variation within diverse and complex morphological environments. Given that I do not intend a full-scale reconstruction, I limit myself to two points: a) the morphosyntax of argument cross-reference in the predicate, and b) the basic segmental shape of the exponents. The hypothesis I advance here is that the geographical cline from the Bantu homeland in the north-west is a proxy for the incremental changes that occurred from earlier to younger clades of the Bantu genealogy as determined, for example, in the recent lexicon-based phylogeny of Grollemund et al. (2015).

Since the expectation is relatively simple, I deem a sample size capable of reflecting such areal trends to be already sufficient. Concretely, forms I assume to be inherited from early Benue-Kwa should still occur detached from the lexical verb close to the Bantu homeland in the north-west, particularly in zones A and B. Further away from the homeland, the retention or loss of more conservative forms is harder to predict, as this depends on the phylogenetic status of an individual language, but if retained, they are likely to then turn up as bound argument cross-reference on the verb. Conversely, presumably innovative forms divergent from those in early Benue-Kwa should be rare or even absent in the north-west but regular as bound verbal cross-reference further away from the homeland. It is important to reiterate that this logic merely expects a rough geographical cline between Bantu homeland and spread zone rather than a clear-cut boundary of the two types of forms, and it does not require an account of language-specific occurrences.

I thus pursue here a methodological shortcut. I undertake an analysis of the data on pronominal argument cross-reference in about 150 Bantu varieties assembled in the appendix of Babaev (2008: 162–179) rather than of a new large dataset with a systematic modern sampling basis, for example, according to the phylogeny in Grollemund et al. (2015). I am fully aware that my dataset underrepresents the Bantu languages in the north-west where the genealogical diversity is

expected to be highest. Furthermore, I restrict myself here to the set of four exponents for speech-act participants, as 3SG/PL forms pertain to the partly separate morphological paradigm of noun classes. My analysis of the language-specific items in Babaev’s dataset is twofold. On the one hand, I classify them according to whether the source lists them as either free morphemes or bound morphemes, i.e. affixes, reflecting the status of person marking vis-à-vis the lexical verb. On the other hand, I undertake a rough cognate judgement in assigning a language-specific form, or a relevant component thereof, to either of two classes: (a) inherited Benue-Kwa form, or (b) other, including the Bantu reconstructions I assume to be widespread innovations. Clearly, this is a rather crude approach from a traditional historical-comparative perspective. However, my investigation does not aim at a genuine reconstruction, but rather at a privative assignment of modern items to two distinct types as prefigured in the assumed reconstructions of Table 4 above, namely early Benue-Kwa (as derived from Proto-Niger-Congo) as opposed primarily to the received Proto-Bantu reconstruction.

Table 5 gives Babaev’s (2008) language coverage separated according to the well-known reference zones, including zone J (cf. Guthrie 1971; Maho 2009).⁵ A number in a cell indicates the number of languages providing data on a given pronominal form in each zone. As can be seen from Table 5, the numbers for languages within an individual zone are not always the same across all person-number features due to possible blanks in Babaev’s data; at the same time, for one language, more than one form may be given there.

Table 5: Bantu languages covered in Babaev’s (2008) cross-reference data

Zone	A	B	C	D	E	F	G	H	J	K	L	M	N	P	R	S	Total
1SG	17	13	12	9	13	7	12	9	6	11	9	3	5	7	4	7	144
2SG	16	13	14	9	14	7	11	7	5	10	9	9	5	7	4	7	147
1PL	14	13	15	9	14	5	15	8	6	10	9	8	6	5	4	7	148
2PL	18	10	15	9	14	6	12	9	6	9	9	9	5	7	4	7	149

In the analysis of a specific person-number form, I sort the language-specific forms for each Bantu zone according to my binary opposition of inherited Benue-Kwa form vs. other innovated form. This serves the purpose of identifying possible areal trends of increase or decrease of the two opposed types. In my two-pronged approach assessing the reconstruction from “top” (= Benue-Kwa) and

⁵Babaev (2008) recognises zone J but fails to reassign Nande JD42 from earlier D to current J, which I correct here. In general, what is labelled D and E in Table 5 is not part of zone J.

“bottom” (= Bantu) simultaneously, a rough geographical picture is already sufficient. Hence, Babaev’s dataset is argued to give a representative picture in spite of his presumably opportunistic language sampling, erratic coverage of paradigmatic items, and incomplete information about them.

2.3.1 Morphosyntactic profile

The first analysis concerns the morphosyntactic status of the cross-reference forms vis-à-vis the lexical verb, for which Table 6 gives the results according to the Bantu zones as long as there is any variation; there is no variation in the south and east so that these zones are no longer distinguished and are lumped under “rest”. Tokens of free forms appear before the slash and those of bound forms after it. It should be borne in mind that my counts reflect the data collation in Babaev (2008), which may well deviate from the full language-specific situation. For example, there could be more diversity within a language, a form given as free may be a more complex STAMP morpheme and thus actually represent an affix rather than an independent pronoun, etc.

Table 6: Free/prefixed predicate cross-reference forms across Babaev’s (2008) data

Zone	1SG		2SG		1PL		2PL		Total
	Subj.	Obj.	Subj.	Obj.	Subj.	Obj.	Subj.	Obj.	
A	15/6	6/2	12/6	5/2	10/3	3/1	11/7	5/2	67/29
B	8/5	0/2	6/7	0/3	2/12	0/2	0/10	0/3	19/49
C	0/13	2/4	1/14	3/5	1/14	2/3	1/18	4/5	14/76
D	0/9	0/5	1/8	1/4	0/9	2/4	0/9	1/3	5/51
H	0/15	1/2	0/8	0/3	0/8	0/3	0/10	0/3	1/52
Rest	0/102	0/40	0/108	0/59	0/102	0/26	0/97	0/48	0/582
Total	23/140	9/50	19/143	9/69	13/140	7/36	12/141	10/61	102/780

Note: Rest = zones E-G, J-S

As expected by my proposed scenario, virtually all variation occurs in north-(west)ern zones that are closer to the Macro-Sudan Belt while argument cross-reference in the zones E-G and J-S is conveyed exclusively by prefixes on the lexical verb. In Table 6, I try to capture the different behaviour of Bantu groups by means of a three-way distinction: the dark shading of a cell means that free forms are more or as frequent as bound markers; light shading symbolises the reverse; and no shading shows that free forms do not exist in the data. The gradual decline

of free forms with greater distance from the Bantu homeland is also observed in numerical terms in the last column for totals. In zone A, free forms predominate; in zones B and C, free forms are still recurrent albeit already in the minority; in zones D and H, free forms are very rare; and in the rest of Bantu, free forms are absent. This is opposed to the picture in the last line for totals across the different pronominal forms where all cells have light shading. The figures here would, according to the problematic “majority rule”, invariably but, I argue, inadequately point towards the veracity of the Bantu reconstruction (cf. also the polar opposite picture in zone A closest to the Bantu homeland).

This overall result does not imply that the assumed fusion between preverbal cross-reference pronouns (or STAMP morphemes) and the stem was a single unitary event, nor that every free form necessarily reflects the original state. There may well be some cases of secondary free forms and, more importantly, person-number markers bound to the verb may have arisen several times independently, obviating the expectation of a single event of innovation. This is because morphological fusion between phonological words is a recurrent natural process in grammaticalisation. Moreover, this process can occur very quickly, as even dialects of a language can differ in this feature. This is reported, for example, for the Pama-Nyungan language Mari in Australia: while its Margany dialect has the conservative state with free forms, as in (8a), the Gunya dialect has verb suffixes, as in (8b).

(8) Mari [Pama-Nyungan, Australian] (Breen 1981: 317, 327)

- a. Margany dialect
ɲaya binda-:lku
1SG sit-PROX:PURP
'I'll stop at home.'
- b. Gunya dialect
binda-ngi-ya
sit-PURP-1SG
'I'm going to sit down.'

There are also straightforward morphosyntactic indications that the verbal template reconstructed by Meeussen (1967) has arisen from earlier more analytical clause structures. As I argue in Güldemann (2011a), a particularly striking piece of evidence is the variable position of object marking. While pre-stem marking in slot -1 of Table 1 is indeed very frequent and thus recognised in the Proto-Bantu reconstruction, some languages have additional or even exclusive

postverbal cross-reference in slot 3 of Table 1. The modern morphotactic variation alone indicates earlier syntactic flexibility in line with Givón's (1971) idea that current morphology reflects past syntax. Since Bantu is overall head-initial and thus more likely to develop postverbal object marking, the prefix slot for objects is quirky in typological terms. However, it can be shown to have arisen from a clausal word-order variant in early Bantu that licensed grammatically conditioned preverbal objects. This kind of word-order variability is not only an areal trait of the Macro-Sudan Belt, rather it is also widely attested in Bantu's closest relatives within Benue-Kwa and thus represents a robust reconstruction for Proto-Bantu before the emergence of a compact predicate (cf. Güldemann 2007; 2008; 2011a). Example (2) above from Ewondo A72a, showing preverbal object pronouns, is therefore a likely syntactic retention.

Again, this syntactic variation need not be tied to a single distinct language stage of early Bantu. It can also exist as a language-internal alternation. A case in point is the Central Sudanic language Ma'di where the Lokai dialect has preverbal objects, as in (9a), but the 'Burulo dialect shows postverbal objects, as in (9b).

(9) Ma'di [Moru-Madi, Central Sudanic] (Blackings & Fabb 2003: 176)

a. Lokai dialect

ámá èbī` jā

1PL.E fish NPST:eat

b. 'Burulo dialect

ámà já ibī

1PL.E eat fish

'We (excluding you) (are) eat(ing) fish.'

There are also phenomena in north-western Bantu languages indicating that subject indexation was not morphologically induced by an obligatory pre-stem slot of a compact predicate. For Kwakum A91, for example, Njantcho Kouangang (2018: 101–116, 273–274) reports that pronominal elements encoding the S/A argument are bound, but they are nevertheless true pronouns replacing full subject noun phrases rather than agreeing with them. The two types of S/A expression, i.e. the ones in (10a) and (10b), are in complementary distribution if the referent is singular. The co-occurrence of a singular noun phrase and a pronoun, as in (10c), is only grammatical if the former is an extra-clausal topic.

(10) Kwakum A91 (Njantcho Kouangang 2018: 273–274)

a. *p^háám H-n-fêH*

1.man PRS-PRS-come

'The man is coming.'

- b. *H-à-n-fèH*
PRS-3SG-PRS-COME
'He is coming.'
- c. *p^hàám H-à-n-fèH*
1.man PRS-3SG-PRS-COME
'[As for] the man, he is coming.'

The obligatory complementariness does not hold for plural S/A arguments because an additional pronoun is optional here, which could be one context from which regular co-occurrence of noun and pronoun, and eventually obligatory bound argument cross-reference on the verb, emerged. This picture can be integrated into Givón's (1976) cross-linguistically relevant historical scenario for the emergence of argument agreement on verbs. That is, languages like Kwakum represent a typologically natural, intermediate stage in the shift from a predicate without obligatory pronominal subject cross-reference to one with full-blown subject agreement in a morphologically compact predicate. If subject pronouns are not obligatory clausal ingredients in the first place, a morphologically prescribed subject slot in the finite verb is hard to entertain.

A relatively late fusion of most of the modern verb prefixes with the lexical verb is also in line with phonological findings about fully agglutinative verb forms. That is, finite verbs are known to involve a word-internal bipartition. Their semantic core is the stem, itself comprising the lexical root with its suffixes, or alternatively the macrostem, which additionally incorporates the pre-stem object marker (cf. Polak 1986: 404–405). Various types of phonological processes with scope over the morphotactic slots from 0/-1 to 2 of Table 1 hold this unit together (cf. e.g. Hyman 2008). To the extent that such phonological processes do not operate further to the left they separate the (macro)stem from the initial prefix complex comprising subject cross-reference and auxiliary-like elements — effectively what Anderson calls a STAMP morpheme in a split predicate structure.

The phonologically-based bipartition of agglutinative verb forms is also reflected by another, admittedly impressionistic observation from my own discourse data on Bantu languages. In natural speech, morphologically unitary verbs can be interrupted by intonation breaks, for example, due to speaker hesitation. The location of such word-internal rupture is regularly at the juncture between the pre-stem and the stem or macrostem. In (11) from Shona S10, “#” marks intonation pauses occurring within verb forms, whereby the pre-stem cluster can, but need not be, repeated. This phenomenon not only supports the verbal bipartition but is also evidence for the internal coherence of the initial STAMP morpheme.

(11) Shona S10 (Tom Güldemann, field notes)

- a. *va-no- # va-no-zvi-pira ku-batsira va-mwe*
 2SBJ-PRS- 2SBJ-PRS-REFL-offer INF-help 2-other
 ‘They are prepared (lit.: offer themselves) to help others.’
- b. *va-mwe va-nhu va-nga-mu- # batsire*
 2-other 2-person 2SBJ-POT-1OBJ- help:IRR
 ‘Other people could/would help him.’
- c. *ndi-no-da ku-zo- # shanda*
 1SG-PRS-want INF-then- work
 ‘I want to work then.’

2.3.2 Segmental form

In the following, I deal with the concrete forms for subject and object cross-reference in the predicate for the eight relevant exponents of speech-act participants, i.e. 1SG/PL and 2SG/PL subject and object indexes. Since there are no appreciable differences between Bantu and Benue-Kwa in the case of the 2SG subject and 1PL subject and object forms (cf. Table 4 above), these do not figure much in the following discussion.

The overall results of my analysis of Babaev’s (2008) data for the remaining relevant forms are given in Table 7. Forms arguably inherited from Benue-Kwa appear before the slash and Bantu-internal innovations occur after it, whereby the heading of Table 7 repeats the competing reconstructions from Table 3 or any other form. The figures after the slash record more generally any forms that differ from Benue-Kwa-like ones. Since they do not only contain likely reflexes of the conventional Bantu reconstructions of Table 3 but also forms that are restricted to individual languages and subgroups, the incidence of Benue-Kwa-like forms vis-à-vis the received Bantu reconstructions is in fact higher. The special case of the 2PL form (and the meaning of “|”) is explained in more detail below.

The picture in Table 7 is overall similar to that in Table 6 in that it is best interpreted in terms of an incremental replacement of Benue-Kwa cognates by Bantu innovations, including those believed to be Proto-Bantu forms, according to the expected geographical pattern. I have again marked the different behaviour of Bantu groups by means of a three-way distinction: the dark shading of a table cell means that old Benue-Kwa forms are more or as frequent as Bantu innovations; light shading symbolises that innovations predominate over old forms; finally, no shading marks that old Benue-Kwa forms no longer exist.

Across the family as a whole, new Bantu forms predominate by a wide margin (see the last line for totals). However, as soon as the data are assessed in geo-

Table 7: Benue-Kwa/Bantu-specific predicate cross-reference forms across Babaev's (2008) data

Zone	1SG		1SG	2PL		Total
	*mi/*Ni+other		*U/*KU +other	*NU/*MU+other		
	Subject	Object	Object	Subject	Object	
A	12/9	5/2	5/2	8/2 7	2/0 5	32/15
B	6/7	1/1	0/3	6/0 4	2/0 1	15/11
C	0/9	2/3	2/4	1/0 13	2/1 3	7/17
D	2/6	0/4	0/3	2/4 3	2/2 0	6/19
E	0/12	0/2	2/11	0/14 -	0/6 -	2/45
H	3/8	1/1	0/3	6/0 6	0/0 3	10/12
K	0/11	0/7	0/5	1/6 1	0/5 0	1/34
L	0/9	0/2	0/1	3/5 1	1/0 0	4/17
R	1/3	0/4	0/4	1/2 1	1/2 1	3/15
S	0/7	0/5	3/5	1/4 2	2/2 2	6/23
Rest	0/50	0/18	0/26	0/49 -	0/15 -	0/158
Total	24/131	9/49	12/67	29/86	12/33	86/366

Note: Rest = zones F, G, J, M, N, P

graphical terms, the picture changes significantly. The Bantu zones A, B, C, D, and H in the north(west) of the family frequently possess forms that are argued here to be retentions from the older Benue-Kwa heritage. It can be expected that this area close to the Bantu homeland harbours languages that are more often conservative than the rest of the family in the colonised area. In the following, I discuss the forms according to the four person categories separately.

I start with the historically most complex case of the 1SG because bound forms with a palatal~alveolar nasal and a close front vowel similar to the received Bantu reconstruction are already recurrent in Benue-Kwa outside Bantu, which led to the reconstruction of such a form for chronolects significantly older than Proto-Bantu. So, it should be clear from the outset that my argument regarding the 1SG form is first of all about the persistence of an original **mi* rather than the absence of **N(i)*.⁶

⁶For the record, there are yet other 1SG forms in Bantu, which complicate the overall picture: see, for example, Bastin (2006) on a form *i*- and Güldemann (2011b) for a fuller list and some discussion. However, most forms are likely to be related to **mi* and/or **Ni* and can thus be argued to derive ultimately from **mi*, which does not alter the general scenario proposed here.

In a survey dedicated to 1SG (Güldemann 2011b), I show that the higher a language (group) is in the conventionally assumed phylogenetic structure of Benue-Kwa the more *m*-forms exist or even predominate, including in Bantu groups in the north-west. While Babaev (2008: 143) concludes “that **me** is a separate subject pronoun not related genetically to ***n(i)**–”, I have presented evidence that **Ni* in fact emerges from (and gradually replaces) inherited **mi*, particularly in the context of bound cross-reference. I even propose that the change of the 1SG exponent from the form **mi* to the form **Ni* may well have occurred multiple times independently across Benue-Kwa and beyond. While this may not seem to be the most economical solution, there are a number of reasons in support of this hypothesis.

A first major factor is that different pronominal categories are not unlikely to fuse with a host in different ways, which mitigates the emergence of a fully symmetrical paradigm of bound pronouns. This has to do with their unequal tendency to occur in verbal constructions and then fuse with other elements as unstressed forms. Mithun (1991: 102) writes from a cross-linguistic perspective (cf. also Givón 1976):

[...] pronominal paradigms do not necessarily become morphologically bound all at once. They may be grammaticalized in predictable stages. Person markers may appear before number markers. Among persons, *first and second person pronouns often become bound before third*. Indefinite third person pronouns may become bound before definite pronouns, and *subjects or ergatives before objects or absolutes*. Number may be distinguished initially for first person, then for second, and only later for third, if at all. (emphasis mine)

Regarding bound argument cross-reference on the predicate in Bantu, this means that the reconstruction of a 1SG prefix does not imply the past existence of a full bound person paradigm. There is indeed ample evidence in Benue-Kwa as a whole not only for the relevant effect of the nominal hierarchy (cf. already Schadeberg 1978 for an extensive discussion concerning Bantu), but also for the greatest likelihood of precisely the 1SG exponent to become bound to its predicate host. That is, the available data suggest that if there is differential argument cross-reference it always includes this paradigmatic item. For example, Green & Igwe (1963: 32) report for Igbo that the 1SG form *mu/mi* partakes in both the incomplete preverbal and postverbal set of partly assimilated subject pronouns and is truncated there to *m*. The Edoid language Engenni (Thomas 1969: 226–228) is a non-Bantu case for the 1SG object form attaching more closely or freely to the

verb stem. In Bantu, this phenomenon is reported for Makaa A83 (Heath 2003: 342, 345), Nzadi B865 (Crane et al. 2011: 158), Rimi F32 (Woolford 2000: 113–115), and across Narrow Bantu in imperative forms (Meeussen 1967: 112; Devos & Van Olmen 2013: 20–21).

In addition to the preference of the 1SG marker to become a bound element before others, it often undergoes sound change, particularly as a proclitic or prefix. For **mi*, this involves in particular the change of the place of articulation in the initial nasal from bilabial *m* to alveolar *n* or palatal *ɲ*, triggered by the quality of the following vowel of the pronoun itself and/or (after vowel loss) by the initial consonant or vowel of the verb stem. Babaev (2008) himself provides evidence that **mV* changes in Benue-Kwa to a bound verb marker and that at least some modern non-bilabial forms are derived from this process, as shown in Table 8 (see also Mieke 2004: 101 for such a hypothesis in genealogically distant Gur). Some Bantu languages even display both forms in the same morphosyntactic context, as shown in (12) for Mbuun B87, where the *mé-* and *N-* 1SG object indexes are interchangeable.

- (12) Mbuun B87 (Bostoen & Mundeke 2011: 77)
a-mpúlúús ba-á-mé/N-leŋ
 2-police 2-PRS-1SG-search
 ‘The police(men) search for me.’

Güldemann (2017: 118–122) shows that the change of a pronoun shape from *mi* to *N(i)* in fact occurred outside Niger-Congo, notably with the 2SG pronoun **mi* in

Table 8: Plausible change from independent 1SG **mV* to bound subject markers

Language	Subgroup	Independent pronoun		Subject proclitic/prefix
Idoma	Idomoid	* <i>mi</i>	>	homorganic <i>N=</i>
Igbo	Igboid	<i>m(u/i)</i>	>	<i>m=</i>
Izere	Platoid	<i>mì</i>	>	<i>mì(-)</i>
Akoose A15C	Narrow Bantu	<i>mè</i>	>	<i>mè- #_V/syllabic N</i> and <i>Ñ-</i> elsewhere
Nen A44	Narrow Bantu	<i>mè</i>	>	<i>mè(-)</i>
Kpa A53	Narrow Bantu	<i>mà</i>	>	<i>m- #_V</i> and <i>ñ- #_C</i>
Bira D32	Narrow Bantu	* <i>mI</i>	>	<i>mi-</i> initially and <i>-m- #V_C</i>

several branches of Central Sudanic. This is significant because these instances are unrelated to those in Benue-Kwa and Bantu in geographical, genealogical, and semantic terms and thus characterise the change as largely phonetically motivated.

There is another, more abstract, argument why *n* from *m* is not an unlikely language change in pronominal forms. Nichols & Peterson (1996: 351) conclude on the basis of a cross-linguistic survey that:

[...] the distribution of *n* is a matter of universal preferences, while that of *m* [...] is less strongly linked to universals and more strongly linked to historical contingencies than that of *n*. *m* is therefore the better potential marker of historical connections.

In a similar vein, Blasi et al. (2016) diagnose a globally observable phono-semantic bias of 1SG pronouns towards the palatal nasal *j*. While forms with exactly this shape are recurrent in Bantu and have been posited as a Proto-Bantu reconstruction (see Table 3 and 4 above), the cross-linguistic findings widen the perspective on the historical evaluation of alveolar and palatal nasals in the Benue-Kwa and Bantu pronoun at issue.

There is also a significant bias in Bantu of largely bound **Ni-* vs. independent **mi* regarding their morphosyntactic contexts. Babaev (2008: 143) observes in this respect:

Statistically, the number of **ni*-forms throughout the [Bantu] family is extremely high in the subject markers, lower in the object, even lower in the possessive markers, and quite rare in the independent stressed pronouns. The share of **me* grows respectively.

While this author wants to reconstruct such a distributional cline to Proto-Bantu and even higher genealogical levels, it can be interpreted inversely. That is, the shift from independent **mi* to bound **Ni-* reflects the expected hierarchy of the innovative fusion of a pronoun with its host as steered by such factors as likely topicality and accompanying de-accentuation and eventual sound change. Insofar as the four grammatical contexts differ with respect to the information status of pronouns and hence their tendency towards fusion, the distribution observed by Babaev arguably reflects where bound **Ni-* would have started its existence and where it encroached upon last (or, as a reviewer observes, its possible successive extension as a bound form to new paradigms).

In general, the potential early existence of a 1SG **Ni-* that was bound to the verb alongside an independent pronominal form **mi* is not evidence for a full-fledged

bound verbal argument cross-reference paradigm. Rather, this coexistence is a reflex of various universal tendencies that converge in recurrently producing **Ni* from **mi*. The free 1SG pronoun **mi* is thus a robust Proto-Bantu reconstruction as well.

The historical assessment of 2PL exponents is also complex in that there are several problems for a superficial cognate identification for both the assumed common Benue-Kwa form in **nV^{close}* and the received Proto-Bantu form in **mU*. For one thing, there are modern Bantu forms that could be cases of denasalisation~fortition of **nV^{close}* to **lV^{close}/dV^{close}* and **mU* to **BU*. Forms with an initial bilabial voiced plosive could reflect the human 3PL marker **ba* of class 2 as a polite form or its incorporation as a (human) plural marker. All such ambiguous forms, whether candidate reflexes of **nU* or **mU*, are found after the vertical bar in the values of Table 7.

It is worth having a closer look at the situation in the zones where forms in arguably inherited *n* are attested. In Table 9, I repeat the values for 2PL from Table 7 but sub-classify them based on whether the initial consonant of the actual forms is *n* or *l/d*, which I consider as possibly related to **nU*, or in *m*, which are likely to derive from **mU*. There are also forms with voiced labial plosives (represented by abstract *B*, see discussion below). Cells are shaded whenever *n/l/d*-forms outnumber *m/B*-forms. I disregard a few other forms, notably plain vowels. The overall picture after this methodological step does not seem to differ much from that of Table 7. I venture, however, that it is in fact more likely that *l*- and *d*-forms

Table 9: 2PL forms according to initial consonants across Babaev’s (2008) data

Zone	Subject					Total	Object					Total
	<i>n</i>	<i>l/d</i>	<i>m</i>	<i>B</i>	other	<i>n+l</i> + <i>d/m+B</i>	<i>n</i>	<i>l/d</i>	<i>m</i>	<i>B</i>	other	<i>n+l</i> + <i>d/m+B</i>
A	8	2	2	4	1x <i>e</i>	10/6	2	1	0	4	–	3/4
B	6	3	0	1	–	9/1	2	0	0	1	–	2/1
C	1	5	0	6	2x <i>o</i>	6/6	2	1	1	1	1x <i>o</i>	3/2
D	2	0	4	3	–	2/7	2	0	2	0	–	2/2
H	6	6	0	0	–	12/0	0	3	0	0	–	3/0
K	1	1	6	0	–	2/6	0	0	5	0	–	0/5
L	3	0	5	0	1x <i>u</i>	3/5	1	0	0	0	–	1/0
R	1	0	2	1	–	1/3	1	0	2	0	1x <i>ku</i>	1/2
S	1	2	4	0	–	3/4	2	2	2	0	–	4/2

are real reflexes of $*nU$, while there are other sources for B -forms besides the theoretically possible denasalisation of $*mU$.

First, the potential cognates of $*nU$ do not only correspond arguably in the consonant but also in the vowel quality, while this is less often the case for the would-be counterparts of $*mU$. Moreover, there is evidence that the initial plosive in at least some of the B -forms reflect the historical $*b$ of a (human) plural marker $*bV$ that fused with both plural pronouns and that such complex forms further contracted. The development of 2PL forms involving this marker $*bV$ can be schematised as: $*nU > *bV-nU \sim *bV-nV \sim *bV-nU-V > bV-n > bV$, with parallels in the 1PL. Such changes occur in Bantoid languages outside Narrow Bantu, as exemplified by Güldemann (2017: 110) for Mambiloid. Looking at the data in Babaev (2008: 175–177) and elsewhere, it can also be reconstructed in Bantu. The plain $*nU$ aside, there is widely distributed evidence far beyond north-western Bantu for the complex form, for example, *bíní* in Mboshi C25, *bijwé* in Lega D25, *βénú* in Bira D32, *βijwé* in Sukuma F21, and *bènò* in Vili H12L (cf. appendix in Babaev 2008). Moreover, suggestive data for the later stages with a lost postnasal vowel or even without the thematic consonant n exist in zone A with such forms as *bin* in Koonzime A842 and *bí* in Makaa A83. Importantly, there is no evidence for the same scenario with reconstructed $*mU$, which would be expected in view of the old age of bV -prefixation if $*mU$ were as old as $*nU$. Considering all these observations, the preponderance of 2PL $*nU$ can be consolidated in the north-western region of Bantu, which is shown in Table 10, based on Table 9.

Table 10: 2PL forms in north-western Bantu across Babaev's (2008) data

Zone	Subject	Object
	$*NU/*MU$	$*NU/*MU$
A	10/2	3/0
B	9/0	2/0
C	6/0	3/1
H	12/0	3/0

Accordingly, I also venture that 2PL forms with initial n and l encountered in the zones D, K, L, R, and S are just as likely to involve reflexes of my assumed old form $*nU$, partly having undergone consonant fortition. The overall picture for 2PL forms is then that $*nU$ predominates in the north-west as well as in zone

H, is gradually replaced by *m*-forms further south and east, but still occurs there sporadically.⁷

It remains to be investigated what was behind the concrete shift from **nU* to **mU*. One obvious factor could be the vowel quality in that the innovative bilabial *m* is closer to the following rounded vowel. In this sense, the shift would be parallel but inverse to that from **mi* to **Ni* in the 1SG. It is, of course, also possible that other factors contributed to the shift in shape, for example, contact interference from unrelated languages with 2(PL) forms in initial *m* (cf. Güldemann 2017) or paradigm-internal pressure (see below).

There is another Bantu-internal piece of evidence that its 2PL forms in *n* reflect the proto-state. The proto-language can be assumed to have possessed another 2PL form that is semantically and formally related to a marker **nU* for subject and object, namely the post-final verb suffix **(n)i* encoding plural addressee (cf. Meeussen 1967: 111; Schadeberg 1978). This form is another likely cognate of the old Benue-Kwa pronoun **nV^{close}*. Post-final **(n)i* may in fact be much older in Bantu as a bound affix than pre-stem **nU-*, as there are various non-Bantu reflexes of the former attested in the same environments as in Narrow Bantu, as shown in (13) for Tikar and in (14) for Ekpeye.

(13) Tikar [Bantoid, Benue-Kwa, Niger-Congo] (Stanley 1991: 58, 60)

- a. *wu-ê-nì*
kill-IRR-PL.AD
'Kill (him)!'
 - b. *bwi' wu-è-nì*
1PL kill-IRR-PL.AD
'Let us kill (him)!'

(14) Ekpeye [Igboid, Benue-Kwa, Niger-Congo] (Clark 1972: 103)

- a. i. *à-kà*
1PL-say
'We (excl.) said ...'
- ii. *à-kà-nì*
1PL-say-PL.AD
'We (incl.) [= we+you] said ...'

⁷It is significant that the original form is recurrently found in languages that are commonly classified with eastern Bantu languages (see Grollemund et al. 2015) as this may be a linguistic reflex of the previous presence of western Bantu in areas that are genealogically eastern today.

- b. i. *í-kà*
 2SG-say
 ‘You said ...’
- ii. *í-kà-ní*
 2SG-say-PL.AD
 ‘You people said ...’

I turn now to the less problematic picture for the 2SG marking. The subject forms do not require much discussion, as they are cognate with the old Benue-Kwa form. Hence, only arguably deviant object markers with an initial posterior consonant need to be considered. Object forms in Babaev’s data where the securely inherited back vowel segment is preceded by a consonant, namely a voiced velar fricative, first turn up sporadically in zone B. The voiceless velar plosive as reconstructed for Proto-Bantu **kU* only starts to unambiguously occur in languages of zone C. It is possible to view the overall variation as reflecting a sound change **k > ɣ > ∅* (cf. Pacchiarotti & Bostoen 2020 for this diachronic sound shift in West-Coastal Bantu). Nonetheless, I think that at this stage it is still open season to test the relevance of a presumably earlier, reverse fortition scenario of **∅ > ɣ > k* (see below for a possible motivation). Given that the form without an initial consonant is the older form in the higher-order groups, I propose to explain the Bantu form in *k* also as a Bantu-internal innovation. For the record, 2SG subject prefixes in *k-* are unlikely to be related to the innovative object prefix. In particular, the recurrent *ku-*form in zones E and G (but also in other areas) derives from the fusion of a pre-initial *ka-*prefix with the inherited subject prefix **U* (cf. Güldemann 1996: §4.5.3 for some discussion).

I conclude the discussion with a short evaluation of the 1PL markers. Forms with an inherited *t-* (or its other reflexes with such initial consonants as *r, l, d, s, z, c, h*) clearly predominate over all other forms, for subjects 128 vs. 25 and for objects 31 vs. 12 attestations. Since the more frequent vowel quality is back rather than front (for subjects 90 vs. 38 and for objects 25 vs. 6), the most likely Proto-Bantu form is indeed **tU*, as previously proposed (see Table 3).

3 Towards a revised reconstruction

I have assembled empirical comparative evidence and cross-linguistic arguments that the received Proto-Bantu reconstruction of a full-fledged and universal paradigm of bound argument cross-reference on verbs is not supported by the available data from in- and outside the family. My revised proposal for speech-act

participants involves two principal differences to the traditional approach. First, there is only sufficient evidence for a *bound* prefix in the 1SG, which was presumably restricted to specific contexts, while the principal marking of predicate arguments was by means of more independent forms that are directly related to those of the general pronoun paradigm (see Bantu non-verbal in Table 4). Second, three of the eight traditional Bantu reconstructions, namely 2SG object **kU* and 2PL subject and object **mU*, are not necessarily wrong as such but should not be ascribed to the proto-language of traditionally conceived Narrow Bantu, which was still characterised largely by clausal argument cross-reference of the Benue-Kwa type.

My partly new hypothetical proto-forms are summarised in Table 11, occurring before the arrows. As pointed out above, forms close to my reconstructions are not only found in Benue-Kwa but also in languages conventionally subsumed under Bantu. In Babaev’s (2008) survey, they are reported in zone A in 10 out of 17 languages for the 1SG subject, 5 out of 7 for the 1SG object, in 9 out of 16 for the 2SG subject, 2 out of 6 for the 2SG object, 7 out of 14 for the 1PL subject, and 6 out of 18 for the 2PL subject. In zone B, the forms I reconstruct for the 1SG and 2SG subject turn up in 4 and 5 of 13 languages, respectively.

Table 11: Revised reconstruction of argument indexing in Proto-Bantu predicates

Person	Singular		Plural		Role	
1st	<i>*mi/*Ni-</i>	>	<i>*N(i)-</i>	<i>*tU</i>	>	<i>*tU-</i> Subject
			<i>*N(i)-</i>			<i>*tU-</i> Object
2nd	<i>*(B)U</i>	>	<i>*U-</i>	<i>*nU</i>	>	<i>*mU-</i> Subject
			<i>*kU-</i>			<i>*mU-</i> Object

The three bolded items in Table 11 after the arrows are innovative forms of Bantu in spite of their frequency across the family today. I specifically propose that they emerged in tandem with the development of bound cross-reference marking in a morphologically compact predicate. This is supported by plausible motivations for the concrete shape of the two new forms. The 2SG object form **kU-* phonologically enhances the pre-radical object slot as part of the macrostem. That is, compared to the inherited weak form starting in a vowel (or glide), the stronger onset of a velar plosive seals, so-to-speak, this morphological domain off from the emerging pre-stem prefix complex. For the record, this idea also applies to the equally innovative consonant-initial object form **mU-* for 3SG human (see

Table 4 above). The other new form in Table 11, 2PL **mU*- replacing inherited **nU*, can be argued to strengthen the paradigm-internal contrast to the already fused 1SG *(-)*Ni*-, whose consonant is similar and whose distinctive vowel is recurrently lost.

I assume that pronouns referring to verb arguments were still largely independent morphemes, as in (15a), but in certain contexts may have been proclitic to certain hosts like predicate operators within STAMP morphemes, as in (15b) and (15c), or even to verb stems in the simplest form without preceding predicate operators, as in (15d). These patterns may have occurred in combination, as in (15e). While this must be investigated in more detail, Proto-Bantu possibly also possessed predicate patterns where morphemes for object cross-reference and nominalisation attached to the verb, as in (15f) and (15g). All configurations in (15) are, however, split predicates and thus exclude the previously proposed Proto-Bantu reconstruction of the compact highly agglutinative type in Table 1 above.

(15) Major morphosyntactic patterns of predicates in Proto-Bantu

- a. * SBJ OBJ STEM
- b. * [SBJ=TAMP] STEM
- c. * [SBJ=TAMP] OBJ STEM
- d. * [SBJ=STEM]
- e. * [SBJ=TAMP] [SBJ=STEM]
- f. * [SBJ=TAMP] [OBJ=STEM]
- g. * [SBJ=TAMP] [INF-STEM]

My alternative reconstruction brings the profile of Proto-Bantu not only in line with common patterns found in Benue-Kwa (cf. also the discussion in Güldemann 2011a; 2013), but also with the overall pronoun system in Bantu itself. Proto-Bantu would have possessed a paradigm still involving relatively homogeneous pronoun forms in subject, object, possessor, and independent-emphatic contexts. This can be seen from a comparison with the available pronominal reconstructions in Table 4 for forms other than for subject and object cross-reference: they are effectively the same as those in Table 11.

The finding that the forms I consider as innovations occur already in Bantu languages of the north-west is not necessarily evidence for their existence in Proto-Bantu. The genealogical classification of the languages as well as contact-induced changes in this highly diverse area has not been determined conclusively, which

restricts the precise historical assessment of such language-specific forms. According to a reviewer, one could view the coexistence of multiple forms, those inherited from Benue-Kwa and those unique to Bantu, as a reflex of archaic heterogeneity in Proto-Bantu that was simplified later in most of the family. However, this begs the question when/where this variation emerged before the Proto-Bantu stage. As far as I can tell, the heterogeneity of forms at issue here exists *inside* Narrow Bantu rather than in a higher-order group like Bantoid and thus is better explained Bantu-internally.

I think that the present proposal advances the historical reconstruction of Bantu, not because it presents a set of conclusive, fully specified proto-forms assigned to specific positions in a phylogenetic family history, but rather because it contributes to what Bostoen (2019: 325) refers to as “new visions on what is archaic and innovative, especially in Bantu grammar, [that] may also lead to new ideas on internal Bantu classification.” The challenges of a thorough historical-comparative evaluation of the prominent pre-stem verb morphology of Bantu only start to emerge with my alternative hypothesis. If the traditional Bantu reconstruction of a compact predicate involving bound argument marking is, as I argue, a family-internal innovation, the central problems are now where, when, and how it took shape, and related to this, to what extent the individual markers differing according to such features as person, number, and semantic role arose in a package or separately.

It is safe to conclude that, *vis-à-vis* the original forms, a separate prefix or proclitic for the verbal indexation of a 1SG argument has quite a long history, even preceding the Proto-Bantu stage. In view of this, as well as some general cross-linguistic findings, there is no strong case for assuming that all original pronouns in Table 11 changed their morphosyntactic status and shape simultaneously, or in other words, that the full cross-reference paradigm as reconstructed traditionally is the result of a single event of language change. Morphological fusion can be fast under appropriate conditions and can occur several times independently. It is also unlikely that such a full paradigm was relevant from the beginning in all possible predicate contexts. Rather, the morphosyntactic diversity of predicate types entertained under (4c), (5), and (15) persisted, if to a lesser extent, throughout Bantu history, and certain sub-types are constantly re-emerging even today. In line with Anderson (2011; 2012; 2016) and as shown in (6) and (7) above, simple concatenations of a subject marker and an auxiliary in STAMP morphemes in particular have always been an important intermediate step to the morphologically complex verb forms heretofore thought to be as old as Proto-Bantu.

Acknowledgements

This research was presented previously in various versions on the following occasions, and I thank the respective audiences for valuable comments and suggestions: Fourth International Conference on Bantu Languages (B4ntu), held at the Humboldt University of Berlin and the Leibniz-Centre General Linguistics (ZAS), Berlin, 7–9 April 2011; International Symposium “Paradigm change in historical reconstruction: The Transeurasian languages and beyond”, held at the Johannes Gutenberg University Mainz, 7–8 March 2013; “Work in Progress” Series at the Linguistics Department, Max Planck Institute for Evolutionary Anthropology in Leipzig, 19 March 2013; Invited talk at the Linguistics Department, University of California at Berkeley, 29 March 2016; International Conference “Reconstructing Proto-Bantu Grammar”, held at Ghent University, 19–23 November 2018. Thanks are also due to the editors of this volume for their many helpful comments and criticisms that helped me make the contribution hopefully more accessible and coherent.

Abbreviations

AD	addressee	PL	plural
AUX	auxiliary	POT	potential
COP	copulative	PROG	progressive
E	exclusive	PROX	proximal
FV	default final vowel	PRS	present
H	human	PST	past
HAB	habitual	PURP	purposive
INF	infinitive	REFL	reflexive
IPFV	imperfective	SBJ	subject
IRR	irrealis	SG	singular
NEG	negative	SIM	simultaneous
NPST	non-past		
OBJ	object		

Arabic number numbers not followed by SG/PL indicate noun classes

S/A subject/agent (as semantic role)

(S)TAM(P) (subject)/tense/aspect/modality/(polarity)

References

- Anderson, Gregory D. S. 2011. Auxiliary verb constructions in the languages of Africa. *Studies in African Linguistics* 40(1-2). 1–409.
- Anderson, Gregory D. S. 2012. S/TAM/P morphs (portmanteau subject/TAM/polarity morphs) in Niger-Congo languages. Paper presented at the International Congress “Towards Proto-Niger-Congo: Comparison and Reconstruction”, Paris, 18–21 September 2012.
- Anderson, Gregory D. S. 2015. STAMP morphs in Central Sudanic languages. In Angelika Mietzner & Anne Storch (eds.), *Nilo-Saharan – Models and descriptions* (Nilo-Saharan Linguistic Analyses and Documentation 28), 151–167. Cologne: Rüdiger Köppe.
- Anderson, Gregory D. S. 2016. STAMP morphs in the Macro-Sudan Belt. In Doris L. Payne, Sara Pacchiarotti & Mokaya Bosire (eds.), *Diversity in African languages. selected papers from the 46th Annual Conference on African Linguistics* (Contemporary African Linguistics 1), 513–539. Berlin: Language Science Press. DOI: 10.17169/langsci.b121.497.
- Babaev, Kirill V. 2008. Reconstructing Benue-Congo person marking I: Proto-Bantoid. *Journal of West African Languages* 35(1–2). 131–183.
- Bastin, Yvonne. 2006. Un PV *ɪ*- à la première personne du singulier en bantou. *Africana Linguistica* 12. 25–36.
- Blackings, Mairi John & Nigel Fabb. 2003. *A grammar of Ma'di. Part 1* (Mouton Grammar Library 32). Berlin: Mouton de Gruyter.
- Blasi, Damián E., Søren Wichmann, Harald Hammarström, Peter F. Stadler & Morten H. Christiansen. 2016. Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences* 113(39). 10818–10823.
- Bostoen, Koen. 2019. Reconstructing Proto-Bantu. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 308–334. Milton Park, Abingdon: Routledge.
- Bostoen, Koen & Léon Mundeke. 2011. Passiveness and inversion in Mbuun (Bantu B87, DRC). *Studies in Language* 35(1). 72–111.
- Breen, John G. 1981. Margany and Gunya. In Robert M. W. Dixon & Barry J. Blake (eds.), *Handbook of Australian languages*, vol. 2, 275–393. Amsterdam: John Benjamins. DOI: 10.1075/z.hal2.08bre.
- Clark, David J. 1972. A four-term person system and its ramifications. *Studies in African Linguistics* 3(1). 97–105.

- Crane, Thera M., Larry M. Hyman & Simon Nsielanga Tukumu. 2011. *A grammar of Nzadi [B.865]: A Bantu language of the Democratic Republic of Congo* (University of California Publications in Linguistics 147). Berkeley, CA: University of California Press.
- Creissels, Denis. 2000. Typology. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*, 231–258. Cambridge: Cambridge University Press.
- Devos, Maud & Daniel Van Olmen. 2013. Describing and explaining the variation of Bantu imperatives and prohibitives. *Studies in Language* 37(1). 1–57.
- Doke, Clement M. 1927. *Text book of Zulu grammar*. Johannesburg: The University of the Witwatersrand Press.
- Downing, Laura J. 2011. Bantu tone. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 2730–2753. Malden, MA: Wiley-Blackwell.
- Fortune, George. 1955. *An analytical grammar of Shona*. London: Longmans, Green & Co.
- Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's field trip. *Chicago Linguistic Society* 7. 394–415.
- Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In Charles N. Li (ed.), *Subject and topic*, 149–188. New York: Academic Press.
- Good, Jeff & Tom Güldemann. 2006. The Bantu verbal prefixes and S-Aux-O-V order in Benue-Congo. Paper presented at the International conference on “Bantu Grammar Description and Theory”, School of Oriental and African Studies, University of London, 20–22 April 2006.
- Green, Margaret M. & G. Egemba Igwe. 1963. *A descriptive grammar of Igbo* (Veröffentlichungen des Instituts für Orientforschung der Deutschen Akademie der Wissenschaften zu Berlin 53). Berlin: Akademie-Verlag.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Güldemann, Tom. 1996. *Verbalmorphologie und Nebenprädikationen im Bantu: Eine Studie zur funktional motivierten Genese eines konjugationalen Subsystems* (Bochum-Essener Beiträge zur Sprachwandelforschung 27). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Güldemann, Tom. 2003. Grammaticalization. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 182–194. London: Routledge.

- Güldemann, Tom. 2007. Preverbal objects and information structure in Benue-Congo. In Enoch O. Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic* (Trends in Linguistics, Studies and Monographs 191), 83–111. Berlin: Mouton de Gruyter.
- Güldemann, Tom. 2008. The Macro-Sudan Belt: Towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151–185. Cambridge: Cambridge University Press.
- Güldemann, Tom. 2010. Sprachraum and geography: Linguistic macro-areas in Africa. In Alfred Lameli, Roland Kehrein & Stefan Rabanus (eds.), *Language and space: An international handbook of linguistic variation. Volume 2: Language mapping* (Handbücher zur Sprach- und Kommunikationswissenschaft 30.2), 561–585, Maps 2901–2914. Berlin: Mouton de Gruyter.
- Güldemann, Tom. 2011a. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas: With special reference to Africa* (Tokyo University of Foreign Studies, Studies in Linguistics 2), 109–141. Amsterdam: John Benjamins.
- Güldemann, Tom. 2011b. Towards early Bantu by reconstructing a fragment across Benue-Congo: 1st-person singular *mI. Paper presented at the Fourth International Conference on Bantu Languages (B4ntu), Humboldt University of Berlin and the Leibniz-Centre General Linguistics (ZAS), Berlin, 7–9 April 2011.
- Güldemann, Tom. 2013. Morphological fusion in Niger Congo and Bantu: Addressing a historical-comparative problem. Paper presented at the International Symposium “Paradigm change in historical reconstruction. The Transeurasian languages and beyond”, Johannes Gutenberg University Mainz, 7–8 March 2013.
- Güldemann, Tom. 2017. A shared pronominal canon in the Macro-Sudan Belt: Typological, areal and genealogical perspectives. In Raija Kramer & Roland Kießling (eds.), *Mechthildian approaches to Afrikanistik: Advances in language based research in Africa. Festschrift für Mechthild Reh*, 101–146. Cologne: Rüdiger Köppe.
- Güldemann, Tom. 2018. Areal linguistics beyond contact, and linguistic areas in Afrabia. In Tom Güldemann (ed.), *The languages and linguistics of Africa* (The World of Linguistics 11), 448–545. Berlin: Mouton de Gruyter.

- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Guthrie, Malcolm. 1967–71. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages (4 volumes)*. Farnborough: Gregg International.
- Heath, Teresa. 2003. Makaa (A83). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 335–348. London: Routledge.
- Henrici, Alick. 1973. Numerical classification of Bantu languages. *African Language Studies* 14. 82–104.
- Hyman, Larry M. 2007. Reconstructing the Proto-Bantu verbal unit: Internal evidence. In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 201–211. London: Department of Linguistics, School of Oriental & African Studies.
- Hyman, Larry M. 2008. Directional asymmetries in the morphology and phonology of words, with special reference to Bantu. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 46(2). 309–350.
- Hyman, Larry M. 2010. Focus marking in Aghem: Syntax or semantics. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure: A documentation of its diversity across Africa*, 95–115. Amsterdam: John Benjamins.
- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.
- Kamba Muzenga, Jean-Georges. 2003. *Substitutifs et possessifs en bantou* (Société d'études linguistiques et anthropologiques de France 401). Louvain: Peeters.
- Kisseberth, Charles W. & David Odden. 2003. Tone. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 59–70. London: Routledge.
- Maho, Jouni F. 2009. *NUGL online: The online version of the new updated Guthrie list, a referential classification of the Bantu languages*. https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf.
- Marlo, Michael R. 2013. Verb tone in Bantu languages: Micro-typological patterns and research methods. *Africana Linguistica* 19. 137–234.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meinhof, Carl. 1938. Die Entstehung der Bantusprachen. *Zeitschrift für Ethnologie* 70(3–5). 144–152.

- Miehe, Gudrun. 2004. Les pronoms personnels dans les langues gur. In Dymitr Ibrizimow & Guillaume Segerer (eds.), *Systèmes de marques personnelles en Afrique* (Afrique et Langage 8), 97–128. Leuven: Peeters.
- Mithun, Marianne. 1991. The development of bound pronominal paradigms. In Winfred P. Lehmann & Helen-Jo Jakusz Hewitt (eds.), *Language typology 1988: Typological models in the service of reconstruction* (Current Issues in Linguistic Theory 81), 85–104. Amsterdam: John Benjamins.
- Nichols, Johanna & David A. Peterson. 1996. The Amerind personal pronouns. *Language* 72(2). 336–371.
- Njantcho Kouagang, Elisabeth. 2018. *A grammar of Kwakum*. Paris: Institut National des Langues et Civilisations Orientales. (Doctoral dissertation).
- Nurse, Derek. 2007. Did the Proto-Bantu verb have a synthetic or an analytic structure? In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 239–256. London: Department of Linguistics, School of Oriental & African Studies.
- Nurse, Derek. 2008. *Tense and aspect in Bantu* (Oxford Linguistics). New York, NY: Oxford University Press.
- Nurse, Derek & Gérard Philippson. 2003. Introduction. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 1–12. London: Routledge.
- Nurse, Derek & Gérard Philippson. 2006. Common tense-aspect markers in Bantu. *Journal of African Languages and Linguistics* 27(2). 155–196.
- Nurse, Derek, Sarah Rose & John Hewson. 2016. *Tense and aspect in Niger-Congo* (Documents on Social Sciences and Humanities). Tervuren: Royal Museum for Central Africa. <https://www.africamuseum.be/sites/default/files/media/docs/research/publications/rmca/online/documents-social-sciences-humanities/tense-aspect-niger-congo.pdf>.
- Odden, David & Lee S. Bickmore. 2014. Melodic tone in Bantu: Overview. *Africana Linguistica* 20. 3–13.
- Pacchiarotti, Sara & Koen Bostoen. 2020. The Proto-West-Coastal Bantu velar merger. *Africana Linguistica* 26. 139–195.
- Polak, Louise. 1986. Les infixes (“prefixes objets”) du bantou et leur reconstruction. *Africana Linguistica* 10. 365–421.
- Redden, James E. 1979. *A descriptive grammar of Ewondo* (Occasional Papers on Linguistics 4). Carbondale, IL: Department of Linguistics, Southern Illinois University.
- Schadeberg, Thilo C. 1978. Partizipanten am Dialog: Eine grammatische Kategorie, nachgewiesen am Bantu. In Herrmann Jungraithmayr (ed.), *Struktur und Wandel afrikanischer Sprachen (Vorträge vom XX. Deutschen Orientalistentag,*

- Erlangen 1977*) (Marburger Studien zur Afrika- und Asienkunde A17), 151–159. Berlin: Dietrich Reimer.
- Schadeberg, Thilo C. 2003. Historical linguistics. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 143–163. London: Routledge.
- Stanley, Carol. 1991. *Description phonologique et morpho-syntaxique de la langue tikar (parlée au Cameroun)*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Stappers, Leo. 1986. *Substitutiv und Possessiv im Bantu* (Mainzer Afrika-Studien 8). Berlin: Dietrich Reimer.
- Stewart, John M. 1976. *Towards Volta-Congo reconstruction: A comparative study of some languages of Black-Africa* (Inaugural Lecture). Leiden: Leiden University Press.
- Thomas, Elaine. 1969. *A grammatical description of the Engenni language*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Williamson, Kay & Roger M. Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*, 11–42. Cambridge: Cambridge University Press.
- Woolford, Ellen. 2000. Agreement in disguise. In Vicki M. Carstens & Frederick Parkinson (eds.), *Advances in African linguistics: Papers presented at the 28th Annual Conference on African linguistics, held July 18–22, 1997, Cornell University* (Trends in African Linguistics 4), 103–117. Trenton, NJ: Africa World Press.

Chapter 10

On reconstructing the Proto-Bantu object marking system

Benji Wald

University of California at Berkeley

This chapter critically examines the divergent hypotheses about the Proto-Bantu (PB) object marking system proposed by Meeussen (1967) and Polak (1986). It then builds on their insights with additional data and details of analysis and develops a new reconstruction of PB object marking, including its place in a larger system of topicality marking also involving the subject marker.

1 Introduction

Bantu object marking consists of a set of single morphemes, i.e. object markers (OMs), also called object prefixes or infixes in some studies, one or more of which immediately precede the verb root, and index objects of the transitive verb. The OM as a grammatical category contrasts with the independent or free pronoun (PRO), called “substitutive” by Meeussen (1967: 105). Unlike the OM, the PRO may index any of a predicate’s nominal arguments, including its subject. It has the basic syntactic occurrence privileges of other nominals. Also unlike the OM, the PRO tends to be polymorphemic, by reduplication and/or suffixation of a deictic marker, whose shape commonly reflects the Proto-Bantu (PB) vowels **e/*o*, as in Swahili G42d (*ye-*)*ye* < **yu-e* [class 1 - e] ‘singular animate referent’, *wao* < **ba-ba-o* [class 2 - class 2 - o] ‘plural animate referent’.¹ Finally, PRO occurs in all Bantu languages, OM in most but not all of them. One key problem is to establish what the historical relationship between OM and PRO is. For present purposes,

¹Swahili differs from many Bantu languages in using class 1/2 markers for indexing not only humans but also animals (cf. Wald 1975).



the guiding questions are whether PB had an OM system, and if so, how it was organised.

Current Bantu OM systems are highly diversified. They range from languages with no OMs at all (NOMs) to languages in which a virtually unlimited number of OMs can be prefixed to a single verb stem, i.e. multi-OM systems (MOMs). In between, there are those languages with a system allowing a single OM per verb stem (SOMs) and those with what I call partial OM systems in which some objects can be indexed by OMs but others cannot, depending on either their inherent topicality (animacy status) or thematic/semantic role (TR), or both these factors. As I argue in this chapter, OMs are the most complex component of a topic marking system that also includes the subject marker (SM).

This chapter is organised as follows: §2 reviews the divergent hypotheses regarding the PB OM system proposed by Meeussen (1967) and Polak (1986); §3 identifies and discusses the more detailed factors involved in examining the variation within and across Bantu OM systems in order to inform decisions about the nature of the PB system, and about the directions of change from the PB system to the various current-day Bantu systems; §4 examines diversity within the three major types of OM, i.e. the MOM, SOM and NOM systems; §5 discusses alternative historical hypotheses about the PB OM system and the directions of change from PB to the current-day diversity in those alternative hypotheses; §6 makes concluding remarks about what currently appears to be the most promising PB reconstruction, and indicates a number of issues that require further research to either support or cast doubt on that reconstruction.

2 Conflicting hypotheses on the PB OM system

The few previous reconstructions of the PB OM system differ on the relative chronology of SOM and MOM systems. Meeussen (1967: 110) proposed that: “In a verb form there may be more than one infix [= OM], the nearest to the radical corresponding to the object nearest to the verb in comparable constructions (or: the last infix corresponds to the first object) [...]”. It is not completely clear what Meeussen had in mind here, since he does not declare a fixed order for either the OM or corresponding postverbal object sequence. However, it fits the description of some current MOM systems, such as the one in Ganda JE15 and some varieties of Tswana S31, where OMs and postverbal objects are fixed in a “mirror image” relationship (cf. Bearth 2003: 127) according to the grammatical relation (GR). $OM_{[DO]}-OM_{[IO]}-V$ corresponds to $V...NP_{[IO]}-NP_{[DO]}$. One or the other of these orders is common in a wide area of the East Bantu interior

from north to south about which Meeussen was especially knowledgeable. However, as discussed in §4, there are other languages in that area, such as Rwanda JD61, and elsewhere, such as Kwanyama R21, where neither the order of OMs nor that of postverbal object NPs is fixed or conditioned by GR. Meeussen's reconstruction was frankly programmatic. He did not refer to any current-day Bantu languages for evidence. Although his reconstruction is supported by synchronic attestations covering a variably large part of the Bantu domain, they fall short of encompassing the entire Narrow Bantu area. Therefore, Meeussen's reconstruction of PB – understood as representing at minimum the period of unity of all Narrow Bantu languages – has remained problematic.

Polak (1986) pursued Meeussen's program by examining a widespread sample of Narrow Bantu languages. She accepted the notion of a PB OM category (p. 374), but implicitly rejected Meeussen's MOM hypothesis on the basis of the relative rarity of MOM languages in her sample (e.g. pp. 371, 374, 403ff). Thus, she favoured the notion that the current languages lacking the OM category, i.e. many north-western Bantu languages of Guthrie's zones A and B and adjacent areas (zones C, D and H), lost the PB OM. Evidence presented in this chapter questions Polak's assumption of rarity of MOM systems across Bantu, both on the basis of data beyond her sample and, to a lesser extent, failure to account for some full MOM languages within her own sample, e.g. Tswana in the south-eastern part of the Bantu domain and Bangi C32 in the deep northern interior.

Significantly for the MOM issue, one of Polak's most historically relevant findings was that there is an intermediate category of partial MOM systems spread across a large part of the interior Narrow Bantu area (zones C-F, H, L-M): languages, such as Rimi F32 in (1), which allow a sequence of two OMs provided that the second one is what she calls "monophonic" ("monophone" in French), i.e. CV-N- or CV-i- with N- being OM_{1SG} and *i-* being the reflexive (Polak 1986: 403ff). She hypothesises that it is an innovative constraint allowing a double OM sequence to occur only if it does not violate a principle that only a single syllable is exclusively reserved for object marking.

(1) Rimi F32 (adapted from Woolford 2000: 113)

- a. *a-limu va-a-mu-N-tum-i-a*
 2-teacher SM₂-PST-OM₁-OM_{1SG}-send-APPL-FV
 'The teachers sent him to me.'
- b. ** *a-limu va-a-mu-ku-tum-i-a*
 2-teacher SM₂-PST-OM₁-OM_{2SG}-send-APPL-FV
 intended: 'The teachers sent him to you.'

As a non-syllabic homorganic nasal, the Rimi OM_{1SG} *N-* attaches to the initial consonant of the following verb root allowing the syllabic OM₁ *mu-* to occupy the only slot reserved for an OM. In point of fact, OM_{1SG} *N-* fuses in (1) with the initial consonant of the verb root to form the complex pre-nasalised onset *nt-* of the next syllable. The homorganic nasal *N-* is a common form of OM_{1SG} in much of Narrow Bantu, whether or not the monophonic OM principle is in effect. Where it is in effect, it establishes an intermediate category between MOM and SOM languages. Polak (1986: 404) conjectures that full MOM systems arose by loss of the monophonic OM constraint. Research reported in the present chapter shows further distinctions among partial MOM systems in which there are other constraints that limit OM sequences on the basis of the relative inherent topicality status (e.g. animacy) and/or transitivity status (i.e. the grammatical or thematic role) of the indexed objects.

Polak (1986: 374) establishes two other categories relevant to the present chapter: (1) languages without OMs (i.e. “*langues sans infixes*”), which I call NOM languages and which are most concentrated in the north-western part (especially zone A), but occur less frequently in adjacent areas (i.e. zones B-D); (2) languages with incomplete series of OMs (i.e. “*séries incomplètes d’infixes*”), an intermediate category between SOM and NOM systems, which I call partial SOM systems. The partial SOM category is also subject to much cross-Bantu diversity, based largely on the inherent topicality of the indexed object.

In sum, Polak’s categorisation of OM systems forms a continuum that ranges from NOM through SOM to MOM with intermediate/partial systems in-between these three major types. Much of the problem of reconstruction lies in determining the direction of change for the numerous points between the polar NOM and MOM types.

3 Factors of variation in Bantu OM systems

This section discusses a number of recurrent factors involved in differentiating Bantu OM systems.

3.1 Number of OMs allowed in sequence

OMs allowed in a sequence range from none to an indeterminate number, a continuum segmentable into three major types: (1) MOM, (2) SOM, and (3) NOM. There is variation in the tolerance of speakers of MOM languages to allow beyond 2 or 3 OMs in sequence, either on a community or idiosyncratic basis (cf.

Marlo 2015: 1). The more pressing historical issue is whether the earliest version of a MOM system evolved before or after the earliest version of the SOM system.

3.2 Contextual topicality

Contextual topicality (ConTop) is a discourse notion. As I use the term, a topic is an old, given or deducible referent, usually first introduced into a discourse by an NP or pronoun, and marked as topical by the SM or OM in relevant subsequent clauses. Examples (2) and (3) below represent the ConTop function of the OM in two Bantu languages of non-adjacent zones.

- (2) Nkore JE13 (Asimwe 2014: 159)

omu-kazi, ti-tw-a-mu-bugana

1-woman NEG-SM_{1PL}-PST-OM₁-meet

NP_{OBJ}... SM-TM-OM-V

‘The woman, we didn’t meet her.’

- (3) Dzamba C322 (Bokamba 1971: 229)

imu-nkanda, a-mu-kóm-el-aki omo-konzi

3-letter SM₁-OM₃-write-APPL-PST 1-chief

NP_{OBJ}... SM-OM-V

‘The letter, he wrote it to the chief.’

This construction, common to all Bantu OM languages, is often appropriately called “topicalisation”. The pattern consists of an object of any information status, functioning as a topic about which the following clause provides new information, both in the event/state denoted by the verb and the relation of concurrent verbal arguments to each other, as if to answer the question ‘What about the topic?’.

Bantu languages vary in the obligatoriness of the OM in this syntactic/discourse context. In most of Narrow Bantu, the OM obligatorily indexes a topicalised human object. More problematic is the OM indexing of an inanimate object. Quite generally in Niger-Congo, inanimate objects may be omitted as understood in the larger discourse context instead of being referred to anaphorically, regardless of their definiteness. As a decontextualised construct in sentence grammar, an inanimate topicalised object may strongly favour OM reference, as in Dzamba (3) above, but that favourability might be pragmatic in nature rather than grammatically obligatory. Obligatory OM indexing of topicalised objects, regardless of their animacy, seems to be strongest in the south-eastern part of

the Bantu domain, for instance in Nguni S40. Meanwhile, the extreme polarisation between compelling OM indexing of human objects but highly disfavoured OM indexing of inanimate objects is localised to the central east coast and adjacent interior, for example Matuumbi P13 (Odden 1984), Matengo N13 (Yoneda 2011). The more extreme absence of OMs indexing objects of the inanimate noun classes occurs locally in adjacent Makhuwa P31 (Stucky 1985; Katupha 1991; van der Wal 2009). For discussion of how this trend affected Swahili, see Wald (2001).

3.3 Inherent topicality

In contrast to ConTop, Inherent Topicality (InTop) is a feature of the NP itself, independent of its discourse context. It is a relative ranking of topics that has a diverse array of influences across the OM area. Often referred to as the person-animacy hierarchy, a comprehensive arrangement of the relative ranking is represented in (4) below.

- (4) reflexive > 1SG > 2SG > human (> animate) > inanimate

Certain aspects of these relative rankings vary across Bantu. The reflexive (REFL) OM is high on the scale because in most contexts it indexes the subject of the clause in a second role as object, where the subject of a clause is higher ranked than any object, indeed often the only topic in the clause (e.g. with intransitive verbs). Nevertheless, among MOM languages, there is some variation in the relative positions of the OM_{REFL} and OM_{1SG} such that some have fixed OM_{1SG} - OM_{REFL} order and others OM_{REFL} - OM_{1SG} order, regardless of their TRs (cf. Marlo 2014: 91–93). Similarly, in some areas, interpersonals (first and second persons) are not distinguished for relative InTop, because both are equally given as discourse participants, e.g. in Shambaa G23 (Riedel 2009: 140).

Where InTop plays a role, interpersonals are invariably ranked higher than other referents, and humans (or personified animals) higher than inanimates. Where personal plural objects are currently ranked differentially from singulars, the singulars outrank the plurals, but evidence of such ranking varies in Bantu.

Among partial SOM languages, the case of Makhuwa represents a system in which inanimates (unless in the typically human classes 1/2) lack OMs. Only objects of higher InTop can be indexed. Polak (1986: 375) lists a scattering of languages of this type across Narrow Bantu, but with the densest distribution in the north-western part and vicinity (zones A-D).

3.4 Grammatical relation

The Bantu OM systems can be viewed as the most complex component of a topicality marking system that also includes the SM, obligatorily indexing the subject of the clause, often the only topic in the clause. A scheme of a Bantu minimal finite clause is as in (5).

- (5) SM-(TM)-(AUX(#INF))-(OM)-V

The relative ConTop of the SM and OM is indicated by their relative positions in the verb complex such that ConTop declines from left to right, i.e. SM > OM. The SM referent is determined by the lexical verb. As typical of typologically nominative/accusative languages, the subject role is usually highly active or sentient, and indexes the initiator of the event represented by the verb. OMs index additional arguments of transitive verbs.

The number of objects that a verb allows is either lexically determined or associated with one or more valence-raising extensions suffixed to the verb. Extensional objects (EOs) are of two types in terms of GR: (1) causative object (CO); (2) applied object (AO). The CO of V-CAUS is the subject of the root verb, e.g. *they him cook-CAUS* “they made/let/helped him [CO] cook”. The AO does not alter the GRs of the subject and lexically allowed object/s to the root verb, but involves an additional argument in an additional role, e.g. *they him cook-APPL* “they cook for him [AO]”. In most languages, the two valence-raisers can both mark a single verb, increasing the number of objects, e.g. *they him them cook-CAUS-APPL* “they made him cook for them”. These languages vary for whether the CAUS and APPL are meaningfully ordered, or whether the order is fixed/templated regardless of meaning (cf. Hyman 2003b; Good 2005). In all cases the DO maintains its status as the DO of the root verb, e.g. *they him them it cook-CAUS-APPL* “they made him [CO] cook it [DO] for them [AO]”.

3.5 Thematic role

The TR is the semantic interpretation of the GR. Both CO and AO express a range of distinguishable TRs. This allows the same GR to appear more than once with a single verb, indexing objects with different TRs. MOM languages vary in their tolerance for this possibility, particularly in supporting an additional object role with an additional extension. Most widely reported are languages jointly marking both the TRs recipient and beneficiary in examples of double AO (e.g. Tswana, Rwanda), as in: *they it him them send-AO-AO* “they sent it to

him (recipient/dative) for them (beneficiary)”. The available data suggest that in such cases the order of two OMs is templated as recipient-beneficiary.²

TR is especially prominent in passivisation, where the SM maintains its object TR identity in terms of its grammatical interaction with concurrent objects. In some SOM systems which of two concurrent objects can passivise is constrained according to their TRs. Languages of this type are classified as asymmetric. Similarly, many MOM languages fix the double OM order according to TR or GR.

In general, the data suggest that the behaviour of concurrent objects need only be compared for GR, because most available examples are limited to a comparison between the DO and one concurrent object, either IO or EO. The double OM configuration consisting of the DO and a single EO/IO is worthy of special consideration because it is undoubtedly the most frequent multiple OM pattern in the discourse of any MOM language. The high relevance of discourse frequency in OM evolution is discussed in §5.3.

3.6 Time depth of PB

Following Meeussen (1967) and Polak (1986), I limit PB to the assumed period of unity of Narrow Bantu, i.e. those languages conventionally categorised as Bantu in the referential classification of Guthrie (1948; 1971). For my historical reconstruction I refer to the phylogeny of Grollemund et al. (2015), which represents relationships among present-day Bantu languages according to an expansion model of nine successive binary major nodes based on shared innovations in basic vocabulary (their Fig. 1). Their node 1 roughly corresponds to what I consider here to be PB, even though I do not consider Jarawan Bantu languages, which are subsumed with Narrow Bantu under node 1 in the phylogeny of Grollemund et al. (2015). Their node 0 also includes Grassfields Bantu languages, which constitute a branch parallel to all languages incorporated under node 1. The nodes subsequent to node 1, i.e. nodes 2 to 9, are geographically nested, proceeding in a southern direction from the north-west to the south-west with the final major node 9 encompassing the entire eastern Bantu area and part of what they consider to be the south-western Bantu area (i.e. Guthrie’s groups L20-40). In addition to indicating nodes in their phylogeny, Grollemund et al. (2015) also subdivide it into five major subgroups which have distinct colours: (1) North-Western Bantu (NWB), subdivided in NWB Cameroon (between nodes 1 and 2)

²Some languages allow nesting of COs, e.g. ... CO_y-CO_z-laugh-CAUS-CAUS “[X] made him [Y] make her [Z] laugh”, but many languages resist such complications in favour of a circumlocution.

and NWB Gabon (between nodes 2 and 5); (2) Central-Western Bantu (CWB) (between nodes 5 and 6); (3) West-Western Bantu (WWB) (between nodes 6 and 7) also known as West-Coastal Bantu (Vansina 1995; Bostoen et al. 2015; de Schryver et al. 2015; Pacchiarotti et al. 2019); (4) South-Western Bantu (SWB) (between nodes 6 and 9); (5) Eastern Bantu (EB) (under node 9 minus Guthrie's L20-40). As their correspondence to major nodes suggests, not all of the geographically labelled subgroups are discrete branches within the phylogeny. Only NWB Cameroon, CWB and WWB really are. All the others cover several distinct branches with SWB and EB actually forming one superclade subsumed under node 7 with many subclades successively branching off (cf. Pacchiarotti & Bostoen 2020: 156–157). Some of Guthrie's A80-90 languages subsumed under NWB Gabon are also spoken in southern Cameroon.

For convenience, apart from their Guthrie code (cf. Guthrie 1971; Maho 2009), I will label individual Bantu languages discussed in the remainder of this chapter by referring to both the major subgroup they belong to in the phylogeny of Grollemund et al. (2015), i.e. NWB Cameroon, NWB Gabon, CWB, WWB, SWB, and EB, and the numbered node under which they are directly subsumed, i.e. nodes 1-9. For example, Eton A71 is labelled "NWB Cameroon, node 1", Orungu B11b "NWB Gabon, node 3", Bangi C32 "CWB, node 5", Yaka H31 "WWB, node 6", Luba-Kasai L31a "SWB, node 9" and Rimi F32 "EB, node 9". To avoid clashes with the finer geographical distinctions I use discussing the distribution of certain types of OM systems, which I always designate with unabbreviated cardinal directions, I will systematically refer to the geographically labelled subgroups of Grollemund et al. (2015) with the abbreviations NWB, CWB, WWB, SWB and EB.

4 Types of OM systems according to significant factors constraining them

In this section we consider examples of diversity within the three major types of OM systems. The types are arranged by their grammatical properties. The discussion will arbitrarily start with the MOM end of the continuum, where there is maximum complexity, and proceed through SOM types to NOM types.

4.1 MOM systems

MOM types are distributed across most of the Bantu area, but more densely in some areas than others. They are most common in SWB languages branching off

at node 7 in the phylogeny of Grollemund et al. (2015), i.e. Guthrie's L10-20-H21a-H30-40, and also in interior EB languages, especially from the Great Lakes region (zone J). They occur less frequently in CWB and WWB, and only marginally in NWB. They are highly diversified. The primary distinction is between those systems that have free OM order and those that have fixed OM order, in the latter case determined by relative InTop or GR, or both.

4.1.1 Free MOM systems

Free MOM order is common in three separate areas: (1) the south-eastern Great Lakes area of the north-eastern interior (i.e. EB); (2) some varieties of Tswana in the south-eastern part of the Bantu domain (i.e. EB); (3) SWB including some varieties of Umbundu R11 in the north and Kwanyama R21 further south, as exemplified in (6).

(6) Kwanyama R21 “SWB, node 8” (Zimmermann & Hasheela 1998: 100)

- a. *een-gobe, o-nde-di-a-p-a*
10-cow PST-SM_{1SG}-OM₁₀-OM₆-give-FV
OM: IO-DO
'The cattle, I gave them it [water].'
- b. *om-eva, o-nde-a-di-p-a*
6-water PST-SM_{1SG}-OM₆-OM₁₀-give-FV
OM: IO-DO
'Water, I gave it to them [cattle].'

Zimmermann & Hasheela (1998: 100) appeal to topicalisation in their examples to distinguish the alternative orders, but state that “the initial nouns are usually omitted in speech, and are here given only for the sake of clarity”. Kwanyama exemplifies a system based on ConTop. The multiple OM order of ConTop corresponds to the SM...OM order so that relative ConTop among arguments declines from left to right. A peculiarity of the Kwanyama system, common in SWB but extremely rare elsewhere, is that the OM_{1SG} is not part of the system. Instead, a pronominal form is *encliticised* to the verb stem, i.e. *V...-nge*. A salient syntactic feature that Kwanyama OM_{1SG} shares with a wide range of other MOM systems is its fixed position. The difference is that the 1SG object reference obligatorily follows rather than immediately precedes the verb, as if to avoid being fixed in the monophonic pattern, as in Tswana “EB, node 9”, e.g. *go-i-N-kanya* [INF-OM_{REFL}-OM_{1SG}-trust] “to trust (self) to me” (Cole 1955: 234). In both the Kwanyama and Tswana systems, the fixed OM_{1SG} is a constraint on ConTop.

There is diversity of MOM types within Tswana itself. Cole (1955) and Creissels (2006) describe fixed orders in Tswana varieties such as Hurutshe and Kgatla. I observed free orders apart from the OM_{1SG} constraint in the Tswana variety Rolong S31a, as in (7).

- (7) Rolong S31a variety of Tswana “EB, node 9” (fieldwork B. Wald & Sheila Mmusi 1993)
- a. *o-mo-e/e-mo-hir-etse*
 $SM_1-OM_1-OM_9/OM_9-OM_1$ -hire-APPL.PFV
 OM: DO-AO/AO-DO or INAN-HUM/HUM-INAN
 ‘She hired him (driver) for it (car).’ or ‘She rented it for him.’
- b. *ke-a-ho-ba/ba-ho-tl-el-a*
 $SM_1-PRS.PROG-OM_{2SG}-OM_2/OM_2-OM_{2SG}$ -bring-APPL-FV
 OM: DO-AO/AO-DO or 2SG-CLASS.2/CLASS.2-2SG
 ‘I’m bringing them (people) for you.’ or ‘I’m bringing you for them.’

This multiple GR ambiguity would not occur in the varieties described by Cole and Creissels, which would have GR-conditioned order, i.e. DO-EO. Nevertheless, Cole describes an exception to GR order in the context of indexing a CO (object of CAUS). The MOM order involving a CO is not determined by GR but by InTop as NONHUM-HUM. As a result, examples like (8) are ambiguous for GR/TR.

- (8) Hurutshe variety of Tswana “EB, node 9” (adapted from Cole 1955: 431)
- ba-e-m-mola-is-itse*
 $SM_2-OM_9-OM_1$ -kill-CAUS-PFV
 OM: DO-CO/CO-DO
 ‘They let it (the dog) kill him/him kill it.’

The order is fixed as NONHUM-HUM according to InTop ranking, i.e. NONHUM < HUM. In that regard, Hurutshe is intermediate between free and fixed OM systems.

Free MOM systems also occur in the north-eastern part of the Bantu domain, more specifically in the south-eastern vicinity of the Great Lakes, widely separated from the free systems in southern Bantu discussed above. Ranero et al. (2013) describe a fully free order as in (9) for Kuria JE43.

- (9) Kuria JE43 “EB, node 9” (adapted from Ranero et al. 2013: example (12))
- n-a-a-mú-ké/ké-mú-háá-ye*
 $FOC-SM_1-PST-OM_1-OM_7/OM_7-OM_1$ -give-PFV
 ‘She gave it (toy) to him.’

Among the varieties of Bubi A31 “NWB Cameroon, node 1”, an insular NWB language, there seem to be some partially free MOM varieties. Bubi is internally diverse. It will be discussed separately in §4.4, in view of some of its apparently unique and instructive features.

4.1.2 Fixed MOM systems

Within fixed MOM languages, the primary distinction is in orientation, i.e. the direction of OM order. By far the most widely distributed orientation is ascending so that a human OM occurs to the right of a concurrent inanimate, i.e. NONHUM-HUM-V. Data for many languages are limited to cases where the IO or EO is human. Where a concurrent human DO is represented, it follows the same pattern as the inanimate DO. The opposite orientation is much rarer, i.e. fixed HUM-NONHUM-V, but occurs in widely separated areas, as discussed below in §4.1.2.2 and §4.4.

4.1.2.1 Ascending fixed MOM

This type is widely distributed outside of the north-western Bantu area. The Great Lakes region of the interior north-eastern part of the Bantu domain has a variety of subtypes. Rwanda represents a type where InTop is the primary ordering principle as in (10). In an appropriate discourse context, (10b) could also mean ‘He bought her for potatoes’.

(10) Rwanda JD61 “EB, node 9” (adapted from Ngoboka 2005: 62; Yokoyama 2016: 5)

- a. *y-a-mu-ku/**ku-mu-eretse*
SM₁-PST-OM₁-OM_{2SG}/**OM_{2SG}-OM₁-show.PFV
OM: 3SG-2SG/**2SG-3SG
‘He showed her to you/you to her.’
- b. *y-aa-bi-mu/**mu-bi-gur-i-ye*
SM₁-PST-OM₈-OM₁/**OM₁-OM₈-buy-APPL-PFV
OM: NONHUM-HUM/**HUM-NONHUM
‘He bought them (potatoes) for her.’

Only in the absence of an InTop differential is a GR order imposed as in (11)

- (11) Rwanda JD61 “EB, node 9” (Yokoyama 2016: 4)

y-a-ba-mw-eretse

SM₁-PST-OM₂-OM₁-show.PFV

OM: DO-IO/**IO-DO (INTOP: 3PL = 3SG)

‘He showed them to her/**her to them.’

The fixed GR order by which the DO is indexed first in (11) parallels the fixed InTop order in (10) by which the object of lower InTop is indexed first. In contrast to the GR ordering principle of double third-person humans, double nonhuman OM_s are freely ordered, as in (12), just like in complete free MOM systems.

- (12) Rwanda JD61 ‘EB, node 9’ (Zeller & Ngoboka 2015: 212)

a::-bi-yi/yi-bi-ha-ye

SM₁.PST-OM₈-OM₉/OM₉-OM₈-give-PFV

OM: DO-IO/IO-DO

‘He has given them (yams) to it (pig).’

Rwanda is more tolerant of numerous multiple objects than most reported MOM languages. The widely cited example in (13) shows an extensive InTop order corresponding to the order of extensions. In (13), ‘[there]’ refers to a locative OM preceding the OM representing the DO in the original example.

- (13) Rwanda JD61 “EB, node 9” (adapted from Marlo 2015: 4)

...ki-zi-ba-ku-n-som-eesh-eesh-er-er-...

...OM₇=DO-OM₁₀=CO₁-OM₂=CO₂-OM_{2SG}=AO₁-OM_{1SG}=AO₂-read-CAUS-CAUS-APPL-APPL

1 2 3 4 5 1 2 3 4 5

‘[She is also] making them (3) read it (1 = book) with them (2 = eyeglasses) to you (4) for me (5) [there].’

The OM order in Rwanda in (13) is obligatorily fixed by ascending InTop order, not GR or extension order, i.e. NONHUM < HUM (3PL) < 2SG < 1SG. Thus, the Rwanda example in (13) is ambiguous in several regards and could mean ‘...to me for you’, ‘...to you for them’, etc.

Haya JE22, like most MOM languages, has the same InTop order as Rwanda, but it also has a reverse strategy determined strictly by GR as in (14).

- (14) Haya JE22 “EB, node 9” (Duranti 1979: 40–41 for Hanja variety, Rubanza 1988: 123–124 for Ziba variety)

a. GR order

a-ka-ba-bi-leet-el-a

SM₁-PST-OM₂-OM₈-bring-APPL-FV

OM: DO_{HUM}-AO_{NONHUM}/^{**}AO_{HUM}-DO_{NONHUM}

‘She brought them (people) to them (yams).’

b. InTop order

a-ka-bi-ba-leet-el-a

SM₁-PST-OM₈-OM₂-bring-APPL-FV

OM: DO_{NONHUM}-AO_{HUM}/AO_{NONHUM}-DO_{HUM}

‘She brought them (yams) to them (people)/them (people) to them (yams).’

Contini-Morava (1983) describes the same OM order options as in Haya (14) in the variety of Rwanda as spoken in Masisi (DRC), illustrated in (15).

- (15) Masisi (DRC) variety of Rwanda JD61 “EB, node 9” (Contini-Morava 1983: 426)

a-za-mu-ki-h-a

SM₁-PST-OM₁-OM₇-give-FV

OM: DO_{HUM}-AO_{NONHUM}/^{**}AO_{HUM}-DO_{NONHUM}

‘She gave him to it [animal]/^{**}it to him.’

Such violations of InTop order are prohibited in metropolitan Rwanda, as shown in (10b) above. Haya and Rwanda spoken in Masisi (DRC) resemble free MOM languages in that either ascending or descending orientation is possible, but differ from the latter in the ordering principles. InTop or GR are the ordering principles in Haya and Masisi Rwanda rather than ConTop.

The GR order corresponding to ascending InTop order also occurs in the MOM systems of north-eastern Bantu languages of the Great Lakes region, such as Ganda in (16), which lack the InTop order option.

- (16) Ganda JE15 “EB, node 9” (van der Wal 2020: 217)

a. *n-a-gi-ba-gul-i-dde*

SM₁-PST-OM₉-OM₂-buy-APPL-PFV

OM: DO_{NONHUM}-AO_{HUM}/^{**}AO-DO

‘I bought it for them [people].’

b. *n-a-ba-gi-gul-i-dde*

SM_{1SG}-PST-OM₂-OM₉-buy-APPL-PFV

OM: DO_{HUM}-AO_{NONHUM}/** AO-DO

‘I bought them [people] for it.’

InTop does not play an ordering role in MOM systems of this type. However, as in Tswana, the OM_{1SG} is an exception by its fixed position. In this limited respect, it resembles free OM languages by the absence of GR determination of double OM order.

Then again, Nyambo JE21 (17) and Shambaa (18) represent north-eastern Bantu MOM systems determined simultaneously by both InTop and GR. OM sequences that violate either GR or InTop order are prohibited.

- (17) Nyambo JE21 “EB, node 9” (Rugemalira 1991: 205)

a-ka-kú-m-p-a

SM₁-PST-OM_{2SG}-OM_{1SG}-give-FV

OM: DO-IO/** IO-DO

‘He gave you to me/** me to you.’

- (18) Shambaa G23 “EB, node 9” (van der Wal 2020: 207, fn. 5)

*wa-za-zi-wa/** wa-zi-ghul-iyá*

SM₁-PST-OM₁₀-OM₂/** OM₂-OM₁₀-buy-APPL

OM: DO-AO/** AO-DO

‘They bought them (DO farms/** slaves) for them (AO slaves/** farms).’

The examples in (17) and (18) show ascending InTop order, typical of the wider area. However, unlike elsewhere in the wider area, InTop order does not result in role ambiguity, because GR order DO-IO/EO is also imposed.

4.1.2.2 Descending fixed MOM

The descending orientation is relatively rare. It occurs where the EO/IO OM (usually human/animate) is fixed to the left of the DO OM regardless of relative InTop, as in Umbundu from Luanda in (19). There are also free MOM varieties of Umbundu (personal communication from T. Schadeberg for the Bihé variety). Valente (1964: 248) may be describing an intermediate variety in reporting that the most common order is HUM-NONHUM-V. This would be expected pragmatically in a free order language of the Kwanyama type, where the OMs are ordered by ConTop, because human objects are expected to be more often of higher ConTop than inanimates. The Luanda Umbundu examples in (19) show strictly descending order by GR.

(19) Umbundu R11 “SWB, node 8” (fieldwork B. Wald & Maria Chikweka 1969)

- a. *w-a-tu-va/va-tu-kong-is-a*
 SM₁-PST-OM_{1PL}-OM₂/OM₂-OM_{1PL}-choose-CAUS-FV
 OM: CO-DO/**DO-CO
 ‘She had us choose them (people)/them choose us.’
- b. *w-a-tu-va/va-tu-kong-el-a*
 SM₁-PST-OM_{1PL}-OM₂/OM₂-OM_{1PL}-choose-APPL-FV
 OM: AO-DO/**DO-AO
 ‘She chose them for us/us for them.’

Luanda Umbundu also has the fixed position of OM_{1SG} as an exception to its GR orientation, e.g. *oku-lu-N-telek-el-a* (pronounced as *okulunelekela*) [INF-OM₁₁.NONHUM-OM_{1SG}-cook-APPL-FV] ‘to cook it [fish] for me’ (DO-AO) as opposed to descending orientation elsewhere, e.g. *oku-ku/tu-lu-telek-el-a* [INF-OM_{2SG/1PL}-OM₁₁.NONHUM-cook-APPL-FV] ‘to cook it for you/us’ (AO-DO).

Mongo-Liinja C61L “CWB, node 5” from Opala may also be of this type, e.g. *t-w-e-ke-ak-é* [NEG-OM₁-OM₉-tell-PREF-SBJV] “don’t tell it to him” OM_{IO}-OM_{DO} (Motingea Mangulu 2008: 320). However, the description is not sufficiently detailed to determine whether this order is fixed, as in Umbundu from Luanda, or optional, as in a free MOM system.

4.1.3 Partial MOM systems

Partial MOM systems are also diversified. The monophonic OM principle is widely distributed, largely adjacent to more complete MOM areas, i.e. Guthrie’s zones C-N. This covers all languages descending from node 5, i.e. those which emerged after the NWB branches in the Grollemund et al. (2015) phylogeny (i.e. languages from Guthrie’s zone A and groups B10-30) had split off. Except for Bubi, there are no reports of MOM systems in NWB languages.

There are also no reports of partial MOM systems in SWB, only of full MOM systems, except for the Kwanyama-type exclusion of an OM_{1SG} in favour of a 1SG enclitic. Lulua L31b exemplifies the monophonic OM principle in a minimal MOM system (cf. Rimi in (1) above), where even the monophonic principle is optional, so that a concurrent object to the 1SG object can be indexed by an object enclitic (OE) instead of an OM as in (20).

(20) Lulua L31b “SWB, node 9” (Morrison 1906: 38–39)

- a. *w-aku-ci-m-p-a*
 SM₁-PST-OM₇-OM_{1SG}-give-FV
 monophonic: NONHUM-OM_{1SG}
 ‘He gave it to me.’
- b. *w-aku-m-p-a-ci*
 SM₁-PST-OM_{1SG}-give-FV-OE₇
 SOM option: OM_{1SG}-V...NONHUM
 ‘He gave it to me.’
- c. *w-aku-ku-h-eye*
 SM₁-PST-OM_{2SG}-give-OE₁
 IO-V-DO/DO-V-IO
 ‘He gave him to you/you to him.’

Use of enclitics (optional or obligatory) instead of OMs is more densely distributed in interior western Bantu languages spoken north of Luba-Kasai L31a and Lulua L31b. As noted by Polak (1986: 377), the forms of OEs generally resemble PROs rather than OMs. This is especially clear for the class 1 OE in Lulua in (20c) above. The form of the Luba/Lulua class 1 PRO₁ is *ye-ye*, a reduplicated form based of the morphologically complex **yu-e* (**yu-* > *u-* as in the Luba/Lulua SM₁ form; the OM₁ form is *mu-*). Luba/Lulua is predominantly an asymmetric SOM system where the OM is selected by its high InTop relative to the concurrent object, as in the above example: 2SG > 3 SG (cl. 1).

More elaborate partial MOM systems are scattered across the interior eastern Bantu area, as in Bemba M42 in (21).

(21) Bemba M42 “EB, node 9” (Marten & Kula 2012: 245)

- mù-ká-bá-mú-éb-él-á-kó*
 SM_{2PL}-FUT-OM₂-OM₁-tell-APPL-FV-PRO₁₇
 OM: IO-AO
 ‘You (all) will tell them for him.’

Marten & Kula (2012) explicitly state that unless the monophonic OM_{1SG} occurs, multiple OMs in Bemba are restricted to persons (HUM), thus, to concurrent objects of high InTop. A similar restriction also seems to apply to the Mathira variety of Kuyu E51 “EB, node 9”, according to the examples offered by Englebretson et al. (2015: 109), while only the monophonic partial MOM has been reported

for other Kuyu varieties. In both Bemba and the Mathira variety of Kuyu, the ascending InTop and GR order apply, as among the intervening full MOM systems, such as Shambaa “EB, node 9” in (18) and Vunjo-Chaga E622C “EB, node 9”.

Lungu M14 displays a peculiar and apparently unique partial MOM system. It exhibits the common OM_{1SG} monophonic pattern in (22a), but, additionally, a descending MOM pattern for OM_{1PL} in (22b).

(22) Lungu M14 “EB, node 9” (Bickmore 2007: 26)

- a. *yá-kú-!cú-n-fúl-il-à*
SM₂-PRS.PROG-OM₇-OM_{1SG}-wash-APPL-FV
OM: DO-AO
‘They are washing it for me.’
- b. *yá-kú-!tú-cí-fúl-il-à*
SM₂-PRS.PROG-OM_{1PL}-OM₇-wash-APPL-FV
OM: AO-DO
‘They are washing it for us.’

In both respects Lungu resembles the Luanda Umbundu MOM system, except for the apparent fixed GR order even when OM_{1SG} is involved. In this respect, Lungu (22a) conforms to the Nyambo (17) / Shambaa (18) pattern, where both InTop and GR order are obligatory. Lungu (22b) is the most eastern reported example of the descending GR order orientation.

4.2 SOM systems

The primary distinction among SOM languages is between symmetric and asymmetric systems. Most frequently explored is the trans-verbal context of concurrent objects: OM_i-V...NP_j. In symmetric systems the relative InTop of OM_i and NP_j is not constrained. In asymmetric languages OM_i is prohibited from indexing an object of lower InTop than NP_j, e.g. **OM_{NONHUM}-V...NP_{HUM}. van der Wal (2020: 205) observes that MOM systems tend to be symmetric, in contrast to SOM systems. For our purposes, the (verb) internal context SM_i...OM_j-V(...)-PASS provides a more discriminating context for asymmetry. It exposes different degrees of asymmetry between EB languages in the south and EB languages in the north and the centre. Thus, first consider Zulu S42 in (23) as representative of southern EB asymmetric SOM systems. The disjoint marking in (23) is obligatory if the verb is final, i.e. when there is no postverbal constituent. The passivisation

prohibition in (23) also occurs in some MOM languages, for example in some varieties of Tswana (cf. Creissels 2006: 22).

(23) Zulu S42 “EB, node 9” (adapted from Zeller 2012: 229)

*i-ya-**m-phek-el-w-a* (umama)

SM₉-DSJ-**OM₁-COOK-APPL-PASS-FV (mother)

SM_{NONHUM(DO)}...**OM_{HUM(AO)}

intended: ‘It (meat) is being cooked for her/(mother).’

It should be emphasised that the InTop Internal Passive Constraint of (23) is precisely due to a conflict between ConTop (SM > OM) and InTop (NONHUM < HUM) within the topicality ranking system, and not due to the option of OM doubling of a postverbal object within the clause. OM doubling of a postverbal object is characteristic of the entire eastern coast and shallow interior. At the same time, the option of lower InTop passivisation in the context of a concurrent postverbal object of higher InTop is strictly a south-eastern Bantu characteristic, in contrast to coastal and shallow interior central and north-eastern Bantu represented in (24) below. The Internal Passive Constraint of south-eastern coastal/shallow interior Bantu is the SOM analogue of fixed MOM order according to In-Top (§4.1.2.1), e.g. SM_{HUM}...OM_{NONHUM}/**SM_{NONHUM}...OM_{HUM} corresponds to north-eastern Great Lakes MOM: OM_{NONHUM}-OM_{HUM}/**OM_{HUM}-OM_{NONHUM}.

The SOM systems of coastal and shallow interior EB languages in the centre have additional constraints. The single OM constraint extends to passivised verbs so that the passivised subject, having a TR commonly associated with the object of the active verb, prohibits a concurrent OM reference, i.e. only *a single object role* can be indexed as a topic in *any* context. The InTop constraint seen above in Zulu is also characteristic of central EB languages, but also in the context of a concurrent postverbal object of higher InTop. Both the single object role and the asymmetric InTop trans-verbal constraint on a concurrent object occur as far north as Swahili belonging to the EB subgroup which Nurse (1999: 5) calls ‘North-East Coast Bantu’ (NECB). The Swahili example in (24a) illustrates the single object role constraint, while (24b) shows the InTop prohibition when a postverbal object has higher InTop than the concurrent object. In the absence of a concurrent object, a single object of any InTop can be indexed in Swahili by an OM or a passive SM, as shown in (24c). Finally, (24d) illustrates that InTop is a more powerful feature than TR in Swahili, because the object role indexed by an OM or passive SM is ambiguous between DO and AO.

(24) Swahili G42d “EB, node 9” (fieldwork B. Wald)

- a. Single Object Role Constraint (Passive)

*a-li-(**i)-p-ew-a*

SM₁-PST-(**OM₉)-give-PASS-FV

SM_{AO/DO...}**OM-V-PASS

‘She [child] was given it [gift].’

- b. InTop Trans-verbal OM Constraint

*wa-li-m/**i-p-a* (zawadi) *m-toto*

SM₂-PST-OM₁-/**OM₉-give-FV (9.gift) 1-child

**OM_{NONHUM}-V...NP_{HUM}

‘They gave it [gift] to the child.’

- c. InTop Trans-verbal Passive Constraint

(zawadi) *wa-li-i-p-a* / *i-li-p-ew-a* (***m-toto*)

(9.gift) SM₂-PST-OM₉-give-FV / SM₉-PST-give-PASS-FV (**1-child)

**SM_{NONHUM}-V-PASS...NP_{HUM}

‘(Gift) they gave it/it was given (**[(to) the child]).’

- d. InTop Trans-verbal OM / Passive Constraint

wa-li-m-tak-i-a / *a-li-tak-i-w-a* *pesa*

SM₂-PST-OM₁-want-APPL-FV / SM₁-PST-want-APPL-PASS-FV 9.money

SM/OM_{HUM=AO/DO...}V...NP_{NONHUM}

‘They wanted him for (his) money.’ or ‘They wanted money for him.’ /

‘He was wanted for (his) money.’ or ‘He was wanted/wished (to

have/get) money.’

Among Swahili’s closest relatives, Kauma E72b illustrates that there is variation in the NECB Mijikenda languages concerning the single object role constraint. The example in (25) shows the operation of the InTop internal (passive) constraint: 1SG > 2SG. The passive SM must index the object of higher InTop, a constraint shared with Zulu (23) above. The SOM Single Object Role Constraint is relatively new to NECB. Among Swahili’s closest relatives, the other Sabaki languages (E70-73), OM indexing of a second object with passivisation of the first is attested in Southern Mijikenda (e.g. Digo E73 and Duruma E72d) in the early twentieth century (Wald 1994: 261, examples (26)–(27)), but is no longer accepted by later generations, undoubtedly under Swahili influence. Thus, the direction of this local change is secure.

- (25) Kauma E72b “EB, node 9” (fieldwork B. Wald & Chris M. 1993)

InTop Internal Passive Constraint

ni-dza-ku-ger-w-a

SM_{1SG}-PST-OM_{2SG}-give-PASS-FV

SM_{IO}...OM_{DO}/SM_{DO}...OM_{IO}-V-PASS

‘I was given to you/you were given to me (today).’

Van der Wal (2020) notes Shambaa as the only exception in her sample to a generalisation that only SOM languages are asymmetric with respect to the InTop trans-verbal constraint. There are, however, more widespread asymmetries among MOM languages. Both Tswana in southern EB and Rwanda (Kimenyi 1976: 134) in northern EB share the InTop internal passive constraint corresponding to Zulu (23) above. However, Rwanda does not exhibit the Zulu constraint when the OM indexes a concurrent AO (Ngoboka 2005: 88).

Certain languages have a partial SOM. As an effect of InTop, these systems are mostly restricted to human objects. Polak (1986: 375) shows a diverse pattern in interior CWB languages of Guthrie’s zones C and D, i.e. those branching off from the remainder of Narrow Bantu at node 5 in the phylogeny of Grollemund et al. (2015). The most restricted SOM system is Mbesa C51, allowing only the class 1 OM. Grégoire (2003: 366ff) adds to the variety of micro-trends in the CWB clade. For example, among HUM OMs Leke C14 has only the OM_{1SG}, but it has an inventory of NONHUM OMs, while Boa C44 has only OMs of human classes 1/2. Widespread in this general area is alternation within the same language between an OM, when it is available, and either an enclitic or a postverbal PRO; an additional option consists of combining both strategies by indexing an object by both an OM and a postverbal PRO. To the extent that the enclitic/postverbal option is favoured, these languages resemble exclusively NOM systems to their immediate north (cf. §4.3). However, in contrast to those NOM systems, the data are not sufficient to determine if GR/TR plays a role in any of these partial SOM languages.

Makhuwa represents a distinct area where only the OMs of the typically human classes 1/2 occur. Makhuwa is adjacent to a central Eastern Bantu area of HUM-NONHUM polarisation, where human objects favour or obligate OM indexing while the available inanimate OMs are rarely used. In contrast, the partial SOM systems in interior CWB are adjacent to the NOM systems of interior NWB further north.

4.3 NOM systems

In NOM systems, only PROs perform the anaphoric function. InTop does not play a discernible role in NOM systems. Instead, the major factors determining OM order are GR/TR and information status. This latter factor distinguishes PROs from lexical nominals. Lexical nominals contain more information than PROs. NOM languages vary in how GR and information status interact in determining the order of PRO objects with respect to concurrent nominal objects and with respect to each other. The information status constraint, where it occurs, compels PRO-NP/**NP-PRO order as in (26).

- (26) Orungu B11b “NWB Gabon, node 3” (Van de Velde & Ambouroue 2017: 619)
- à-gòl-ín yé á-bà / **á-bà yé*
 SM₁-PST-buy-APPL PRO₁ 6-mango / **6-mango PRO₁
 ‘She bought mangoes for him.’

The information status constraint on concurrent objects, V PRO_i-NP_j/**NP_i-PRO_j, parallels trans-verbal OM_i...NP_j in OM languages. The competing factor is GR order EO/IO-DO (cf. Type 2 = NOM, in Beaudoin-Lietz et al. 2004: 186). Further north in NOM systems of NWB Cameroon languages (node 1), postverbal information status order is optionally violated in favour of GR order EO/IO-DO, as in Basaa A43a (27). There is no parallel in OM systems to the NOM postverbal double-object order NP-PRO.

- (27) Basaa A43a “NWB Cameroon, node 1” (Hyman 2003a: 284)
- mɛ n-lémb-él gwó bɔŋgé / bɔŋgé gwó*
 1SG PST-cook-APPL it 2.child / 2.child it
 V...PRO_{DO}-NP_{AO} / NP_{AO}-PRO_{DO}
 ‘I cooked it [food] for the children.’

A second variable among NWB NOM systems is the position of PRO objects in relation to the verb. One position is postverbal, i.e. after the main verb, just as in non-NWB languages, i.e. SM-(TM)-AUX#(INF)-V...PRO_{OBJ}. The alternative is post-AUX: SM-(TM)-AUX#PRO_{OBJ} (#INF)-V... Within Narrow Bantu, the post-AUX type is unique to NWB. Intermediate types, as in Eton A71 (28), occur in which the post-AUX type is limited to certain AUXs and/or allows either the postverbal or post-AUX option. As in Basaa (27), Eton postverbal order allows NP-PRO to accommodate GR order IO-DO. The only differences with Basaa are the post-AUX options in (28a–28b).

(28) Eton A71 “NWB Cameroon, node 1” (Van de Velde 2008: 302)

- a. post-AUX (preferred order)
mèèy nyí dḥ vé
 1_{SG}.FUT PRO_{1[IO]} PRO_{5[DO]} (INF)give
- b. trans-verbal
mèèy nyí ↓vé dḥ
- c. postverbal
mèèy vé nyí dḥ
 ‘I will give it to him.’

In closely related Atsi A75D post-AUX position is obligatory for some AUX, e.g. the future marker *kḥ*, as in *mḥ-kḥ dḥ ḥ-dzḥ* [1_{SG}-AUX=FUT PRO₅ INF-eat] “I will eat it [mango]”, but the postverbal option occurs for others, e.g. the remote past marker *ngá*, as in *mḥ-ngá ḥ-dzḥ dḥ* [1_{SG}-AUX=PST2 INF-eat PRO₅] “I ate it [mango] (a long time ago)” (Nzang-Bie 2014: 78ff). As in Eton, the post-AUX PRO order is strictly IO-DO, e.g. *mḥ-ngá nyá zḥ ḥ-kólḥ* [1_{SG}-AUX=PST2 PRO₁ PRO₉ INF-lend] “I lent him it [book]” (Nzang-Bie 2014: 81).

Among the post-AUX systems, there are a few NWB NOM systems, for example Nen A44, see (29), where full nominals as well as PROs are allowed in post-AUX position.

(29) Nen A44 “NWB Cameroon, node 1” (Mous 2005: 419)

- mé-ḥò ḥḥó mímé fálábi*
 SM_{1SG}-FUT PRO_{2SG} house build.CAUS
 AUX PRO_{IO}-NP_{DO} V...
 ‘I will build a house for you.’

Mous (2005) argues that Nen represents an innovative system such that its line of development is not relevant to the PB OM hypothesis. Nen represents an extremely localised Narrow Bantu type that will not be pursued further here. A few scattered Bantoid languages also have some version of this feature, for example Vute (North Bantoid) near Nen in Cameroon.

4.4 The Bubi OM systems

A peculiarity of Bubi A31, apparently unique in Bantu, is its split orientation of double OMs. According to Abad (1928), all varieties display split double OM order according to the person of the IO, e.g. most northern and southern varieties agree on the fixed order DO-IO₁ for class 1 (3SG), but the reverse fixed order IO₂-DO

for class 2 (3PL). The south-western Batete variety allows both options for class 1, as in (30a), resembling a free MOM system in this respect. However, in all other instances, Bubi OM order is fixed by GR and person. Some persons are ordered in opposite ways in different varieties. Examples (30b–30c) illustrate that IO_{2SG}-DO order in southern varieties (including Batete) corresponds to DO-IO_{2SG} in northern varieties.

(30) Bubi A31 “NWB Cameroon, node 1” (Abad 1928: 45)

a. Batete variety

o mo ma / ma mo mbi

SM_{2SG} OM₁ OM₆ / OM₆ OM₁ give.PST

OM: IO-DO/DO-IO

‘You gave them (the palms) to him.’

b. Southern varieties

a o ma mbi

SM₁ OM_{2SG} OM₆ give.PST

OM: IO-DO

‘He gave them (the palms) to you.’

c. Northern varieties³

a b’ o pei

SM₁ OM₆ OM_{2SG} give.PST

OM: DO-IO

‘He gave them (the palms) to you.’

In the imperative (non-negative), enclitics of the same form as the OMs occur. In that case, all varieties agree on the IO-DO order, like double object PROs in the NWB mainland NOM systems. However, even in this position the peculiar order DO-IO₁ persists across varieties (Abad 1928: 88). The obligatory postverbal position in the imperative is noteworthy. The same position is obligatory for PRO objects in Nen “NWB Cameroon, node 1”, and may be more widespread among post-AUX NOM systems. Data are lacking for the Bantu A70 languages. However, it seems likely that the imperative is generally restricted to postverbal PRO objects, because the imperative provides no post-AUX context among NOM systems. Bubi (30) resembles a MOM rather than a post-AUX NOM system in the apparent absence of an obligatory AUX preceding the OMs. Nevertheless, in contrast to the affirmative imperative, the negative imperative has a negative AUX to trigger post-AUX position for the OMs, cf. *bëëla-lö* [sing-OE₅] ‘sing it!’

³Note the characteristic Northern Bubi denasalisation inducing cl. 6 **ma* > *ba*.

vs. *wë-lö-béél-è* [2SG.NEG-OM₅-sing-SBJV] ‘don’t sing it!’ (Bolekia Boleká 1991: 151).

The forms of the Bubi OMs are problematic for historical analysis. They are all monosyllabic but vary within varieties with respect to the vowel used, e.g. class 2 *ba/bo/be* ‘them’. The *-e/o* forms are suggestive of the PB deictic suffixes appended to PRO, but alternative explanations are conceivable. The vowels could also reflect one or more former AUXs or TMs with which the preceding SM fused, and then were transferred to the OM forms, just like the more limited central Bantu reanalysis of the SM/OM_{1SG} *ndi-* < **N-di* [SM_{1SG}-COP/AUX] (cf. Polak 1986: 379).⁴ There is nothing in the current Bubi varieties to suggest that the OMs are perceived as polymorphemic. A point in favour of a MOM (OM) rather than NOM (PRO) analysis is that the 1SG SM/OM is apical as in OM systems rather than bilabial as in the NWB NOM systems, i.e. PB **ni/(N)-* SM/OM_{1SG} vs. PB **mí-* PRO_{1SG}.⁵

In sum, Bubi fits the major criteria for a (M)OM system with respect to the monosyllabicity and morphological simplicity of its OMs. However, it resembles the NOM systems in the obligatory postverbal position of its OMs in the imperative context.

5 Historical object marking hypotheses

This section examines a number of hypotheses about the nature of the PB OM system in light of the types we have examined in §4 above. In the background of this discussion is the understanding that the PB period is a lower limit to the age of the PB OM system that can be reconstructed by comparing the diverse current Narrow Bantu languages. The system may be much older, because the current situation may preserve defining features that have been lost elsewhere in Bantoid or even Benue-Congo. Alternatively, any form of the OM system may be a post-PB development so that some type of NOM directly reflects earlier

⁴Similarly, in the “NWB Cameroon, node 1” SOM languages Mbonge A121 and Kpe A22, the SM_{1SG} has the form *na-* suggesting **n-a-* [SM_{1SG}-TAM], also one of the forms of the Bubi SM_{1SG}.

⁵With regard to the nature of the boundary between the OM and the following verb, Bubi standard orthography follows Spanish convention in representing the preverbal OMs as separate words like Spanish preverbal clitics, e.g. preverbal: Bubi <*a ñe ri bbi*> [he me it gave] = Spanish <*me lo dió*> [me it gave.he] ‘he gave it to me’; but as enclitics suffixed to the verb when they are postverbal, e.g. postverbal: Bubi <*mbañelo*> = *mba-ñe-lo* = Spanish <*démelo*> = *dé-me-lo* [give-me-it] ‘give it to me’ (Abad 1928: 88). The issue of whether the Bubi OMs are indeed separate words, as they would be as object PROs in a NOM system, cannot be pursued further here.

PB object indexing systems. This is a primary issue to be discussed. It is one of numerous questions of direction of change. The NOM issue is: Did current OM systems evolve from NOM systems, or vice-versa? The discussion will begin with the PB OM hypothesis, because Meeussen (1967) and Polak (1986) agreed on some version of this hypothesis. They disagreed on whether the particular PB OM system was MOM or SOM (cf. §2 above).

5.1 The PB OM hypothesis

As stated immediately above, there are two fundamental types of OM hypotheses: SOM and MOM. Polak (1986) favoured a SOM hypothesis for reasons discussed in the present section. In doing so, she preferred an OM hypothesis over a NOM hypothesis, the latter discussed in §5.2. The relative merits of Polak's SOM hypothesis and some form of MOM hypothesis are then discussed in §5.3.

Polak (1986: 374) generally appeals to the geographical distribution of current OM languages to posit the OM as a feature of PB. Some form of OM system, full or partial, occurs in all zones. Polak acknowledged that it was troubling that NWB (zone A and vicinity) is almost devoid of OM systems, but she mentioned Jö as having an almost full OM system, specifying its proximity to Duala A24, a NOM language of the type exemplified by Basaa A43a in (27) above. Representative of the Jö area are the full SOM systems of Mbonge A121 (Friesen 2002) and Kpe A22 (Hawkinson 1986). Polak's suggestion implies that they represent a *relic area*. According to the Grollemund et al. (2015) phylogeny, these languages show their closest lexical affinities to the NOM languages of zone A surrounding them.

A contrary hypothesis would be that the area reflects post-PB OM systems that originated further south, subsequently transported to their current area and consequently undergoing relexification through contact with the surrounding area. In the absence of any supporting evidence for the relexification hypothesis, the relative simplicity of the relic hypothesis is preferable.⁶ For further reference, this area is called the “NWB Cameroon, node 1, full SOM” area.⁷

⁶Another archaic feature of this area, shared with Bubi, is the initial apical nasal for SM/OM_{1SG} reflecting PB **n(i)*-. In the surrounding NOM area, the SM_{1SG} has an initial bilabial nasal reflecting the PB PRO_{1SG} **mi*- as does wider Benue-Congo for the most part. In the NOM area of Narrow Bantu and adjacent Bantoid languages, even the SM, where it survives, has the initial bilabial nasal of PRO.

⁷Polak's (1986) Map 2 represents this full OM within “NWB Cameroon, node 1” as a small north-western portion of zone A surrounded by systems left blank on the map as NOMs. The map in Beaudoin-Lietz et al. (2004: 180) shades the surrounding NOM systems (“Type 2” in their terminology).

Partial SOM systems are explicitly considered to be due to loss by Polak (1986), thus positing a specific historical direction: full > partial SOM systems. According to this hypothesis, the change at issue is the loss of OMs from the full PB set, so that some objects cannot be indexed by OMs. In this context it is instructive to consider the partial OM systems of Makhuwa and NWB as independent innovations in widely separated areas. They have in common that InTop plays a major role in favouring loss of some or all of the inanimate OMs in both areas. They differ in the predominant nature of adjacent systems.

As discussed in §4.2, Makhuwa and NWB represent distinct cases of inanimate OM loss in terms of the nature of adjacent systems. Makhuwa has lost the OMs of all classes except classes 1/2 (typically human). It is surrounded by languages that maintain full SOM systems, but with prohibitions against inanimate OM indexing in preference to a concurrent human object, as exemplified in Swahili (24) above.

In contrast to the Makhuwa area, Polak's (1986) Map 2 shows that the north-western Bantu partial OM area is much larger and adjacent to numerous distinct types of systems along its southern and eastern borders, including other partial OMs and NOMs. The area is attested in zones A-D with some further southern extension into zone H. In other words, it occurs in several early branches: "NWB Cameroon, node 1", "NWB Gabon, nodes 2-4", "CWB, node 5" and "WWB, node 6". Among other partial OM systems there are some that have also lost human objects, including the interpersonal OMs (cf. §4.2). The logical conclusion to this trend is the loss of all OMs, resulting in NOM systems.

The preceding account follows from a direction of increasingly constraining the OM system, until it is completely lost. This is not a likely outcome for the Makhuwa area. Preferential OM indexing of human objects is characteristic of the entire area surrounding it; no further movement towards loss is indicated. In contrast, the north-western Bantu NOM adjacency to partial OMs offers a model for further evolution towards losing the remaining OMs. The particular paths taken by the zone C languages from only HUM OMs to NOMs remain unclear and problematic at present. Some partial OM systems suggest phonological influence, e.g. the monophonic principle, but also the loss of the initial consonant from the surviving human OMs in parts of zone C.

5.2 The PB NOM hypothesis

The opposite direction from NOM to OM is a currently disputed position advanced by Güldemann (2011; Güldemann (2022 [this volume])). It implies that PB had a NOM system of the form AUX# PRO_{OBJ} V, where multiple pronominal

objects were allowed. For the most part, Güldemann appeals to typology rather than current Narrow Bantu for support. He proposes that starting from a hypothetical pre-PB VO system, e.g. systems like Orungu (26) or Basaa (27) above, only the PROs among postverbal NPs came to be preposed to the verb, as in Eton (28), representative of the A70 group and various other groups in the vicinity, e.g. Maande A46 (cf. Mous 2005). Romance is a well-documented case to serve as a typological model for the posited direction of change, VO_{PRO} > O_{PRO}-V (cf. Wald 1994: 250). Romance also serves as a typological model for the phonological condensation of the preverbal PROs to monosyllables, i.e. PRO-(#)V > OM-V. In the Bantu analogue, there are grammatical consequences to the reduction, such that the loss of the deictic markers suffixed to PRO, leaves only the class and interpersonal markers as the forms of the OM. This model is plausible but problematic for direct evidence. So, in relating this proposal to current Bantu, Güldemann (2011) offers Ewondo A72a in (31) as partially preserving this system from its PB origin.

(31) Ewondo A72a “NWB Cameroon, node 1” (Redden 1979: 167)

a-kad mə dzɔ vá
 SM₁-TM/AUX PRO_{1SG=IO} PRO_{9=DO} give
 ‘He usually gives it to me.’⁸

Here I have substituted a double object example for Güldemann’s single object example, as a reminder that number of objects is not an issue in this change. The account does not rule out reduction of each preverbal pronominal object to a single syllable (or less), resulting in a MOM system, as suggested by the variation in the formal ambiguity between OM and PRO forms discussed for Bubi (cf. §4.4).

The problem with Ewondo (31) as a direct reflection of a hypothetical PB NOM system is that among its closest relatives, an INF intervenes between the last object PRO and the verb root, i.e. PRO(#)/OM INF-V. In Atsi A75D INF is explicit, as discussed under Eton (28) above. In Eton, the INF often manifests as a floating tone downstepping a high tone verb immediately following the object PRO, as in (28b) above (e.g. Van de Velde 2008: 272). Such a floating tone is a commonly attested feature of north-western Bantu in the wake of the loss of various syllabic grammatical morphemes retained in other Bantu areas. Most likely the Ewondo system evolved from the system still reflected in Atsi and Eton, but at some point lost all trace of the INF. The PRO_{OBJ} INF-V order of A70 contrasts with the INF-OM-V order of OM systems throughout Bantu, including the “NWB Cameroon,

⁸Ewondo AUX *kad* < PB **jikad* ‘dwell; be; sit; stay’ (BLR 3441) (Bastin et al. 2002).

node 1, full SOM” area. Thus, it is doubtful that Ewondo (31) directly reflects the PB situation.

Along with the criticisms formulated by Hyman (2011), a major problem of Güldemann’s dependence on typology is the timing of the $V-O_{\text{PRO}} > O_{\text{PRO}}-V$ change relative to PB. It conflicts with the relic hypothesis for the “NWB Cameroon, node 1, full SOM” area, discussed in §5.1 above. Among possible resolutions to this conflict is one in which Güldemann’s typologically inferred reconstruction projects back to an earlier stage than PB, and that MOMs had already arisen by the PB period, and were subsequently widely restricted to SOM systems, by processes comparable to the reduction of SOMs from full to partial advocated by Polak (1986).

5.3 The PB MOM hypothesis

Polak (1986) rejects the MOM hypothesis for PB. Her main argument is that SOMs are more common across the entire Bantu area. Clear-cut MOMs currently have a more limited distribution, all south of the greater NWB area (nodes 1–4). She suggests that the monophonic partial MOM type was a transition to the greater elaboration of SOMs to MOMs. This contrasts with her positing of partial SOMs as a transition between full SOMs and NOMs. At first glance, the “NWB Cameroon, node 1, full SOM” systems seem to support the chronological priority of SOM to MOM. If the direction of change was $\text{MOM} > \text{SOM}$, then the reduction to SOM in the isolated “NWB Cameroon, node 1, full SOM” area and the reduction to SOM in a large part of EB (especially coastal but expanding deep into the central EB interior) seem to be independent innovations. Zone C is a transitional area for either direction of change. It is unusually diverse in containing full and partial SOMs and MOMs in proximity to NOMs. MOM systems are attested as far into the north-western interior as Bangi C32 “CWB, node 5” (Whitehead 1899). Particularly in the proximity of partial SOMs and MOMs, zone C suggests that the same process of reduction that Polak (1986) posits for SOMs also applies to MOMs. By this account, there is a single direction of change towards reduction of OM complexity for both SOMs and MOMs, so that SOMs represent an intermediate stage in the change of MOM to NOM systems. The “NWB Cameroon, node 1, full SOM” area independently follows the same line of development of reduction under the same conditions but stops at the full SOM stage.

An internal motivation can be offered for the above hypothesised persistent direction of change to reduced OM systems. It follows the principle of discourse utility, measured by the higher frequency of use of single than multiple OMs in MOM systems. Uses of single OMs are far more frequent than multiple OMs

in discourse. By the discourse utility principle, the complexity of the system is reduced by restricting the system to SOM. From this point of view, it is appropriately termed the principle of discourse economy, referring to a less complex, thus more economical system with fewer grammatical options.

There are many historical junctures suggested by the data where the discourse economy principle may account for a reduction in the number of OM indexing options allowed by a hypothetical PB MOM system. They arise in reconsidering Meeussen's original suggestion that PB had a MOM system. The foremost problem is the issue of the hypothetical type of PB MOM system, given that current Bantu has numerous distinct MOM systems, as discussed in §4.1 above.

The primary alternatives are free vs. fixed MOMs. The fixed ascending MOM seems to be the most widely distributed type, dominating the EB MOM area and extending far westward into zones C and H, i.e. the CWB and WWB clades. At the same time, free MOMs are concentrated, apart from Tswana in the south of the EB domain, in two widely separated areas: (a) SWB as far north as Umbundu and (b) the south-eastern area of Great Lakes Bantu in the north of the EB domain, as in Kuria (9). Is the agreement between these two areas a case of independent development, or relics of an older previously more widespread system, possibly the PB system?

Taking a PB version of the free MOM as the starting point offers some advantages over alternatively hypothesising a PB fixed MOM alternative. An immediate advantage is that the free MOM system can be seen as a pivot between the two subsequent fixed orientations, ascending and descending, as suggested by the variation in Bubi (30a). It is also a first step in accounting for the use of both orientations in the type represented by Haya (14a–14b). The direction of change MOM > SOM simplifies the account of subsequent developments. For example, taking a version of the Kwanyama (6) free MOM type as the PB point of departure, one of the most widespread subsequent changes is fixing the order of OMs according to InTop instead of ConTop.

The principle of discourse economy comes into play in this change. The change from ConTop to InTop reduces the complexity of the free MOM by eliminating less frequently used discourse options, particularly with respect to distinguishing humans from other objects. Most often in discourse human objects are indexed regardless of the type of system. Therefore, in the free MOM system, human objects will be indexed more often than inanimates whether or not there is a concurrent inanimate object. HUM objects will be indexed leftmost in sequence in a descending system: HUM-NONHUM, and rightmost in an ascending system NONHUM-HUM. These orders are both options in the free MOM system but become obligatory and decontextualised in the fixed MOM systems.

Meanwhile, in contradiction to the PB MOM hypothesis above, Polak (1986: 404) conjectures that the earliest version of the MOM system arose outside of NWB in an assumed more innovative Narrow Bantu area in which the languages have a “general tendency [...] to lengthen words”. She seems to be referring to agglutination here, as opposed to the more isolating tendencies of NWB, as seen in NOM systems. The problem with this assumption is that there is little doubt that PB already had an agglutinative system including the transitivity-raising suffixes CAUS and APPL. The extension sequence CAUS-APPL occurs throughout Bantu, including in the SOM languages of the “NWB Cameroon, node 1” clades, such as Mbonge *di-kab-is-él-é* [INF-share-CAUS-APPL-FV] “to sell (lit. let share) [something DO] [to someone AO]” (Friesen 2002: 97). By the same criterion of distribution across Bantu that Polak (1986) invokes to justify positing the PB OM, the sequence CAUS-APPL can be posited for PB, where each extension is associated with an object, expressed or implied. This, then, seems like sufficient motivation for developing a MOM system at the PB stage – had it not already existed. Certain “NWB Cameroon, node 1” languages allow the double OM sequence OM-REFL. Such is the case, for instance in Kpe, where the OM_{REFL} *a-* then replaces the vowel of the preceding OM, thus maintaining the monosyllabic OM slot, e.g. *na-ma-l-a-ké-én-é* [SM_{1SG}-PST-OM₁₁-OM_{REFL}-cut-INSTR-PFV] “I cut myself with it [knife]” (Hawkinson 1986: 152).⁹

A final point in favour of the notion that MOM systems were formerly more common in the “NWB Cameroon, node 1, full SOM” area is the nature of the full SOM system in those languages. It is a symmetric system both with respect to the trans-verbal multiple object context: OMi-V...NPj.OBJ and the internal passive subject context SMi.OBJ...OMj-V...PASS. There is no constraint on which of two objects can be assigned higher ConTop, just as in the free MOM system. As van der Wal (2020: 206) observes (especially in her Table 3), MOM languages tend to be symmetric with respect to the trans-verbal multiple object context: OMi-V...NPj.OBJ and the internal passive subject context SMi.OBJ...OMj-V...PASS. There is no constraint on which of two objects can be assigned higher ConTop. SOMs of this type tend to be closer to MOM areas than asymmetric SOMs. In van der Wal’s sample, symmetric SOM systems are widely dispersed across EB but also include Mongo C61 from the “CWB, node 5” clade, which she classifies as partial MOM (termed 1+). In contrast, her asymmetric SOMs are largely coastal EB, along with some partial MOM systems as far west as Ruund L53 “SWB, node

⁹Kpe represents a wider NWB area in which the APPL extension was replaced by a reflex of the PB *-an to incorporate an instrument as an object argument of the verb (cf. Wald 1997). A reflex of the PB APPL *-id continues in this area in other uses.

8” and Yaka “WWB, node 6”. Her solitary example of an asymmetric MOM is Shambaa “EB, node 9”.

Shambaa asymmetries involve both InTop and GR, cf. (18), where MOM order is fixed according to both InTop and GR. The same factors play a role in restricting the use of NONHUM OM indexing in Shambaa single-OM trans-verbal contexts. When there is an unindexed concurrent HUM object, a NONHUM DO cannot be OM indexed (but a NONHUM AO can), i.e. IO/EO/**DO=OM_{NONHUM}-V...DO/**IO/EO=NP_{HUM}. This limitation is one step less severe than the asymmetric trans-verbal constraint of SOM systems in Shambaa’s vicinity, as exemplified for Swahili (24b–24d). In those systems there is no GR condition on OM indexing of NONHUM objects, only the InTop condition, i.e. **OM_{NONHUM}-V...NP_{HUM}. In contrast to this situation in northern and central EB, the symmetry of the “NWB Cameroon, node 1, full SOMs” suggests that they, like other symmetric SOMs, were formerly in the proximity of MOM systems (later replaced by the current NOM systems), and/or that they formerly had MOM systems themselves, subsequently replaced by SOM systems according to the discourse economy principle while retaining the symmetry of their previous MOM state.

6 Conclusions

This section summarises the PB MOM hypothesis preferred above and indicates problems requiring further investigation for support or refutation of the hypothesis.

The hypothesis proposed in this chapter is that PB hosted a free MOM topic marking system consisting of an obligatory SM and one or more OMs in sequence. Subsequent local innovations altered the use of this system of ranking objects by changing the OM ordering principle from ConTop > InTop. This is the earliest indication of the post-PB discourse economy principle applied to indexing objects. Ultimately the line of evolution driven by this principle reduced OM indexing to partial HUM SOM systems, as in Mbesa “CWB, node 5” (cf. §4.2), and then the complete loss of the system. As early as the partial MOM systems, PRO had been compensating for restrictions on OM indexing, as in Luba/Lulua “SWB, node 9” (20). This was a change from the PB uses of PRO, e.g. focus uses iconic to their overt morphological complexity.

It remains unclear that the final NOM state still involves topicality, either ConTop or InTop, other than the minimal topicality bestowed by PRO as an anaphor. In any case, the predominant NOM state is ordering of multiple objects as IO/EO-DO. This order applies to both postverbal and post-AUX types of NOMs (see

(26)–(29)). Thus, GR seems to be the dominant principle determining order. An intermediate stage is suggested by Orungu “NWB Gabon, node 3” in (26), where the invariant postverbal order V...PROi-NPj corresponds to the symmetric order OMi-V...NPj as reflected in the “NWB Cameroon, node 1, full SOM” systems. GR order IO/EO-DO moves the PRO further from its OM analogue, as in Basaa “NWB Cameroon, node 1” in (27), and persists in the subsequent change to post-AUX position, as in Eton “NWB Cameroon, node 1” in (28).

More generally, the origin of GR in OM ranking according to the PB MOM hypothesis remains unresolved. In a free MOM system like the one in Kwanyama in (6), GR plays no role. How and at what stage did GR become a factor in OM indexing according to the hypothesis? So far the data are insufficient to answer this question decisively. As a rare “SWB, node 8” example of fixed descending OM, the Luanda variety of Umbundu (19) displays only GR ordering, not InTop. Decisive evidence of a previous InTop stage is yet to be uncovered.

A similar problem occurs in the northern EB languages from the Great Lakes area with fixed MOM systems, such as Ganda in (16). Only GR seems to play a role. In the northern EB case there are distinct adjacent MOM types that reveal further details of an interplay between InTop and GR order. Haya in (14) accepts free OM order, but it imposes constraints on its interpretation. Descending and ascending orders are both fixed but distinct. Descending order is determined strictly by GR/TR, and ascending order by InTop. How this state arose is unclear. One possibility is that GR preceded InTop, so that InTop introduced GR ambiguity with the understanding that the discourse context would easily resolve most such ambiguity. Ganda supports this possibility by showing no influence of InTop on its GR OM order. On the other hand, Rwanda represents a system in which InTop operates in spite of GR. It presents a model for the contrary hypothesis that InTop preceded GR/TR historically. Chronological ordering of these systems and its implications for the PB MOM hypothesis remains unresolved.¹⁰

A more general problem of data affecting the PB MOM hypothesis is the rarity of descriptions of the less favoured discourse cases, e.g. *double-human* and *cross-animate* object examples. Double-human objects are more often attended to, e.g. “he showed her to them”. In most reported systems, the human DO is treated in the same way as an inanimate DO whether by InTop or GR. Rwanda in (11) shows that GR only plays a role in its system when human objects of equal InTop are OM indexed. The Rolong variety in Tswana (7b) shows no grammatical effect of GR at

¹⁰Ganda is among languages that have been tested for the cross-animate context. It maintains GR ordering, often producing anti-pragmatic interpretations in cases where the human is pragmatically expected to be the IO/EO; e.g. “she mailed him (the man) to it (the letter)”.

all in a free MOM system.¹¹ Data are lacking for the SWB systems. Cross-animate examples are more often neglected, corresponding to their rarity as discourse contexts, i.e. where the IO/EO is NONHUM and the DO is HUM, e.g. “she hired him (the driver DO) for it (the car AO)”. The cross-animate context is often crucial in deciding whether an OM sequence is ordered by InTop or GR.¹²

A final problem challenging all PB object-marking hypotheses is evidence from other East Benue-Congo languages, if not beyond. Preverbal object indexing systems restricted to anaphors, most often monosyllabic, occur in other branches of East Benue-Congo. In close proximity to the “NWB Cameroon, node 1, full SOM” systems are the OM systems of the Ogonoid and upriver Cross languages, e.g. Ibibio across the eastern Nigerian border. The surrounding postverbal NOM systems, even within the Cross branch, are similar to the postverbal NOM systems predominant in NWB. The coastal Cross area looks like a continuation of the “NWB Cameroon, node 1, full SOM” area as a relic area, similarly adjacent to NOMs. A comparable situation occurs in the widely separated north-west Nigerian area of the Jos Plateau, where some languages of the Kainji branch of East Benue-Congo also display similar systems, e.g. Kaje, Izere (cf. Blench & Kaze 2019: 12ff). As a much more distant branch of East Benue-Congo than Cross, Kainji suggests the possibility of a much more archaic status to some version of the PB OM system.¹³ The general issue of the historical relationship between SM-AUX-(OM)-OM-V...and current NWB SM-/#AUX (PRO) PRO V remains unresolved and continues to challenge any version of the PB OM hypothesis.

Acknowledgements

Special thanks go to Sara Pacchiarotti, Koen Bostoen, and two anonymous referees for valuable comments leading to many improvements in this chapter. I also thank Jenneke van der Wal, Nobuko Yoneda, Thilo C. Schadeberg and Daisuke Shinagawa for their communications and generous offering of relevant information. It is my pleasure to more generally thank all the participants of the

¹¹However, in some instances, the Rolong speaker expressed a decontextualised preference for an order corresponding to the obligatory order of prestigious fixed MOM varieties like Hurutshe.

¹²Willems (1970: 116) presents a fortuitous cross-animate example for Luba-Kasai *bà-bù-n-ship-él-è* [SM₂-OM₁₄-OM_{1sg}-kill-APPL-SBJV] “they would kill me for it [reason]”, revealing the InTop order AO_{NONHUM}-DO_{HUM} contrasting with the GR order DO-EO.

¹³However, Izere suggests a system in which the OM became suffixed to certain AUXs rather than prefixed to the following verb root. Thus, following AUX use of the COP *sen*, the Izere 1SG OM form is *ní*, transparently cognate with the PB 1SG SM/OM **ni*, but in past contexts the Izere 1SG OM has the form *tí*, as if suffixed to an older AUX **tV*. The other persons are similarly formed with an initial *t-* in past contexts.

November 2018 Ghent workshop on “Reconstructing Proto-Bantu Grammar” for informative and stimulating presentations. Finally, I thank Larry M. Hyman for encouraging me to participate in that workshop.

Abbreviations

AO	applied object	NOM	no OM system
APPL	applicative	NONHUM	non-human referent
ASP	aspect	NP	noun phrase
AUX	auxiliary	NWB	North-Western Bantu
CAUS	causative	OBJ	object
cl.	class	OE	object enclitic
CO	causative object	OM	object marker
ConTop	contextual topicality	PASS	passive
COP	copula	PB	Proto-Bantu
CWB	Central-Western Bantu	PFV	perfective
DO	direct object	PL	plural
DSJ	disjoint	PREF	pre-final
EB	Eastern Bantu	PRO	free pronoun
EO	extensional object	PROG	progressive
FOC	focus marker	PRS	present
FUT	future	PST	past
FV	final vowel	REFL	reflexive
GR	grammatical relation	RM	relative marker
HUM	human referent	SBJV	subjunctive
INAN	inanimate referent	SG	singular
INF	infinitive marker	SM	subject marker
INSTR	instrument	SOM	single-OM system
InTop	inherent topicality	SWB	South-Western Bantu
IO	indirect object	TAM	tense/aspect/mood
MOM	multi-OM system	TM	tense marker
NECB	North-East Coast Bantu (subgroup of EB)	TOP	topic
NEG	negative	TR	thematic role
		V	verb
		WWB	West-Western Bantu

*X historical reconstruction of X **X rejection of intended synchronic
X by L1 speakers

References

- Abad, Isidoro. 1928. *Elementos de la gramática Bubi*. Madrid: Editorial del corazón de Maria.
- Asiimwe, Allen. 2014. *Definiteness and specificity in Runyankore-Rukiga*. Stellenbosch: Stellenbosch University. (Doctoral dissertation).
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bearth, Thomas. 2003. Syntax. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 121–142. London: Routledge.
- Beaudoin-Lietz, Christa, Derek Nurse & Sarah Rose. 2004. Pronominal object marking in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 175–188. Cologne: Rüdiger Köppe.
- Bickmore, Lee S. 2007. *Cilungu phonology* (Stanford Monographs in African Languages). Stanford, CA: Center for the Study of Language & Information.
- Blench, Roger M. & Bitrus Kaze. 2019. *A dictionary of the Izere language of Fobur*. Cambridge: McDonald Institute for Archaeological Research, University of Cambridge.
- Bokamba, Georges D. 1971. Specificity and definiteness in Dzamba. *Studies in African Linguistics* 2(3). 217–237.
- Bolekia Boleká, Justo. 1991. *Curso de lengua bubi* (Colección Ensayos 8). Malabo: Centro Cultural Hispano-Guineano.
- Bostoen, Koen, Bernard Clist, Charles Doumenge, Rebecca Grollemund, Jean-Marie Hombert, Joseph Koni Muluwa & Jean Maley. 2015. Middle to late Holocene paleoclimatic change and the early Bantu expansion in the rain forests of Western Central Africa. *Current Anthropology* 56(3). 354–384.
- Cole, Desmond T. 1955. *An introduction to Tswana grammar*. London; New York: Longmans; Green.
- Contini-Morava, Ellen. 1983. Ranking of participants in Kinyarwanda: The limitations of arbitrariness in language. *Anthropological Linguistics* 25(4). 425–435.
- Creissels, Denis. 2006. Tswana verb morphology and the lexical integrity principle. *Lingue e linguaggio* 5(1). 49–66.
- de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87–162 + supplementary material online.

- Duranti, Alessandro. 1979. Object clitic pronouns in Bantu and the topicality hierarchy. *Studies in African Linguistics* 10(1). 31–45.
- Englebretson, Robert, Lisa Jeon, Jessica Li, Samantha Mauney, Anaí Navarro & Jonas Wittke. 2015. *A basic sketch grammar of Gikūyū* (Rice Working Papers in Linguistics 6). Houston, TX: Rice Linguistics Society, Rice University.
- Friesen, Lisa. 2002. *Valence change and Oroko verb morphology (Mbonge dialect)*. Grand Forks, ND: University of North Dakota. (MA thesis).
- Good, Jeff. 2005. Reconstructing morpheme order in Bantu: The case of causativization and applicativization. *Diachronica* 22(1). 3–57.
- Grégoire, Claire. 2003. The Bantu languages of the forest. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 349–370. London: Routledge.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas: With special reference to Africa* (Tokyo University of Foreign Studies, Studies in Linguistics 2), 109–141. Amsterdam: John Benjamins.
- Güldemann, Tom. 2022. Predicate structure and argument indexing in early Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 387–421. Berlin: Language Science Press. DOI: [10.5281/zenodo.7575831](https://doi.org/10.5281/zenodo.7575831).
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Hawkinson, Ann Katherine. 1986. *Bakweri verb morphology*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).
- Hyman, Larry M. 2003a. Basaa (A43). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 257–282. London: Routledge.
- Hyman, Larry M. 2003b. Suffix ordering in Bantu: A morphocentric approach. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2002*, 245–281. Dordrecht: Kluwer Academic Publishers.

- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.
- Katupha, Jose Mateus Muaria. 1991. *The grammar of Emakhuwa verbal extensions: An investigation of the role of extension morphemes in derivational verbal morphology and in grammatical relations*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Kimenyi, Alexandre. 1976. *A relational grammar of Kinyarwanda*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Maho, Jouni F. 2009. *NUGL online: The online version of the new updated Guthrie list, a referential classification of the Bantu languages*. https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf.
- Marlo, Michael R. 2014. Exceptional patterns of object marking in Bantu. *Studies in African Linguistics* 43(2). 85–123.
- Marlo, Michael R. 2015. On the number of object markers in Bantu languages. *Journal of African Languages and Linguistics* 36(1). 1–65.
- Marten, Lutz & Nancy C. Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30(2). 237–253.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Morrison, William McC. 1906. *Grammar and dictionary of the Buluba-Lulua language as spoken in the Upper Kasai and Congo Basin*. New York, NY: American Tract Society.
- Motingea Mangulu, André. 2008. Esquisse du parler des m̄nḡo-liinja d’Opala (bantou C60-70). *Annales Æquatoria* 29. 261–364.
- Mous, Maarten. 2005. The innovative character of object-verb word order in Nen (Bantu A44, Cameroon). In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics, with special focus on Bantu and Mande. Essays in honour of Yvonne Bastin and Claire Grégoire* (Collection Sciences Humaines 169), 411–424. Tervuren: Royal Museum for Central Africa.
- Ngoboka, Jean Paul. 2005. *A syntactic analysis of Kinyarwanda applicatives*. Durban: University of KwaZulu-Natal. (MA thesis).
- Nurse, Derek. 1999. Towards a historical classification of East African Bantu languages. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 1–41. Stanford, CA: Center for the Study of Language & Information.
- Nzang-Bie, Yolande. 2014. La transitivité en atsi, un dialecte du fang parlé au Gabon. *Journal of West African Languages* 41(1). 73–93.

- Odden, David. 1984. Formal correlates of focusing in Kimatuumbi. *Studies in African Linguistics* 15(3). 275–299.
- Pacchiarotti, Sara & Koen Bostoen. 2020. The Proto-West-Coastal Bantu velar merger. *Africana Linguistica* 26. 139–195.
- Pacchiarotti, Sara, Natalia Chousou-Polydouri & Koen Bostoen. 2019. Untangling the West-Coastal Bantu mess: Identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica* 25. 155–229.
- Polak, Louise. 1986. Les infixes (“prefixes objets”) du bantou et leur reconstruction. *Africana Linguistica* 10. 365–421.
- Ranero, Rodrigo, Michael Diercks & Mary Paster. 2013. *One of these things is not like the others: Epiphenomenal clitics in Kuria*. Unpublished ms. Claremont, CA: Pomona College.
- Redden, James E. 1979. *A descriptive grammar of Ewondo* (Occasional Papers on Linguistics 4). Carbondale, IL: Department of Linguistics, Southern Illinois University.
- Riedel, Kristina. 2009. *The syntax of object marking in Sambia: A comparative Bantu perspective*. Leiden: Leiden University. (Doctoral dissertation).
- Rubanza, Yunus Ismail. 1988. *Linear order in Haya verbal morphology: Theoretical implications*. East Lansing, MI: Michigan State University. (Doctoral dissertation).
- Rugemalira, Josephat M. 1991. What is a symmetrical language? Multiple object constructions in Bantu. In Kathleen Hubbard (ed.), *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: Special Session on African Language Structures*, 200–209. Berkeley, CA: Berkeley Linguistics Society, University of California.
- Stucky, Susan U. 1985. *Order in Makua syntax*. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign. (Doctoral dissertation).
- Valente, José Francisco. 1964. *Gramática umbundu; a língua do centro de Angola*. Lisbon: Junta de Investigações do Ultramar.
- Van de Velde, Mark. 2008. *A grammar of Eton* (Mouton Grammar Library 46). Berlin: Mouton de Gruyter.
- Van de Velde, Mark & Odette Ambouroué. 2017. The origin and use of a relative clause construction that targets objects in Orungu (Bantu, Gabon). *Studies in Language, International Journal sponsored by the Foundation “Foundations of Language”* 41(3). 615–637.
- van der Wal, Guenever Johanna. 2009. *Word order and information structure in Makhuwa-Enahara*. Leiden: Leiden University. (Doctoral dissertation).

- van der Wal, Jenneke. 2020. The AWSOM correlation in comparative Bantu object marking. In Peter W. Smith, Johannes Mursell & Katharina Hartmann (eds.), *Agree to agree: Agreement in the minimalist programme* (Open Generative Syntax 6), 199–234. Berlin: Language Science Press. DOI: 10.5281/zenodo.3541755.
- Vansina, Jan. 1995. New linguistic evidence and ‘the Bantu expansion’. *Journal of African History* 36(2). 173–195.
- Wald, Benji. 1975. Animate concord in Northeast Coastal Bantu: Its linguistic and social implications as a case of grammatical convergence. *Studies in African Linguistics* 6(3). 267–314.
- Wald, Benji. 1994. East Coast Bantu and the evolution of constraints on passivisation. *Sprache und Geschichte in Afrika* 15. 211–315.
- Wald, Benji. 1997. Instrumental objects in the history of topicality and transitivity in Bantu. In Rose-Marie Déchaine & Victor Manfredi (eds.), *Object positions in Benue-Kwa: Papers from a workshop at Leiden University, June 1994* (Holland Institute of Generative Linguistics Publications 4), 221–253. The Hague: Holland Academic Graphics.
- Wald, Benji. 2001. Substratal and superstratal influences on the evolution of Swahili syntax: Central East Coast Bantu and Arabic. *Sprache und Geschichte in Afrika* 16/17. 455–522.
- Whitehead, John. 1899. *Grammar and dictionary of the Bobangi language; as spoken over a part of the Upper Congo, West Central Africa. Compiled and prepared for the Baptist Missionary Society’s Mission in the Congo Independent State by John Whitehead*. London: Baptist Missionary Society & Kegan Paul, Trench, Trübner & Co.
- Willems, Emile. 1970. *Le tshiluba du Kasayi : pour débutants*. 4th revised edition. Luluabourg: Mission de Scheut.
- Woolford, Ellen. 2000. Agreement in disguise. In Vicki M. Carstens & Frederick Parkinson (eds.), *Advances in African linguistics: Papers presented at the 28th Annual Conference on African linguistics, held July 18–22, 1997, Cornell University* (Trends in African Linguistics 4), 103–117. Trenton, NJ: Africa World Press.
- Yokoyama, Tomohiro. 2016. Structure and features: The ordering of object markers in Kinyarwanda. Paper presented at Bantu 6, University of Helsinki.
- Yoneda, Nobuko. 2011. Word order in Matengo (N13): Topicality and informational roles. *Lingua* 121(5). 754–771.
- Zeller, Jochen. 2012. Object marking in isiZulu. *Southern African Linguistics and Applied Language Studies* 30(2). 219–235.

10 *On reconstructing the Proto-Bantu object marking system*

- Zeller, Jochen & Jean Paul Ngoboka. 2015. On parametric variation in Bantu, with particular reference to Kinyarwanda. *Transactions of the Philological Society* 113(2). 206–231.
- Zimmermann, Wolfgang & Paavo Hasheela. 1998. *Oshikwanyama grammar*. Windhoek: Gamsberg Macmillan.

Chapter 11

Agreement on Proto-Bantu relative verb forms

Mark Van de Velde

LLACAN - Langage, Langues et Cultures d’Afrique (CNRS, INaLCO, EPHE)

This chapter argues that Meeussen’s (1967) reconstruction of a Direct and an Indirect relative clause construction in Proto-Bantu (PB) is untenable, because there exists no scenario of morphosyntactic change that can lead from that reconstructed state of affairs to the relative clause constructions attested in contemporary Bantu. Although typologically unusual and widely attested across Bantu, relative verb forms that agree with the relativised noun phrase are not reflexes of a proto-construction with the same properties, but are the result of recent, parallel evolutions driven by a mechanism called the Bantu Relative Agreement (BRA) cycle. The only logically possible starting point from which the currently attested typological variation in Bantu relative clause constructions could have evolved is one in which relative verbs agreed with their subject. This conclusion has consequences for the reconstruction of the PB verbal template, which must have lacked a Pre-initial position.

1 Introduction

In his *Bantu Grammatical Reconstructions* (BGR), Meeussen (1967: 113, 120) reconstructs two relative clause constructions in Proto-Bantu (PB), called Direct and Indirect.¹ As for their verb forms, he only reconstructs their behaviour as agreement targets and the tone of their Final morpheme, stating that any other formal characteristics were “not within reach of reconstruction” at the time of writing.

¹Following common practice in the typological literature, names for Bantu-specific grammatical forms and categories such as *Direct relatives*, *Final* and *Pronominal prefix* are capitalised. The meaning and use of these terms will be discussed in this introduction.



The PB Direct and Indirect constructions, illustrated in (1), should therefore be interpreted as morphosyntactic templates, rather than full syntactic reconstructions.

(1) Partial PB reconstructions (Meeussen 1967: 113)

- a. *mò-ntò jò-dim-à i-pià*
1-person PPR1-cultivate-FV 5-garden
'a person who cultivates the garden'
- b. *i-pià dí-dim-á mò-ntò*
5-garden PPR5-cultivate-FV 1-person
'the garden that the person cultivates'
- c. *mò-ntò jò-tó-dim-ìd-á i-pià*
1-person PPR1-VPR1PL-cultivate-APPL-FV 5-garden
'the person for whom we cultivate the garden'

Direct relative clause constructions are used for subject relatives (1a) and for non-subject relatives when the relative verb has a lexical subject (1b) in Meeussen's PB. Their verb form has a prefix of the Pronominal prefix (PPR) paradigm that indexes the relativised noun phrase. Still according to Meeussen, the PB Indirect construction is used for non-subject relatives when the subject of the relative verb is not lexical (1c). Their verb form is characterised by a succession of two agreement prefixes. The first is a Pronominal prefix that indexes the relativised NP and the second is a prefix from the Verbal prefix (VPR) paradigm that indexes the subject of the relative verb (Meeussen 1967: 120). As can be seen in (1b), the relative verb precedes its lexical subject in Meeussen's reconstruction of PB non-subject relative clauses; see also Hamlaoui (2022 [this volume]).

Direct and Indirect relative constructions of the type exemplified in (1) are attested throughout the Bantu area. In contrast, I am not aware of any occurrence in the Benue-Congo languages outside of Narrow Bantu. Moreover, these constructions are formally highly unusual. In the position where other Bantu finite verb forms have a prefix that indexes the subject, the verb of the Direct construction has a prefix that indexes the relativised NP and is taken from a paradigm of agreement markers normally found on adnominal modifiers, whereas the verb of the Indirect construction has a succession of two agreement markers. Their unusual character, their omnipresence in Narrow Bantu and their absence elsewhere in Benue-Congo make the Direct and Indirect templates seemingly perfect candidates for reconstruction in PB. However, I argue that they should not be reconstructed in PB, nor in the proto-language of any of Bantu's major

genealogical subgroupings, such as those corresponding to the numbered nodes in the classification by Grollemund et al. (2015).

This conclusion is based on two observations. First, there are languages in all major subgroups of Bantu that have subject and/or non-subject relative clause constructions in which the relative verb starts with a vpr that indexes its subject. This type of agreement, here called TYPE SBJ and illustrated in (2), is typologically trivial and not reconstructed in PB by Meeussen.

- (2) Shi JD53 (Polak-Bynon 1975: 260)
áa-ba-lume Ludúunge a-a-rhum-íre
 AUG2-2-man 1.Ludunge vpr1-RPST.PFV-send-RPST.PFV
 ‘the men whom Ludunge sent’

Second, I will show that this widely attested type cannot be a reflex of the Direct relatives reconstructed by Meeussen, because there is no scenario of morphosyntactic change that can replace an adnominal agreement marker used to index the relativised noun phrase by a subject agreement marker.

In contrast, I will show that there exists a scenario of morphosyntactic change, the Bantu Relative Agreement (BRA) cycle (Van de Velde 2021), that can generate the full extent of observed variation in agreement types of Bantu relative clause constructions when the starting point is a PB morphosyntactic template in which the relative verb agrees with its subject, as in (2).

Before moving on to the main topic of this chapter, I will make a number of general methodological observations in §2. §3 provides an overview of the contemporary constructional variation in the domain of relative clauses that a successful reconstruction needs to account for. §4 then shows how the BRA cycle can account for this variation if we assume that PB relative verb forms indexed their subject by means of a prefix of the vpr paradigm. In contrast, §5 argues that there is no path from the reconstruction proposed in BGR to the current situation. §6 provides arguments for the assumption that the BRA cycle must not have been active yet at the PB stage. §7 explores some of the consequences of this chronology for the typological profile of the pre-stem domain in PB and Proto-Benue-Congo. A brief conclusion is given in §8.

2 Methodological preliminaries

2.1 Paradigms, functions and positions of verbal morphemes

When Bantuists analyse and gloss a verb form such as that in (1c), we can approach the first two prefixes *jú-tú-* from three different perspectives, viz. func-

tional, positional and paradigmatic. From a functional point of view, the prefix *tʰ-* is used to index the subject of the relative verb. It could therefore be glossed *SP*, short for *subject prefix*. Likewise, the prefix *jʉ-* is used to mark agreement with the relativised NP, so could be glossed *REL*, for instance. A second way to characterise these two prefixes is to situate them in the morphological slot-filler template of the Bantu verb. Using terminology introduced by Meeussen, the prefix *jʉ-* occupies the Pre-initial slot of the verb and *tʰ-* the Initial slot. Consequently, these morphemes can alternatively be glossed as *preIN* and *IN* respectively. A simplified version of Meeussen's slot-filler template is provided in (3). A number 1 in the second row means that maximally one morpheme can fill the position, whereas *n* stands for one or more. Brackets are used to mean that the position can be left empty, depending on the verb form. *IN* is short for Initial, *FO* for Formative, *IF* for Infix (the name of a prefix position, i.e. not an actual infix), *EXT* for Extension and *FV* for Final (vowel).

(3) Simplified version of Meeussen's slot-filler template

<i>preIN</i>	<i>IN</i>	<i>postIN</i>	<i>FO</i>	<i>IF</i>	<i>ROOT</i>	<i>EXT</i>	<i>FV</i>	<i>postFV</i>
(1)	(1)	(1)	(n)	(n)	1	(n)	1	(1)

Finally, the prefixes can be characterised in terms of the formal paradigms to which they belong. The *jʉ-* prefix belongs to the morphological paradigm called *Pronominal prefixes* (*PPR*) in Meeussen (1967), whereas *tʰ-* is a form from the *Verbal prefixes* (*VPR*) paradigm. Meeussen (1967) reconstructs five paradigms of class markers in PB, viz. *Nominal prefixes* (*NPR*), *Numeral prefixes* (*EPR*), *Pronominal prefixes* (*PPR*), *Verbal prefixes* (*VPR*) and *Object prefixes* (*OPR*). These are shown in Table 1.

In most circumstances the distinction between these three perspectives has little relevance for glossing. There is a general preference for functional labels, no doubt because they are the most transparent and universal. Positional and paradigmatic labels are hardly ever used to gloss verb forms. They are often actively discouraged by reviewers and editors, who point out that they are idiosyncratic (restricted to Bantu philology) and potentially misleading. The positional label Initial, for instance, is not necessarily used for the first morpheme of a verb form, nor is the *FV* always the last morpheme, and the so-called Infix is prefixed to the root, not infixed. The same is true for paradigmatic labels, where *Pronominal prefixes* do not always show up on pronouns and the term *Verbal prefixes* is arbitrarily assigned to only one of several paradigms of morphemes that can be prefixed to a verb stem. However, it is important to bear in mind that these positional and paradigmatic labels are not descriptive terms, but names for language-

Table 1: The Proto-Bantu class marker paradigms (Meeussen 1967: 97) [abridged]

	NPR	EPR	PPR	VPR	OPR
1SG	-	-	-	<i>n̄-</i>	<i>n̄-</i>
2SG	-	-	-	<i>ò-</i>	<i>kò-</i>
1PL	-	-	-	<i>tò-</i>	<i>tò-</i>
2PL	-	-	-	<i>mò-</i>	<i>mò-</i>
cl. 1	<i>mò-</i>	(<i>ò?</i>)	<i>jò-</i>	<i>ó-, á-</i>	<i>mò-</i>
cl. 2	<i>bà-</i>	<i>bá-</i>	<i>bá-</i>	<i>bá-</i>	<i>bá-</i>
cl. 3	<i>mò-</i>	(<i>ó?</i>)	<i>gò-</i>	<i>gó-</i>	<i>gó-</i>
cl. 4	<i>mì-</i>	(<i>í?</i>)	<i>gí-</i>	<i>gí-</i>	<i>gí-</i>
cl. 5	<i>ì-</i>	<i>dí-</i>	<i>dí-</i>	<i>dí-</i>	<i>dí-</i>
cl. 6	<i>mà-</i>	(<i>á?</i>)	<i>gá-</i>	<i>gá-</i>	<i>gá-</i>
cl. 7	<i>kì-</i>	<i>kí-</i>	<i>kí-</i>	<i>kí-</i>	<i>kí-</i>
cl. 8	<i>bì-</i>	<i>bí-</i>	<i>bí-</i>	<i>bí-</i>	<i>bí-</i>
cl. 9	<i>n-</i>	(<i>ì-</i>)	<i>jì-</i>	<i>jí-</i>	<i>jí-</i>
cl. 10	<i>n-</i>	<i>í-</i>	<i>jí-</i>	<i>jí-</i>	<i>jí-</i>
...

or family-specific categories, which is conventionally signalled by their initial capitalisation. Hence, it makes perfect sense to write that not all verbal prefixes are Verbal prefixes.

However, when discussing agreement on relative verb forms, it is essential to distinguish between the three above-mentioned perspectives as clearly as possible, since we are interested in determining which slot in the verbal template is occupied by a marker from which paradigm, indexing which element in the syntactic context. Examples in this chapter will mostly be glossed using positional labels, because their assignment is the least dependent on analysis.

There are obviously strong correlations between the three alternative ways in which a verbal morpheme can be characterised, which is definitely another reason why the distinction is rarely made. For instance, morphemes in the Initial position tend to be Verbal prefixes indexing the subject in non-relative verb forms, but there are two complications. First, it is not always clear whether the participant that is indexed by the Initial morpheme is best analysed as a subject, e.g. in some of the so-called inversion constructions. Second, in many languages, including Meeussen's Proto-Bantu, there is strictly speaking more than one paradigm of Verbal prefixes. Paradigms of agreement prefixes are normally minimally

differentiated in the Bantu languages, with only a minor formal distinction in a couple of classes. Since Meeussen reconstructs two prefixes for class 1 in his vpr paradigm, viz. *á-* and *ú-*, his PB actually has two paradigms of Verbal prefixes, which could be abbreviated as A-vpr and U-vpr. The notion of *Verbal prefixes* is a useful cover term for the A-vpr and U-vpr paradigms in Proto-Bantu and other Bantu languages with similar minor paradigmatic distinctions.

Turning to relative verb forms such as those illustrated in (1), there is only a partial correlation between position in the verbal template on the one hand and paradigm and function on the other. The Initial slot can be occupied by a vpr that indexes the subject of the relative verb (1c) or by a ppr that marks agreement with the relativised NP (1a, 1b). The Pre-initial slot, if present, is always occupied by a ppr marking agreement with the relativised NP. Paradigm and macro-function correlate by definition: Verbal prefixes are used to index an argument of the verb (which is normally the subject, but could be a topic, hence “macro” function). Pronominal prefixes are used to mark class agreement in a relation of adnominal modification between the relativised NP and the relative verb. In subject relatives, where the relativised NP and the subject of the relative verb are co-referential, the choice of paradigm shows us which kind of syntactic relation is being marked: verb-argument (vpr) or head noun-modifier (ppr).

2.2 Distributional criteria versus scenarios

Assessing the validity of reconstructions proposed in BGR is complicated by the lack of an explicit presentation of data and methodology. Some discussion of the methods used for reconstructing grammar can be found in publications on specific grammatical topics by other members of the *Lolemi* programme,² such as Nsuka-Nkutsi’s (1982) work on relative clause constructions. These methodological remarks give an indication of the decision-making process that may have led to the reconstructions proposed in BGR. It is clear, for instance, that the geographical distribution of currently attested phenomena played a major role, such that forms or patterns that have a very wide or a highly discontinuous distribution were readily recognised as retentions. Moreover, grammatical quirks that are attested in only a handful of non-adjacent languages also made it into PB, such as Burssens’ rule changing the word-final *HL sequence of a head noun

²The *Lolemi* programme was a large research project at the Royal Museum for Central Africa led by A.E. Meeussen, which started in the early 1960s and aimed at using all grammatical descriptions of Bantu languages available at that time for the historical-comparative study of Bantu morphology and syntax.

into *HH when immediately followed by a connective relator, a possessive pronoun or a relative verb form with an initial *H (Meeussen 1967: 106; Nsuka-Nkutsi 1982: 58). This is no doubt because it was deemed unlikely that such a seemingly random grammatical phenomenon could emerge several times independently. Finally, when alternative candidates for reconstruction have a comparable geographical distribution, there appears to have been a tendency for reconstructing the more complex or elaborate situation and to assume that the more likely diachronic evolution in Bantu is simplification.

Two things are lacking in this approach. One is the pursuit of a detailed and credible diachronic scenario that can lead from a proposed reconstruction to the totality of currently attested patterns. The other is awareness of recurrent morphosyntactic changes that can have occurred independently at different times and places, such that cognate morphemes with a similar function and morphosyntactic behaviour can be the outcomes of parallel evolutions, rather than reflexes of the same proto-form. I will briefly illustrate this with two aspects of Meeussen's PB reconstruction that may need to be reconsidered, viz. the reconstruction of an augment and that of a full paradigm of possessive pronouns.

The augment is a prefix or proclitic that precedes the class prefix of nouns and some adnominal or nominalised modifiers (de Blois 1970). Formally, it is typically either identical to the Pronominal prefix (PPR) or it consists of the vowel of the PPR. Its function, if any, differs from language to language. Often, one can only list the conditions in which it does or does not appear, and the former tend to be far more numerous than the latter. Augments can be found all over the Bantu speaking area. Their loss is also well documented in many languages, because they often leave formal traces, such as so-called "latent augments" (de Blois 1970; Grégoire & Janssens 1999). This is probably why Meeussen (1967: 99) reconstructs an augment in PB, more precisely as a weak demonstrative in prenominal position that functioned as an anaphoric marker in specific syntactic contexts. However, the pre-posing of demonstratives is still a common process in Bantu, including in languages where the noun usually has an augment, which tends to be deleted in the presence of a prenominal demonstrative (Van de Velde 2005). Such prenominal demonstratives are very similar to the augment as reconstructed by Meeussen and arguably represent a new cycle of augment creation. Moreover, there are several Bantu languages in which the augment appears to be a relatively recent innovation, or that have (traces of) an older augment coexisting with a more recently developed one (Van de Velde 2019: 254–255), as in Nyakyusa M31 in (4). Nyakyusa has two paradigms of augments, one with a vocalic shape (4a) and one with a CV- shape (4c). The first one is part of the default form of the noun and has no clear semantic value, whereas the more recent CV-shaped augment is

an anaphoric marker, in line with Meeussen's reconstructed augment. Both are cognate with the proximal demonstrative *ɔ-jɔ* in (4b).

- (4) Nyakyusa M31 (Bastian Persohn, p.c.)
- a. *ɔ-mu-ndɔ* 'the person'
 - b. *ɔ-mu-ndɔ ɔ-jɔ* 'this person'
 - c. *jɔ-mu-ndɔ* 'the very person'

The Nyakyusa data in (4) show that augments can emerge and disappear repeatedly. Since the demonstrative modifiers from which they develop are at least partially cognate, augments in different Bantu languages are cognate as well, without necessarily being reflexes of a single PB paradigm. The recurrent nature of augment creation and erosion makes it impossible to know whether an augment existed at any given proto-stage, much less at which state it was in its grammatical evolution.

The second illustration concerns the paradigm of possessive pronouns. Pronominal forms are extremely unstable in Bantu, with morphological material constantly being added and deleted (as shown, for instance, in Kamba Muzenga (2003), and Idiatov (2022 [this volume])). Consequently, Meeussen (1967: 107) points out that it is very difficult to reconstruct specific proto-forms. Instead, he tentatively provides one out of a number of alternative reconstructions for the forms that could have made up the PB paradigm of possessive pronouns. We are here less interested in the proto-forms of the pronouns than in the structure of their paradigm, and more precisely in the question of how many forms it contained. Among contemporary Bantu languages, there is a typological distinction between those with a full and those with a reduced paradigm. Languages with a full paradigm have a possessive stem for all the nominal classes to which a possessor can belong. The Mituku D13 examples in (5) provide a partial illustration of a full paradigm: possessors expressed by means of a noun of class 3, 4, 12 or 13 are each indexed by means of a different possessive stem (bolded in the examples).

- (5) Mituku D13 (Stappers 1973: 32)
- a. *meli y-a**ɔ*** 'its roots' (of a tree, cl. 3)
 - b. *meli y-**ay**ɔ* 'their roots' (of trees, cl. 4)
 - c. *beópi b-**ák**ɔ* 'its wings' (of a bat, cl. 12)
 - d. *beópe b-**át**ɔ* 'their wings' (of bats, cl. 13)

In contrast, languages with a reduced paradigm have only two stems for third person possessors: one for the singular and one for the plural. In the Mwera P22 examples in (6), the human class 1 possessor is indexed by means of the same pronominal stem as the class 14 possessor, the one also used for all other 3SG possessors.

- (6) Mwera P22 (Harries 1950: 59)
- a. *meyo g-a:kwe* ‘her eyes’ (the woman’s, cl. 1)
 - b. *kunona kw-a:kwe* ‘its tastiness’ (the beer’s, cl. 14)

Both types of paradigm are found throughout the Bantu area (Van de Velde Forthcoming), so that current geographical distributions do not provide a clear hypothesis for reconstruction.³ Meeussen reconstructs a full paradigm, perhaps due to a general preference for the more complex of alternative reconstructions. However, in terms of diachronic scenarios, the path from a reduced to a full paradigm is much more likely than the reverse path. Possessive pronouns for third person possessors of class 2 upwards in full paradigms are transparent genitive (aka connective) constructions with a personal pronoun in the modifier position. This can be seen in (5), where the possessive stems consist of the genitive linker *a*, followed by a class marker and the personal pronoun stem *ɔ*. The scenario for the emergence of full paradigms is therefore trivial. In contrast, we would expect much less formal transparency in full paradigms if they had been handed down from PB. Moreover, there is no obvious reason why so many Bantu languages would have reduced their original paradigm. The hardest thing to explain in a scenario of paradigmatic reduction that must have repeated itself independently on numerous occasions is the uniform all-or-nothing nature of the typological distinction. All examples of reduced paradigms known to me have six members (one for each person and number) and all full paradigms have as many third person forms as they have noun classes, on top of first and second person forms. A plausible scenario of paradigmatic reduction would have resulted in partially reduced systems in at least some languages, e.g. along lines of animacy.

Now that more Bantu descriptive and comparative studies are available, we can and should be more attentive to attested patterns of morphosyntactic change in an attempt to verify whether plausible diachronic scenarios lead from proposed PB reconstructions to the current morphosyntactic variation. I will do this for relative clause constructions in the remainder of this chapter.

³Full paradigms may be absent in zones A and B, i.e. the far North-West. This needs to be verified. If they are, this would strengthen the case for reconstructing a reduced paradigm in PB, as pointed out by Koen Bostoen.

3 Typological variation in Bantu relative clauses

According to much of the literature starting with Nsuka-Nkutsi (1982), the three types of agreement patterns in Table 2 can be found on the relative verb in contemporary non-subject relative clause constructions.

Table 2: Agreement patterns on the relative verb in contemporary non-subject relative clause constructions

Type	Agreement pattern
SBJ	agreement with the subject only
NPreI-SBJ	agreement with the relativised NP and the subject (Meeussen's Indirect)
NPreI	agreement with the relativised NP only (Meeussen's Direct)

Type SBJ agreement is illustrated in (2), type NPreI-SBJ in (1c) and type NPreI in (1a–1b). These three agreement types strongly correlate with the choice of a paradigm of agreement markers. Agreement of type SBJ tends to be expressed by means of a Verbal prefix and agreement of type NPreI by a Pronominal prefix. Consequently, agreement of type NPreI-SBJ is normally expressed by a PPR-vPR-succession. Since Pronominal prefixes are typically used on adnominal modifiers to mark agreement with their head noun, this correlation is not surprising.

This general picture has to be clarified and completed on three accounts. First, as will be illustrated below, an additional marker of agreement with the relativised NP can occur in types NPreI and NPreI-SBJ, giving rise to two more agreement types, namely type NPreI-NPreI and type NPreI-NPreI-SBJ. Second, contrary to what appears to be generally assumed in the literature, all types of agreement can be found in subject relatives as well as in non-subject relatives. Third, in many relative clause constructions across the Bantu domain, agreement markers on the verb belong to a morphological paradigm that formally differs from the paradigms of both Verbal prefixes and Pronominal prefixes and that is found exclusively in relative verb forms. I will use the term *Relative prefixes* (RPR) to refer to such paradigms of agreement markers dedicated to relative verb forms. Nsuka-Nkutsi (1982) did not recognise a separate RPR paradigm because he relied on a binary distinction, identifying every paradigm of agreement markers as PPR as soon as it diverges from the vPR paradigm. As we saw in §2.1, the most important distinction is between paradigms that contain first and second person forms and those that do not. Dedicated RPR paradigms tend to be of the latter type.

Indeed, in subject relative clauses, the distinction between adnominal NPrel agreement and SBJ agreement is easiest to see with first or second person relativised NPs, because paradigms of adnominal agreement markers only have third person forms. Example (7) illustrates agreement type NPrel-SBJ in a subject relative clause. The first person plural form is indexed twice on the relative verb, once as its relativised NP (by a third person plural prefix of class 2 in Pre-initial position) and once as its subject (by a first person plural prefix in Initial position).

- (7) Yao P21 (Sanderson 1922: 73)
uwe [*u-tu-li* *ŵa-yao*]
 1PL preIN2-IN1PL-be 2-yao
 ‘we who are Yao’

In contrast, the second person plural pronoun in (8a) is indexed twice on the relative verb as its relativised NP and never as its subject, illustrating agreement of type NPrel-NPrel. Both agreement prefixes, *a-* and *ba-*, are class 2 forms, i.e. third person forms. Both differ from the second person plural prefix *mu-* seen in the following main verb *mu-raire*. Example (8b) illustrates agreement of type NPrel-NPrel-SBJ. It also shows how agreement prefixes in relative verbs can be formally distinct from those of both the vPR and the pPR paradigms. In Nkore-Kiga JE13/14 non-subject relative clause constructions, prefixes that would have an /a/ in the vPR or pPR paradigm have an /u/ in the rPR paradigm, hence *a-bu-* in (8b), instead of *a-ba-*. Something similar can be observed in (7). In Yao the class 2 rPR in the Pre-initial slot is *u-*, instead of the *a-* we would have found in the class 2 form of other paradigms.

- (8) Nkore-Kiga JE13/14 (Taylor 1985: 23)
- a. *imwe* [*a-ba-tuura* *aha*], *mu-raire* *buhooro*
 2PL preIN2-IN2-live here IN2PL-sleep well
 ‘You, who live here, how are you (lit. did you sleep well)?’
- b. *a-ba-ntu* *a-bu-tu-twire* *omu n-si*
 AUG2-2-person preIN2-preIN2-IN1PL-live.PFV LOC 9-country
y-aabo
 pPR9-their
 ‘the people in whose country we live’

Individual languages can have multiple constructions that belong to different agreement types. Moreover, individual constructions can show a split in agreement type depending on properties of their agreement controllers. With respect

to his Indirect type (= type NPrel-SBJ), Meeussen (1971) points out that in some constructions it only appears when the subject of the relative verb is of the first or second person. With a third person subject, these constructions are of the Direct type (i.e. with type NPrel agreement). Following Meeussen, I will refer to these as Luba-type constructions, and to “normal” type NPrel-SBJ constructions that do not show such a split as Lega-type constructions. Luba-type constructions can be found in the East of the DRC and in Eastern Angola. The Mituku D13 non-subject relative clause construction in (9) is an example. In (9a) the relative verb has a first person subject and agreement is of type NPrel-SBJ: the relativised NP is indexed on the verb by the prefix *ɔ́-* and the subject by the prefix *ntɔ́-*. In (9b) the relative verb has a third person subject (expressed by means of the postverbal independent pronoun *bô*) and agreement of type NPrel.

(9) Mituku D13 (Stappers 1973: 59)

- a. *mɔ́-ntɔ́ ɔ́-tɔ́-túma*
1-person preIN1-IN1PL-send
‘the person we send’
- b. *mɔ́-ntɔ́ ɔ́-túma bô*
1-person IN1-send they
‘the person they send’

Finally, relative clause constructions of all agreement types across Bantu can involve one or more optional or obligatory relativisers. There is formal variation between the attested relativisers, which is due to the great number of their possible sources (different types of demonstratives, personal pronouns, connective relators, etc.) and to the fact that many of them are clearly recent innovations. Since relativisers immediately precede the relative verb in many cases, it is often impossible to determine in a non-arbitrary way whether one is dealing with an independent relativiser or with an agreement prefix that indexes the relativised NP, a recurrent ambiguity that is inherent to the BRA cycle. This indeterminacy can be illustrated by the alternative ways in which Nkore-Kiga non-subject relatives such as (10) have been analysed in the literature.

(10) Nkore-Kiga JE13/14 (Taylor 1985: 22)

- a-ka-cumu ku w-aakozesa*
AUG12-12-pen preIN12 IN2SG-used
‘a pen you used’

Taylor writes *ku* separately from the verb, apparently analysing it as an independent relativiser, whereas Nsuka-Nkutsi (1982: 124) treats it as a prefix of the relative verb. In (10) Taylor's analysis is reflected in the orthography and Nsuka-Nkutsi's in the glosses.

4 From PB agreement type SBJ to the present

The goal of this section is to show how the BRA cycle can generate every construction attested in contemporary Bantu if we start from a proto-language that had subject and non-subject relative clause constructions with type SBJ agreement. Translated into Meeussen's PB, this starting point looks like (11). Note that in (11a) the Verbal prefix indexes the noun *mò-ntò* 'person' as the subject of the relative verb, not as the relativised NP. The construction in (11a) does not differ from a non-relative clause construction and as such is ambiguous between the readings 'the person who cultivates' and 'the person cultivates'. The examples in (1) and (11) are only partial reconstructions that concentrate on agreement. There may have been a relativiser and/or prosodic or morphological differences between relative and non-relative constructions. That being said, many instances of morphosyntactic ambiguity between relative and non-relative constructions exist in the contemporary Bantu languages as well.

(11) Meeussen's (1967: 113) PB examples reanalysed

- a. *mò-ntò á/ó-dim-á i-pià*
 1-person vpr1-cultivate-FV 5-garden
 'a person who cultivates the garden'
- b. *i-pià á/ó-dim-á mò-ntò*
 5-garden vpr1-cultivate-FV 1-person
 'the garden that the person cultivates'
- c. *mò-ntò tó-dim-id-á i-pià*
 1-person vpr1PL-cultivate-APPL-FV 5-garden
 'the person for whom we cultivate the garden'

The great majority of contemporary instances of type SBJ agreement can be considered direct reflexes of the proto-situation illustrated in (11). The other agreement types are the result of the BRA cycle, schematised in Figure 1. The stages will be commented on and illustrated in what follows. Figure 1 is illustrative rather than exhaustive, in that it only schematises the BRA cycle applied to

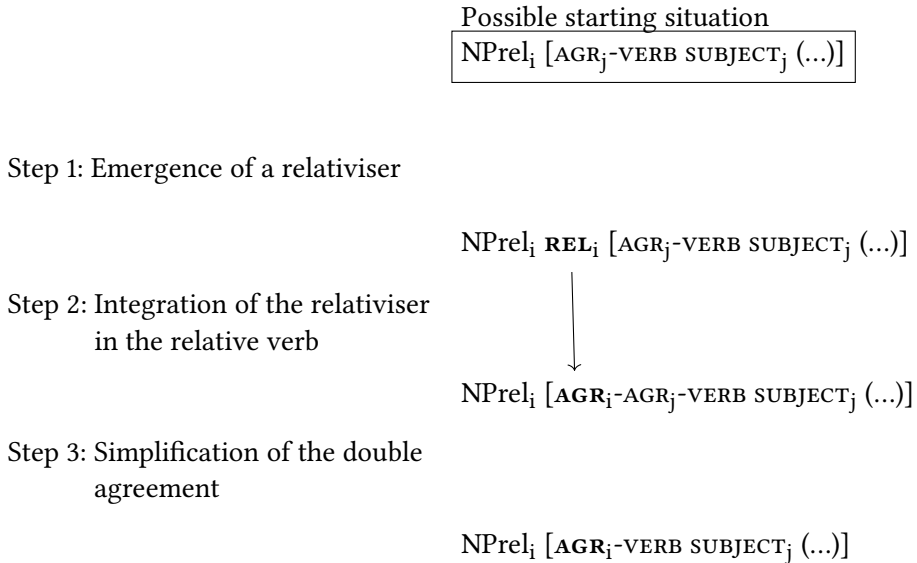


Figure 1: Illustration of a possible BRA cycle

non-subject relatives with a postverbal lexical subject. Subscripts i and j signal a relation of agreement between two elements. REL is short for relativiser.

The first stage of the BRA cycle involves the emergence of a relativiser in-between the relativised NP and the relative clause, which can originate in a demonstrative, a personal pronoun, a connective relator, or another element. Whatever its origin, the relativiser agrees with the relativised NP. The overwhelming variety of origins and forms of this relativiser, and its random distribution in the Bantu domain,⁴ make it clear that its presence is in most cases a recent innovation and therefore that the BRA cycle is often and easily initiated in the Bantu languages. The first stage of the BRA cycle can be illustrated with the Chokwe K11 example in (12).

- (12) Chokwe K11 (Kawasha 2008: 50)
- ly-onda [lízé a-a-mbách-ile pwo]*
- 5-egg REL5 IN1-TNS-carry-RPST 1.woman
- ‘the egg which the woman carried’

Relative verb forms with agreement of type NPrel-SBJ (i.e. “Indirect” relatives) are the result of the second stage in the BRA cycle: the gradual integration of

⁴For a detailed discussion of the origins, use and distribution of relativisers in Bantu, see Nsuka-Nkutsi (1982: 1–93).

an erstwhile independent relativiser into the verb. Evidence for this stage can be found in the sometimes unexpected shape of the prefix in Pre-initial position that indexes the relativised NP, due to the fact that it is usually a reflex of a morphologically complex relativiser, rather than of a Pronominal prefix. The unexpected *bu-* shape (versus expected *ba-*) of the Pre-initial in the Nkore-Kiga example (8b) may illustrate this, although its origin is currently not clear. Moreover, in non-subject relative clause constructions with a lexical subject, there is a very strong correlation between agreement of type NPrel-SBJ and the postverbal position of the lexical subject and between type SBJ agreement and a preverbal subject. The straightforward historical explanation in terms of the BRA cycle is that a preverbal lexical subject hampers the integration of a relativiser into the relative verb form (see also Hamlaoui (2022 [this volume])).

Relative verbs with agreement of type NPrel (“Direct” relatives) are the result of the third and last stage of the BRA cycle, viz. the reduction of the succession of two agreement prefixes to a single one. This can happen through merger or through the deletion of one of the prefixes. In theory, when a PPR-VPR succession of prefixes is simplified through the deletion of one of them, the surviving prefix can be the second, i.e. the one that indexes the subject. It is impossible to know whether this may have happened in the history of a construction with agreement of type SBJ, but there are some rare examples of reduction through merger in which the newly forged agreement marker indexes the subject. For instance, the initial *a* of the class 2 RPR *abá* in (13a) from Mbagani L22 is very likely a reflex of the initial *a* that also shows up on (optional) relativisers (13b). Crucially, the resulting subject index is the reflex of a prefix that had been there from the start, accreted by an invariable initial element.⁵

(13) Mbagani / Binji L22 (van Coillie 1948: 272)

- a. *di-kamá* [*abá-bátúlé·la*]
 5-foot REL.IN2-cut_off DEM
 ‘the foot that they cut off’
- b. *di-kamá* [(*a*)*di abá-bátúlé·la*]
 5-foot REL5 REL.IN2-cut_off DEM
 ‘the foot that they cut off’

⁵The Nguni S40 languages have a non-subject relative clause construction with a Relative Prefix of the shape (l)V(-)VPPr-, in which the quality of the first vowel (here represented as V) is determined by that of the vowel of the Verbal prefix. This appears to suggest that this Relative prefix originates in a form that contained a succession of two prefixes that both index the subject, which is not obviously compatible with the BRA cycle. However, this Nguni Relative prefix is similar to the one found in Mbagani. Its initial *a* comes from a demonstrative stem *la* and undergoes anticipatory assimilation.

The BRA cycle also accounts for the minority patterns mentioned in §3, such as NPrel-NPrel agreement, which is the result of successive applications of the cycle. It explains why dedicated paradigms of *rpr*'s have emerged in many languages, either as reflexes of relativisers, or of mergers between two prefixes; and why there is no fundamental distinction in agreement types between subject and non-subject relatives. A relativiser can appear before relative verbs of any agreement type, because the BRA cycle can be re-initiated while constructions are halfway or fully through a previous cycle. The BRA cycle also makes perfect sense of constructions of the Luba-type, which have agreement type NPrel-SBJ when their subject is of the first or second person, but agreement type NPrel elsewhere. These constructions are halfway between stage 2 and stage 3 of the cycle. The reason why reduction has not taken place with subject agreement prefixes of the first and second person is that the non-lexical subject in these constructions is expressed by means of a postverbal pronoun, whose paradigm lacks first and second person forms in languages with Luba-type constructions (Nsuka-Nkutsi 1982: 42, 222).

The fact that BRA cycles are easily started and that they can evolve fast is illustrated by languages that have multiple alternative relative clause constructions that can be shown to be at different stages of a BRA cycle. Van de Velde (2022) illustrates this with examples from Punu B43, taken from Blanchon (1980).

5 No path from Meeussen's PB to the present

As a reminder, Meeussen's (1967) partial reconstruction of relative clause constructions has three features that are relevant for relative verbs as agreement targets:

1. a distinction between a Direct (*ppr*-) and an Indirect (*ppr*- *vpr*-) construction;
2. the *ppr* indexes the relativised NP and the *vpr* the subject of the relative verb;
3. the Indirect construction was used for non-subject relatives if the relative verb has a pronominal/grammatical (i.e. non-lexical) subject and the Direct construction elsewhere.

The picture we find in contemporary Bantu differs considerably from this reconstructed situation. First, there are some additional agreement types, namely

type SBJ, NPrel-NPrel and NPrel-NPrel-SBJ. Moreover, subject relatives and both types of non-subject relatives (with lexical versus grammatical subject) can belong to any of the attested agreement types. Third, a wide variety of relativisers has emerged, distributed randomly over the Bantu domain, as well as a number of dedicated RPR paradigms in individual languages. For the sake of the argument, we will assume in this section that Meeussen's partial reconstructions are valid. Starting from that assumption, we will try and identify paths of morphosyntactic change that can lead from that reconstruction to the morphosyntactic variation that is currently attested in Bantu. As will become clear, this turns out to be impossible.

Nevertheless, if we take Meeussen's reconstruction as the starting point, the BRA cycle could account for much of the needed morphosyntactic change. For instance, the evolution from Indirect to Direct relatives involves the type of prefix reduction found in stage 3 of the BRA cycle. The constant emergence of new relativisers corresponds to stage 1 of the cycle, and type NPrel-NPrel agreement corresponds to stage 2 of a BRA cycle that has a 'direct' relative as its starting point.

However, since the BRA cycle cannot generate a VPR that indexes the subject of the relative verb, contemporary constructions with agreement of type SBJ are problematic, and so are constructions with agreement of type NPrel-SBJ that are used for subject relatives or non-subject relatives with a lexical subject. I will address the problems arising from Meeussen's reconstruction in order of increasing complication. The least complicated are non-subject relatives with a pronominal subject, as these are reconstructed as Indirect, so that a VPR that indexes the subject of the relative verb is already present from the start. All we need to assume is the simplification of the original PPR-VPR- succession of prefixes through the loss of the PPR in the constructions that today have agreement of type SBJ, as schematised in (14).

- (14) BGR attested agreement type
-
- relativised NP_x [PPR_x-VPR_y-VERB...] → relativised NP_x [VPR_y-VERB...]

Although such an evolution is in theory possible, it is impossible to show that it has taken place, because it leaves no traces. We do know that when reduction of a PPR-VPR succession takes place through merger, the resulting RPR tends to be a continuation of the PPR, in that it indexes the relativised NP. The Mbagani example in (13) is one of the few clear counterexamples that I could find. This may point to a tendency for the PPR to survive and the VPR to be deleted, and

constructions with a relativiser.⁶ He proposes a functional motivation for this correlation: analogical levelling is more likely to occur if it does not lead to ambiguity between a relative and a non-relative construction, an ambiguity that is lifted by the presence of a relativiser. There are two problems with an explanation in terms of analogical levelling. First, analogical change is not as rigidly systematic as needs to be assumed in this case. Second, while it could explain formal changes in parts of paradigms of agreement markers, it cannot explain a change in agreement controller. This second problem dismisses the hypothesis of analogical levelling for non-subject relatives with a lexical subject. If the inherited PPR in the Initial position of their relative verb were to acquire the shape of a VPR, this prefix would still index the relativised NP, rather than the subject of the relative clause, which is *not* the type of agreement we find in the contemporary Bantu constructions with an Initial VPR. Example (16a) repeats Meeussen's pseudo-PB example from (1b). Analogical levelling of the PPR and VPR paradigms would bring about no changes whatsoever, as the two paradigms were identical from the outset for class 5 controllers. However, what we find in the contemporary constructions with a VPR in their Initial position is a reflex of (16b), repeated from (11b), i.e. my proposal for reconstruction translated in Meeussen's PB.

(16) Meeussen's (1967: 113) PB examples reanalysed

- a. *i-pià dí-dim-á mò-ntò*
 5-garden PPR5-cultivate-FV 1-person
 'the garden that the person cultivates'
- b. *i-pià á/ó-dim-á mò-ntò*
 5-garden VPR1-cultivate-FV 1-person
 'the garden that the person cultivates'

Turning to the first problem with analogical levelling, we will now see why analogical levelling does not work for subject relatives either. According to Meeussen's (1967: 97) PB reconstruction, the VPR and PPR paradigms differ from each other in that the PPR paradigm lacks first and second person forms and that it has a low tone in classes 1 and 9, where the Verbal prefixes are high. Moreover, there is a segmental difference in class 1, where the PPR is **jò-* and the VPR **á-* or **ó-*. Therefore, if we take the reconstructed PB situation as a starting point, three formal changes are needed in order for the PPR and the VPR paradigms to

⁶Note that what Nsuka-Nkutsi counts as Pronominal prefixes includes dedicated Relative prefixes. According to my counts using his sample, 42% of subject relative clause constructions have a VPR. This percentage goes down to 28% in constructions that lack a relativiser, but it goes up to 67% in constructions with a relativiser.

collapse in the third person: a tone change in class 1, a tone change in class 9 and a segmental change in class 1. These three changes are each very minor and individually plausible, but in the context of analogical change they have to be counted as separate evolutions. In contrast, the formal changes needed for a merger of both paradigms are by no means minor in the case of first and second person controllers. In the singular, a PPR **jò-* would have to change to VPR **n-* (first person) or **ò-* (second person). In the plural, the PPR **bá-* has to change to VPR **tò-* (first person) or **mò-* (second person). All in all, these are seven formal changes, of which four are minor (including 2SG) and three are radical. Now, if we look at the geographical distribution of subject relative clause constructions in Nsuka-Nkutsi's sample, we find that those with a VPR Initial are found in every Guthrie zone (albeit marginally in zones E, H and J) and those with a PPR Initial in all zones except N and S. Whatever the direction of change one wishes to assume, one has to conclude that changes must have taken place recently. Otherwise, we would find more clustering along regional and genealogical lines. In other words, the exact same set of seven formal changes motivated by analogical levelling should have produced itself dozens of times independently. This is by no means plausible and it is not what we find in languages where analogical levelling can be shown to have taken place. In Cuwabo P34, for instance, the PPR and VPR paradigms have collapsed in their third person forms (17), but relative clauses with a relativised NP of the first or second person have a class 1 prefix in the Initial slot of the relative verb (18), showing that we have agreement of type NPrel and a PPR in Initial position.

(17) Cuwabo P34 (Guérois & Creissels 2020: 478)

- a. *Múyaná oṅgúlíhá nígagádda.*
mú-yaná o-ní-gul-ih-a ní-gagádda
 1-woman IN1-IPFV-buy-CAUS-FV.CJ 5-dry_cassava.H1D
 'The woman is selling dry cassava.'
- b. *múyaná oṅgúlíha nígagádda*
mú-yaná o-ní-gul-ih-a ní-gagádda
 1-woman IN1-IPFV-buy-CAUS-FV.REL 5-dry_cassava
 'the woman who is selling dry cassava'

(18) Cuwabo P34 (Guérois & Creissels 2020: 480)

- nootéene íy' óom̃vívéeéívatákùl' áapa*
ni-oté=ene íyo o-ní-viveéri va-tákulu ápa
 1PL-all=INT 1PL.PRO IN1-IPFV-live.REL 16-9A.house 16.DEM.I
 'all of us who live in this house'

Another example of a language where analogical levelling has taken place is Orungu B11b. Here, too, the resulting picture differs considerably from what we find in languages where relative verbs have a *vpr* that indexes the subject. The tonal differences between the *ppr* and *vpr* paradigms have disappeared in Orungu, leaving only a segmental distinction in class 1 and in the first and second person forms, which are absent in the *ppr* paradigm (Van de Velde & Ambouroué 2017). Moreover, the choice between a *ppr* and a *vpr* is free in class 1 in some relative clause constructions, suggesting that partial analogical levelling is still ongoing.

The near impossibility of full analogical levelling across the entire paradigm having taken place independently dozens of times can be contrasted with the trivial nature of the changes that make up the BRA cycle and that can easily explain the evolution from a construction with a *vpr* that indexes the subject to one with a *ppr* indexing the relativised NP in a relative verb form. Likewise, in order to explain a certain correlation between the presence of a relativiser and a *vpr* in Initial position, Nsuka-Nkutsi had to make the awkwardly functionalist claim that ambiguity avoidance would have blocked analogical levelling time and again in the absence of a relativiser. Compare this to the straightforward explanation that can be found in the application of the BRA cycle to a situation in which most languages have inherited relative verbs with a *vpr* that indexes the subject in Initial position: some have never started a BRA cycle (type SBJ, *vpr*-, no relativiser); some are in Stage 1 (type SBJ, *vpr*-, relativiser), some are in Stage 2 (type NPrel-SBJ, *ppr-vpr*) and some are in Stage 3 (type NPrel, *ppr*). The latter two can have a relativiser too, but this implies that they have started a second BRA cycle, which is less common.⁷

To conclude, there is no scenario of morphosyntactic change that can lead to the contemporary typological variation in Bantu relative clause constructions when starting from Meeussen's PB reconstruction. In contrast, if we reconstruct relative verb forms with a Verbal prefix that indexes the subject of the relative verb, the BRA cycle can handle the full catalogue of currently attested constructions without problems. Its strength is that it consists of small, trivial changes, all of which are widely attested, often in one and the same language. Meeussen

⁷In fact, Nsuka-Nkutsi (1982) does not recognise a subject relative clause construction with type NPrel-SBJ agreement. This strongly suggests that his analytical choices were influenced by the absence of Indirect subject relatives in the Bantuist tradition. It might explain, for instance, why Nsuka-Nkutsi (1982: 98) recognises an augment in the verb of subject relative clause constructions more than three times as often as in non-subject relatives, where the same morpheme would have more readily been analysed as a *ppr*. Likewise, the distinction between an independent relativiser and an agreement prefix that indexes the relativised NP is often vague, and indeed fully arbitrary.

(1967: 120) observes that it is not clear whether the Indirect construction he reconstructs was of the Lega-type or of the Luba-type. Remember that the latter is a hybrid of Direct and Indirect relatives: Direct in the case of a third person subject, Indirect elsewhere. In the absence of a clear scenario for morphosyntactic change, the evolution from either to the other is puzzling. However, as we saw in §4, the Luba-type is clearly an innovation as compared to the Lega-type in view of the BRA cycle and neither can be reconstructed in PB.

6 Which path from pre-PB to BGR?

The crucial argument against the reconstruction proposed in BGR with a Direct (PPR-) and an Indirect (PPR-VPR-) relative verb form is that no path has been identified that could lead from that reconstructed state of affairs to the current situation. In this short section, for the sake of the argument we will also assume, as we did in §5, that the reconstructions in BGR are right, and ask how this PB situation could have come about.

According to Meeussen (1967: 113), there are no clear indications for reconstructing morphological differences between relative verb forms and their non-relative counterparts, other than their agreement prefixes in (Pre-)Initial position and the tone of their Final morpheme (although see Meeussen 1971 for the latter). Therefore, the adnominal nature of the agreement marked by the PPR in BGR's PB relative verb forms is a participial characteristic of otherwise fully finite verb forms, which is typologically unusual and in need of an explanation. This need is strengthened by the fact that, to my knowledge, relative verb forms do not show agreement with the relativised NP in any of the Benue-Congo languages outside of Narrow Bantu, so that BGR's Direct and Indirect constructions must be Bantu innovations. This brings up a question similar to the one discussed in the previous section: which scenario of morphosyntactic change could have led from the most likely (typologically usual and universally attested) pre-Bantu situation in which relative verbs agree with their subject to BGR's Direct and Indirect constructions?

An obvious candidate for such a scenario is the BRA cycle. The question is then how old the BRA cycle is. If we wish to assume that it was already active at the PB or pre-PB stage, then we also have to assume that a BRA cycle had created a Direct and an Indirect construction in PB, while the pre-Bantu construction with type SBJ agreement continued to exist. The problem with this latter assumption is that every agreement type is currently attested in languages of almost every Guthrie zone. Therefore, one would also have to make the extremely unlikely

assumption that this tripartite distinction (VPR-, PPR-VPR-, PPR-) has continued to exist for many centuries, surviving in all branches at every split of the Bantu tree, which is only imaginable if there had been for some reason a long cross-Bantu pause in the activity of the BRA cycle. It is therefore far more straightforward to assume that the BRA cycle was not yet active in PB.

7 Consequences for the PB verbal template

Since the BRA cycle consists of a succession of small steps, each of which must be independently motivated, the question is which innovation exactly could have activated the BRA cycle in early Bantu. As far as I can see, there are two options. First, it could be the tendency for relativisers to emerge, corresponding to stage 1 of the cycle. Second, it could be the tendency of verb forms to attract and integrate morphological material at their left hand side (stage 2 of the BRA cycle).

The first option is unlikely to be a feature that could set apart Narrow Bantu from the other Benue-Congo languages. The emergence of relativisers from all kinds of sources is typologically very common (cf. e.g. Hendersy 2012 for an overview of the multiple sources of relativisers in the languages of the world). Besides being typologically common, the emergence of relativisers is also widely attested in contemporary Benue-Congo languages outside of Narrow Bantu. In a small sample of 25 languages covering the major sub-branches of Benue-Congo, the great majority of relative clause constructions are introduced by a relativiser.⁸ In slightly more than half of these cases, this relativiser is invariable. Elsewhere it agrees with the relativised NP. Most languages in which relativisers are invariable lack noun classes. Agreeing relativisers can have different sources. As in the Bantu languages, they can originate in a demonstrative (as in Bafut, Southern Bantoid, Tamanji 2009), in a personal pronoun (as in Noone, Southern Bantoid, Hyman 1981) or in another element (e.g. *-yī* in Kuche, Plateau, Wilson 1996). It is therefore by no means unlikely that PB had one or more agreeing relativisers.

⁸The six Southern Bantoid languages in my sample (Noni, Mungong, Medumba, Bafut, Ejagham and Mundabli) have an agreeing relativiser, which in Mundabli follows the relative verb. The three Northern Bantoid languages are typologically maximally diverse: Vute has no relativiser, Wawa a non-agreeing one and Tikar an agreeing relativiser. The three Edoid languages have a non-agreeing relativiser (Engenni, Degema and Bini); some Plateau languages have a non-agreeing relativiser (Migili, Fyem, Birom), some an agreeing relativiser (Tyap, Kuche). The Delta Cross (Obolo, Ibibio, Eleme) and Jukunoid (Kuteb, Mbembe) languages have either no relativiser or an invariable one. The two Kainji languages in the sample (C'lela, Cicipu) have an agreeing relativiser. Finally, Oko has an invariable relativiser.

In contrast, as far as I know, there are no Benue-Congo languages outside of Narrow Bantu that have relative clause constructions of agreement type NPrel-SBJ or NPrel, or that otherwise show signs of the integration of an original relativiser into the relative verb form. Their relative verb forms either agree with their subject, or show no agreement at all. Many have subject markers that are analysed as separate pronouns, rather than prefixes. What sets apart most of Narrow Bantu, then, is a tendency for verbs to morphologise formerly independent relativisers.

This conclusion is relevant for the reconstruction of the typological profile of PB verb forms, especially the much debated issue of whether their pre-stem domain was rather synthetic or rather analytical, or whether it may have cyclically shifted between these typological profiles (Nurse 2007; Hyman 2011; Güldemann (2022 [this volume])). As has been pointed out by Hyman (2011) regarding Niger-Congo verb forms, there is evidence for both accretion and breakdown in their pre-stem domain and the main difficulty for reconstruction is to determine at which stage any given proto-language was. However, the dead-end nature of NPrel agreement in relative verb forms strikes me as an argument in favour of reconstructing a more analytical profile for the PB and Proto-Benue-Congo pre-stem domain. As pointed out in the previous sections, there is a clear path from type SBJ agreement to type NPrel agreement, but not for the inverse evolution. Therefore, if there had been a strong tendency for integrating agreeing relativisers (or any other preverbal syntactic material) into verb forms at a pre-PB stage, we would expect to find traces of NPrel(-SBJ) agreement in at least some branches of Benue-Congo and we would expect to find the contemporary variation in Bantu relative clause constructions to be compatible with the reconstruction in BGR. A possible hypothesis is that the emergence of a Pre-initial position in the verbal template is an innovation that took place at node 2 in the internal classification of Bantu proposed in Grollemund et al. (2015), i.e. excluding most of zone A. Indications for this hypothesis can be found in the near-absence of agreement of type NPrel-SBJ in relative verb forms in zone A (Nsuka-Nkutsi 1982: 217) and in the overall absence of Pre-initial negative markers in zone A (Kamba Muzenga 1981: 130–132). As for the latter, Kamba Muzenga (1981: 132) remarks “*L’emploi d’une postinitiale en zone A et dans une partie de la zone B peut s’expliquer sans doute par le fait que ces langues ont perdu l’usage de la préinitiale de la conjugaison* [The use of a Post-initial in zone A and parts of zone B could be explained by the fact that these languages have lost the use of a Pre-initial in conjugation – my translation].” However, unless clear indications for the loss of a Pre-initial position in zone A languages come up, the more likely hypothesis is that these languages have never developed one.

Finally, it would be interesting to apply this reasoning to other branches of Niger-Congo. The Atlantic languages, for instance, tend to have little morphological material prefixed to the verb root, versus a lot of suffixation. They also usually have type SBJ agreement on relative verb forms. An exception on both accounts is Bijogo, where categories such as negation, tense and phasal polarity are expressed by means of verbal prefixes. Interestingly, non-subject relative clauses are of the Luba-type in Bijogo. They have agreement of type NPrel-SBJ if the subject is of the first (19a) or second person or of class σ -, and agreement of type NPrel elsewhere (Segeer 2002). Subject relatives have agreement of type NPrel, as can be seen in examples where the relativised NP is a pronoun of the first or second person (19b).

(19) Bijogo (Atlantic, Niger-Congo) (Segeer 2002: 179, 176)

- a. *e-we i-na-rɔrak-ɔ*
 E-goat E.IPFV-SM1SG-look_for-REL
 ‘the goat I am looking for’
- b. *amɔ ɔ-bajokam-mɔ*
 you O.PFV-be_late-REL
 ‘you (sg) who are late’

This appears to confirm the idea that there is a correlation between rich verbal morphology in the pre-stem domain and agreement of relative verbs with the relativised NP within Niger-Congo.

8 Conclusions

The reconstruction of Direct and Indirect relative clauses proposed in BGR (Meeussen 1967) is untenable, because there exists no scenario that could lead from that reconstruction to the current situation. The best reconstruction is one of a default situation in which relative verbs have the same agreement properties as non-relative verb forms. Despite their typological rarity and their wide distribution across Bantu, all contemporary attestations of constructions with agreement of types NPrel (Direct), NPrel-SBJ (Indirect) and previously undetected types such as NPrel-NPrel must be due to relatively recent parallel evolutions. However counterintuitive this conclusion may seem, it is not that unfamiliar for Bantuists. Consider, for instance, what Schadeberg had to say about Spirantisation:

The languages which [have] undergone Spirantization, or both Spirantization and 7>5 [vowel shift], are not genetic subgroups or branches of Bantu. I think this is a safe statement to make, even if the details of the genetic subclassification of Bantu are, after many decades of research, still rather hazy. Historical-comparative studies of (presumably) genetic subgroups of Bantu, even small ones, again and again end up reconstructing a consonantal system prior to Spirantization and a seven-vowel system prior to 7>5. A commonly used argument is the observation that the precise results of Spirantization differ even between closely related languages. Even synchronic descriptions of languages which have undergone both changes sometimes posit the situation as found prior to these changes for the underlying representation in order to account for regular allomorphic alternations. An example is Louise Polak-Bynon's grammar of Shi (D.53). (Schadeberg 1994: 81)

The mechanism that has driven the many parallel local evolutions from type SBJ agreement to other agreement types can be clearly identified as the Bantu Relative Agreement (BRA) cycle. The next obvious question to be asked and answered is which grammatical change may have activated the BRA cycle itself in Bantu. My favourite hypothesis for answering that question is that function words in preverbal position started morphologising at some PB stage, presumably the common ancestor of the languages under node 2 in Grollemund et al.'s (2015) internal classification. Due to this innovation, formerly independent relativisers started having the potential of being integrated into relative verb forms as prefixes in a new Pre-initial position.

Acknowledgements

This work is partially supported by a public grant overseen by the French National Research Agency (ANR) as part of the program "Investissements d'Avenir" (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris – ANR-18-IDEX-0001. I wish to thank Sara Pacchiarotti, Rozenn Guérois, Koen Bostoën, Dmitry Idiatov and two anonymous reviewers for their generous comments.

Abbreviations

In what follows the starred forms are terms for positions in Meeussen's morphological template of the Bantu verb; those marked by a degree sign are names of paradigms of agreement markers:

11 Agreement on Proto-Bantu relative verb forms

1, 2, 3, E, O ...	noun classes	INT	intensive
1SG, 2PL ...	first and second person	IPFV	imperfective
	singular/plural markers	L	low tone
		LOC	locative
		NP	Noun Phrase
APPL	applicative	NPr	nominal prefix
AUG	augment	NPreI	relativised NP
BGR	Bantu Grammatical Reconstructions	OPr	object prefix
		PB	Proto-Bantu
		PFV	perfective
BRA	Bantu Relative Agreement	PL	plural
		postFV	Post-final*
CAUS	causative	postIN	Post-initial*
CJ	conjoint	PPR	pronominal prefix°
cl.	class	preIN	Pre-initial*
DEM	demonstrative	PRO	pronoun
PRE	numeral prefix	REL	relative / relativiser
EXT	extension*	RPr	relative prefix°
FO	formative*	RPST	remote past
FV	final (vowel)*	SBJ	subject
H	high tone	SM	subject marker
H1D	first high tone deletion	TNS	tense
		V	vowel
IF	infix*	VPr	verbal prefix°
IN	initial*		

References

- Blanchon, Jean A. 1980. La relativisation. In François Nsuka-Nkutsi (ed.), *Éléments de description du punu* (Linguistique et sémiologie. Travaux du centre de recherches linguistiques et sémiologiques de Lyon), 115–128. Lyon: Université Lyon II.
- de Blois, Kornelis Frans. 1970. The augment in the Bantu languages. *Africana Linguistica* 4. 85–165.
- Grégoire, Claire & Baudouin Janssens. 1999. L'augment en bantou du nord-ouest. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (CSLI Lecture Notes 99), 413–429. Stanford, CA: Center for the Study of Language & Information.

- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guérois, Rozenn & Denis Creissels. 2020. The relative verb forms of Cwabo (Bantu P34) as contextually oriented participles. *Linguistics* 58(2). 463–491.
- Güldemann, Tom. 2022. Predicate structure and argument indexing in early Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 387–421. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575831.
- Hamlaoui, Fatima. 2022. On subject inversion in Proto-Bantu relative clauses. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 495–535. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575837.
- Harries, Lyndon. 1950. *A grammar of Mwera, a Bantu language of the eastern zone, spoken in the south-eastern area of Tanganyika Territory* (Bantu Grammatical Archives 1). Johannesburg: Witwatersrand University Press.
- Hendery, Rachel. 2012. *Relative clauses in time and space. A case study in the methods of diachronic typology* (Typological Studies in Language 101). Amsterdam: John Benjamins.
- Hyman, Larry M. 1981. *Noni grammatical structure: With special reference to verb morphology* (Southern California Occasional Papers in Linguistics 9). Los Angeles, CA: Department of Linguistics, University of Southern California.
- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.
- Idiatov, Dmitry. 2022. Diachronic typology and the reconstruction of non-selective interrogative pronominals in Proto-Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 667–737. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575843.
- Kamba Muzenga, Jean-Georges. 1981. *Les formes verbales négatives dans les langues bantoues* (Annales – Série in-8° – Sciences humaines 106). Tervuren: Royal Museum for Central Africa.
- Kamba Muzenga, Jean-Georges. 2003. *Substitutifs et possessifs en bantou* (Société d'études linguistiques et anthropologiques de France 401). Louvain: Peeters.
- Kawasha, Boniface. 2008. Relative clauses and subject inversion in Chokwe, Kaonde, Lunda and Luvale. *Africana Linguistica* 14. 37–61.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

- Meeussen, Achiel Emiel. 1971. Relative clauses in Bantu. *Studies in African Linguistics Supplement* 2. 3–10.
- Nsuka-Nkutsi, François. 1982. *Les structures fondamentales du relatif dans les langues bantoues* (Annales – Série in-8° – Sciences humaines 107). Tervuren: Royal Museum for Central Africa.
- Nurse, Derek. 2007. Did the Proto-Bantu verb have a synthetic or an analytic structure? In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 239–256. London: Department of Linguistics, School of Oriental & African Studies.
- Polak-Bynon, Louise. 1975. *A Shi grammar: Surface structures and generative phonology of a Bantu language* (Annals – Series in-8° – Human Sciences 86). Tervuren: Royal Museum for Central Africa.
- Sanderson, Meredith. 1922. *A Yao grammar*. 2nd revised edition. London: Society for Promoting Christian Knowledge.
- Schadeberg, Thilo C. 1994. Spirantization and the 7-to-5 vowel merger in Bantu. *Belgian Journal of Linguistics* 9. 73–84.
- Segerer, Guillaume. 2002. *La langue bijogo de Bubaque (Guinée Bissau)* (Afrique et Langage 3). Louvain: Peeters.
- Stappers, Leo. 1973. *Esquisse de la langue mituku* (Annales – Série in-8° – Sciences humaines 80). Tervuren: Royal Museum for Central Africa.
- Tamanji, Pius N. 2009. *A descriptive grammar of Bafut – a Grassfields Bantu language of Cameroon* (Grammatical Analyses of African Languages 35). Cologne: Rüdiger Köppe.
- Taylor, Charles. 1985. *Nkore-Kiga* (Croom Helm Descriptive Grammars). London: Croom Helm.
- van Coillie, Gustaaf. 1948. Korte Mbagani-spraakunst: De taal van de “Babindji” in Kasayi (Deel 1). *Kongo-Overzee* 14(5). 257–279.
- Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics, with special focus on Bantu and Mande. Essays in honour of Yvonne Bastin and Claire Grégoire* (Collection Sciences Humaines 169), 425–441. Tervuren: Royal Museum for Central Africa.
- Van de Velde, Mark. 2019. Nominal morphology and syntax. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 237–269. Milton Park, Abingdon: Routledge.
- Van de Velde, Mark. 2021. The Bantu relative agreement cycle. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 59(4). 981–1015.

- Van de Velde, Mark. 2022. The AMAR mechanism: Nominal expressions in the Bantu languages are shaped by apposition and reintegration. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 60(3). 899–931.
- Van de Velde, Mark. Forthcoming. Adnominal possession. In Lutz Marten, Nancy C. Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages* (Oxford Guides to the World's Languages 2). Oxford: Oxford University Press.
- Van de Velde, Mark & Odette Ambouroue. 2017. The origin and use of a relative clause construction that targets objects in Orungu (Bantu, Gabon). *Studies in Language, International Journal sponsored by the Foundation "Foundations of Language"* 41(3). 615–637.
- Wilson, Janet Evelyn. 1996. *A phonological grammar of Kuche*. Arlington, TX: University of Texas at Arlington. (MA thesis).

Chapter 12

On subject inversion in Proto-Bantu relative clauses

Fatima Hamlaoui

University of Toronto

This chapter concentrates on the canonical position of lexical subjects in Proto-Bantu non-subject relative clauses. Based on both the geographical and the genealogical distribution of different word orders (Subject Verb-only, Verb Subject-only and Subject Verb / Verb Subject), I propose that the Verb Subject (VS) order is an innovation that came into use only after the split between the North-Western Cameroonian branch of Grollemund et al.'s (2015) classification and the rest of the tree, that is, node 2 or 3. I thus argue for a revision of Meeussen's (1967) and Nsuka-Nkutsi's (1982) claim that Proto-Bantu (node 1) non-subject relative clauses were characterised by a VS order. After expanding Nsuka-Nkutsi's sample from over a hundred Narrow Bantu languages to a total of 167 languages (151 Narrow Bantu and 16 other Niger-Congo languages), we observe that VS-only is still the most frequent word order. However, the Subject Verb (SV) order is dominant in the major clades of Grollemund et al. (2015) located in the north-western Bantu area (20 out of 22 languages in our sample), that is, in the languages that are both closer to the Bantu homeland and more similar to the Niger-Congo languages outside of Narrow Bantu in our sample. SV-only is also found in a significant portion of our sample in the Eastern branch (28 out of 57 languages). Together, these facts suggest that the SV order might be more ancient than previously thought. If this scenario is correct, Bantu zone A languages would not have lost VS due to their evolution from more syntheticity to more analyticity, but they would never have had it at all.

1 Introduction

Except for a few notable exceptions, such as Nen A44 (Mous 2003: 304), basic word order in present-day narrow Bantu languages is SVO (Bearth 2003).



Discourse-driven word order is also considered to be typical of the Bantu family and, according to Schadeberg (2003: 152–153), characteristic of Proto-Bantu. In many Bantu languages indeed, various permutations of major constituents yield grammatical sentences. In particular, a lot has been written on so-called “inversion constructions”, in which a logical subject, i.e. the highest thematic role selected by the verb, occupies a postverbal position. Depending on the language, the logical subject either controls subject agreement with the verb, as in (1), or does not control it, as in (2). These changes in word order are most often attributed to communicative needs and, in particular, the expression of information-structural notions such as focus and topic (e.g. Marten 2014, Hamlaoui Forthcoming).

- (1) Matengo N13 subject inversion (Yoneda 2011: 756)
ílásí ju-a-hémála Kinû:nda
8.potatoes SM1-PST-buy 1.Kinunda
(Talking about potatoes) ‘(These) potatoes, Mr. Kinunda bought (them).’
- (2) Luguru G35 subject inversion (Mkude 1974: 133)
ibalua i-andika mwalimu
9.letter SM9-write.PST 1.teacher
‘The teacher wrote the letter.’

Interestingly, at least some (rather Western) Bantu languages do not seem to share this property. Instead, they have been reported to display a much more rigid word order, where surface positions primarily express argument relations and the highest thematic role must be realised as a canonical subject. This is the case of the North-Western Bantu language Basaa A43a, for instance (Hamlaoui & Makasso 2015). Alternatively, other languages such as Mbuun B87 (Bostoen & Mundeke 2011; 2012) and Sikongo H16a (De Kind 2014) display other types of discourse-driven constituent re-orderings, primarily involving the preverbal domain.

In the present chapter, we are interested in Proto-Bantu (PB) word order. In particular we explore the issue of word order in relative clauses, an area in which variation is found and which has not yet been extensively explored. As relative clauses are a rather traditional part of grammatical descriptions and have generally attracted considerable attention, a critical mass of data is now available and the time seems ripe for us to try to reconstruct their PB word order.

Note that the interest in PB relative clauses is not new. Meeussen (1967) dedicates two sections to the topic: one on relative tenses (i.e. verb forms in relative clauses) and one on their syntax. What is of particular interest to us in this

chapter is the location of full subjects (also called “free subjects”, as opposed to subject markers, either agreeing with them or referring to them anaphorically) with respect to the verb of the relative clause, as this seems to be a major locus of variation across Bantu. Within relative clauses, some languages indeed have subjects strictly precede the relative verb, as in (3), whereas others have them strictly follow it, as in (4).

- (3) Eton A71 relative clause (SV-only) (Van de Velde 2017: 53)

ímákíd η[↓]kúnjkúmá á-↓kúzgí kálâdâ ménê ùjàb vǎ
í-m-ákíd Ñ-kúnjkúmá à^H-küz^L-gì^H kálâdâ mǎ-nè ù-ǫjàb vǎ
 AUG-6-market 3-chief SM1-PST-buy-G-NF book SM6-be 3-far here
 ‘The market where the chief bought the book is far from here.’

- (4) Makonde P23 relative clause (VS-only) (Manus 2010: 182)

aviilá vy-á vy-á-súum-a vá-dyóóko ví-díkídiiki
 8.DEM SM8-REL 8-PRES-buy-FV 2-children 8-small
 ‘Those that the children are buying are small.’

Other languages have relative clauses in which the relative subject is sometimes preverbal and sometimes postverbal, as in (5). The reason for this alternation is not always well understood.

- (5) Swahili G42d relative clause (SV/VS) (Givón 1972: 291)

- a. *ki-tabu amba-cho m-toto a-me-ki-ona*
 7-book say-REL7 1-child SM1-PRF-OM7-see
 ‘the book that the child has seen’
- b. *ki-tabu amba-cho a-me-ki-ona m-toto*
 7-book say-REL7 SM1-PRF-OM7-see 1-child
 ‘the book that the child has seen’

In the specific case of Nzadi B865, illustrated in (6), subject doubling can sometimes be observed. We will come back to this particular case later on and, following Hyman (2012), classify it with languages that display a postverbal subject-only order.

- (6) Nzadi B865 relative clause (SVs) (Hyman 2012: 103)

- a. *èsúú (ná) (ηgʻ) ò món àkáàr mwǎàn*
 day (that) (which) PST see women child
 ‘the day that the women saw the child’

- b. *èsúú (nà) (ɲg') àkáár ò món bõ mwààn*
day (that) (which) women PST see they child
'the day that the women saw the child'

Postverbal subjects are so widespread that Meeussen (1967: 120) explicitly reconstructs PB relative clauses as having full subjects following the verb, as shown in (7).

- (7) PB reconstruction (Meeussen 1967: 120)
i-pía dí-dim-á mu-ntu/ba-ntu
5-garden SM5-cultivate-FV 1-person/2-person
'the garden which the person(s) cultivate(s)'

Nsuka-Nkutsi (1982) offers a thorough overview of morphosyntactic features in Bantu relative clauses. Based on a survey of what we have counted to be 107 languages from 16 Bantu zones (including Tervuren's J zone), he observes that VS is indeed the most frequent word order in his sample and reaches the same conclusion as Meeussen, that VS is the basic word order in object relative clauses and the one characterising PB.¹ As pointed out to us by a reviewer, both of them use the "majority wins" principle (Campbell 1998: 117ff; Dimmendaal 2011: 12). Not referring to the internal genealogical classification of the Bantu languages is however potentially problematic, in that some Bantu languages are more closely related to each other than others (cf. e.g. Grollemund et al. 2015 and references therein). Following Campbell (1998: 114), one would have to make sure that the languages considered do not come from the same branch of the family and thus have an immediate parent that is itself a daughter of PB that might have undergone a separate change. Ideally, one would have to look into the distribution of SV and VS across all major branches of the Bantu family.

In the present chapter, we re-examine both Meeussen's and Nsuka-Nkutsi's proposal that VS is the basic word order in PB relative clauses, using an expanded language set of 167 Niger-Congo languages of which 151 are Narrow Bantu. We have chosen to include 16 Niger-Congo languages that are not Narrow Bantu because, following Nurse (2007), we find it important to try to provide some broader perspective: how does Meeussen and Nsuka-Nkutsi's proposal that VS is the basic word order in relative clauses fit into the larger Niger-Congo picture? Considering that the Niger-Congo phylum counts 1553 languages in the latest Ethnologue (Eberhard et al. 2022), it is not possible yet to provide a sample that

¹Unfortunately we have not found a count of the languages discussed in Nsuka-Nkutsi (1982). The numbers appearing in the present chapter are our own.

would be representative of it. Our sample is primarily based on the literature that was available to us at the time of writing. We concentrate on relative clauses in non-Bantu Niger-Congo SVO languages, and the ones we have included are listed in §3.1. Among them, 8 are Southern Bantoid, that is, both geographically and genealogically close to Narrow Bantu. Almost all 16 languages show a strict SV word order in relative clauses.

The chapter is structured as follows. In §2, we first give a brief overview of the basic properties of non-subject relative clauses and lay out the existing proposals as to the motivation for VS in present-day Bantu relative clauses. Harvesting data from a number of grammatical sketches and descriptions published more recently (among others Henderson 2006; Downing et al. 2010; Atindogbé & Grollemund 2017), we expand Nsuka-Nkutsi's database to look at the frequency and distribution of each of the three attested patterns of variation, both across Guthrie's (1971) zones and the major branches of the lexicon-based phylogenetic classification in Grollemund et al. (2015). We show, in §3, that although VS is still the most common word order in our sample, its frequency is not much higher than the SV order. Moreover, the geographical distribution of both VS and SV questions the idea that SV developed later. Among other things, we see that in our database, 20 out of our 21 Bantu zone A languages only display SV. Of those 20 languages, 15 belong to Grollemund et al.'s (2015) North-Western Cameroon branch, the one closest to the Bantu homeland. In this respect, Bantu zone A languages also seem more similar to our sample of (non-Bantu) Bantoid ($n = 8$) and (non-Bantoid) Niger-Congo languages ($n = 8$). One of the questions that arises is whether these Bantu languages (i.e. zone A/North-Western Cameroon) that show the SV-only pattern and other Narrow Bantu languages had a common ancestor that had VS, or whether VS is a word order that emerged only after the split between them, i.e. at the level of node 2 or 3 of Grollemund et al.'s (2015) classification. In the latter case, VS would not be the word order characterising relative clauses in the common ancestor to all present-day Bantu languages, but potentially only to a subset of the major branches of the family. We consider this hypothesis in §4.² Remaining agnostic about the degree of agglutinativity of PB verb structure (Güldemann 2003; Hyman 2007; Nurse 2007; Güldemann 2011; Hyman 2017), we explore the possibility that the near absence of VS in present-day Bantu zone A languages, as represented in the sample, is related to their

²According to Harris & Campbell (1995: 27), many scholars maintain the view that syntactic change affects main clauses before subordinate clauses. If this is correct and if VS is indeed an innovation at node 2 or 3 of the Bantu tree, VS was thus probably found in main clauses before appearing in subordinate clauses, in particular if VS was motivated by considerations relevant to main clauses such as, for instance, information-structural ones.

more analytic morphology and in particular the absence of headmarking found in other Bantu languages. If this hypothesis is correct, we believe that a correlation should be found between the lack of pre-stem object markers (i.e. the lack of OM-V order) and the absence of VS in particular Bantu languages. Based on a sub-sample of 162 languages (146 of which are Narrow Bantu languages), we show that such a statistically significant correlation is indeed found. Among the Bantu languages that do not have pre-stem object markers, VS-only represents the minority, and SV/VS has so far been found (in non-restrictive relative clauses) in only one language. Finally, we propose that the lack of pre-stem object markers could have broader consequences for the syntax of North-Western Bantu languages and could explain why a Bantu zone A language like Basaa significantly differs from other Bantu languages as to how it expresses information-structural notions, most particularly with regard to the lack of connection between focus and postverbal position. §5 concludes the chapter.

2 Bantu relative clauses

2.1 Some basic properties

Bantu languages vary widely as to how they form their relative clauses, and various aspects of relative clause formation have thus been the object of extensive investigation – see Van de Velde (2022 [this volume]), and Cheng (Forthcoming). Here we give only a brief overview of the main issues concerning non-subject relative clauses, also called “indirect relative clauses”.

Bantu relative clauses follow their head noun and they typically have a relative marker which agrees with it in noun class features. The morphosyntactic nature of the relative marker and its position with regard to the subject vary considerably. According to Cheng (Forthcoming: 3), clause-initial relative markers, as in (8), are a common strategy.

- (8) Venda S21 object relative clause (Zeller 2004: 81)
munna ane nngwa dza-mu-pandamedza
1.man REL1 10.dog SM10-OM1-chase
‘the man whom the dogs are chasing’

In many Bantu languages, the relative marker is either a demonstrative pronoun or based on one. This is the case in Venda S21 in (8), as well as in Bemba M42, Chewa N31b (Cheng Forthcoming) and in Basaa A43a in (9).

- (9) Basaa A43a possessive relative clause (Jenks et al. 2017: 22)

í-m-ààngé nú ↓ngwó jé↓é ì-βí-kògól mê
 AUG-1-child REL1 9.dog POSS9 SM9-PST2-bite me
 ‘the child whose dog bit me’

In other languages, a bound relative marker appears prefixed to the verb. As shown in (10) and (11), with Zulu S42 and Lega D25 respectively (Cheng Forthcoming), the subject of the relative clause either precedes or follows the verb depending on the language.

- (10) Zulu S42 object relative clause (Zeller 2004: 79)

in-cwadi isi-tshudeni esi-yi-funda-yo
 9-letter 7-student REL7-OM9-read-REL
 ‘the letter that the student is reading’

- (11) Lega D25 object relative clause (Carstens 2005: 233)

bi-tondo bí-ku-ténd-a úzo mwána ta-bí-lí bi-sóga
 8-word REL8-PROG-say-FV DEM1 1.child NEG-SM8-be SM8-good
 ‘The words that that child is saying are not good.’

Quite a few languages also seem to display a verb-final or suffixed relative marker. This is the case of Kwakum A91 in (12), which has an additional relative marker at the end of the clause, but also of geographically more distant languages such as Chewa and Zulu (see (10)) (Cheng Forthcoming).

- (12) Kwakum A91 possessive relative clause (Hare 2018: 13)

ai mon paam mo fanç-e kam-e bulaw-e i
 3SG COP child man REL-PRO father-3.SG.POSS like-3SG.a_lot REL
 ‘He is a boy whose father loved him a lot.’

Relative markers can also be found within the verb, as in Swahili in (13) (Cheng Forthcoming).

- (13) Swahili G42d object relative clause (Ngoyani 2001: 61)

vi-tabu a-li-vyo-nunu-a Juma ni ghali
 8-book SM1-PST-REL8-buy-FV Juma COP expensive
 ‘The books Juma bought are expensive.’

Another difference between Bantu languages depends on which item the relative marker agrees with and whether the relative verb shows subject agreement

as well. As seen in (10), the relative marker of Zulu seems to agree with the subject of the relative clause rather than with the head noun. In contrast, in Lega in (11), the relative marker agrees with the head noun and the relative verb shows no agreement with the (postverbal) subject, ‘that child’.

Finally, in object relative clauses, Bantu languages vary as to whether the relative verb displays an object marker which agrees in noun class features with the head noun, as it does in Chewa in (14).

- (14) Chewa N31b object relative clause (Downing & Mtenje 2011: 76)
a-lendó a-méné á-ná-wa-bweretsérá m-phátsoo-wo
2-visitor 2-REL SM2-PST2-OM2-bring_for 10-gift-REL2
a-koondwa
SM2.PRF-be_happy
‘The visitors who they brought the gifts for are happy.’

Let us now turn to the issue that is central to this chapter: the position of full subjects in relative clauses.

2.2 Possible motivations for inverted embedded subjects

Although a lot of work has been done on Bantu inverted subjects in simple sentences, comparatively little has been done on the topic in embedded clauses. Only a few proposals have been made.

Givón (1972) and Demuth & Harford (1999) have proposed that the nature of the complementiser, and particularly its status as a bound morpheme, motivates syntactic operations that result in the VS order. Givón (1972: 289) proposes a “universal principle of pronoun (or subordinator) attraction in relativization”, by which the relative marker needs to be immediately adjacent to the head noun modified by the relative clause. Concentrating on object relative clauses and based on data from Swahili, Givón proposes that whenever the relative pronoun is a disyllabic free morpheme, it can be extracted from the canonical position of the argument or modifier it corresponds to and be made adjacent to the head noun, with no other necessary changes in word order. This is the case in *amba*-relative clauses in Swahili (already illustrated in (5a) and repeated below for convenience), in which subject postposing is optional: subject-verb inversion is possible, but not necessary to achieve the adjacency between head noun and relativiser.

- (5a) Swahili G42d relative clause (Givón 1972: 291)

ki-tabu amba-cho m-toto a-me-ki-ona
 7-book say-REL7 1-child SM1-PRF-OM7-see
 ‘the book that the child has seen’

In contrast, when the relativiser is a bound morpheme (i.e. bound to the verb), subject postposing is necessary to achieve adjacency between it and the head noun, resulting in the VS order. This pattern can be illustrated with a different type of relative clause also found in Swahili, in which the relativiser is a verbal prefix. According to Givón, in (15), the subject can only be postverbal.

- (15) Swahili G42d object relative clause (Givón 1972: 291)

ki-tabu a-li-cho-ki-ona m-toto
 7-book SM1-PST-REL7-OM7-see 1-child
 ‘the book that the child saw’

Givón’s proposal finds further support in Takizala’s (1972) Hungan H42 data. In this language, VS is obligatory whenever an overt relativiser is present, as in the pseudo-cleft in (16).

- (16) Hungan H42 pseudo-cleft sentence (Givón 1972: 292)

(kiim) ki-a-swiim-in Kipes zoon kwe kit
 (7.thing) REL7-SM1-buy-PST Kipese yesterday is 7.chair
 ‘(the thing) what Kipese bought yesterday is a chair’

In the cleft sentence in (17), in contrast, which Takizala analyses as involving a relative clause with no overt relativiser, no subject postposing is observed.

- (17) Hungan H42 cleft sentence (Takizala 1972: 269)

kwe kit Kipes ka-swiim-in zoono
 it’s 7.chair Kipese SM1-buy-PST yesterday
 ‘It’s a/the chair (that) Kipese bought yesterday.’

Subject postposing is also found in object *wh*-questions, but only when the optional relativiser appears, as shown in (18) and (19).

- (18) Hungan H42 *wh*-question (Takizala 1972: 293)

na Kipes ka-mweene?
 whom Kipese SM1-see.PST
 ‘Whom did Kipese see?’

- (19) Hungan H42 clefted *wh*-question (Takizala 1972: 293)
na wu-u-mweene Kipes?
whom that-AG-see.PST Kipese
'Who (is it) that Kipese saw?'

Demuth & Harford (1999) show that Givón's proposal also finds support in Southern Sotho S33 and Shona S10 data. In the former language, in (20), the relativiser is a disyllabic free morpheme, and subject inversion is ungrammatical, while in the latter, in (21), the relativiser is a monosyllabic bound morpheme and subject inversion is obligatory.

- (20) Southern Sotho S33 object relative clause (Demuth & Harford 1999: 49)

- a. *di-kobo tseo ba-sadi ba-di-rekileng kajeno*
10-blanket REL10 2-woman SM2-OM10-bought today
'the blankets which the women bought today'
- b. * *di-kobo tseo ba-di-rekileng ba-sadi kajeno*
10-blanket REL10 SM2-OM10-bought 2-woman today
'the blankets which the women bought today'

- (21) Shona S10 object relative clause (Demuth & Harford 1999: 50)

- a. *mbatya dza-v-aka-sona va-kadzi*
10.clothes REL10-SM2-TAM-sew 2-woman
'the clothes which the women sewed'
- b. * *mbatya dza va-kadzi v-aka-sona*
10.clothes REL10 2-woman SM2-TAM-sew
'the clothes which the women sewed'

In their view, the difference between the two languages lies in the fact that in Shona, a prosodic constraint that requires words to be minimally disyllabic triggers verb movement over the subject towards the relativiser. This prosodic constraint has no effect on the syntax of Northern Sotho, as the disyllabic relativiser satisfies it without the need for the verb to raise over the subject, hence the absence of VS in this language.

In sum, the VS word order has been attributed to morpho-phonological properties of the relativiser and its tight relation to the verb, to which it either needs or does not need to attach depending on the language, resulting in obligatory VS order or SV(/VS) order, respectively.

Counterevidence to this analysis is provided by Kawasha (2008) and Letsholo (2009). In Chokwe K11 and Luvale K14, Kawasha (2008: 50) shows that inverted subjects are obligatory even when the relativiser is a free morpheme. This is shown in (22) and (23), for Chokwe and Luvale respectively.

- (22) Chokwe K11 object relative clause (Kawasha 2008: 50)

ly-onda lízé a-a-mbách-ile pwo
 5-egg REL5 SM1-TNS-carry-REM 1.woman
 ‘the egg which the woman carried’

- (23) Luvale K14 object relative clause (Kawasha 2008: 50)

chi-twámó chízé a-a-neh-á-nga mu-kwézé
 7-chair REL7 SM1-TNS-bring-FV-PST 1-youngster
 ‘the chair that the youngster brought’

Examples (22) and (23) are comparable to Givón’s Swahili *amba*-relative clauses in which the relative subject is postverbal: as verb movement is not necessary for the relativiser and the head noun to be adjacent, it is unclear what motivates the VS order. A stronger argument, however, comes from Letsholo (2009: 144), who shows that the affix status of the relativiser does not force inversion in Kalanga S16. As visible in (24), the subject remains preverbal in this language.

- (24) Kalanga S16 object relative clause (Letsholo 2009: 144)

nlúmé bo-Néo wa-bá-ka-bóna wá-énda
 1.man 2a-Neo REL1-SM2A-PST-see SM1-leave
 ‘The man that Neo and others saw left.’

Nsuka-Nkutsi (1982: 256) succinctly puts forward an alternative proposal, according to which VS in relative clauses finds its origin in the emphatic postverbal subjects of simple sentences:

[...] il y a de très nombreux cas dans les langues bantoues (et même dans d’autres langues du monde) où, à partir d’une construction emphatique dont l’utilisation devient de plus en plus répandue, on arrive à une phrase admise comme normale.

[...] there are a great many cases in Bantu languages (and even in other languages) where, from an emphatic construction whose use becomes more widespread, we arrive at a sentence considered as normal. (my translation)

He proposes that object relative clauses with a preverbal subject derive, in their turn, from the fronting of the subject before the relative verb, returning to the most widespread word order within Bantu main clauses.

It seems to be an open question whether relative clauses should be expected to be influenced by information-structural considerations, as they generally seem not to participate in the main information-structural articulation of sentences but rather to be embedded within constituents whose information-status is relevant at the sentential level. In languages with overt topic markers, such as for instance Korean and Japanese, these are reported to rarely occur in relative clauses (Kuno 1973; Song 2014).

Hardly any studies have explored the possibility that changes in word order in Bantu relative clauses might be due to information structure. Hyman (2012: 105), however, reports that in Nzadi B865, which has both preverbal and postverbal full subjects, there is no known pragmatic difference between the two possible word orders. One of the few studies which directly addresses the role of information structure in the ordering of the constituents of relative clauses is the description of Mungbam and Mundabli (Southern Bantoid) by Lovegren & Voll (2017). The authors explicitly state that focus-induced changes in word order (i.e. subject inversion) are similar in main and relative clauses in both languages.

In sum, few proposals have been made regarding the origin and motivation of VS in Bantu embedded clauses and they all can be challenged by empirical evidence. More investigation is needed to establish the motivation for VS in embedded clauses, and why other Niger-Congo languages, outside of Narrow Bantu, do not seem to resort to this strategy as much as languages belonging to the major Narrow Bantu branches outside of the North-West. Let us now turn to the frequency and distribution of the three possible patterns: SV-only, VS-only and SV/VS.

3 Exploration of an expanded sample

3.1 Geographical distribution of SV-only, VS-only and SV/VS

In Bantu non-subject relative clauses, the postverbal location of full subjects has long been noted and treated as a common and widespread phenomenon. According to Nsuka-Nkutsi (1982: 77), VS is the most represented type of object relative clauses. More recent studies have identified additional Bantu languages in which VS is either allowed or compulsory in relative clauses (among others Demuth & Harford 1999; Kawasha 2008; Kisseberth 2010; Hyman 2012).

Interestingly, recent studies have also dedicated more attention to lesser-studied language zones, in particular in North-Western Bantu and among closely related Benue-Congo languages from Cameroon (Downing et al. 2010; Atindogbé & Grollemund 2017). What emerges from these studies is that, in contrast, many of these languages do not allow postverbal subjects in relative clauses. Hamlaoui & Makasso (2015) show that Basaa actually does not accept any type of subject inversion. North-Western Bantu languages are however generally underrepresented in the study of inversion constructions. By way of illustration, out of 46 languages, Marten & van der Wal's (2014) typological study of Bantu inversion constructions includes only one North-Western Bantu language (i.e. Basaa A43a), one Central-Western Bantu language (i.e. Dzamba C322) and two West-Western Bantu languages (i.e. Mbuun B87 and Nzadi B865). All others belong to South-Western and Eastern Bantu, which constitute one single superclade in Grollemund et al. (2015). Based on available studies, it is hard to know whether a construction that is considered typical of the Bantu family as a whole is also typical of the Bantu languages that are geographically closest to the ancestral homeland and which are known for showing a higher degree of diversity than those further removed (Bearth 2003).

Nsuka-Nkutsi's (1982) sample is less biased towards South-Western and Eastern Bantu languages. On the contrary, the best represented group is zone C (n = 18), i.e. Central-Western Bantu, followed by zone D (n = 11), i.e. Central-Western and Eastern Bantu, and zones A, B, L and S (n = 9), i.e. North-Western, West-Western, South-Western and Eastern Bantu. The number of languages found in each zone is shown in a lighter colour in Figure 1. The number of languages found after expanding the sample with data and observations harvested from more recent grammatical sketches and studies appears in a darker colour. In our expanded sample, languages from the north-western part of the Bantu domain remain well represented, with zone A, B, C and D languages constituting 44% of the total sample.

Distinguishing between languages for which only SV, only VS or both SV and VS is reported, Figure 2 provides the position of lexical subjects in non-subject relative clauses for each of the Bantu zones. Both SV-only and VS-only are found in most zones, but in varying proportions.

Some zones seem to have a majority of SV-only languages (i.e. zones A, M, R and S), while others predominantly have VS-only languages (i.e. zones B, C, D, H, K and L). In zones F and P, only VS is found and in zones E, JD and JE, subjects are reported to be only preverbal. As our sample only contains a small number of languages for some zones (i.e. between 4 and 6 languages for zones E, JD and JE),

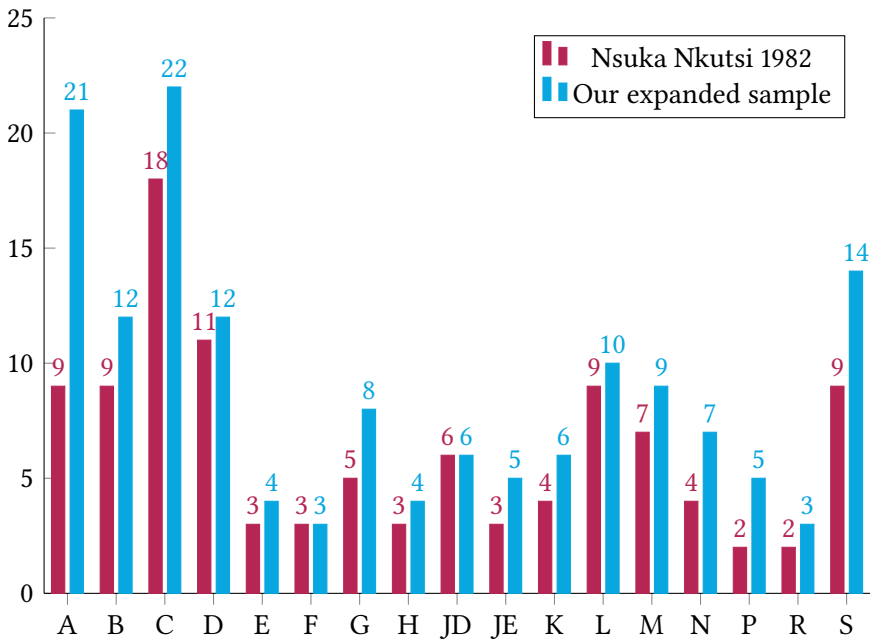


Figure 1: Number of relative clauses in Nsuka-Nkutsi (1982) versus our expanded sample

it is not presently possible to safely conclude that these numbers are representative of what is actually found in each of these zones as a whole (which might not be truly problematic from a reconstruction point of view though). However, our sample for zone A presently counts 21 languages and only one of them (i.e. Kwakum) has, to the best of our knowledge, been reported to allow VS (David M. Hare, p.c.) in some restricted contexts (see §4.2). Table 1 summarises the number of languages for each of the three patterns observed in our sample of Narrow Bantu languages.

Based on the information available in existing descriptions, a handful of languages ($n = 15$, 10%) variably allow SV and VS. These languages are however found across the Bantu domain (zones A, B, G, H, L, M, N, R and S), in languages belonging to the North-Western, West-Western, South-Western and Eastern clades of Grollemund et al.'s (2015) classification, i.e. in all except Central-Western. They are also found in two of the outgroup Bantoid languages, i.e. Mungbam and Mundabli (Lovegren & Voll 2017).

Nzadi, a West-Western language discussed in detail in Hyman (2012), has been reported to display a singular pattern, in which a full preverbal subject phrase

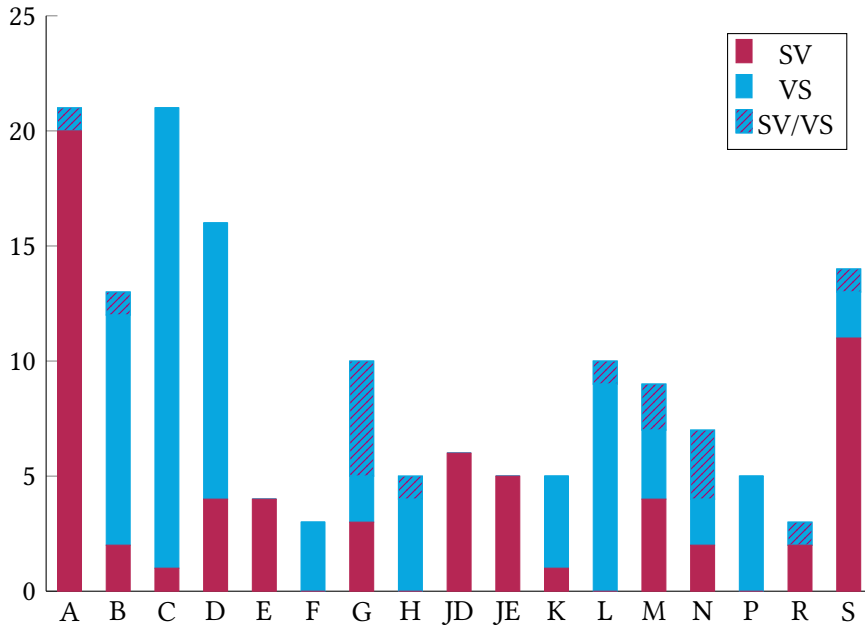


Figure 2: Distribution of SV, VS and SV/VS relative clause word orders across Bantu zones

Table 1: Total number of languages for each observed pattern (151 Narrow Bantu languages)

Word order	Number of languages
SV-only	63
VS-only	72
SV/VS	15
SVs	1
Total	151

can only appear if it is resumed by a postverbal subject pronoun, thus yielding a subject doubling configuration (noted SVs), see example (6). According to Hyman, the grammatical subject is probably the postverbal one, actually making Nzadi a type of VS language, and this is how we have treated it in our statistics.

As we are interested in PB, it is crucial for us to also have a broader perspective into the Niger-Congo phylum (Nurse 2007). Our set of outgroup languages is admittedly very modest, with only 16 languages, and will need to be expanded, but for the time being, it allows us to have an idea of what is found outside Narrow Bantu. For the sake of comparison with Bantu languages, we limit our examination to other SVO languages. Our sample thus includes other South Bantoid languages ($n = 8$): Ejagham (Ekoid) (Watters 1981), Bafut and Medumba (Eastern Grassfields) (Tamanji & Achiri-Taboh 2017), Kenyang (Mamfe) (Tabe & Atindogbé 2017), Mungbam and Mundabli (Beboid/Yemne-Kimbi) (Lovegren & Voll 2017), and Wawa and Vute (Mambiloid) (Martin 2017, Thwing 2017). It also includes a few more distant Niger-Congo languages ($n = 8$): Buli (Gur) (Schwarz 2006), Lelemi, Ewe and Asante Twi (Kwa) (Allan 1973; McCracken 2013; Dzamehie 1995), Pulaar (North Atlantic) (Ba 2015), Zande (Ubangi) (Pasch & Mbolifouye 2011), and Moro and Lumun (Kordofanian) (Rose et al. 2014; Smits 2017).

As in Nurse (2007), our choice was primarily guided by the availability of a reasonable description. Interestingly, 14 out of the 16 outgroup languages display a strict SV order in their relative clauses. The only two languages that, to the best of our knowledge, depart from this pattern are Mungbam and Mundabli, which both show a SV/VS word order (Lovegren & Voll 2017).

In sum, what we observe after expanding Nsuka-Nkutsi's (1982) language sample is that VS remains the most common pattern found across the Bantu family, with 48% (72/151, cf. Table 1) of the languages represented in the Narrow Bantu sample. SV is however not far behind, with 42% (63/151). Geographically speaking, SV-only and VS-only seem equally widespread, except when it comes to Bantu zone A languages, in which SV-only is by far the most common pattern found so far. In this respect, Bantu zone A languages seem more similar to their Bantu relatives outside Narrow Bantu, which also mostly display the SV-only word order. This might be surprising considering the diversity that generally characterises North-Western Bantu languages and the fact that Guthrie's zone A does not correspond to a specific branch of the Bantu family tree. Instead, following the classification offered by Grollemund et al. (2015), our zone A languages spread over 2 different major branches (i.e. North-Western Cameroon and North-Western Gabon).

3.2 Genealogical distribution of SV-only, VS-only and SV/VS

Returning to the “majority wins” principle, used by Meeussen and Nsuka-Nkutsi, we have already noted in §1 that it should be used with caution, and that it is necessary to examine how the patterns are distributed over the major branches of the Bantu family to draw conclusions as to the order that characterised PB. Table 2 gives an overview of the distribution of our three patterns across major branches of the Bantu family tree offered by Grollemund et al. (2015).

Table 2: Distribution of SV-only, VS-only and SV/VS patterns across major branches of Bantu (subset of 124 Narrow Bantu languages)

Major Bantu branch	Number of languages	SV	VS	SV/VS
North-Western Cameroon	15	15	0	0
North-Western Gabon	7	5	1	1
Central-Western	30	3	27	0
West-Western	12	1	10	1
South-Western	3	0	3	0
Eastern	57	28	19	10
Total	124	52	60	12

Table 2 shows that SV-only is no more circumscribed to specific major branches of the Bantu family tree than VS-only is. In the subset of 124 languages that we could assign to the classification of Grollemund et al. (2015), SV-only is the most common pattern found both in the geographical North-West (i.e. the North-Western Cameroon and North-Western Gabon branches) and in the East and South (i.e. the Eastern branch). These results appear compatible with the idea that VS could be a later development and thus question Meeussen and Nsuka-Nkutsi’s idea that Proto-Bantu displayed a VS-only order in relative clauses. We come back to this conclusion in §4. Let us first examine some of the proposals laid out in §2 in the light of our expanded database and see whether VS in main and embedded clauses necessarily correlate.

3.3 Possible correlation with sentential VS

Our set of languages does not allow us to provide any direct evidence in favour of or against the claims in §2.2 regarding possible motivations for embedded inverted subjects. Together with the data found in Marten & van der Wal (2014),

it however allows us to check whether there are correlations between possible word orders at the sentential level and in relative clauses, and thus to get a better idea of whether there could be something specific to either domain that triggers (or licenses) the VS word order. In monoclausal sentences, it is fairly well established that the main motivation for inverted subjects is either focusing or detopicalisation (e.g. Marten 2014; Hamlaoui Forthcoming). If information structure plays a role in relative clause word order, we expect VS at the sentential level to go together with VS at the embedded level. If, on the other hand, a constraint such as the one proposed by Givón (1972) and Demuth & Harford (1999) is at play, VS should be found at the embedded level without necessarily being possible at the sentential level.³ By checking correlations, we can also see whether any geographical clustering emerges as to the use (or absence) of VS in one or both syntactic domains.

Checking against the set of 37 languages that are found both in Marten & van der Wal's (2014) and our database, we find that five languages seem to have VS in relative clauses only: Mbuun B87, Matuumbi P13, and Makhwa P31; and possibly Bembe D54 and Gciriku K332. This suggests that there might indeed be something specific to relative clauses that either forces or licenses a non-canonical word order and might help in explaining the predominance of VS-only over SV-only in a family in which the canonical order is SVO. We also find that nine languages have VS at the sentential level but not in relative clauses: Chaga E60, Nande JD42, Soga JE16, Bukusu JE31c, Tumbuka N21, Herero R30, Zulu S42, Sindebele S44, and possibly Rwanda JD61. Together, these results indicate that VS can occur in simple sentences but not relative clauses, and vice versa, and thus that the two processes can function independently.

Further checking our language set against Marten & van der Wal's, we find that 14 languages have VS in both simple sentences and embedded clauses: Nzadi B865, Dzamba C322, Kagulu G12, Makwe G402, Swahili G42d, Rundi JD62, Bemba M42, Ndendeule N101, Chewa N31b, Nsenga N41, Yao P21, Shona S10, and possibly Lega D25, and Swati S43. Our results indicate no clear correlation between word order in embedded and non-embedded clauses. For the sake of completeness, we find seven languages that have inverted subject neither in simple nor

³As noted by an anonymous reviewer, considerations of information structure could be different in main vs. embedded clauses. In the absence of evidence that in some languages information structure influences word order only in embedded and not in main clauses, our rationale is that if, in a particular language, information structure determines word order in embedded clauses, the most economic hypothesis is that main clauses are subject to the same rules/constraints rather than different ones. Our prediction is thus that if information structure is a key factor in the word order of embedded clauses, the likelihood is higher of finding VS in both embedded and main clauses in a particular language.

in embedded clauses: Basaa A43a, Kuyu E51, Tharaka E54, Lozi K21, Tswana S31, Southern Sotho S33 and Xhosa S41. No geographical clustering seems to arise, either, when it comes to the distribution of the combination (or absence) of VS in simple sentences and relative clauses.

Using our expanded database of Narrow Bantu languages and checking it against Marten & van der Wal's, we have provided a brief overview of the relation between VS in simple sentences/main clauses and in embedded clauses. What we have seen is that knowing possible word orders in one syntactic domain does not allow one to predict possible word orders in the other, suggesting that there could be distinct motivations for departing from the canonical word order in each of these domains. We now turn to our proposal regarding word order in PB relative clauses.

4 Word order in relative clauses

4.1 An alternative hypothesis?

So far, we have seen that neither the frequency and distribution of the VS-only and SV-only patterns nor the motivation for VS clearly allow us to conclude which order characterised PB. What is striking, however, is the uniformity of our North-Western Cameroon Bantu languages of zone A, which show the SV-only word order and are, in this respect, more similar to the non-Narrow Bantu languages of our sample, which also tend to show a strict SV order. Several scenarios can be considered. We have seen that Bantu zone A languages actually spread over two distinct major branches of the Bantu family according to the classification offered by Grollemund et al. (2015): North-Western Cameroon Bantu, in which 15 out of 15 languages show SV-only, and North-Western Gabon Bantu, in which five out of seven languages show SV-only, one shows VS-only and one SV/VS.

Given that North-Western Cameroon Bantu, which is a sister to the remainder of Narrow Bantu, only has SV, a word order also attested in nearly all other clades, it is most parsimonious to reconstruct SV to PB and to consider VS as a later innovation. The VS word order would then have emerged at node 2 or 3 in the phylogeny of Grollemund et al. (2015), or several times independently as a parallel innovation. Considering SV as the most archaic word order in non-subject relative clauses also ties with its prevalence in the closest Benue-Congo relatives outside Narrow Bantu.

Another scenario would consist in treating the languages showing VS-only as the more conservative ones, as suggested by both Meeussen and Nsuka-Nkutsi.

The present-day SV languages could have “shifted back” to SV from VS, an alternative that is considered by both Nsuka-Nkutsi (1982: 78) and more recently by Hyman (2012: 104) for SVs relative clauses in Nzadi. North-Western Cameroon languages, in particular, could also have developed the SV-only word order through contact with other Southern Bantoid languages in their vicinity. Nonetheless, at present, we do not have enough evidence of intensive language contact and multilingualism to substantiate such a claim.

Interestingly, something else could be at the source of the difference in word order between higher clades in the tree, i.e. the north-western part of the Bantu domain, and the lower ones elsewhere: the difference between analytic and synthetic verbal morphology that distinguishes the former from the latter. Recall Givón’s observation regarding relative clauses in Swahili: the more archaic pattern consists of a bound relativiser together with the VS-only word order, whereas the more innovative pattern (*amba*-relative clauses) consists of a free relativiser together with an SV/VS order.

In the Bantu literature, there is presently no consensus on the direction in which morphological typology in the Bantu family as a whole evolved. Whereas Hyman (2007; 2017) and Nurse (2007) defend the view that North-West Bantu languages generally went from being morphologically more synthetic to being more analytic, Güldemann (2003; 2011) argues for the opposite scenario. How would a particular verbal morphology relate to word order? In our view, what is crucial is the head-marking property typical of the more synthetic type of Bantu languages. Morphologically synthetic languages tend to show the discourse-driven, flexible word order considered typical of Bantu languages (Bearth 2003). As underlined in Schadeberg (2003: 152), subject and object concord is “the primary means to identify the arguments that function as subject and object”. In contrast, a more analytic language like Basaa does not have subject and object markers: it only has a single paradigm of personal pronouns, and their surface location (i.e. before or after the verb) is the only indicator of their grammatical function (Hyman 2003). Instead of a primarily discourse-driven word order, Basaa displays a so-called “indirect role marking” syntax (Noonan 1992), where surface position primarily encodes grammatical relations (and not information-structural status). As Bantu languages are SVO and thus preferably encode grammatical subjects preverbally and grammatical objects postverbally, it is not surprising to find an SV-only word order in Basaa relative clauses. As the Benue-Bantu languages outside Narrow Bantu in our sample also tend to show a more analytic morphology (Nurse 2007), this would be consistent with their showing SV-only word order in relative clauses too.

In our view, another property that generally restricts the possibility for VS in more analytic Bantu languages has to do with subject agreement and the fact that Bantu subjects generally need to precede the verb in order to agree with it. As outlined in Meeussen's example in (25), relative clauses with postverbal subjects typically have a verb whose subject agreement features are controlled by the head noun. If both nouns are animate and nothing morphologically distinguishes subjects from objects, VS relative clauses result in systematic argument structure ambiguities that are generally avoided by natural languages (Wasow 2015).

(25) Reconstruction of Proto-Bantu relative clause (Meeussen 1967: 114)

mu-ntu ju-dim-id-a ba-genj
1-person SMI-cultivate-APPL-FV 2-stranger

i. 'the person who cultivates for the strangers (subjective)'

ii. 'the person for whom the strangers cultivate (objective)'

In the absence of other morphosyntactic marking, a rigid word order can serve the purpose of reducing ambiguities in argument structure (see also Vennemann 1973 regarding the loss of case and the related change from SOV to SVO in English, and a recent discussion in Harris & Campbell 1995). If this is indeed the case, it would be expected that the more analytic Bantu languages generally should show little to no optionality in word order and a strong preference for SV (i.e. the canonical order) in relative clauses. (See however footnote 4.)

As to the more synthetic languages, their head-marking property could generally allow them to have postverbal subjects in main clauses and thus display a VS-only or even SV/VS word order in relative clauses (as seen for instance in the case of Swahili, which is a more synthetic Bantu language). Note however that nothing (aside from other, independent morphophonological considerations, as for instance discussed in §2.2) should in principle force synthetic languages to display these orders. They can also simply display a strict SV word order.

If we are on the right track, a correlation should be found between the analytic verbal morphology of particular Bantu languages and the absence of VS in their relative clauses.⁴ How to determine the level of analyticity/syntheticity of a particular language's verbal morphology is not a trivial question. To test our hypothesis, we additionally collected data on the type of (weak) object shown by the languages of our sample, and in particular whether they retained the object

⁴Note however that some analytic languages have ways of encoding grammatical relations other than a strict word order, for instance through distinct pronoun paradigms as in Nen A44 (Maarten Mous, p.c.). These languages might thus show a more flexible syntax than a language like Basaa A43a.

prefix slot at all (Meeussen 1967: 109). The basic idea was for us to be able to distinguish languages that qualify as being of the head-marking type, which tend to have an object prefix, from the ones, such as Basaa, which lack this property and have postverbal object pronouns instead. To do so, we would ideally need to look at two aspects of object marking: how a (weak) object is encoded (affix vs. pronoun) and where it is encoded (preverbally or postverbally). This would at least yield the four following types of languages: oV (typical, synthetic Bantu-type), V o (Basaa-type), Vo and o V.

As stated in Polak (1986: 371), citing Gregersen (1967), the distinction between pronouns and agreement affixes (“*éléments d’accord*”) is often difficult in Bantu. Additionally, as many Bantu languages only allow for one object prefix, when several objects are pronominalised, they follow the verb, either as free pronouns (“*un substitutif qui suit le verbe*”) or as suffixes/enclitics, meaning that many languages have both an object prefix and a postverbal pronoun/enclitic.

The overall picture is also slightly more complex in that among languages with an object suffix, several types are attested. By way of illustration, in Suku H32 (Polak 1986: 376), objects referring to humans are encoded with a prefix and other weak objects are encoded by means of a suffix, with a few exceptions with non-human indirect objects, which can be encoded as prefixes (Piper 1977). In a typology of weak object marking, Suku would thus classify as oV/Vo.

Object pronouns and enclitics are, according to Polak (1986: 377), often morphologically similar, but behave differently in terms of the tonal and segmental processes to which they are subjected. Some languages, such as Myene B11, alternate between the two types of postverbal weak objects, further complicating the typology.

In her study, Polak (1986) distinguishes only between object prefixes (so-called infixes), (postverbal) autonomous pronouns and enclitics, and notes that enclitic objects do not seem to exist in the East (zones E, G, N, P, S). She does not mention cases of preverbal autonomous objects as found in some Bantu zone A languages (and discussed in the next subsection).

As most preverbal object markers are prefixes and as it is difficult to truly distinguish postverbal object pronouns from object enclitics without having access to (often not-yet existing) much more detailed studies of individual languages, we have so far distinguished only the three following types: languages that have pre-stem object markers only (oV), languages that have both pre-stem object markers and pre- or postverbal object pronouns or enclitics (oV/Vo) and languages that only have postverbal object pronouns or enclitics (Vo). The last group is the crucial one to our hypothesis. We are aware that one could argue that some of these languages might not have an object pronoun but rather an enclitic or a suffix

that would militate in favour of classifying them as less analytic. We leave the detailed analysis of these languages open for future research.

4.2 Analytic verbal morphology and (absence of) VS order

Collecting data from existing studies on object markers as well as from grammatical sketches (Polak 1983; 1986; Beaudoin-Lietz et al. 2004; Marlo 2014), our hypothesis can be tested on a sample of 162 languages: our 16 outgroup languages and 146 Narrow Bantu languages.⁵ With the exception of Moro (Kordofanian) (Jenks & Rose 2015), which displays both pre- and post-stem object markers, the rest of our outgroup languages have a strict Vo order.

Figure 3 shows the distribution of our object-marking types (Vo-only, oV-only and oV/Vo) across the Bantu zones, while Table 3 shows the distribution of these types across the major sub-branches of Bantu.

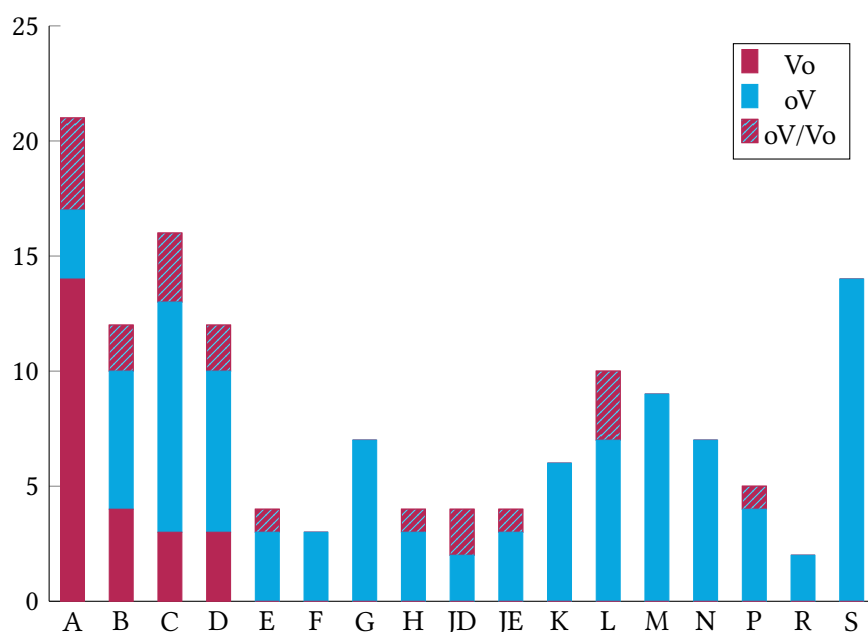


Figure 3: Distribution of Vo, oV and oV/Vo object markers across Bantu zones (146 languages)

⁵Information regarding object marking in the following languages could not be found: Bakutu C61A, Konda C61E, Yela C74, Konzo JD41, and Soga JE16.

Table 3: Distribution of Vo-only, oV-only and oV/Vo patterns across major branches of Bantu (subset of 120 Narrow Bantu languages)

Major Bantu branches	Number of languages	Vo	oV	oV/Vo
North-Western Cameroon	15	9	3	3
North-Western Gabon	7	5	0	2
Central-Western	27	5	17	5
West-Western	12	3	8	1
South-Western	3	0	3	0
Eastern	56	0	52	4
Total	120	22	85	13

Unsurprisingly, languages with no pre-stem object markers are found only in the north-western part of the Bantu domain (zones A, B, C and D), which is consistent with the fact that they are the most analytic in terms of verbal morphology (Nurse 2007; Hyman 2017). In our sample, other Bantu languages seem to show a pre-stem object marker-only pattern, while others show both pre- and postverbal object marking (with the object prefix sometimes only limited to reflexive markers).

According to the descriptions we accessed, three North-Western Cameroon Bantu languages exhibit what we have classified as oV/Vo: Nen A44, Duala A24 and Tuki A601. In the case of Nen, note that oV is actually different from what is found in typical Bantu languages, as the object is here a preverbal pronoun rather than a prefix. In examples from Mous (1997: 126), a second object can even appear between the verb and a preverbal object pronoun. Just like Nen, Duala is also classified by Nurse (2007: 254) as belonging to the more analytical type of languages. According to Polak (1986: 374) the only pre-stem object prefix left in Duala is a reflexive that is about to disappear. Tuki however seems to have a more agglutinative morphology, with a pre-stem object marker rather than a preverbal object pronoun (Bilola 2013).

So far we have three languages in the oV-only category among North-Western Cameroon Bantu languages: Bubi A31, Maande A46, Gunu A622. As stated above, our information might be incomplete and some or all of these languages might also have postverbal object pronouns or enclitics. What is crucial for us is whether they can be said to belong to the more analytic type of languages. We believe that this is the case for Maande (Wilkendorf 2001) and this is how it is classified by Nurse (2007: 253). This is also the case for Gunu according to Nurse and based

on data from Orwig (1991). It is however unclear for Bubi, which might be of the more agglutinative/synthetic type (Clarke 1848). Maande and Gunu illustrate the fact that a more analytic morphology is not necessarily exclusive with a pre-stem object marker in Bantu. As acknowledged in previous studies on the topic, languages sit on a continuum. More work is needed in this area to establish a more fine-grained typology.

Figure 4 shows subject-verb word order in relative clauses as a function of the type of object marking in our 146 Bantu languages as well as in our 16 outgroup languages ($n = 162$). Languages with only pre-stem object markers (oV) conform to what we have observed in §2, in that the VS word order is the most frequent ($n = 51$). They also show a considerable number of languages with only SV ($n = 38$). Interestingly, 13 out of our 17 languages displaying a flexible word order with SV/VS are found in the oV-only group.

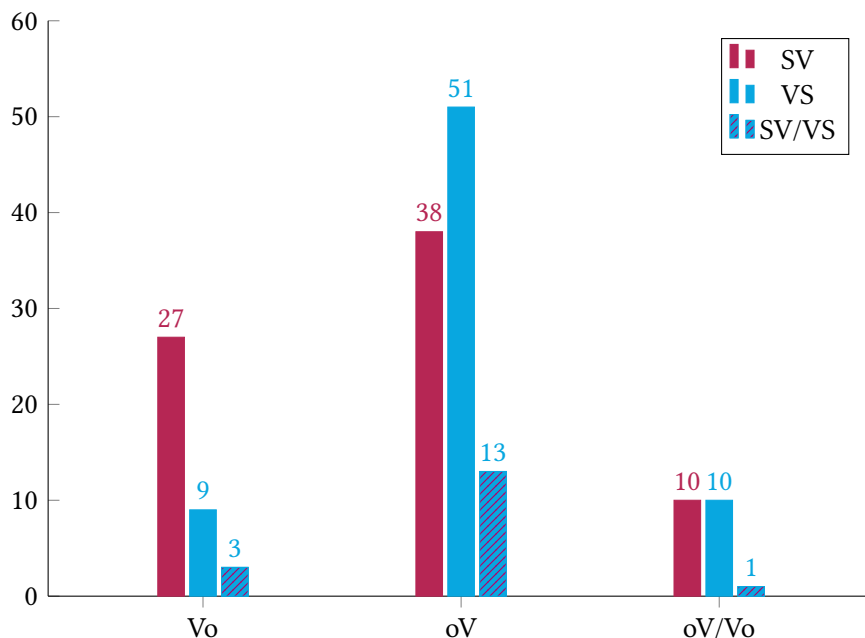


Figure 4: Relative clause subject position by object marking type (162 languages)

Languages with both pre-stem objects and pronouns/enclitics (oV/Vo) have as many languages displaying a VS-only word order ($n = 10$) as languages displaying an SV-only word order ($n = 10$). Only one language allows both VS and SV.

The Vo group is the one that interests us most in connection with our hypothesis, as it is the one in which pre-stem object markers are absent and for which we expect the word order to be much less flexible. What we observe is that in this group the tendencies are reversed: SV is the predominant pattern ($n = 27$), followed by VS ($n = 9$) and three languages that allow both SV and VS.

To investigate whether there is a relationship between subject-verb word order and object marking in our sample of 162 languages, a chi-square test of independence was conducted. The result of this test was significant, chi-square (4) = 12.52, $p < .05$. However, the effect size of this relationship (i.e. the strength of this effect) was weak, Cramér's $V = .196$.⁶ Examination of standardised residuals indicates that of the 39 languages that display Vo, 69.2% also display SV, while of the 102 languages that show oV, only 37.2% show SV. At the same time, the proportion of languages that use oV and VS is more than two times higher than the proportion of Vo languages that use VS (50.0% (51/102) vs. 23.0% (9/39)). These results thus tend to confirm our hypothesis that the verbal morphology specific to languages in the north-western part of the Bantu domain might lie at the origin of the preference for the SV order in our sample for this area. Our contention is that this morphological typological difference probably goes hand in hand with radical syntactic differences and a general lack of word order flexibility compared to more typical, morphologically synthetic Bantu languages.

Although our results generally fit with our prediction that Vo-only languages should favour the SV-only order, 9 of our Vo-only languages still favour VS: Myene B11, Duma B51, Mbede B61, Ndumu B63, Mboshi C25, Soko C52, Kele C55, Mbole D11, and Enya D14. As these languages are surrounded by oV/Vs, an effect of contact cannot be excluded and could explain why they retained VS despite the systematic argument-structural ambiguity associated with this word order. Here we can examine one of these Vo/Vs languages more closely, i.e. Mboshi, to show why it is actually not a problem for our hypothesis. In the existing literature on this language, some of its relative clauses indeed display the above-mentioned type of argument-structural ambiguity, so that subject and object cannot be identified with certainty (Beltzung et al. 2010). This is illustrated in (26) (the first line is the phonetic form of the sentence while the second line is its phonological form).

⁶Note that we still find statistical significance if we conflate our oV-only and oV/Vo categories: chi-square (2) = 11.04, $p < .05$. This is important in case further examination of the languages we have classified as oV-only revealed that some also have postverbal pronouns/enclitics in addition to the pre-stem object marker. The crucial group remains the Vo-only group.

- (26) Mboshi C25 (Beltzung et al. 2010: 22)
- ndzɔyi yebomí obengi*
 N-dzɔyi ye-ye-bom-i í mo-bengi
 1a-elephant REL1-SM1-kill-FV CONJ.H 1-hunter
 i. ‘the elephant that killed the hunter’
 ii. ‘the elephant that the hunter killed’

Interestingly, this language seems to have an alternative strategy to disambiguate this structure: an auxiliary (/di/) can be used to impose a fixed word order, which yields two different word orders depending on the interpretation of the sentence. In (27a) the object must follow the auxiliary+verb complex, while in (27b) the subject must precede the lexical verb.

- (27) Mboshi C25 (Beltzung et al. 2010)
- a. *ɔmbɔri móódze lóbéré Jean*
 mo-mbɔri mo’-mo’-di-i lá Co-bér-a Jean
 3-gendarme REL3-H-SM3-H-AUX-FV with INF-hit-FV Jean
 ‘the gendarme who hit Jean’
- b. *ɔmbɔri móódze Jean lóbéré*
 mo-mbɔri mo’-mo’-di-i Jean lá Co-bér-a
 3-gendarme REL3-H-SM3-H-AUX-FV Jean with INF-hit-FV
 ‘the gendarme whom Jean hit’

The avoidance of argument-structural ambiguities might have motivated a shift “back” to a strict SV order. Mboshi is not the only present-day Bantu language that allows relative clauses which are ambiguous from an argument-structural perspective. Based on the data in (27), the question however arises as to whether, in speakers’ productions, ambiguous relative clauses are not already supplanted by other structures with a strict SV word order and/or richer morphological marking of argument relations.

Finally, Kwakum A91 also shows a rather unexpected pattern with respect to our predictions. This language indeed shows a strict Vo order but allows both pre- and postverbal subjects in some of its relative clauses. According to David M. Hare (p.c.), although only preverbal subjects are allowed in object relative clauses of the type in (28), i.e. a restrictive clause with a transitive verb, both orders are acceptable in (29), i.e. a non-restrictive clause with an intransitive verb.⁷

⁷At the time of writing, data on restrictive relative clauses with intransitive verbs and non-restrictive relative clauses with transitive verbs were not available. We refer the interested reader to David M. Hare’s future work on Kwakum relative clauses.

(28) Kwakum A91 (David M. Hare, p.c.)

- a. *ní'á kum baki mo Emanu mé jaŋse*
1.SG.PST2 find hoe REL Emanu PST2 lose
'I found the hoe that Emanu lost.'
- b. **ní'á kum baki mo mé jaŋse Emanu*
1.SG.PST2 find hoe REL PST2 lose Emanu
'I found the hoe that lost Emanu.'

(29) Kwakum A91 (David M. Hare, p.c.)

- a. *ní'á kwalyε ɔ AbongMbang ndɔɔ mbɔnjo je njilɔ yi*
1.SG.PST2 arrive LOC AbongMbang REL Makaa 3.PL live REL
'I arrived in AbongMbang, where the Makaa live.'
- b. *ní'á kwalyε ɔ AbongMbang ndɔɔ je njilɔ mbɔnjo yi*
1.SG.PST2 arrive LOC AbongMbang REL 3.PL live Makaa REL
'I arrived in AbongMbang, where live the Makaa.'

Kwakum is thus similar to two of our outgroup languages, Mungbam and Mundabli, which also only display postverbal objects and, according to Lovegren & Voll (2017), allow both VS and SV in their relative clauses. One significant difference between these languages and a language like Basaa is that Mungbam, Mundabli and Kwakum show different pronoun paradigms for different grammatical functions. Changes in word order would thus not result in as much argument-structural ambiguity in the latter languages. More research is however necessary to determine the full range of contexts (e.g. relative clause types, verb types, information structure) in which VS is licit in Mungbam, Mundabli and Kwakum and whether it results in the type of systematic ambiguity that other languages seem to avoid.

4.3 Further possible effects of verbal morphology on the lack of VS: Expressing focus

Languages in the north-western part of the Bantu domain are generally seen as much more diverse than those further South and East (Bearth 2003). We have seen that when it comes to word order in relative clauses, our Bantu zone A languages show a rather uniform pattern, with 20 out of 21 languages displaying only the SV order. Another way they tend to differ from other Bantu languages, it seems, is in the association for an item in being postverbal and being focused. In many Bantu languages, there is a strong relation between focus and either

an immediately postverbal position or the right edge of the clause. Such a connection is also common in other language families, like Romance and Chadic for instance, to the extent that a number of generalisations have been formulated as to the natural connection between being postverbal and being focused. Such a generalisation is found, for instance, in Fiedler et al. (2010: 255):

Whenever a subject is not to be interpreted as topic, but as focus, it must occur in the prototypical focus position, that is, in a postverbal position at the right edge of VP.

This particular relation between syntax and information structure is not shared by every Bantu language. In Basaa, for instance, constituents are focused *in situ*, and there is no general connection between being focused and being postverbal (or any non-canonical word order to express focus) (Hamlaoui & Makasso 2015). We propose that the absence of VS order is, in this context as well, related to the more analytic morphology and, more specifically, to the lack of pre-stem object markers. Our contention is that by lacking the pre-stem object markers, analytic languages like Basaa lack the opposition between weak (i.e. discourse-given or anaphoric) and strong (i.e. discourse-new or focused) objects visible in (typical) more synthetic Bantu languages. According to Güldemann (2003: 185), who discusses the functional contrast between Bantu pre- and postverbal objects, “the postverbal position is associated with the pragmatic function to present new, asserted information. An object concord, however, most often refers to something given and extrafocal which would disfavor its place after the verb.”

In more analytic languages such as Basaa, the canonical position of objects is thus more restricted to the postverbal domain. In the absence of other focus marking devices (e.g. prosodic prominence), the postverbal position becomes information-structurally neutral and thus not reserved to non-anaphoric/focused objects in opposition to anaphoric/non-focused objects, which appear elsewhere. In this context, there is no reason for equating postverbal with focused and, as a consequence, no reason for placing other focused items, such as subjects, after the verb. Instead of being a “natural” field for focus, the postverbal domain might thus simply be the neutral location of full objects in a number of SVO languages which also tend to have preverbal object markers (Bantu) or object proclitics (Romance). In the absence of preverbal object markers, as in a number of Bantu zone A languages, the association between being focused and being postverbal simply does not hold.

An interesting question is why other Bantoid languages, like many Grassfields languages, display a strong connection between being focused and being postverbal despite the fact that they are generally considered more analytic as well. Many

of these languages however seem to have grammatical properties that are not necessarily shared by some Bantu zone A languages (e.g. the availability of expletive subjects and case within the pronominal system), which might explain why they display a more flexible, information-structure-driven word order despite their analytic morphology. Additionally, it might be because, instead of lacking a preverbal morphological slot for weak (defocused/anaphoric) objects, they actually have a full-blown preverbal syntactic slot for them, as is the case for instance in Mungbam and Mundabli (Lovegren & Voll 2017: 21). Just like in more synthetic languages, the word order of these analytic languages can be primarily discourse-driven. We leave this question, and in particular the direction of the change in typological morphology within Bantoid, open for future research.

5 Conclusion

After expanding Nsuka-Nkutsi's (1982) sample to a total of 167 languages (151 Narrow Bantu and 16 other Niger-Congo languages), VS is still the most frequent word order in Bantu relative clauses. However, Bantu zone A languages predominantly show an SV-only word order. We have questioned the claim by Meeussen (1967) and Nsuka-Nkutsi (1982), that Proto-Bantu, understood here as node 1 in Grollemund et al. (2015), had a VS-only word order. What we see when examining both the geographical and the genealogical distribution of different word orders is that SV-only is the dominant pattern in the major clades of Grollemund et al. (2015) situated in the north-western Bantu area: 20 out of 22 languages, cf. Table 2 in §3.2 and the phylogenetic tree in Appendix A. These languages are both closer to the Bantu homeland and more similar to the Niger-Congo languages outside of Narrow Bantu in our sample, as the latter languages also predominantly show SV-only order in their relative clauses. Even if the SV-only word order found in these areas is innovative, as what we believe would be a natural consequence of the shift in morphology argued for by Nurse (2007) and Hyman (2017) (i.e. from synthetic to analytic verbal morphology), SV-only is also found in a significant portion of our sample in the major Eastern branch (28 out of 57 languages), together with a more typical, synthetic verbal morphology. The VS order could thus be an innovation that came into use only after the split between the major North-Western Cameroonian branch of Grollemund et al.'s (2015) classification and the rest of the tree, i.e. node 2 or 3. If this is correct, Bantu zone A languages would not have lost the VS order shown by a common ancestor to them and the rest of the Bantu family, but rather, they would not have had it at all (see for instance Ehret (1972) for a similar perspective on other features of North-Western

Bantu languages). Research on a larger set of Bantu zone A languages might shed some light on whether there is any evidence for VS in the relative clauses of this zone and, particularly, in the North-Western Cameroon clade of the Bantu tree.

As we have seen, due to some morphosyntactic properties typical of Bantu languages, VS sometimes leads to systematic argument structure ambiguities which languages generally tend to avoid. We have mentioned evidence of this from two languages, Swahili and Mboshi, whose basic word order in relative clauses seems to be VS. Interestingly, these languages also have alternative relative clause structures (with *amba* and the copula *di*, respectively) which either allow SV (Swahili) or impose it (Mboshi). In a language in which VS relative clauses are ambiguous, the introduction of SV might lead to the eventual loss of VS if there are no functional (e.g. information-structural) differences between the two alternatives and thus a possible shift “back” from VS to SV comparable to the one considered by Nsuka-Nkutsi (1982) and Hyman (2012).

We have also mentioned the case of the Grassfields speech varieties Mungbam and Mundabli which, rather against expectations considering the prevalence of the SV-only pattern in our outgroup sample, show a SV/VS word order (Lovegren & Voll 2017). These languages also show a VS order in main clauses when the subject is focused, and generally associate focus with the immediately after the verb position. According to Lovegren & Voll (2017), the same information structure-motivated alternations in word order are found in main and relative clauses, making SV and VS functionally different in these languages. As suggested by a reviewer, it is possible that these Bantoid languages have developed this alternation in constituent order as an independent innovation.

Independently of the direction of the analytic vs. synthetic morphological shift, we have proposed that what distinguishes our SV-only languages in the north-western part of the Bantu domain from other Bantu languages is the fact that they primarily, or even exclusively, encode grammatical relations through word order. Using Basaa as a reference, we have argued that due to the lack of devices such as object concord (commonly found in Bantu languages from the East and South) and distinct paradigms of pronouns (as in some Bantu languages in the North-West and Grassfields languages), the VS word order would lead to systematic argument structure ambiguities.

Finally, we have put forward the idea that the above-mentioned differences in object marking morphology between Bantu zone A and other languages could have further consequences for their syntax. In particular, we have proposed that the lack of association between being focused and being postverbal might be related to the general lack of contrast between preverbal weak/anaphoric objects (object prefixes) and full new/focused objects. One of the questions that remains

open is why Grassfields languages, which also display a more analytic morphology, still maintain the contrast between neutral preverbal subjects and focused postverbal ones.

Acknowledgements

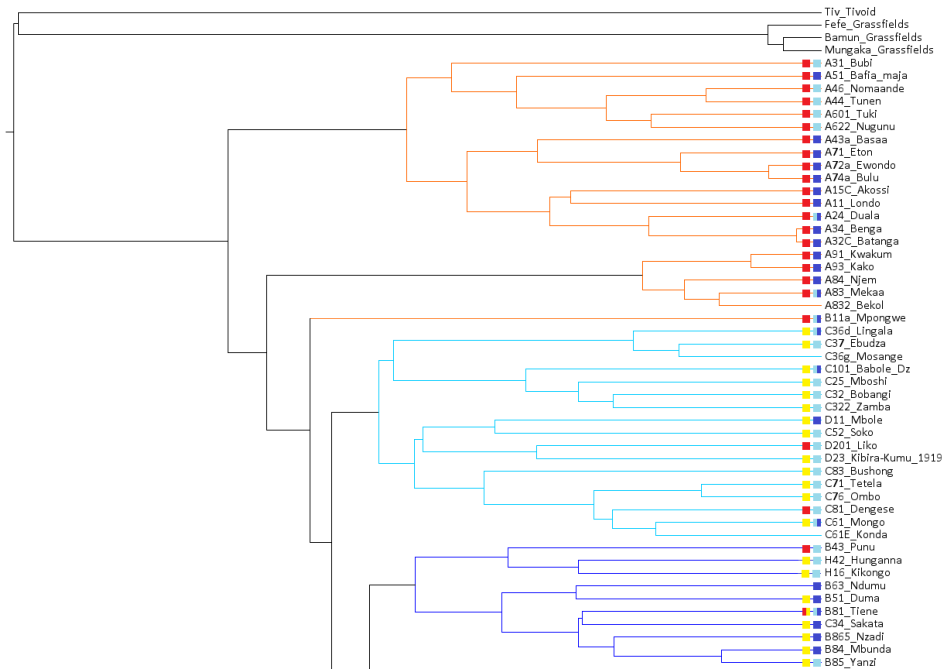
Many thanks go to the organisers of the workshop on “Reconstructing Proto-Bantu Grammar” and the editors of this volume, and in particular to Koen Bostoen for extensive feedback on previous versions of this chapter. I am very grateful to Rebecca Grollemund and two anonymous reviewers for helpful comments, as well as to Jacky Maniacky for generously providing me with a copy of François Nsuka-Nkutsi’s precious dissertation. Heartfelt thanks also go to Frank Collins for proofreading this chapter. The usual disclaimers apply.

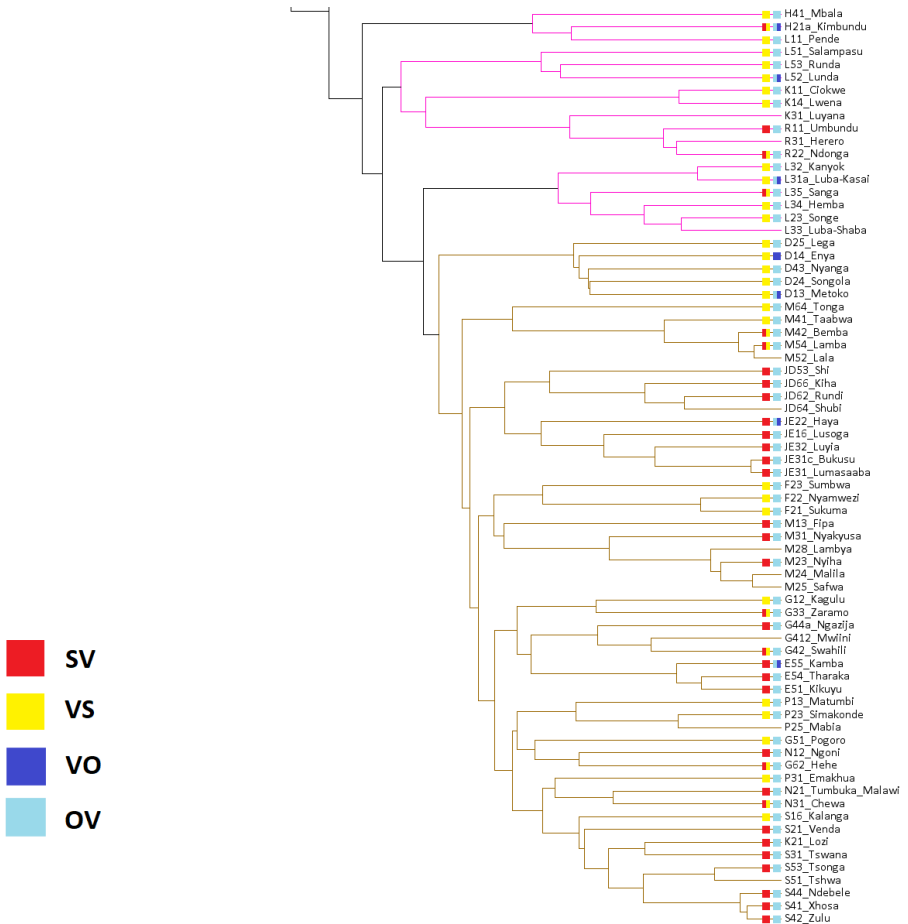
Abbreviations

AG	agent
AUG	augment
AUX	auxiliary
CONJ	conjunctive (tone)
COP	copula
DEM	demonstrative
FV	final vowel
G (‘g-form’)	reflex of the PB pre-final morpheme (Van de Velde 2017)
INF	infinitive
LOC	locative
NEG	negation
NF	non-final marker (Van de Velde 2017)
OM	object marker
PFV	perfective
PL	plural
POSS	possessive
PRES	present
PRO	pronoun
PROG	progressive
PST	past
REL	relative
REM	remote

SG	singular
SM	subject marker
TAM	tense/aspect marker
TNS	tense
1, 2, 3 ...	noun classes

Appendix A Word order and object marking across branches of the Bantu phylogenetic tree (107 languages)





[Figure produced by R. Grollemund, using the phylogenetic tree presented in Grollemund et al. (2015) as a base.]

References

Allan, Edward Jay. 1973. *A grammar of Buem: The Lelemi language*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).

Atindogbé, Gratién Gualbert & Rebecca Grollemund (eds.). 2017. *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58). Berlin: De Gruyter Mouton.

Ba, Ibrahima. 2015. Factive relative clauses in Pulaar. *Kansas Working Papers in Linguistics* 36. 84–99.

- Bearth, Thomas. 2003. Syntax. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 121–142. London: Routledge.
- Beaudoin-Lietz, Christa, Derek Nurse & Sarah Rose. 2004. Pronominal object marking in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 175–188. Cologne: Rüdiger Köppe.
- Beltzung, Jean-Marc, Annie Riailand & Georges Martial Embanga Aborobongui. 2010. Les relatives possessives en mbochi (C25). In Laura J. Downing, Annie Riailand, Jean-Marc Beltzung, Sophie Manus, Cédric Patin & Kristina Riedel (eds.), *Papers from the Workshop on Bantu Relative Clauses* (ZAS Papers in Linguistics 53), 7–31. Berlin: Zentrum Allgemeine Sprachwissenschaft.
- Bilola, Edmond. 2013. *The syntax of Tuki: A cartographic approach* (Linguistik Aktuell / Linguistics Today 203). Amsterdam: John Benjamins.
- Bostoen, Koen & Léon Mundeke. 2011. Passiveness and inversion in Mbuun (Bantu B87, DRC). *Studies in Language* 35(1). 72–111.
- Bostoen, Koen & Léon Mundeke. 2012. Subject marking, object-verb order and focus in Mbuun (Bantu, B87). *Southern African Linguistics and Applied Language Studies* 30(2). 139–154.
- Campbell, Lyle. 1998. *Historical linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- Carstens, Vicki M. 2005. Agree and EPP in Bantu. *Natural Language & Linguistic Theory* 23(2). 219–279.
- Cheng, Lisa L.-S. Forthcoming. Relative clauses. In Lutz Marten, Nancy C. Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages* (Oxford Guides to the World's Languages 2). Oxford: Oxford University Press.
- Clarke, John. 1848. *Introduction to the Fernandian tongue*. 2nd revised edition. Berwick-on-Tweed: Daniel Cameron.
- De Kind, Jasper. 2014. Pre-verbal focus in Kisikongo (H16a, Bantu). In Fatima Hamlaoui (ed.), *Proceedings of the Workshop BantuSynPhonIS: Preverbal domain(s)* (ZAS Papers in Linguistics 57), 95–122. Berlin: Zentrum Allgemeine Sprachwissenschaft.
- Demuth, Katherine & Carolyn Harford. 1999. Verb raising and subject inversion in Bantu relatives. *Journal of African Languages and Linguistics* 20(1). 41–61.
- Dimmendaal, Gerrit J. 2011. *Historical linguistics and the comparative study of African languages*. Amsterdam: John Benjamins.
- Downing, Laura J. & Al Mtenje. 2011. Prosodic phrasing of Chichewa relative clauses. *Journal of African Languages and Linguistics* 32(1). 65–112.

- Downing, Laura J., Annie Riailand, Jean-Marc Beltzung, Sophie Manus, Cédric Patin & Kristina Riedel (eds.). 2010. *Papers from the Workshop on Bantu Relative Clauses* (ZAS Papers in Linguistics 53). Berlin: Zentrum Allgemeine Sprachwissenschaft.
- Dzameshie, Alex K. 1995. Syntactic characteristics of Ewe relative clause constructions. *Research Review (NS)* 11(1-2). 27–42.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2022. *Ethnologue: Languages of the world*. 25th edn. Dallas, TX: SIL International.
- Ehret, Christopher. 1972. Bantu origins and history: Critique and interpretation. *Transafrican Journal of History* 2(1). 1–9.
- Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz & Malte Zimmermann. 2010. Subject focus in West African languages. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives* (Oxford Linguistics). Oxford: Oxford University Press.
- Givón, Talmy. 1972. A note on subject postponing. *Studies in African Linguistics* 3(2). 289–299.
- Gregersen, Edgar A. 1967. *Prefix and pronoun in Bantu* (Indiana University Publications in Anthropology and Linguistics Memoir 21). Baltimore, MD: Waverly Press.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Güldemann, Tom. 2003. Grammaticalization. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 182–194. London: Routledge.
- Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas: With special reference to Africa* (Tokyo University of Foreign Studies, Studies in Linguistics 2), 109–141. Amsterdam: John Benjamins.
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Hamlaoui, Fatima. Forthcoming. Inversion constructions. In Lutz Marten, Nancy C. Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages* (Oxford Guides to the World's Languages 2). Oxford: Oxford University Press.

- Hamlaoui, Fatima & Emmanuel-Moselly Makasso. 2015. Focus marking and the unavailability of inversion structures in the Bantu language Bàsàá (A43). *Lingua* 154. 35–64.
- Hare, David M. 2018. Relative clauses in Kwakum (A91). Unpublished ms.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective* (Cambridge Studies in Linguistics 74). Cambridge: Cambridge University Press.
- Henderson, Brent M. 2006. *The syntax and typology of Bantu relative clauses*. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign. (Doctoral dissertation).
- Hyman, Larry M. 2003. Basaá (A43). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 257–282. London: Routledge.
- Hyman, Larry M. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville, MA: Cascadilla Press.
- Hyman, Larry M. 2012. Post-verbal subject in the Nzadi relative clause. *Journal of African Languages and Linguistics* 33(1). 97–117.
- Hyman, Larry M. 2017. Multiple argument marking in Bantoid: From syntheticity to analyticity. In Walter Bisang & Andrej Malchukov (eds.), *Unity and diversity in grammaticalization scenarios* (Studies in Diversity Linguistics 16), 67–95. Berlin: Language Science Press. DOI: 10.5281/zenodo.823230.
- Jenks, Peter, Emmanuel-Moselly Makasso & Larry M. Hyman. 2017. Accessibility and demonstrative operators in Basaá relative clauses. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 17–46. Berlin: De Gruyter Mouton.
- Jenks, Peter & Sharon Rose. 2015. Mobile object markers in Moro: The role of tone. *Language* 91(2). 269–307.
- Kawasha, Boniface. 2008. Relative clauses and subject inversion in Chokwe, Kaonde, Lunda and Luvale. *Africana Linguistica* 14. 37–61.
- Kisseberth, Charles W. 2010. Phrasing and relative clauses in Chimwiini. In Laura J. Downing, Annie Riailand, Jean-Marc Beltzung, Sophie Manus, Cédric Patin & Kristina Riedel (eds.), *Papers from the Workshop on Bantu Relative Clauses* (ZAS Papers in Linguistics 53), 109–144. Berlin: Zentrum Allgemeine Sprachwissenschaft.
- Kuno, Susumu. 1973. *The structure of the Japanese language* (Current Studies in Linguistics 3). Cambridge, MA: The MIT Press.

- Letsholo, Rose. 2009. The 'forgotten' structure of Ikalanga relatives. *Studies in African Linguistics* 38(2). 131–154.
- Lovegren, Jesse & Rebecca M. Voll. 2017. Relative clause constructions in two Yemne-Kimbi languages. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 169–208. Berlin: De Gruyter Mouton.
- Manus, Sophie. 2010. The prosody of Símákonde relative clauses. In Laura J. Downing, Annie Riailand, Jean-Marc Beltzung, Sophie Manus, Cédric Patin & Kristina Riedel (eds.), *Papers from the Workshop on Bantu Relative Clauses* (ZAS Papers in Linguistics 53), 159–186. Berlin: Zentrum Allgemeine Sprachwissenschaft.
- Marlo, Michael R. 2014. Exceptional patterns of object marking in Bantu. *Studies in African Linguistics* 43(2). 85–123.
- Marten, Lutz. 2014. The preverbal position(s) in Bantu inversion constructions: Theoretical and comparative considerations. In Fatima Hamlaoui (ed.), *Proceedings of the Workshop BantuSynPhonIS: Preverbal Domain(s)* (ZAS Papers in Linguistics 57), 136–159. Berlin: Zentrum Allgemeine Sprachwissenschaft.
- Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318–368.
- Martin, Marieke. 2017. Relative clauses in Wawa. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 233–260. Berlin: De Gruyter Mouton.
- McCracken, Chelsea. 2013. Relative clauses in Asante Twi. *Rice Working Papers in Linguistics* 4. 1–28.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Mkude, Daniel J. 1974. *A study of Kiluguru syntax with special reference to the transformational history of sentences with permuted subject and object*. London: University of London. (Doctoral dissertation).
- Mous, Maarten. 1997. The position of the object in Tunen. In Rose-Marie Déchaine (ed.), *Object positions in Benue-Kwa: Papers from a Workshop at Leiden University, June 1994* (Holland Institute of Generative Linguistics: HIL Publications 4), 127–137. The Hague: Holland Academic Graphics.
- Mous, Maarten. 2003. Nen (A44). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 283–306. London: Routledge.

- Ngoyani, Deo. 2001. Evidence for head raising in Kiswahili relative clauses. *Studies in African Linguistics* 30(1). 59–73.
- Noonan, Michael. 1992. *A grammar of Lango* (Mouton Grammar Library 7). Berlin: De Gruyter Mouton.
- Nsuka-Nkutsi, François. 1982. *Les structures fondamentales du relatif dans les langues bantoues* (Annales – Série in-8° – Sciences humaines 107). Tervuren: Royal Museum for Central Africa.
- Nurse, Derek. 2007. Did the Proto-Bantu verb have a synthetic or an analytic structure? In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 239–256. London: Department of Linguistics, School of Oriental & African Studies.
- Orwig, Carol. 1991. Relative time reference in Nugunu. In Stephen C. Anderson & Bernard Comrie (eds.), *Tense and aspect in eight languages of Cameroon* (SIL International Publications in Linguistics 99), 147–162. Dallas, TX: Summer Institute of Linguistics & the University of Texas at Arlington.
- Pasch, Helma & François Mbolifouye. 2011. I am subordinate to Gbudwe, but your sovereign: Using a subject pronoun in object position in order to claim power. *Afrikanistik Online* 8. <https://www.afrikanistik-aegyptologie-online.de/archiv/2011/2907>.
- Piper, Klaus. 1977. *Elemente des Suku: Zur Phonologie und Morphologie einer Bantusprache*. Leiden: Rijksuniversiteit te Leiden. (Doctoral dissertation).
- Polak, Louise. 1983. Le réfléchi en bantou. *Africana Linguistica* 9. 271–304.
- Polak, Louise. 1986. Les infixes (“prefixes objets”) du bantou et leur reconstruction. *Africana Linguistica* 10. 365–421.
- Rose, Sharon, Farrell Ackerman, George Gibbard, Peter Jenks, Laura Kertz & Hannah Rohde. 2014. In-situ and ex-situ wh-question constructions in Moro. *Journal of African Languages and Linguistics* 35(1). 91–125.
- Schadeberg, Thilo C. 2003. Historical linguistics. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 143–163. London: Routledge.
- Schwarz, Anne. 2006. *Sentence-medial adverbials in Buli (Gur, Northern Ghana)*. Berlin: Humboldt University of Berlin.
- Smits, Helena Johanna. 2017. *A grammar of Lumun, a Kordofanian language of Sudan*. Leiden: Leiden University. (Doctoral dissertation).
- Song, Sanghoun. 2014. Information structure of relative clauses in English: A flexible and computationally tractable model. *Language and Information* 18(2). 1–29.

- Tabé, Florence A. E. & Gratien Gualbert Atindogbé. 2017. Kenyang relative clauses. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 89–118. Berlin: De Gruyter Mouton.
- Takizala, Alexis. 1972. Focus and relativization: The case of Kihung'an. *Studies in African Linguistics* 3(2). 252–286.
- Tamanji, Pius N. & Blasius Achiri-Taboh. 2017. A prolegomenon to the syntax of the relative clause in the Eastern Grassfields Bantu borderland. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 145–168. Berlin: De Gruyter Mouton.
- Thwing, Rhonda. 2017. Relative clauses in Vute grammar and discourse. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 209–232. Berlin: De Gruyter Mouton.
- Van de Velde, Mark. 2017. The augment as a construct form marker in Eton relative clause constructions. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 47–66. Berlin: De Gruyter Mouton.
- Van de Velde, Mark. 2022. Agreement on Proto-Bantu relative verb forms. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 465–494. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575835.
- Vennemann, Theo. 1973. Topics, subjects, and word order: From SXV to SVX via TVX. In John M. Anderson & Charles Jones (eds.), *Historical linguistics: Proceedings of the first International Conference on Historical Linguistics, Edinburgh, 2nd–7th September 1973*, vol. 2 (North-Holland Linguistic Series 12b), 339–376. Amsterdam: North-Holland Publishing.
- Wasow, Thomas. 2015. Ambiguity avoidance is overrated. In Susanne Winkler (ed.), *Ambiguity: Language and communication*, 29–48. Berlin: De Gruyter.
- Watters, John R. 1981. *A phonology and morphology of Ejagham, with notes on dialect variation*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Wilkendorf, Patricia. 2001. *Sketch grammar of Nɔmaándé: Sections 1–4*. Yaoundé: SIL Cameroon.
- Yoneda, Nobuko. 2011. Word order in Matengo (N13): Topicality and informational roles. *Lingua* 121(5). 754–771.

Zeller, Jochen. 2004. Relative clause formation in the Bantu languages of South Africa. *Southern African Linguistics and Applied Language Studies* 22(1-2). 75–93.

Chapter 13

Predicate partition for predicate-centred focus and Meeussen’s Proto-Bantu “advance verb construction”

Tom Güldemann^{a,b} & Ines Fiedler^a

^aHumboldt University of Berlin ^bMax Planck Institute for Evolutionary Anthropology in Leipzig

Meeussen’s (1967: 121) extensive grammatical reconstructions for Proto-Bantu contain a so-called “advance verb construction” that is comprised of an infinitive followed by a finite form of the same verb (typologically commonly called “cognate verb”) and conveys a marked type of information structure (IS) in which a predicate component is highlighted pragmatically. While Güldemann (2003: 335–337) already characterised this construction to pertain to the IS subdomain of so-called “predicate-centred focus”, he had to leave open some important structural and functional details. Since then, much more relevant data have become available, both inside and outside of Bantu. In this chapter, we attempt to specify Meeussen’s (1967) proposal about his “advance verb construction” and its “relatives” by providing a cross-linguistic perspective of the relevant domain, presenting and analysing a wide range of relevant structures from across the Bantu family, and finally discussing the results of this comparative family survey regarding both the synchronic variation and the diachronic dynamics of change.

1 Introduction

Meeussen’s (1967: 121) extensive grammatical reconstructions for Proto-Bantu (PB) also contain a remark on a so-called “advance verb construction”, which he describes as follows:



A peculiar kind of sentence, with twice the same verb, the first occurrence being an infinitive, is attested frequently, and will have to be ascribed to Proto-Bantu. The meaning varies between stress of « reality », stress of « degree », and even « concession »: *kutákũna báátákũnide*, « they chewed as (much as) they could »; « (as for chewing) they did chew, (but ...) ».

The construction's generalised structure is [**Verb_{non-finite}**][Cognate_Verb_{finite}].¹ Example (1) from Sundi H131K illustrates this construction, showing one of its functions: a marked type of information structure (henceforth IS) in which one predicate component, here the state of affairs 'to read', is highlighted pragmatically. It is often used for the expression of what we call here contrastive state-of-affairs (SoA) focus as opposed to the simple predicate structure *ndyèká-tá:ngà* 'I am going to read', which lacks such a function.

(1) Sundi H131K (Hadermann 1996: 161)

kù-tá:ng-à ndy-èká-tá:ng-à
15INF-read-FV 1SG-NEAR.FUT-read-FV
'I am going to READ.'

While only of minor importance in the large body of grammatical forms proposed for PB, the above pattern has been of considerable interest in the typological discussion about syntax and IS (see §2 below). It is thus worthwhile to combine a more theoretical linguistic question with the rich data of a well-known and close-knit language family and thus advance both strands of research. For Bantu, this is particularly desirable as the reconstruction of complex morphosyntactic structures is still at its beginning.

Based primarily on the geographically restricted comparative treatment of the phenomenon in Bantu languages of zones B and H by Hadermann (1996), Güldemann (2003: 335–337) already characterised the construction in (1) to pertain to the IS subdomain of so-called “predicate-centred focus” (henceforth PCF), but he had to leave open some important structural and functional details when writing:

¹This construction must be distinguished from a superficially very similar one whose structure is [INFINITIVE COGNATE_RELATIVE_VERB], as reported by Koni Muluwa & Bostoen (2014: 132–133) for Nsong B85d, by Mufwene (1987; 2013) for Kituba H10A, by Mufwene (1987) and Meeuwis (2013) for Lingala C30B, and by Guérois (2015: §10.1.6) for Cuwabo P34. Since the finite verb is a *modifier* of the infinitive, one is confronted here with a noun phrase rather than an asserted clause. It also has information structural effects and thus belongs in the wider domain at issue. However, sparsity of relevant information as well as lack of space does not allow us to include it in our discussion.

Two structural interpretations of the fronted-infinitive pattern are conceivable. [...] The first analysis, which accounts in a straightforward way for the focus function, is that the initial infinitive is a preposed focus constituent in the form of a nominal term and the following finite verb is the predicate. The second possibility is more complex, involving some form of functional reanalysis. That is, the construction may have originally had a topic-focus organization, best paraphrased as ‘As for VERBing, (I assert that) X VERBs’, and this has yielded the conventionalized reading ‘X does VERB’. Such a pattern is parallel to a similar German expression, which is typically followed by an adversative clause. In a sentence like *Spielen tut er, aber ihm fehlt ein eigenes Instrument*. ‘He does play [lit.: to play, does he], but he needs an instrument of his own.’, a clear contrast holds between the two clauses. Important for the present discussion is that this contrast is not only conveyed by the conjunction *aber* ‘but’, but also by the structure [infinitive + dummy verb + subject] in the initial clause by virtue of its focus on the predicate.

Since then, much more data have become available, both inside and outside of Bantu. Given this background and building on the first typological overview by Güldemann et al. (2014), the goal of the present chapter is to flesh out Meeussen’s (1967) partly vague characterisation of his “advance verb construction” in semantic and formal terms, in particular by relating it to its “relatives” in a much larger constructional space, and to fine-tune its reconstruction to PB both structurally and functionally. In §2, we provide a cross-linguistic survey of the domain. In §3, a wide range of relevant structures from across Bantu are reported, presented and discussed. In §4, we discuss the results of this comparative survey in terms of synchronic morphosyntactic and semantic-functional variation. In §5, by way of conclusion, we consider the construction’s diachronic dynamics and reassess its reconstruction with respect to PB, the ancestor of Narrow Bantu as conventionally delimited by Guthrie (1948; 1971).

2 IS-sensitive verb preposing from a wider perspective

What we call here predicate-centred focus (PCF) subsumes roughly non-term focus in opposition to nominal “term focus”, as per Dik (1997) (cf. also Hyman & Watters’s (1984) related concept of “auxiliary focus”), whereby focus is conceived here as a phenomenon on the level of a simple sentential assertion rather than larger discourse units.

The principal types of PCF and their relationships are given in Figure 1, followed by aligned English examples with preceding typical discourse contexts.

Polarity and Tense/Aspect/Modality (TAM) focus are not necessarily the only subtypes belonging to the umbrella concept of operator focus.

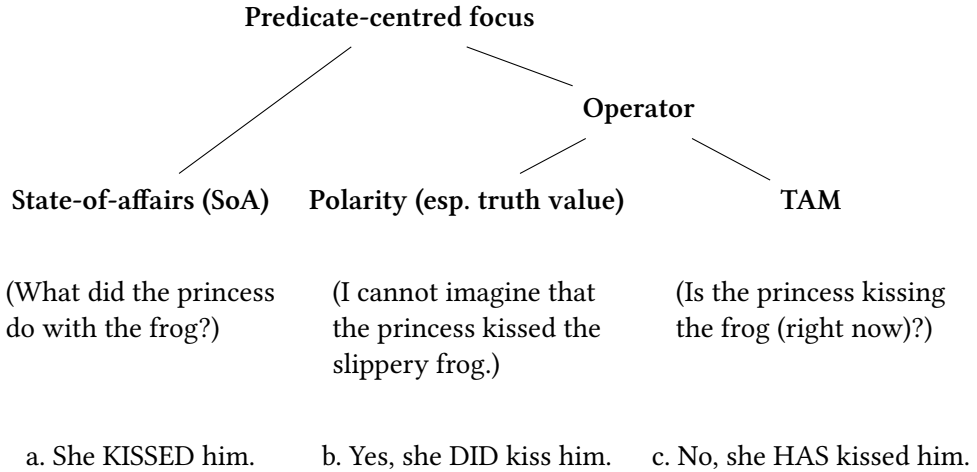


Figure 1: Basic typology of predicate-centred focus (PCF)

What follows is a cross-linguistically informed survey of structures where the predicate is partitioned or dissected into its two IS-relevant components pertaining to the SoA expression on the one hand and to the assertion on the other hand. A construction targeting pragmatically the former component renders SoA focus, while one oriented to the latter renders different types of operator focus.

A major formal mechanism of dissecting the predicate is the apparently tautological double use of the same verb called variously “predicate cleft”, “verb doubling”, “cognate object construction”, etc.² While the available literature on such structures is extensive, analyses largely deal with language-specific cases without providing a cross-linguistically representative picture. Such a systematic typology will be proposed in Güldemann (In preparation); see Güldemann et al. (2010) for a first publicly available version. The diversity of the wider domain of IS-related predicate partition is established according to various parameters summarised in Table 1 and discussed subsequently.

The first crucial distinction under I in Table 1 is triggered by the variable pragmatic role of the non-finite verb. In the case of preposed verb doubling in (1), the initial verb can either be the focus of the utterance, the case commonly called

²The terms “cognate” verb and verb “doublet” are used interchangeably merely to refer to the mutual lexical relationship without any conviction that either verb is basic and/or copied by the other.

Table 1: Some variation parameters of predicate partition/dissection

I	Pragmatic status of non-finite verb	focus vs. topic
II	Position of non-finite verb vis-à-vis clause	preposed vs. in-situ vs. postposed ^a
III	Lexical relation of finite verb vis-à-vis non-finite verb	doublet vs. light verb

^aVerb postposing plays a marginal role in Bantu and is only referred to briefly in §4.1.

“predicate cleft”, or it can be the topic, as foreshadowed in the above quotation from Güldemann (2003). Güldemann (In preparation) argues that the difference between the two patterns correlates robustly with two distinct PCF subtypes, namely SoA focus in the first vs. operator focus in the second case.

There are languages that possess both options and thereby distinguish two principal PCF types, as holds for Amharic illustrated in (2) and (3). While (2) shows a cleft structure with focus on the initial verbal noun and conveys SoA focus, (3) displays a verbal noun in topic function and accordingly renders truth value focus.³

- (2) Amharic [Semitic, Afro-Asiatic] (Andreas Wetter, p.c.)

SoA focus

mākina-w-n māt'aggän nä-w yä-t'aggän-ä
 car-DEF-OBJ repair:VN COP-3M.SG REL-repair-3M.SG
 [FOC] < I [BG]

‘He REPAIRED the car.’ [lit.: It is repairing the car that he repaired.]

- (3) Amharic [Semitic, Afro-Asiatic] (Andreas Wetter, p.c.)

Truth focus

māt'aggän-əs t'aggän-o-all
 repair:VN-TOP repair:CONV-3M.SG-AUX:3M.SG
 [TOP] < I [FOC]

‘He DID repair (the car).’ [lit.: As for repairing, he repaired.]

³For the sake of a better understanding of the IS configuration, these and most other examples are accompanied, i.e. usually followed, by a schema with underlying IS fields; these possibly involve segmental indices (I) that encode the IS status of the constituent in their scope as well as arrows that mark the scope direction (cf. Güldemann 2016 for a similar presentation of IS constructions).

This first distinction between “preposed verb focus doubling (= PrepFocDoubling)” and “preposed verb topic doubling (= PrepTopDoubling)” is summarised in Table 2. In this and following tables, “verb” refers to the non-finite verb, if not stated otherwise, in line with the explanation around Table 1.

Table 2: Preposed verb focus doubling vs. preposed verb topic doubling

Verb position:	Preposed	
IS status of verb:	Focus	Topic
Verb doubling:	PrepFocDoubling	PrepTopDoubling
IS function:	SoA	Operator

The second distinction within IS-sensitive predicate partition, given under II in Table 1, concerns the position of the non-finite verb. With pre- or postposing the non-finite verb we imply its *ex-situ* (aka extra-clausal) position, as opposed to an *in-situ* (aka intra-clausal) position. In the focus case, this type of syntactic variation corresponds with the existence of distinct IS field positions reserved for focus constituents.

Compare in this regard (4) and (5) from two closely related Bongo-Bagirmi languages, which both encode SoA focus. Example (4) from Mbay is an instance of PrepFocDoubling, parallel to (2) from Amharic; (5) on the other hand, from Bagirmi, represents a case of *in-situ* verb doubling (= InFocDoubling), where the non-finite form *tádà* follows the verb phrase with its object or, if the object is an initial topic marked by *ná*, the finite verb directly.

- (4) Mbay [Bongo-Bagirmi, Central Sudanic] (Keegan 1997: 148)

SoA focus

nà ndūsē lā ndūsē yé

but INF:worm_eaten FOC worm_eaten BG

[FOC] < I [BG] < I

(A: Your wood is bad. B: No, the wood is fine.) ‘It’s just that it’s WORM-EATEN.’ [lit.: It’s worm-eaten that it’s worm-eaten.]

- (5) Bagirmi [Bongo-Bagirmi, Central Sudanic] (Jacob 2010: 129)

SoA focus

Boukar táđ djùm téη táđà

PN PFV:do gruel millet INF:do

[BG] [FOC]

(or: *djùm téη ná, Boukar táđ táđà*)

(Did Boukar cook millet gruel or did he eat it?) ‘Boukar COOKED millet gruel.’ [lit.: Boukar cooked (millet gruel) COOKING.]

Including the new pattern in (5), abbreviated here as InFocDoubling, the extended range of verb-doubling structures is given in Table 3.

Table 3: Preposed verb focus/topic doubling vs. in-situ verb focus doubling

Verb position:	Preposed		In-situ
IS status of verb:	Focus	Topic	Focus
Verb doubling:	PrepFocDoubling	PrepTopDoubling	InFocDoubling
IS function:	SoA	Operator	SoA

So far, the diversity pertained to constructions that all displayed the co-occurrence of a finite and a non-finite verb of the same lexical type. However, this is not a necessary ingredient of the domain at issue. Dissecting the predicate for the expression of PCF without any change of IS reading can also be achieved by combining a non-finite lexical verb with a finite verb that is auxiliary-like, what is called here a light-verb structure.

A language that recruits this and all previous strategies is Hausa. Example (6) demonstrates the expression of truth value focus by means of verb topic preposing, whereby the version in (6a) is a case of PrepTopDoubling, while in the version in (6b) the preposed verb topic is followed by a finite light verb ‘do’. Example (7) is a light verb structure with verb focus preposing.

- (6) Hausa [Chadic, Afro-Asiatic] (after Jaggar 2001: 542)

Truth focus

a. *sàyé-n àbinci kòò, sùn sàyaa*

buy:VN-GEN food moreover 3PL.PFV buy

- b. *sàyé-n àbinci kòò, sùn yi*
 buy:VN-GEN food moreover 3PL.PFV do
 [TOP] < I [FOC]
 ‘Buying food moreover, they bought/did.’ [they DID ...]

(7) Hausa [Chadic, Afro-Asiatic] (Green 2007: 60)

VP focus

- sàyé-n àbinci nèè, sukà yi*
 buy:VN-GEN food FOC 3PL.PFV.DEP do
 [FOC] < I [BG]
 ‘They BOUGHT FOOD.’

The two light-verb options, PrepTopLight and PrepFocLight, increase the inventory of structures with IS-sensitive predicate dissection even further, as shown in Table 4.

Table 4: Verb focus/topic doubling vs. verb focus/topic light-verb structure

Verb position:		Preposed	In-situ
IS status of verb:	Focus	Topic	Focus
Verb doubling:	PrepFocDoubling	PrepTopDoubling	InFocDoubling
Light verb structure:	PrepFocLight	PrepTopLight	?
IS function:	SoA	Operator	SoA

Finally, in a language like German, where the two separated predicate components can be manipulated quite freely by means of prosody, the light-verb structure can also be employed in-situ. When emphasising the light verb *tun* ‘do’ suprasegmentally, the IS reading is truth value focus, irrespective of whether the non-finite verb is a preposed topic (= PrepTopLight), as in (8a), or an in-situ complement (= InTopLight), as in (8b) (cf. also English *do*-support).

(8) German [Germanic, Indo-European] (personal knowledge)

Truth focus

- a. *Lesen TUT er*
 read:INF does he
 [TOP] [FOC]

b. *er TUT lesen*

he does read:INF
[FOC] [BG]

[lit.: As for reading, he DOES.] > ‘He DOES read (but ...).’

Shifting prosodic emphasis to the non-finite lexical verb results in SoA focus, again independent of whether this verb is a preposed focus (= PrepFocLight), as in (9a), or an in-situ focus (= InFocLight), as in (9b). Recall that the disambiguation in the IS reading between (8a) and (9a), as well as between (8b) and (9b), is merely achieved by prosody.

(9) German [Germanic, Indo-European] (personal knowledge)

SoA focus

a. *LESEN tut er*

read:INF does he
[FOC] [BG]

b. *er tut LESEN*

he does read:INF
[BG] [FOC]

[lit.: READING he does.] > ‘He READS (rather than sleeps).’

Table 5 presents a fuller range of constructions with a dissected predicate in PCF expression, including reference to the examples above. It displays an overall symmetrical setup where only one pattern is not yet attested, the InTopDoubling pattern, which would be the counterpart of the InTopLight structure illustrated by (8b) from German.

3 PCF with non-finite verbs in Bantu

According to §2, Meeussen’s (1967) “advance verb construction” is embedded in a larger family of related structures, which provides a better background for evaluating the former. One central result of our survey is the existence of two basic morphosyntactic schemas in Bantu-like languages with a basic word order SBJ-V-OBJ, namely ex-situ infinitive fronting, as in [I], and an in-situ counterpart, as in [II].

[I] [Verb_{non-finite} [SBJ Verb_{finite} (Other)]]

[II] [SBJ Verb_{finite} (Other) Verb_{non-finite} (Other)]

Table 5: Dissected predicate constructions for PCF

Verb position:	Verb preposing		In-situ	
IS status of verb:	Focus	Topic	Focus	Topic
Verb doubling:	PrepFocDoubling	PrepTopDoubling	InFocDoubling	?
<i>Example:</i>	<i>Amharic (2)</i> <i>Mbay (4)</i>	<i>Amharic (3)</i> <i>Hausa (6a)</i>	<i>Bagirmi (5)</i>	
Light verb structure:	PrepFoc-Light	PrepTop-Light	InFoc-Light	InTop-Light
<i>Example:</i>	<i>Hausa (7)</i> <i>German (9a)</i>	<i>Hausa (6b)</i> <i>German (8a)</i>	<i>German (9b)</i>	<i>do-support</i> <i>German (8b)</i>
IS function:	SoA	Operator	SoA	Operator

In both patterns, it is not trivial to ascertain the exact structure and function of the entire construction without information about the pragmatic status of the non-finite verb, which can be marked by segmental and/or supra-segmental encoding. This partly lacking information is at the basis of the inconclusive characterisation of the ex-situ pattern by both Meeussen (1967) and Güldemann (2003: 335–337). Due to the availability of much more data on Bantu and our cross-linguistically informed perspective, we survey the domain across a large set of languages that are known from the literature to possess them. We organise the data according to five geographical clusters. The full list of languages, including those that are so far isolated cases outside these clusters, can be found in Appendix A.

We intentionally start out in the north-west, from which the family emanated, as this area is not unlikely to host the structural diversity the modern cross-Bantu profile emerged from. The wider areal and genealogical background of the Bantu homeland is the Macro-Sudan Belt (see Güldemann 2008), which hosts a large amount of language diversity but at the same time is dominated by Niger-Congo, the genealogical higher-order group to which Bantu belongs. While this part of West Africa harbours the full range of constructions in Table 5, the available literature focusses in particular on PrepFocDoubling (aka “predicate clefts”) because this has been transferred so often into Atlantic and West African creoles. Some such works are Bynoe-Andriolo & Yillah (1975), Goodman (1985: 125–126), Gilman (1986: 39–40), Mufwene (1987), and Manfredi (1993), which in fact deal

not only with West African languages but also mention some Narrow Bantu languages such as Lingala C30B, Kuyu E51, Kituba H10A, Kongo H16, and Makhuwa P31.

3.1 Grassfields and Bantu zone A

The immediate genealogical context of Bantoid and north-western Bantu seems to be characterised by the (co)existence of InFocDoubling and PrepFocDoubling. Some languages are only reported for possessing the first structure, for example Ngwe, as in (10). See also Ibrahīm (2007) for the Ngiemboon variety of Bamileke and Makaa A83.

- (10) Ngwe [Grassfields, Mbam-Nkam, Bamileke] (Nkemnji 1995: 138)

SoA focus

Atem a kɛ̀? nčúū akendòŋ čúū

PN 3SG PST1 ?:boil plantains boil

[BG] [FOC]

‘Atem BOILED plantains.’

In Limbum, InFocDoubling and PrepFocDoubling exist side by side, whereby we lack information about possible interpretational differences. See Bassong (2014: §V) for the same situation in Basaa A43a. This variation arises from the availability of both an in-situ and an ex-situ focus position. Regarding the first case, (11a) shows in-situ term focus, while the variant of InFocDoubling for SoA focus is given in (11b). In (12), the same opposition between term and SoA focus holds respectively for the negative cleft structures in (12a) and (12b) – the second being a case of PrepFocDoubling.

- (11) Limbum [Grassfields, Mbam-Nkam, Nka] (Ndamsah 2012: ex. (11b), resp. ex. (11a))

a. Term focus

Nf̃ɔ̃ t̃f̃ē yē á byē:

PN PROG eat FOC food

[BG] I > [FOC]

‘It is food that Nfor is eating.’ [Nfor eats FOOD.]

b. SoA focus

Nfɔ̃ tfē būmī á búmí
 PN PROG sleep FOC sleeping
 [BG] I > [FOC]

‘It is sleeping that Nfor is sleeping, not ...’ [Nfor SLEEPS, not ...; last verb in citation form: Gratiana Ndamsah, p.c.]

(12) Limbum [Grassfields, Mbam-Nkam, Nka] (Ndamsah 2012: ex. (3a), resp. ex. (3b))

a. Term focus

á Nfɔ̃ tfé é tfē būmī kā?
 FOC PN REL PRO PROG sleep NEG
 I > [FOC] I > [BG]

‘It is not Nfor who is sleeping.’

b. SoA focus

á būmí tfé Nfɔ̃ tfē būmī kā?
 FOC sleep REL PN PROG sleep NEG
 I > [FOC] I > [BG]

‘It is not sleep that Nfor is sleeping.’ [Nfor is not SLEEPing.]

Tuki A601, finally, is a language that seems to use only cleft-like PrepFocDoubling for SoA focus, as in (13b), which again also serves to express term focus, as in (13a)

(13) Tuki A601 (Bilola 1997: 111, resp. 110)

a. Term focus

nambari owu Mbara a-nu-enda-m n(a) adongo
 tomorrow FOC PN.1 1-FUT-go-? to village
 [FOC] < i [BG]

‘It is tomorrow that Mbara will go to the village.’ [Mbara will go to the village TOMORROW.]

b. SoA focus

o-suwa owu Puta a-nu-suwa-m tsono raa
 INF-wash FOC PN.1 1-FUT-wash-? clothes her
 [FOC] < I [BG]

‘Putu will WASH her clothes.’

3.2 Bantu zone J

The alternation between InFocDoubling and PrepFocDoubling is not restricted to Bantu in the north-west but found elsewhere, notably in interlacustrine Bantu of zone J. The diversity in this language group is even greater, because it concerns two additional parameters.

For one thing, verb doubling, at least in the in-situ pattern, has recourse to different verbal nouns, which is associated with distinct focus subtypes. The default infinitive with class 15 **kù-* preceded by the conjunction **na* ‘and’ when following the finite verb encodes additive SoA focus. In opposition to this, the parallel pattern with the verbal noun occurring in class 14 (marked by the reflex of PB **bù-*) conveys restrictive SoA focus, as in (14a).⁴ This effect is most likely related to the use of class 14 in Ganda JE15 to express single points in time with particular reference to the noun *obu-dde* ‘occasion, time of day’ (Ashton et al. 1954: 211, 278), which seems to imply here ‘once’ and hence restrictive focus ‘only’.⁵

The example pair in (14) from Ganda JE15 exemplifies this contrast between restrictive and additive SoA focus in (14a) and (14b), respectively. An interesting point of variation of InFocDoubling in Bantu zone J compared to that illustrated above in (5) for Bagirmi and in (10) for Ngwe is that the non-finite verb can precede the object. We call this pattern “*Postverbal* InFocDoubling” as opposed to “*Final* InFocDoubling” in the other case. Example (15) shows that at least additive SoA focus is not only conveyed by InFocDoubling, as in (14b), but is also possible with PrepFocDoubling.

(14) Ganda JE15 (Jenneke van der Wal & Saudah Namyalo, p.c.)

a. Restrictive SoA focus

<i>w-a-gúl-a</i>	<i>bu-gúzi</i>	<i>kí-tábó</i>
2SG-PST-buy-FV	14-buy:NOM	7-book
[BG]	[FOC]	[BG]
‘You just/only BOUGHT the book.’		

⁴Note that this nominalisation involves the change of the final vowel to agentive *-i* (cf. Schadeberg & Bostoen 2019: 188), which can trigger (agent noun) spirantisation of the final stem consonant (cf. Bostoen 2008), as in (14a).

⁵The structural potential for such a possible alternation between two types of verbal nouns in InFocDoubling seems to be quite old in Bantu, as Watters (1981: 246–247) describes a very similar alternation in the Ekoid Bantu language Ejagham.

III [Verb_{non-finite}] [(SBJ) Verb_{finite} (Other)]

(17) Soga JE16 (Nabirye 2016: 380)

Term focus

oku-zimb-a tw-a-zimb-anga ma-yumba ga nnanka
 15INF-build-FV 1PL-PST-build-HAB 6-house 6:GEN certain_kind
 [TOP] [BG] [FOC]

‘As for building [houses], we always built houses of a CERTAIN KIND.’

Other zone J languages also possess PrepTopDoubling, used here, as expected, for truth and other types of operator focus. Asiimwe & van der Wal’s (2019) new data for Nkore-Kiga JE13/14 strongly suggest that this language possesses this pattern and the two versions of InFocDoubling, for which see also Taylor (1985: 77–220a/b). While the authors do not disambiguate the status of an initial infinitive as a topic or focus, Jenneke van der Wal (p.c.) excludes the existence of PrepFocDoubling. Personal communication from Jean Paul Ngoboka also confirms the existence of PrepTopDoubling in Rwanda JD61, as shown in (18); the pronominal element *byo* is an explicit topic marker of class 8, which is the canonical agreement in the language for infinitives of class 15, here *kurya* ‘to eat’.

(18) Rwanda JD61 (Jean Paul Ngoboka, p.c.)

Truth focus

ku-ry-á byó a-ra-ry-á
 15INF-eat-FV TOP 1-DJ-eat-FV
 [TOP] < I [FOC]

‘He DOES eat.’ [As for eating, he EATS.]

3.3 Bantu zones B and H

Bantu languages of the Kongo cluster commonly display structures with preposed infinitives. The feature was first surveyed by Hadermann (1996) and analysed by Güldemann (2003) as generically pertaining to the PCF domain. More recently, this trait has been described extensively by De Kind et al. (2015).

The structure encountered predominantly is PrepFocDoubling, as illustrated previously with (1) above from Sundi H131K. While overall comparable to the pattern across Bantu, some languages of the Kongo cluster display certain morphological specificities. For one thing, the fronted non-finite verb doublet often lacks an overt nominalising prefix, but this reflects a historical change independent of our domain (see Bostoen & de Schryver 2015). Moreover, the subject concord on

the out-of-focus finite verb referring to a class 1 referent has the marked form *ka-* rather than unmarked *u-*.

The PrepFocDoubling pattern with its specific SoA focus interpretation is associated with a more general trend towards a preverbal focus position (cf. Hadermann 1996) that derives ultimately from an original cleft-like focus construction (De Kind et al. 2015). From a functional-semantic perspective, however, it is noteworthy that one can diagnose a developmental cline away from SoA focus toward general PCF (subsuming SoA *and* operator focus) and then, in line with observations by Güldemann (2003), to temporal predicate meanings, first to focus-sensitive progressive and finally to a proximal future, as illustrated by the following examples. While the expected function of SoA focus holds for (19) from Woyo H16dK (West Kongo),⁶ (20) from Ndibu H16bZ (Central Kongo) appears to involve emphasis on the truth value in the domain of operator focus. The encroachment of general PCF on the progressive domain seems to apply to (21) from Kamba H112A (North Kongo) because Hadermann (1996: 160) cites Bouka (1989: 237) who observes that the relevant form *sàlá kàmú:sàlá*, as opposed to the canonical progressive form *wàmú:sàlá*, serves to “*renforcer l’idée de répétition dans le déroulement de l’action*” [“reinforce the idea of repetition in the unfolding of the action” (our translation)]. Example (22) from Fiote H16d (West Kongo), however, is likely to represent a case of a plain progressive, as the predicate occurs in a dependent clause which, by default, does not involve focality. Finally, example (23) from Yaka H31 (Kongoid) is an instance of future meaning.⁷

- (19) Woyo H16dK (West Kongo) (De Kind et al. 2015: 119)

SoA focus

zeng-a ba-Ø-zeng-eza wao

INF:cut-FV 2-PRS-cut-PFV 2PRO

‘(What ... they did to the tree?) They CUT it.’

- (20) Ndibu H16bZ (Central Kongo) (De Kind et al. 2015: 120)

Truth focus

mon-a mbwene N-kenda za zula ...

INF:see-FV 1SG:see:PRF 10-affliction 10:GEN 7.people

‘I have surely seen the affliction of that people ...’

⁶The Kongo subgroups indicated refer to the phylogenetic classification of the Kikongo Language Cluster (KLC) by de Schryver et al. (2015).

⁷De Kind et al. (2015: 130) discuss two possibilities for the emergence of a future reading of this construction: it develops a) directly from the present progressive as observed elsewhere in Bantu, or b) from the inflected unmarked verb via analogy to simple zero-marked verbs that can get future interpretation in some South Kongo varieties.

- (21) Kamba H112A (North Kongo) (Hadermann 1996: 160)

PCF~PROG
sàl-á kà-mú:-sàl-á
 INF:work-FV 1-PROG-work-FV
 ‘He is working.’

- (22) Fiote H16d (West Kongo) (De Kind et al. 2015: 125)

PROG
kadi vov-a lu-Ø-vov-ang-a mu N-pamba
 because INF:speak-FV 2PL-PRS-speak-IPFV-FV INE 9-vanity
 ‘[...] because you are speaking in the air.’

- (23) Yaka H31 (Kongoid) (De Kind et al. 2015: 131)

FUT
vuumbuk-a yi-Ø-vuumbuk-a
 INF:dress-FV 1SG-PRS-dress-FV
 ‘I’ll dress myself.’

The good state of description of PrepFocDoubling in the Kongo cluster adds another point of structural variation to the domain. While all previous examples lack an independent expression for the *S/A* referent, its possible presence raises the question of its syntactic position. In a structure that is still close to a cleft, one expects that the *S/A* is part of the extra-focal clause domain and thus appears immediately before the finite verb and hence after the initial verbal noun, as in (12b) from Limbum and (13b) from Tuki A601. It is conceivable, however, that the *S/A* constituent occurs before an uninterrupted, syntactically tighter sequence of the two verbs, so that the non-finite verb is no longer initial but preverbal. We reformulate the morphosyntactic variation regarding the *S/A* position before or after the preposed infinitive with reference to the *non-finite verb* position as an opposition between “*Initial PrepFocDoubling*”, as in [I]a and (24) from Vili H12L (West Kongo), vs. “*Preverbal PrepFocDoubling*”,⁸ as in [I]b and (25) from Zali H16cZ (West Kongo).⁹

⁸The syntactic status of the *S/A* in this pattern is ambiguous as it could be an external topic or an internal subject topic. Since the necessary information is normally insufficient, we keep using the syntactically neutral semantic label *S/A*.

⁹This variation is the mirror image of the distinction between *Postverbal* and *Final InFocDoubling* mentioned briefly in §3.2 above.

[Ia] [Verb_{non-finite} [S/A Verb_{finite}]]

(24) Vili H12L (West Kongo) (De Kind et al. 2015: 117)

SoA focus

ko **kú-tél-à** *ń-cétù* **ù-à-ń-tél-à**

no! 15INF-call-FV 1-woman 1-PRF-1SG.OBJ-call-FV

[FOC] [BG]

(Has the woman beaten Pierre?) ‘No, the woman has (only) CALLED him.’

[Ib] [S/A] [Verb_{non-finite} Verb_{finite}]

(25) Zali H16cZ (West Kongo) (De Kind et al. 2015: 114)

PROG

i-búlu **zawúl-a** *ci-Ø-zawúl-a*

7-cattle INF:RUN-FV 7-PRS-RUN-FV

[TOP] [FOC] [BG]

‘The cattle is running.’

The data on the Kongo cluster available to us contain only a single example of Initial PrepFocDoubling, exemplified in (24), without much information as to whether this reflects real rarity or is coincidental. There is, however, indirect evidence that Preverbal PrepFocDoubling, as in (25), is indeed the predominant pattern, which we argue to be the reflex of a stronger degree of grammaticalisation of that construction away from its original nature as a cleft.

For one thing, the position of the S/A constituent before the preposed focal infinitive and outside the earlier background clause appears to be entrenched in a more general syntactic phenomenon. That is, the infinitive is analysed by Hadermann (1996: 158–159) as occurring in a preverbal focus position:

Cependant, Grégoire (1993) a montré que l’antéposition de l’objet n’est pas exceptionnelle en zones B, C, H et K, c’est-à-dire au Nord-Ouest du domaine bantou. L’apparition de l’ordre SOV est, selon elle, liée à « l’expression de la focalisation portant sur l’objet du verbe transitif » [...] ou à « l’emploi d’une forme composée de la conjugaison, [...] » [...]

Nevertheless, Grégoire (1993) has shown that the preposing of the object is not exceptional in zones B, C, H and K, i.e. in the North-West of the Bantu domain. The occurrence of the SOV order is, according to her, linked with “the expression of the focalisation bearing on the object of the transitive verb” [...] or with “the use of a compound form of the conjugation, [...]” [...] (our translation)

This is unusual for canonical Bantu languages and even opposed to the more general Benue-Congo trait of a preverbal *extrafocal* position (cf. Güldemann 2007). The following example from Nzebi B52 clearly illustrates the preverbal focus position that applies both to nominal terms, as in (26a), and the verbal noun in PrepFocDoubling, as in (26b). Nzebi is not part of the Kongo cluster, but belongs to the same major branch of the Bantu family, i.e. West-Coastal Bantu (Pacchiarotti et al. 2019).

(26) Nzebi B52 (Hadermann 1996: 162)

a. Term focus

bà-kà:sá bá-n'á; péndá bâ:vádà
 2-woman 2-DEM groundnut 2:cultivate
 [TOP] [FOC] [BG]

‘These women, they cultivate GROUNDNUTS.’

b. PROG

bà-kà:sá bá-n'á; vádá bâ:vádá péndà
 2-woman 2-DEM INF:cultivate 2:cultivate groundnut
 [TOP] [FOC] [BG]

‘These women, they ARE CULTIVATING groundnuts.’

There is another indication of increased grammaticalisation of preverbal PrepFocDoubling in West-Coastal Bantu. That is, its syntactic pattern tying the two predicate components closer together correlates with the shift away from pragmatic constituent-oriented IS functions (namely SoA focus derived directly from term focus) toward semantic predicate-centred tense/aspect notions of progressive and future, as mentioned above and illustrated again in (26b).

It was said in §2 (cf. Table 1) that another option in the focus fronting of infinitives concerns the finite verb: it can also be a light verb rather than being lexically identical with the verbal noun. This variant of the PrepFocLight structure, as exemplified in (8a) above from German, occurs repeatedly in the Kongo cluster and elsewhere in West-Coastal Bantu and can be schematised as in [IV].

[IV] [SBJ (OBJ)] [**Verb**_{non-finite} (Other) Auxiliary~Light_Verb_{finite}] Other]

Such a structure, which in Bantu turns out to be like an inverted version of an auxiliary periphrasis, was already associated with the domain at issue by Güldemann (2003: 336–337). Thus, (27) from Shona S10 shows an instance of a well-known progressive form based on locative periphrasis, which is frequent both

inside Bantu and also more generally in the world's languages (cf. Bybee & Dahl 1989). Example (28) from Kuria JE43 demonstrates a predicate with largely cognate morphological material but the inverse word order.

- (27) Shona S10 (personal knowledge)
ndi-ri ku-taur-a
 1SG-be 15INF-talk-FV
 'I am talking.'
- (28) Kuria JE43 (Güldemann 2003: 336) < (Sillery 1936: 20)
ku-tun-a n-di
 15INF-see-FV 1SG-be
 'I am (in the act of) seeking.'

De Kind et al.'s (2015) discussion of their Kongo Bantu data confirms the proposed affinity between a structure as in (28) and focus fronting more generally in that both share behavioural properties in opposition to the canonical [AUXILIARY VERB] structure exemplified in (27). The closer alignment of the PrepFocLight structure with plain auxiliary periphrasis in turn correlates with formal and functional observations. In opposition to PrepFocDoubling, it is only attested with an infinitive immediately preceding the finite auxiliary and with tense/aspect meaning. The following examples from Sundi H131K (North Kongo) in (29) and Tsootso H16hZ (South Kongo) in (30) illustrate these facts¹⁰ as well as some variation with respect to the auxiliary, i.e. *di* as in (29) vs. *(i)na* in (30), and the nature of the nominalising prefix, i.e. infinitive class 15 in (29) vs. locative~inessive class 18 in (30).

- (29) Sundi H131K (North Kongo) (Hadermann 1996: 166)
 PROG
bùkù kù-tá:ng-à dyò kà-di
 5.book 15INF-read-FV 5.PRO 1-be
 [TOP] [FOC] [BG]
 'He is reading the book.'

¹⁰The object marker *dyò* in (29) is best analysed as a weak anaphoric pronoun, possibly even enclitic, rather than a full noun phrase.

- (30) Tsootso H16hZ (South Kongo) (Hadermann 1996: 164)

PROG

mw-à:nà mù-sákán-á kéna

1-child 18INE-joke-FV 1:be

[TOP] [FOC] [BG]

‘The child is joking.’

3.4 Bantu zones E and F

Bantu languages of zone E were among the first mentioned in the literature in connection with predicate clefts. Thus, the early paper on African-based creoles by Bynoe-Andriolo & Yillah (1975: 234) had already reported the feature for Kuyu E51. This language is not the only one possessing this and related constructions. The closely related Tharaka E54 is another language with PrepFocDoubling.¹¹ This is illustrated in (31), whereby the example (31b) seems to suggest an additional reading of operator focus. We assume that this is independent of the fact that the finite predicate is a nominal predication.

- (31) Tharaka E54 (Abels & Muriungi 2008: 704)

- a. SoA focus

i-kû-gûr-a Maria *a-gur-ire* nyondo

FOC-15INF-buy-FV PN.1 1-buy-PRF 9.hammer

I > [FOC] [SBJ BG]

‘Maria BOUGHT the hammer.’ (she did not borrow it)

- b. ? Truth focus

i-ku-nog-a Maria *a-rî mù-nog-u*

FOC-15INF-tire-FV PN.1 1-be 1-tire-ADJ

I > [FOC] [SBJ BG]

‘Maria is really tired.’ (she is not kidding!)

As opposed to PrepFocDoubling in zones B and H, languages of zone E display overt signs of a cleft-like syntactic bisection involving an identificational and focus marker before the infinitive and sometimes even traces of dependent clause-marking in the finite background clause, which suggests a historically young age of the phenomenon.

¹¹According to information by Landman & Ranero (2014: 406), the construction may also exist in Kuria JE43, although the situation remains unclear, as the authors only give a single example of a fronted focalised nominalisation of an entire verb phrase, which changes the IS configuration.

Similar to zones B and H, one can observe an alternation between initial and preverbal PrepFocDoubling, whereby the first seems more salient, which again would suggest a younger historical age. Within the framework of our project on PCF, Morimoto (2017) carried out more detailed research on the ubiquitous use of the focus proclitic *nĩ* in Kuyu E51, including in predicate clefts (cf. also Schwarz 2003). An interesting observation was that her informant produced a progressive form that not only involved a canonical progressive verb prefix but also a PrepFocDoubling structure, as given in (32). It may well be significant that this token displays the *preverbal* variant of the construction, as opposed to the initial one attested so far in contexts of SoA focus, as in (33), which seems to replicate a trend described in §3.3 toward a motivated form-meaning covariation.

(32) Kuyu E51 (Morimoto 2017: 165)

PROG

fafa w-anyú nĩ gũ-kiny-á a-rá:-kiny-a (reu)

1.father 1-2PL.POSS FOC 15INF-arrive-FV 1-PROG-arrive-FV now

[S/A] I > [FOC] [BG]

‘Your father is arriving (now) [as we speak].’

(33) Kuyu E51 (Schwarz 2003: 96)

SoA focus

a. *ne atea Abdul e-k-irɛ na mae?*

FOC what PN.1 1-do-PFV COM 6.water

I > [FOC] [SBJ BG]

‘(What did Abdul do with the water?)’

b. *ne ko-nyu-a Abdul a-nyu-irɛ mae*

FOC 15INF-drink-FV PN.1 1-drink-PFV 6.water

I > [FOC] [SBJ BG]

‘He DRANK the water.’

As already observed by Güldemann (2003: 337–338), the relevant Bantu area also hosts languages that display structures labelled in §3.3 above as PrepFoc-Light with a fronted infinitive followed by an auxiliary, cf. Sillery (1936: 20) for Kuria JE43 and Whiteley (1960: 57, 61–62) for Gusii JE42, both involving forms with imperfective meaning. Gibson (2012: §3.3–3.5), Gibson (2019) and Roth & Gibson (2019: 300–302) add Ngoreme JE401, Simbiti JE431, Rangi F33, and Mbugwe F34, of the geographically close zone F, to the list of relevant languages where the phenomenon turns up in the immediate future with auxiliary

íise and the general future with auxiliary *rí* and is expectedly largely restricted to PCF-sensitive contexts such as polar questions and affirmative main clauses.

3.5 Bantu zone K

Another hotbed of Bantu languages with fronted infinitive doubling is zone K. Such structures are attested so far in Luvale K14 (Horton 1949: 209), Kwangali K33 (Westphal 1958: 94), Manyo including Gciriku K332 (Möhlig 1967: 206), Mbukushu K333 (Fisch 1977: 95, 103), Fwe K402 (Gunnink 2016; 2018: §11.1.2; 2019; p.c.), and both Zambian Totela K41 and Namibian Totela K411 (Crane 2019: 684–685; p.c.).

In Fwe and Totela, the syntactic analysis is sufficiently clear in order to assign the phenomenon to the PrepFocDoubling type and in both languages the expected SoA reading is indeed the most salient. Gunnink’s extensive analysis of the construction in Fwe provides other important details. Thus, only the preverbal variant is grammatical and the S/A argument occurs either clause-initially or after the finite verb. This is compatible with the finding that the compact sequence of non-finite and finite verb can in addition to SoA focus also express progressive, as shown in (34) and (35), respectively. Crane (2019; p.c.) also reports this for Namibian Totela. In the Zambian Fwe variety, the construction is even obligatory in sentences without a postverbal constituent and thus behaves similarly to PCF-sensitive “disjoint” verb forms in other Bantu languages.

- (34) Fwe K402 (Gunnink 2019: 73)

SoA focus

ka-ri ndí-aku-rir-a ku-fek-a
 NEG-be 1SG.REL-PST.IPFV-cry-FV 15INF-laugh-FV
 [FOC]

ndí-aku-fek-a

1SG.REL-PST.IPFV-laugh-FV

[BG]

‘I was not crying, I was LAUGHING.’

- (35) Fwe K402 (Gunnink 2018: 352)

PROG

e-N-tí ku-hór-a í-shi-hor-á
 AUG-9-tea 15INF-cool-FV 9.REL-PERS-cool-FV

‘The tea is still cooling down.’

Most other instances of such constructions in zone K are hard to analyse conclusively as to whether the underlying pattern is PrepFocDoubling or PrepTopDoubling. For one thing, there is very little information about the syntax of the language-specific structures. In functional terms, the available examples are usually without discourse context and on their own can be interpreted recurrently as conveying truth value focus, which is expected for PrepTopDoubling rather than PrepFocDoubling. The treatment in Mbukushu K333 is a typical case: while (36) conveys progressive, (37) focusses on the assertion.

- (36) Mbukushu K333 (Fisch 1977: 95)

PROG

ku-w-a thi-na_ku-w-a thi-tondo

15INF-fall-FV 7-PRS-fall-FV 7-tree

‘*Der Baum fällt gerade.*’ [‘The tree is falling right now.’]

- (37) Mbukushu K333 (Fisch 1977: 103)

Truth focus

ku-yend-a tu-na_ku-yend-a

15INF-go-FV 1PL-PRS-go-FV

‘*Wir gehen ja schon.*’ [‘We DO go, don’t we.’]

Given that such authors as Horton (1949), Westphal (1958), and Möhlig (1967: 206; p.c.) even appear to analyse the initial infinitive as an extraposed topic, the structures could well be cases of PrepTopDoubling. However, generalised PCF including truth value focus can emerge from PrepFocDoubling, too (see §4.2 below), so that a conclusive assessment requires more detailed information on both form and function.

4 Summary and discussion

The data presented and discussed above show that Meeussen’s (1967: 121) “advance verb construction” is not an isolated structure, but is best appreciated when analysed within a larger cross-linguistically relevant family of constructions, which are characterised by the partition of the predicate for the expression of PCF, *and* within its wider areal context in and beyond Narrow Bantu. In the following, we discuss the variation that emerged in terms of structural properties (§4.1) as well as semantic-functional aspects (§4.2).

4.1 Morphosyntactic variation

In terms of morphosyntax, we started out in §1 above with Meeussen’s characterisation, which involves three crucial structural ingredients, namely:

1. two lexically identical verbal constituents, whereby
2. one is non-finite and the other is finite, and
3. the former syntactically precedes the latter.

However, there are a number of closely related constructions across the Bantu family that diverge from the above pattern in each of the three properties as well as various other points, which we present systematically in the following.

One type of variation that is not prefigured by Meeussen’s characterisation but widely attested across Narrow Bantu concerns the position of the possible constituent that refers to the S/A argument of the verb. Focusing on the position of the fronted non-finite verb, we speak of initial PrepFocDoubling if the S/A noun phrase occurs after the initial non-finite verb but before the finite one, while if preceding both we call the pattern preverbal PrepFocDoubling, as shown for Kuyu in (38) and (39), respectively.

[Ia] [Verb_{non-finite} [SBJ Cognate_Verb_{finite}]]

(38) Kuyu E51 (Mugane 1997: 148)

SoA focus

nĩ kũ-nyu-a Kamau a-nyu-ire njohi ny-ingĩ

FOC 15INF-drink-FV PN.1 1-drink-PFV 9.beer 9-lot

I > [FOC] [BG]

‘Kamau DRANK a lot of beer.’

[Ib] [S/A] [Verb_{non-finite} [Cognate_Verb_{finite}]]

(39) Kuyu E51 (Morimoto 2017: 165)

PROG

mw-aná nĩ kũ-rey-a a-rá:-rey-á mbó:so

1-baby FOC 15INF-eat-FV 1-PROG-eat-FV 10.bean

[TOP] I > [FOC] [BG]

‘The child is eating beans.’

A second if minor difference to Meeussen's prototype concerns the above feature 2, in that in some languages the non-finite verb is not an infinitive of class 15, but rather a verbal noun of another class (notably 14 and 18) or a bare verb stem without any inflection. The latter case is shown again in (40) by an example of PrepFocDoubling in Solongo H16aM (South Kongo).

- (40) Solongo H16aM (South Kongo) (De Kind et al. 2015: 118)

SoA focus

kin-a be-kin-ang-a
 dance-FV 2-dance-IPFV-FV
 [FOC] [BG]

(No, they're not fighting.) 'They're DANCING.'

A third but major deviation, also stipulated by Meeussen as feature 3 above, is that some languages possess a structure where the infinitive is placed in an in-situ focus position *after* rather than before the finite verb. This is labelled here for short InFocDoubling, the simple pattern being exemplified again in (41) from Lingala C30B. Examples (42) also from Lingala and (43) from Zulu S42 show special variants with focus-sensitive markers before the infinitive. The former displays a restrictive marker 'only, just' and would have encoded originally restrictive SoA focus, while the latter has an additive marker 'also' (< comitative **na*) and would have encoded additive SoA focus. Both patterns have, however, widened their functional range to operator-like PCF meanings such as truth and intensity.

[II] [Cognate_Verb_{finite} Verb_{non-finite}]

- (41) Lingala C30B (Meeuwis 2013: ex. 60–151)

SoA focus

a-défis-áki yó yangó ko-défis-a, a-kabél-áki yó té
 3SG-lend-PST 2SG 3SG.IAN 15INF-lend-FV 3SG-offer-PST 2SG NEG
 [BG] [FOC]

'She LENT it to you, she didn't give it.'

- (42) Lingala C30B (Joseph Koni Muluwa, p.c.)

Truth focus

a-bongís-áki káka ko-bongis-a
 1-repair-PST RES.F 15INF-repair-FV
 [BG] I > [FOC]

(Having heard that somebody washed and polished his car, A asks: And he did not fix it? B replies:) 'He just REPAIRED/DID repair (it).'

- (43) Zulu S42 (Michel Lafon, p.c.)

Operator focus

ngi-ya-sab-a *no-ku-sab-a*

1SG-PCF-be_scared-FV ADD_F-15INF-be_scared-FV

[BG] I > [FOC]

‘I am so scared.’

The fourth type of variation is again covert in Meeussen’s description but is crucial for the general topic. His quite vague semantic-functional characterisation says nothing specific about the IS status of the different major constituents, in particular of the nature of the non-finite (preposed) verb. That is, PrepFocDoubling with this verb as the focus needs to be distinguished from PrepTopDoubling where the verb is a topic, triggering a different IS interpretation. Another illustrating example of the latter is (44) from Makhuwa P31.

[III] [**Verb**_{non-finite}] [Cognate_Verb_{finite}]

- (44) Makhuwa P31 (Asiimwe & van der Wal 2019)

Truth focus

o-rampelel-a *ki-naa-rampelel-a*

15INF-swim-FV 1SG-PRS.DJ-swim-FV

[TOP] [FOC]

(Don’t you know how to swim?) ‘I do know how to swim.’ [As for swimming, I DO swim.]

A final major variation relates to the above feature 1: finite verb and non-finite verb need not be lexically identical, but the former can be a generic auxiliary or another type of light verb – a phenomenon independent of other factors. The light-verb counterpart of PrepFocDoubling is PrepFocLight, as illustrated in (45) from Ntandu H16g (East Kongo).

[IV] [**Verb**_{non-finite} Auxiliary~Light_Verb_{finite}]

- (45) Ntandu H16g (East Kongo) (De Kind et al. 2015: 143)

Truth focus

nde yezu mu Ø-zing-a ka-ina

that PN.1 LOC INF-live-FV 1-to_be

[BG] [FOC] [BG]

‘... that Jesus IS (indeed) alive.’ (lit.: ... that Jesus in LIVING is.)

The InFocDoubling pattern has its relevant counterpart in an InFocLight structure. This is shown in (46) from Matengo N13, akin to English *do*-support.

[V] [Light_Verb_{finite} Verb_{non-finite}]

(46) Matengo N13 (Yoneda 2009: 160)

SoA focus

Maria ju-a-tend-aje kú-telek-a

PN.1 1-PST-do-CJ 15INF-cook-FV

[BG] [FOC]

(What did Maria do?) ‘Maria COOKed.’ (lit.: Maria did COOKING.)

While no case in Bantu of a possible counterpart of PrepTopDoubling, specifically PrepTopLight, has come to our knowledge so far, there is nevertheless a third light-verb structure that takes the form of a pseudo-cleft. Since the non-finite verb occurs in a final or postposed position, we use the short label PostFocLight. We only encountered it so far in Shona S10, as illustrated in (47), but it may well exist in more languages.

[VI] [Light_Verb_{finite}] [Verb_{non-finite}]

(47) Shona S10 (Peggy Jacob, field notes)

SoA focus

cha-a-it-a *ne-bhínzi* *ku-dzi-bik-a*

7:REL-1:DEP:PROX.PST-do-FV with-10.beans ID:15INF-10OBJ-cook-FV

[BG] I > [FOC]

(The woman ate the beans, didn’t she?) ‘She COOKed the beans.’ (lit.: What she did with the beans is COOKING them.)

The PostFocLight pattern is not attested with a PostFocDoubling counterpart and we assume that this is unlikely to exist at all. It would simply be awkward to already use the lexical element in the initial background domain whose meaning is to be focused on, in the subsequent assertion domain – that is, some nonsensical counterpart of (47) like ‘What she *cooked* with the beans is COOKING them.’

Table 6 gives the eight major morphosyntactic types that emerge theoretically from the basic parameters discussed above. Since two are not (yet) attested, the following Table 7 only presents the structure schemas of the six relevant patterns.

The above discussion does not exhaust the variation possible. A full picture requires a more fine-grained analysis for most language-specific cases recorded

13 Predicate partition and “advance verb construction”

Table 6: Dissected predicate constructions for PCF across Bantu

Verb position:	Verb preposing		Verb in-situ	Verb postposing
IS status of verb:	Topic	Focus	Focus ^a	Focus
Verb doublet:	[III] PrepTop-Doubling	[I] PrepFoc-Doubling	[II] InFocDoubling	[Ø] PostFoc-Doubling
<i>Meaning:</i>	‘(As for) <i>verbing</i> , <i>he VERBED.</i> ’	‘(It is) <i>VERBING</i> (that) <i>he verbed.</i> ’	‘ <i>He verbed VERBING.</i> ’	‘ <i>What-he verbed-is VERBING.</i> ’
<i>Example:</i>	<i>Makhuwa</i> (44)	<i>Solongo</i> (40)	<i>Lingala</i> (41)	Ø
Light verb:	[?] PrepTop-Light	[IV] PrepFoc-Light *	[V] InFoc-Light	[VI] PostFoc-Light
<i>Meaning:</i>	‘(As for) <i>verbing</i> , <i>he DID.</i> ’	‘(It is) <i>VERBING</i> (that) <i>he did.</i> ’	‘ <i>He did VERBING.</i> ’	‘ <i>What he did is VERBING.</i> ’
<i>Example:</i>	?	<i>Ntandu</i> (45)	<i>Matengo</i> (46)	<i>Shona</i> (47)
Primary function:	Operator (truth) focus	SoA focus	Various PCF types	SoA focus

Notes: VERB IN UPPERCASE = FOCUS; Ø = not expected to occur; ? = not (yet) attested; * = finite verb is not ‘do, make’.

^aRecall from §1, particularly (8a) from German, that the non-finite verb can in principle also have a background status, which, however, is not clearly attested yet in Bantu.

Table 7: Structure schemas of dissected predicate constructions for PCF in Bantu

No.	Label	Structure schema
[I]	PrepFocDoubling	[Verb _{non-finite} [Cognate_Verb _{finite}]]
[II]	InFocDoubling	[Cognate_Verb _{finite} Verb _{non-finite}]
[III]	PrepTopDoubling	[Verb _{non-finite}] [Cognate_Verb _{finite}]
[IV]	PrepFocLight	[Verb _{non-finite} Auxiliary~Light_Verb _{finite}]
[V]	InFocLight	[Light_Verb _{finite} Verb _{non-finite}]
[VI]	PostFocLight	[Light_Verb _{finite}] [Verb _{non-finite}]

above. Further potentially diverse parameters relate to the formal expression of the IS status of the non-finite verb beyond its mere position (e.g. (supra)-segmental or no marking), to the encoding of the out-of-focus domain(s), or to the possibility of fronting more than just a finite verb.

4.2 Semantic-functional variation

The insufficient information about the last points of possible structural variation leads us to the assessment of the semantic-functional variability in the domain at issue. We restrict the discussion to PrepFocDoubling and PrepTopDoubling, as the situation is more complete here.

On several occasions, we have referred to the considerable difficulties to determine the functional distinction of SoA vs. operator focus in verb preposing structures recruited for PCF. One major reason for this is that PrepFocDoubling and PrepTopDoubling structures that lack segmental focus and/or topic marking look superficially identical. In general, there is a considerable risk of misinterpretation when having to trust short treatments of such cases, which in future calls for a more detailed analysis by language specialists in terms of their prosodic and morphosyntactic properties as well as their semantic-pragmatic effects.

Problems not only surface in Meeussen's description but also in many later works dealing with such structures. An informative case is the contradictory interpretation of an example from Ntandu H16g (East Kongo) provided by Lubasa (1974) in a different thematic context without much discussion. It is repeated in (48) in its original form in the first two lines, followed by our annotation as well as the two different schemas of IS interpretation in terms of PrepFocDoubling as per Gilman (1986) and PrepTopDoubling as per Mufwene (1987).

(48) Ntandu H16g (East Kongo)

'He/she wants to see.' (lit.: see – he wants he sees)

tálá ká-zól-ele ka-talá (Lubasa 1974: 22)

see 1-want-PRF 1-see

[FOC] [BG] (Gilman 1986: 39)

[TOP] [FOC] (Mufwene 1987: 81, fn. 12)

Mufwene (1987: 81, fn. 12) explains in more detail:

[...], it is not obvious either that, strictly speaking, all the cleft-related focus constructions invoked from African languages involve Clefting. For instance, Gilman (1986: 39) discusses them quite cautiously under the rather

vague term of ‘front-focusing’. The [... above] example from his paper seems more to involve TOPICALIZATION than Clefting, though it certainly involves nominalization of the verb by prefix-deletion (which is common in a number of Bantu languages). (use of uppercase is ours)

However, the original source of Lubasa (1974) gives (48) in connection with another formally related example under (49) that clearly involves focus fronting. This strongly favours an analysis in terms of PrepFocDoubling, which is in line with the general situation in zone H (see §3.3, cf. also the subject concord *ka-* typical for cleft-like focus structures).

(49) Ntandu H16g (East Kongo) (Lubasa 1974: 22)

Term focus

mw-ááná ká-túm-ini

1-child 1-send-PRF

[FOC] [BG]

‘It is a child that he/she has sent.’

There is also another reason why certain structures in Bantu and beyond may be hard to pin down in functional terms. That is, a particular construction can start out in a restricted subdomain of PCF (cf. Figure 1 of §1 for the distinction of SoA vs. operator focus) but over time expand in use within the wider PCF domain. As an example, we present in (50) the multifunctional fronting construction in Aja that is used for term focus and, in the case of PrepFocDoubling, all major types of PCF.

(50) Aja [Gbe, Benue-Kwa, Niger-Congo] (Fiedler 2010)

[FOC] (< I) [BG]

a. Term focus

āyú (yí) é qù

bean FOC 3SG eat

‘She ate BEANS [not ...].’

b. SoA focus

óò, qà (yí) é qà

no! cook FOC 3SG cook

(The woman ate the beans.) ‘No, she COOKED them.’

c. Truth focus

óò, *nyó* (yi) é *nyóvi*
 no! be_beautiful FOC 3SG be_beautiful
 (She is not beautiful.) ‘No, she IS beautiful.’

d. TA focus

óò, *xó-i* á *xó-i*
 no! hit-3SG.OBJ 3SG.FUT hit-3SG.OBJ
 (The woman has hit Peter.) ‘No, she WILL hit him.’

While a conclusive identification of the PCF type remains a central challenge regarding the semantic-functional variation of the structural domain, we have also described above other possible and recurring meaning changes that should be taken into account. We refer in particular to the grammaticalisation of PCF into the marking of progressive that subsequently can progress further into the marking of future or general imperfective. This development was dealt with extensively by Güldemann (2003) and the above data add several more cases to the initial data set.

We try to capture the major functional changes of preposed verb doubling in Bantu in the semantic map of Figure 2. As can be expected in grammaticalisation, the general historical trajectory goes from pragmatics to semantics. The data available to us do not clarify whether operator focus can also directly develop into progressive. Further research is also needed regarding other semantic readings of the structure, for example, of intensity.

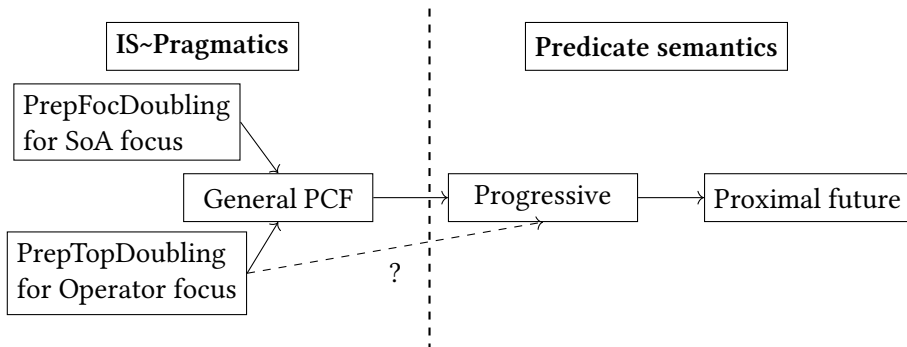


Figure 2: Semantic map for verb preposing constructions across Bantu

5 Historical assessment and conclusions

The above synchronic survey attests to the considerably increased documentation and understanding of infinitive fronting that was described only briefly and hence quite vaguely by Meeussen (1967) under the label “advance verb construction”. Its historical assessment may still be partly premature due to an incomplete knowledge about the full distribution of this family of constructions across the Bantu area. Nevertheless, we offer here a first, albeit preliminary, attempt on the basis of the above data and some cross-linguistic considerations.

A first observation can be made regarding the alternation of the position of the non-finite verb. Extra-clausal verb postposing is very rare, followed by the occasional but widely distributed option with the verb in in-situ position, while preposing is recurrent and very widespread (see Appendix A). However, in north-western Bantu and Bantoid (cf. §3.1), in-situ position and preposing appear to be equally prominent in the form of InFocDoubling and PrepFocDoubling, which matches the overall picture in the adjacent parts of the Macro-Sudan Belt. PrepFocDoubling only comes to predominate clearly across Bantu further away from the family homeland. We interpret this biased distribution of the two patterns to reflect the early coexistence of both with a later recurrent shift from the syntactically simple InFocDoubling to the more marked PrepFocDoubling. The cases of the former further south(east), including the variation in the form of the non-finite verb, could reflect either its long existence and hence sporadic retention in Narrow Bantu or its structurally latent presence connected to its universal availability. Regarding a possibly old age, it is worth considering that the quite specific pattern of InFocDoubling for additive SoA focus and other derived functions involving a focus-sensitive marker preceding the non-finite verb, such as comitative **na* in Narrow Bantu, has a wide albeit disperse geographical distribution. It occurs in the Nigeria-Cameroon border zone, for instance in the Ekoid Bantu language Ejagham (Watters 1981: 246–247), it also exists in the interlacustrine Bantu zone J languages (see §3.2), and it turns up again in the southernmost parts of the continent with Zulu (Doke 1927: 367; Michel Lafon, p.c.).¹² There is yet another possible argument for InFocDoubling being an old retention. In footnote 1 we mentioned another structure: [INFINITIVE COGNATE_RELATIVE_-VERB]. Its equivalent in English is something like “VERBing that I verb” and thus

¹²It is impossible to say whether this represents parallel independent innovation or a direct link between Nguni S40 and Great Lakes Bantu J. The latter is certainly possible, as the two groups display other affinities regarding both linguistic and non-linguistic traits (cf. Güldemann 1996: 112–113; 1999a: 77; 1999b: 175, fn. 10; 2019: 299–300).

a nominalisation directly derived from the InFocDoubling pattern. The observation that this derived structure exists in at least zones B, C, H and P is compatible with the assumption that its base pattern was also present in early clades of the family tree.

Regarding another recurrent variation within PrepFocDoubling, that between a post-infinitive and a clause-initial S/A constituent or, in our terms, between initial and preverbal PrepFocDoubling, we more firmly suggest a historical change from the former to the latter. The shift of the S/A position is associated with a shift away from a bisected cleft-like to a monoclausal syntactic structure, tightening the bond between the two verbs and re-establishing a more compact predicate constituent. This formal shift correlates in an expected way with the functional change from various PCF types within the IS domain to the encoding of such temporal meanings as progressive and proximal future pertaining to predicate semantics, as observed by Güldemann (2003) and De Kind et al. (2015). It would be useful to test systematically whether initial PrepFocDoubling never develops these semantic readings.

Summarising the above observations, we propose two historical clines in (a) and (b), which link the situation in the modern languages to PB. This clade is conceived here as by Guthrie, Meeussen and their contemporaries and is thus a little lower than the ancestral node 0 in the Bantu family tree of Grollemund et al. (2015), which includes Grassfields Bantu.

- (a) **InFocDoubling* > **initial PrepFocDoubling* > preverbal PrepFocDoubling
- (b) **SoA focus* > general PCF > progressive > proximal future

The states marked in italics are proposed as PB reconstructions (and possibly of earlier ancestral stages). The cline under (a) presents the formal and the one under (b) the corresponding functional development. As InFocDoubling and initial PrepFocDoubling recurrently coexist in languages, both can be ascribed plausibly to PB.

An important issue that still remains unclear is whether PB possessed in addition to PrepFocDoubling also PrepTopDoubling, which Meeussen's (1967) admittedly indeterminate account wants to suggest. While several instances of this construction exist in Bantu and are geographically quite widespread, various caveats cast doubt on reconstructing it for PB. One is that some cases of preposed verb doubling with an operator rather than SoA focus reading could be instances of a construction conveying today generalised PCF but having emerged from a PrepFocDoubling structure that grammaticalised beyond narrow SoA focus. Furthermore, the clearer cases of PrepTopDoubling have an overall eastern Bantu

distribution further away from the north-western homeland and may thus have appeared later. Finally, one needs to consider that the construction as such recurs cross-linguistically, so that it is well possible that such cases reflect multiple independent events of innovation. Opting for the latter scenario, Meeussen’s (1967) reconstruction would have to be qualified regarding its semantic-functional characterisation. Given his intimate knowledge of Bantu one wonders in fact which particular Bantu language(s) steered him to propose the quite specific IS reading in terms of PrepTopDoubling.

A more general synchronic and diachronic question that is worthwhile investigating in the future concerns the important role of the structural domain at issue for the marking of PCF and the dynamics holding between different relevant constructions, including their diverse functional effects. For one thing, this concerns languages described above that have recourse to more than one of the six patterns listed in §4.1 (see Appendix A). It also raises the issue of the relationship between PCF-sensitive predicate partition and other relevant marking strategies, in particular the conjoint/disjoint alternation that is equally pervasive in the Bantu family (cf. e.g. Güldemann 1996: §4.3; van der Wal & Hyman 2017). Two preliminary observations emerge in this respect from the above survey. First, the conjoint/disjoint alternation in the traditional narrow sense of segmental and/or supra-segmental marking pertaining to simplex verb forms appears to have a more restricted geographical distribution than the syntactic complex dealt with here. Second, there are relatively few languages like Rwanda JD61, Matengo N13, Makhuwa P31, and Zulu S42 that possess both basic strategies. Future research must show whether these findings can be substantiated and, if so, how they can be explained.

Acknowledgements

The present research was carried out within the project “Predicate-centered focus types in African languages” as part of the Collaborative Research Centre (SFB) 632 “Information structure: The linguistic means for structuring utterances, sentences and texts”, financed by the German Research Foundation (DFG). We are very grateful to the DFG for their generous funding of the project. This research was presented on previous occasions: Berlin-Ghent Workshop “Information Structure in Bantu Languages” held at the Humboldt University of Berlin, 10–11 December 2013; International Workshop “BantuSynPhonIS: Preverbal Domains” held at the Leibniz-Centre General Linguistics (ZAS), Berlin, and the Humboldt University of Berlin, 14–15 November 2014; Final Conference of the SFB 632 “Advances in Information Structure Research 2003–2015” (Poster ses-

sion) held at the Humboldt University of Berlin, 8–9 May 2015; International Conference “Reconstructing Proto-Bantu Grammar” held at Ghent University, 19–23 November 2018. We are grateful for helpful comments by the respective audiences. We also thank Yukiko Morimoto for her collaboration in the initial stages of our research on this topic, Thera M. Crane, Sebastian Dom, Rozenn Guérois, Hilde Gunnink, Joseph Koni Muluwa, Michel Lafon, Wilhelm J. G. Möhlig, Minah Nabirye, Jean Paul Ngoboka, and Jenneke van der Wal for additional language-specific information, and three reviewers for helpful suggestions for the final version of this chapter. Finally, we gratefully acknowledge English proofreading by Gianna Marks.

Abbreviations

ADD_F	additive focus	M	masculine
ADJ	adjective	NEG	negative
AUG	augment	NOM	nominalisation
AUX	auxiliary	OBJ	object
BG	background	PASS	passive
CJ	conjoint	PCF	predicate-centred focus
CONV	converb	PERS	persistive
COP	copula	PFV	perfective
DEF	definite	PL	plural
DEM	demonstrative	PN	proper name
DEP	dependent	POSS	possessive
DJ	disjoint	PRF	perfect
EMPH	emphatic	PRO	pronoun
FOC	(non-specific) focus	PROG	progressive
FUT	future	PROX	proximal
FV	default final vowel	PRS	present
GEN	genitive	PST	past
HAB	habitual	REL	relative
I	index	RES_F	restrictive focus
IAN	inanimate	S/A	subject/agent (as semantic role)
ID	identification	SBJ	subject (as grammatical relation)
INE	inessive	SBJV	subjunctive
INF	infinitive	SG	singular
IPFV	imperfective		
LOC	locative		

SoA	state-of-affairs	TOP	topic
STAT	stative	V	verb
TA(M)	tense/aspect/(modality)	VN	verbal noun

Arabic number numbers not followed by SG/PL indicate noun classes

<, > mark the scope direction of IS indices

Appendix A Predicate partition and PCF in (Narrow) Bantu

Variety	II	I	III	V	IV	VI	Source(s)
Ngiemboon [GF]	✓						Ibrahim (2007)
Ngwe [GF]	✓						Nkemnji (1995)
Limbum [GF]	✓	✓					Ndamsah (2012)
Basaa A43a	✓	✓					Bassong (2014)
Tuki A601		✓					Biloa (1997)
Makaa A83	✓						Ibrahim (2007; 2010)
Punu B43		✓					Hadermann (1996)
Nzebi B52		✓					Hadermann (1996)
Nsong B85d		✓	?				Koni Muluwa & Bostoen (2014; p.c.)
Lingala C30B	✓						Mufwene (1987); Meeuwis (2013), Joseph Koni Muluwa (p.c.)
Kituba H10A	✓						Joseph Koni Muluwa (p.c.), Mufwene (2013)
Kamba H112A		✓					Hadermann (1996)
Vili H12L		✓					De Kind et al. (2015)
Sundi H131K		✓			✓		Hadermann (1996); De Kind et al. (2015)
Kongo H16		✓			✓		Hadermann (1996); De Kind et al. (2015)
Kuyu E51		✓					Bynoe-Andriolo & Yillah (1975); Morimoto (2017)
Tharaka E54		✓					Abels & Muriungi (2008)
Rangi F33					✓		Gibson (2012; 2019)
Mbugwe F34					✓		Roth & Gibson (2019)

Variety	II	I	III	V	IV	VI	Source(s)
Rwanda JD61			✓				Jean Paul Ngoboka (p.c.)
Nkore-Kiga JE13/14	✓		✓				Taylor (1985); Asiiimwe & van der Wal (2019)
Ganda JE15	✓	✓					Jenneke van der Wal & Saudah Namyalo (p.c.)
Soga JE16	✓		✓				Nabirye (2016)
Ngoreme JE401					✓		Roth & Gibson (2019)
Gusii JE42					✓		Whiteley (1960)
Kuria JE43		?			✓		Sillery (1936); Landman & Ranero (2014)
Simbiti JE431					✓		Roth & Gibson (2019)
Luvale K14		?	?				Horton (1949)
Kwangali K33		?	?				Westphal (1958)
Manyo K332		?	?				Möhlig (1967, p.c.)
Mbukushu K333		?	?				Fisch (1977)
Fwe K402		✓	?				Gunnink (2016, 2018, 2019, p.c.)
Totela K41(1)		✓					Thera M. Crane (p.c.)
Ndendeule N101				✓			Tom Güldemann (field notes)
Ngoni of Tanz. N12				✓			Tom Güldemann (field notes)
Matengo N13				✓			Yoneda (2009)
Yao P21				✓			Hetherwick (1902); Sanderson (1922)
Makhuwa P31			✓				Asiiimwe & van der Wal (2019)
Shona S10						✓	Peggy Jacob (field notes)
Zulu S42	✓						Doke (1927), Michel Lafon (p.c.)
Copi S61	✓	✓	?				Asiiimwe & van der Wal (2019)

Abbreviations used in this table: GF = Grassfields; ✓ = present; ? = possibly present; I = PrepFocDoubling; II = InFocDoubling; III = PrepTopDoubling; IV = PrepFocLight; V = InFocLight; VI = PostFocLight.

References

- Abels, Klaus & Peter Muriungi. 2008. The focus marker in Kĩitharaka: Syntax and semantics. *Lingua* 118(5). 687–731.
- Ashton, Ethel O., Enoch M. K. Mulira, E. G. M. Ndawula & Archibald N. Tucker. 1954. *A Luganda grammar*. London: Longmans, Green & Co.
- Asiimwe, Allen & Jenneke van der Wal. 2019. Puzzle 1: Predicate doubling in Rukiga, Copi, and Makhuwa. Paper presented at the workshop “BaSIS Brainstorm”, Leiden University, 8–9 November 2019.
- Bassong, Paul R. 2014. *Information structure and the Basa’a left peripheral syntax*. Yaoundé: University of Yaoundé I. (Doctoral dissertation).
- Bilola, Edmond. 1997. *Functional categories and the syntax of focus in Tuki* (LINCOM Studies in African Linguistics 2). Munich: LINCOM Europa.
- Bostoen, Koen. 2008. Bantu spirantization: Morphologization, lexicalization and historical classification. *Diachronica* 25(3). 299–356.
- Bostoen, Koen & Gilles-Maurice de Schryver. 2015. Linguistic innovation, political centralization and economic integration in the Kongo kingdom: Reconstructing the spread of prefix reduction. *Diachronica* 32(2). 139–185 + 13 pages of supplementary material online.
- Bouka, Léonce-Yembi. 1989. *Éléments de description du kaamba : parler bantou de la République du Congo (groupe koongo, H17b)*. Brussels: Université libre de Bruxelles. (MA thesis).
- Bybee, Joan L. & Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13(1). 51–103.
- Bynoe-Andriolo, Esla Y. & M. Sorie Yillah. 1975. Predicate clefting in Afro-European creoles. In Robert K. Herbert (ed.), *Proceedings of the 6th Conference on African Linguistics* (Working Papers in Linguistics 20), 234–239. Columbus, OH: Department of Linguistics, Ohio State University.
- Crane, Thera M. 2019. Totela (K41). In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 645–691. Milton Park, Abingdon: Routledge.
- De Kind, Jasper, Sebastian Dom, Gilles-Maurice de Schryver & Koen Bostoen. 2015. Event-centrality and the pragmatics-semantics interface in Kikongo: From predication focus to progressive aspect and vice versa. *Folia Linguistica Historica* 36. 113–163.
- de Schryver, Gilles-Maurice, Rebecca Grollemund, Simon Branford & Koen Bostoen. 2015. Introducing a state-of-the-art phylogenetic classification of the Kikongo Language Cluster. *Africana Linguistica* 21. 87–162 + supplementary material online.

- Dik, Simon C. 1997. *The theory of functional grammar, part 1: The structure of the clause*. 2nd revised edition (Functional Grammar Series 20). Berlin: Mouton de Gruyter.
- Doke, Clement M. 1927. *Text book of Zulu grammar*. Johannesburg: The University of the Witwatersrand Press.
- Fiedler, Ines. 2010. Predicate-centered focus in Aja. Unpublished manuscript. Berlin.
- Fisch, Maria. 1977. *Einführung in die Sprache der Mbukushu, Ost-Kavango, Namibia* (Wissenschaftliche Forschung in Südwestafrika 15). Windhoek: S.W.A. Wissenschaftliche Gesellschaft.
- Gibson, Hannah C. 2012. *Auxiliary placement in Rangi: A Dynamic Syntax perspective*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Gibson, Hannah C. 2019. The grammaticalisation of verb-auxiliary order in East African Bantu: From information structure to tense-aspect. *Studies in Language* 43(4). 757–799.
- Gilman, Charles. 1986. African areal characteristics: Sprachbund, not substrate? *Journal of Pidgin and Creole Languages* 1(1). 33–50.
- Goodman, Morris F. 1985. Book review: Bickerton, D. 1981. Roots of language. *International Journal of American Linguistics* 51(1). 109–137.
- Green, Melanie. 2007. *Focus in Hausa* (Publications of the Philological Society 40). Oxford: Wiley-Blackwell.
- Grégoire, Claire. 1993. L'ordre SOV du nen: exception ou généralisation d'une tendance répandue en bantou du nord-ouest? Paper presented at the 23rd Colloquium on African Languages and Linguistics, Leiden University, 30 August – 1 September 1993.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guérois, Rozenn. 2015. *A grammar of Cuwabo (Mozambique, Bantu P34)*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Güldemann, Tom. 1996. *Verbalmorphologie und Nebenprädikationen im Bantu: Eine Studie zur funktional motivierten Genese eines konjugationalen Subsystems* (Bochum-Essener Beiträge zur Sprachwandelforschung 27). Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Güldemann, Tom. 1999a. Head-initial meets head-final: Nominal suffixes in eastern and southern Bantu from a historical perspective. *Studies in African Linguistics* 28(1). 49–91.

- Güldemann, Tom. 1999b. Toward a grammaticalization and typological account of the ka-possessive in Zulu. *Journal of African Languages and Linguistics* 20(2). 157–184.
- Güldemann, Tom. 2003. Present progressive vis-à-vis predication focus in Bantu: A verbal category between semantics and pragmatics. *Studies in Language* 27(2). 323–360.
- Güldemann, Tom. 2007. Preverbal objects and information structure in Benue-Congo. In Enoch O. Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic* (Trends in Linguistics, Studies and Monographs 191), 83–111. Berlin: Mouton de Gruyter.
- Güldemann, Tom. 2008. The Macro-Sudan Belt: Towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151–185. Cambridge: Cambridge University Press.
- Güldemann, Tom. 2016. Maximal backgrounding = focus without (necessary) focus encoding. *Studies in Language* 40(3). 551–590.
- Güldemann, Tom. 2019. The linguistics of Holocene High Africa. In Yonatan Sahle, Hugo Reyes-Centeno & Christian Bentz (eds.), *Modern human origins and dispersal. Words, bones, genes, tools* (DFG Center for Advanced Studies Series 2), 285–313. Tübingen: Kerns.
- Güldemann, Tom. In preparation. Dissecting predicates for focus: Towards a typology of predicate clefting, verb doubling, and co.
- Güldemann, Tom, Ines Fiedler & Yukiko Morimoto. 2014. The verb in the preverbal domain across Bantu: Infinitive “fronting” and predicate-centered focus. Paper presented at the International Workshop “BantuSynPhonIS Preverbal Domains”, ZAS Berlin and Humboldt University of Berlin, 14–15 November 2014.
- Güldemann, Tom, Ines Fiedler, Yukiko Morimoto & Kirill Prokhorov. 2010. Preposed verb doubling and predicate-centered focus. Paper presented at the International Conference on “Information Structure 2010”, University of Potsdam & Humboldt University of Berlin, 8–10 July 2010.
- Gunnink, Hilde. 2016. The fronted infinitive construction in Fwe. Paper presented at the 6th International Conference on Bantu Languages, University of Helsinki, 20–23 June 2016.
- Gunnink, Hilde. 2018. *A grammar of Fwe: A Bantu language of Zambia and Namibia*. Ghent: Ghent University. (Doctoral dissertation).
- Gunnink, Hilde. 2019. The fronted-infinitive construction in Fwe. *Africana Linguistica* 25. 65–87.

- Guthrie, Malcolm. 1948. *The classification of the Bantu languages* (Handbook of African Languages). London: Oxford University Press (for the International African Institute).
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Hadermann, Pascale. 1996. Grammaticalisation de la structure infinitif + verbe conjugué dans quelques langues Bantoues. *Studies in African Linguistics* 25(2). 155–169.
- Hetherwick, Alexander. 1902. *A handbook of the Yao language*. London: Society for Promoting Christian Knowledge.
- Horton, Alonzo E. 1949. *A grammar of Luvale*. Johannesburg: Witwatersrand University Press.
- Hyman, Larry M. & John R. Watters. 1984. Auxiliary focus. *Studies in African Linguistics* 15(3). 233–273.
- Ibrahim, Njoya. 2007. *Identificational vs. information focus in Makaa and Njembɔn: Interaction between syntax and semantics*. Yaoundé: University of Yaoundé I. (MA thesis).
- Ibrahim, Njoya. 2010. Reduplication in Makaa (A83). *Afrika und Übersee: Sprachen, Kulturen* 91(2). 259–298.
- Jacob, Peggy. 2010. On the obligatoriness of focus marking: Evidence from Tar’Barma. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure: A documentation of its diversity across Africa* (Typological Studies in Language 91), 117–144. Amsterdam: John Benjamins.
- Jaggar, Philip J. 2001. *Hausa* (London Oriental and African Language Library 7). Amsterdam: John Benjamins.
- Keegan, John M. 1997. *A reference grammar of Mbay* (LINCOM Studies in African Linguistics 14). Munich: LINCOM Europa.
- Koni Muluwa, Joseph & Koen Bostoen. 2014. The immediate before the verb focus position in Nsong (Bantu B85d, DR Congo): A corpus-based exploration. *ZAS Papers in Linguistics* 57. 123–135.
- Landman, Meredith & Rodrigo Ranero. 2014. On focus marking in Kuria. Paper presented at the 45th Annual Conference on African Linguistics, University of Kansas, 17–19 April 2014.
- Lubasa, N’ti Nseendi. 1974. *Locativity in Kikóongo and English*. Lubumbashi: National University of Zaire. (MA thesis).
- Manfredi, Victor. 1993. Verb focus in the typology of Kwa/Kru and Haitian. In Francis Byrne & Donald Winford (eds.), *Focus and grammatical relations in Creole languages. Papers from the University of Chicago Conference on Focus*

- and *Grammatical Relations in Creole Languages* (Creole Language Library 12), 3–51. Amsterdam: John Benjamins.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeuwis, Michael. 2013. Lingala structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Möhlig, Wilhelm J. G. 1967. *Die Sprache der Dciriku*. Cologne: University of Cologne. (Doctoral dissertation).
- Morimoto, Yukiko. 2017. The Kikuyu focus marker *nĩ*: Formal and functional similarities to the conjoint/disjoint alternation. In Jenneke van der Wal & Larry M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu* (Trends in Linguistics. Studies and Monographs 301), 147–174. Berlin: Walter de Gruyter.
- Mufwene, Salikoko S. 1987. An issue on predicate-clefting: Evidence from Atlantic creoles and African languages. In Philippe Maurer & Thomas Stolz (eds.), *Varia Creolica* (Bochum-Essener Beiträge zur Sprachwandelforschung 3), 71–89. Bochum: Studienverlag Dr. N. Brockmeyer.
- Mufwene, Salikoko S. 2013. Kikongo-Kituba structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Mugane, John M. 1997. *A paradigmatic grammar of Gikūyū* (Stanford Monographs in African Languages). Stanford, CA: Center for the Study of Language & Information.
- Nabirye, Minah. 2016. *A corpus-based grammar of Lusoga*. Ghent: Ghent University. (Doctoral dissertation).
- Ndamsah, Gratiana. 2012. Negation and the structure of Limbum. Paper presented at the 7th World Congress of African Linguistics (WOCAL 7), University of Buea, Cameroon, 20–24 August 2012.
- Nkemnji, Michael. 1995. *Heavy pied-piping in Nweh*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Pacchiarotti, Sara, Natalia Chousou-Polydouri & Koen Bostoen. 2019. Untangling the West-Coastal Bantu mess: Identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica* 25. 155–229.
- Roth, Tim & Hannah C. Gibson. 2019. Differential diagnoses of language change in Ngoreme, a Bantu language of Tanzania. *Africana Linguistica* 25. 281–319.
- Sanderson, Meredith. 1922. *A Yao grammar*. 2nd revised edition. London: Society for Promoting Christian Knowledge.

- Schadeberg, Thilo C. & Koen Bostoen. 2019. Word formation. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 2nd edn. (Routledge Language Family Series), 172–203. Milton Park, Abingdon: Routledge.
- Schwarz, Florian. 2003. Focus marking in Kikuyu. *ZAS Papers in Linguistics* 30. 41–118.
- Sillery, Anthony. 1936. Notes for a grammar of the Kuria language. *Bantu Studies, A Journal Devoted to the Scientific Study of Bantu, Hottentot and Bushmen* 10(1). 9–29.
- Taylor, Charles. 1985. *Nkore-Kiga* (Croom Helm Descriptive Grammars). London: Croom Helm.
- van der Wal, Jenneke & Larry M. Hyman (eds.). 2017. *The conjoint/disjoint alternation in Bantu* (Trends in Linguistics Studies and Monographs 301). Berlin: De Gruyter Mouton.
- Watters, John R. 1981. *A phonology and morphology of Ejagham, with notes on dialect variation*. Los Angeles, CA: University of California at Los Angeles. (Doctoral dissertation).
- Westphal, Ernst O. J. 1958. *Kwangari: An index of lexical types*. London: School of Oriental & African Studies, University of London.
- Whiteley, Wilfred H. 1960. *The tense system of Gusii* (East African Linguistic Studies 4). Kampala: East African Institute of Social Research.
- Yoneda, Nobuko. 2009. Matengo-no doushi-katsuyoukei-to shouten [Verb inflection and focus in Matengo]. *Journal of Swahili and African Studies* 20. 148–164.

Chapter 14

Proto-Bantu existential locational construction(s)

Maud Devos^a & Rasmus Bernander^b

^aRoyal Museum for Central Africa, Tervuren ^bUniversity of Helsinki

This chapter proposes a Proto-Bantu reconstruction of existential constructions based on a convenience sample of 180 Bantu languages, which points towards “existential locationals” (ELs) as a suitable base for comparison. ELs include inverse-locational predications as well as expressions of generic existence. We develop a detailed typology of ELs through a careful examination of the morphosyntactic variation which their building blocks display across Bantu. This typology clearly singles out two types of ELs with high frequencies and Bantu-wide distributions, which are reconstructable to at least node 5 in the phylogenetic tree of the Bantu family of Grollemund et al. (2015). Both display locative subject markers and “figure inversion” in relation to plain locational constructions. The difference between the main types lies in the selection of the copula: either a locative or a comitative one. North-Western and Central-Western Bantu languages show few reflexes of the suggested reconstructions. Instead, they often have non-inverted ELs which are cross-linguistically uncommon or, less frequently, ELs involving expletive inversion. The non-dedicated EL can be considered a retention of the original structure or a (contact-induced) innovation. Our preference goes to the second hypothesis assuming that a severe reduction of (locative) noun classes and ensuing (locative) agreement triggered a more rigid word order and consequently non-inverted ELs or inverted expletive ELs exempt of locative marking.

1 Introduction

1.1 On existential locationals and related notions in Bantu languages

Existential sentences or in short existentials have been defined as “specialized or non-canonical constructions which express a proposition about the existence or



the presence of someone or something” (McNally 2011: 1830, see also Bentley et al. 2013: 1). Existence is part of the semantic space LOCATION-EXISTENCE-POSSESSION (Lyons 1967) of which the English examples in (1) are typical instances.

- (1) “Test sentences” for semantic space LOCATION-EXISTENCE-POSSESSION (Koch 2012: 545)
- a. The boy has a book.
 - b. The book is on the table.
 - c. There is a book on the table.
 - d. There are many lions in Africa.
 - e. There are many unhappy people.

Whereas (1a) and (1b) are clear instances of respectively possession and location, there is some variation in the way the remaining three sentences are conceptualised. Although all three sentences are commonly designated as “existentials”, many authors consider only (1d) and (1e) as expressions of existence (e.g. Lyons 1967; Hengeveld 1992; Koch 2012). Conversely, sentences like (1c), in which the ground is an obligatory part of the predication, are characterised as “locational”. Koch (2012) distinguishes between (1b) and (1c) in information-structural terms. He considers (1b) as an instance of “thematic location”, because the located figure is the theme of the predication. In (1c), however, the pragmatic roles are inverted: the located figure is the rheme and the predication is characterised as expressing “rhematic location”. Creissels (2019c) rather uses the term “inverse-locational” predication for (1c) as opposed to “plain locational” predication for (1b) to reflect a change in perspectivisation: in (1c) the ground rather than the figure constitutes the perspectival centre. As for existentials proper, Koch (2012) makes a distinction between “bounded existence” (1d) and “generic existence” (1e). The latter is characterised by the absence of a nominal ground, whereas the former includes a nominal ground which specifies the locative context in which the statement of existence holds (Koch 2012: 538). In expressions of bounded existence, the relation between the figure and the ground is of a habitual rather than of a temporary and accidental nature, as in (1c) (Czinger 2002). Koch (2012) thus argues for a threefold distinction between thematic location, rhematic location and existence (including both bounded and generic subtypes). Still, on the basis of a 19-language sample with a bias towards Africa and Europe, he concludes that languages tend to reduce this conceptual diversity. Most languages display a constructional split between expressions of thematic location on the one hand and expressions of rhematic location and existence on the other hand.

A few languages have one construction type for expressions of location (whether thematic or rhematic) and another one for expressions of existence. In-between are those languages that cover the domain of location and existence by a single construction.

We originally framed our research in Creissels' (2019c) typology of inverse-locational predication and were thus particularly interested in those locational predications which involve an alternative way of encoding the prototypical figure-ground relationship, i.e. the ground rather than the figure is the perspectival centre. However, it quickly became clear that examples including a nominal ground were not always available. We therefore decided to include expressions of generic existence and also presentational clefts (Lambrecht 1988; 2001) or other presentationals (Gast & Haas 2011) which are used to "call the attention of an addressee to the hitherto unnoticed presence of some person or thing in the speech setting" (Lambrecht 1994: 39) and constitute a common extension of inverse-locational predication (Creissels 2019a). They are typically found at the beginning of a story and are thus easily retrievable. Nyamwezi F22 is one of the few languages for which we have examples of a plain/thematic locational (2a), an inverse/rhematic locational (2b), a bounded existential (2c), a generic existential (2d) and a presentational presentative (2e).

(2) Nyamwezi F22 (Maganga & Schadeberg 1992: 212, 200, 116, 209, 199)

a. plain/thematic location (cf. 1b)

ʊ-m-zó!gá gweén' ʊʒgo suúmvwá gʒ-βi m-kaayá
 AUG-3-pot 3.same 3.DEMII ought SM3-be.SBJV 18-9.house
 'that pot ought to be in the house'

b. inverse/rhematic location

aa-li=mo mʳnh' ʊʒ-ɳw-iilaále
 SM1-COP=LOC18 1.person AUG-18-farm
 'there is a person on the farm'

c. bounded existence (cf. 1d)

m-bʊ-holáanzi zi-lí=mó ɳóombe ɳiingí
 18-14-Holland SM10-COP=LOC18 10.cattle 10.many
 'there are many cows in Holland'

d. generic existence (cf. 1e)

zi-lí=hó ɳhaangála jáá-mbík' iifílí
 SM10-COP=LOC16 10.maize_beer 10.CONN-10.type 10.two
 'there are two types of *kangala* (maize beer)'

e. presentational presentative¹

βáa-li βá-lí=hó βáánhw' aafso βa-ka-lim' iilaále
 SM2.REM-COP SM2-COP=LOC16 2.people 2.DEMII SM2-NARR-farm 5.farm
 'there once were some people who cultivated a farm'

Nyamwezi has a single construction for the expression of what Koch (2012: 591) tentatively refers to as “existential location”, i.e. the semantic space involving expressions of inverse/rhematic location and existence. Moreover, the presentative (2e) shares the same construction.² In contrast to plain/thematic locationals this shared existential-presentational construction is characterised by a change in word order and “double” agreement on the verb. In (2b–2d) the figure follows the verb which displays double agreement: its subject marker agrees with the figure as in (2a) but it also takes a locative enclitic which agrees with the nominal ground in (2b–2c). The locative enclitic is also present in the absence of a nominal ground (2d–2e). In the latter case it can be interpreted as an exophoric agreement marker referring to an implicit ground or as a non-referential expletive marker. As will be discussed in §2.3, the distinction is not always an easy one to make.

We gathered data from 180 Bantu languages. Table 1 shows for how many languages we found three, two or only one (conceptual) type of locational/existential expression. As we only encountered five clear examples of bounded existence, we will not consider this existential subtype in our chapter. Intuitively, we expect expressions of bounded existence to pattern with expressions of inverse/rhematic location and probably also generic existence, as is the case in Nyamwezi (2b–2d). However, cross-linguistic data show that this should not be taken for granted. Somali, for example, uses *yaall* ‘be’ in expressions of (plain/thematic and inverse/rhematic) location, but expressions of bounded and generic existence involve *jiri* ‘exist’ (Koch 2012: 540, 542). Liko D201 also seems to make a distinction between locationals and existentials. Present tense “inverse locationals”/“plain locationals” select a suppletive form of the verb *ik* ‘be’ which is identical to the subject prefixes (de Wit 2015: 395). In generic existentials as well as expressions of bounded existence, the ‘insistive’ enclitic =*tɔ* is obligatorily added to this suppletive form. See the examples in (3).

¹Note that we use the term “presentative” to refer to a speech event, whereas “presentational” refers to the construction used to encode a presentative utterance (see also Gast & Haas 2011: 1). In the same vein, “location” and “existence” are conceptual notions, whereas “locational” and “existential” refer to their respective encoding constructions (or predications).

²The complex verb construction serves to set the story in the remote past.

(3) Liko D201 (André Ndagba, p.c.)

a. inverse location

bo-miki ba ka ndabɔ
 2-child SM3PL:be PREP 9.house

‘there are children in the house/the children are in the house’

b. bounded existence

bo-kpwíngi ba=tv ka Aflika
 2-lion SM3PL:be=INS PREP Africa

‘there are lions in Africa’

c. generic existence

ma-kpɔmɔka ma-pɔpɔ a=tv
 6-thing 6-strong SM3SG:be=INS

‘there are problems’

Further research is needed to determine whether more instances of split lexicalisations or other divergences between locationals and existentials occur in Bantu languages. Some preliminary findings are presented in §1.2.

Table 1: Quantification of conceptual types of existential expressions in our sample of 180 Bantu languages

Conceptual type(s) of existential expression(s)	number of languages	%
inverse/rhematic location & generic existence & presentative	24	13%
inverse/rhematic location & generic existence (only)	22	12%
inverse/rhematic location & presentative (only)	21	12%
generic existence & presentative (only)	2	1%
inverse/rhematic location (only)	76	42%
generic existence (only)	12	7%
presentative (only)	19	11%
negatives (only)	4	2%
	180	100%

1.2 Inverse location and generic existence

Languages for which we have examples of inverse/rhematic location and generic existence tend to be like Nyamwezi in that they use a single construction for both. In comparison to plain/thematic locationals this shared construction is either non-canonical (4b–4c), dedicated (2b–2d) or identical (5a–5b). Non-canonical constructions differ only in word order from the plain/thematic locational construction. In (4a) the figure functions as the subject and occurs in preverbal position. This order is inverted in inverse/rhematic locationals (4b) and existentials (4c), but the postverbal figure still functions as the subject triggering subject agreement on the copula. Inverse locationals and existentials thus have a non-canonical (VS instead of SV) word order – see also Bearth (2003) and van der Wal (2015) on basic/default or canonical word order in Bantu. Bantu languages are known to have flexible word order (van der Wal 2015: 19) and subject inversion is not restricted to expressions of inverse/rhematic location or generic existence (Marten & van der Wal 2014). We therefore do not consider the Swahili G42d inverse/rhematic locational (4b) or existential (4c) as dedicated (see §2.5.1 for further elaboration on the different existential constructions in Swahili) and we use the term non-canonical instead. However, the Nyamwezi examples in (2b–2d) are dedicated to the expression of inverse location and existence because they include locative morphology absent in the plain/thematic locational (2a).³ As will be discussed in §2.4 dedicated constructions are often characterised by the presence of a(n additional) locative proform. Finally, Lingala C30B in (5) is an example of what Koch (2012) refers to as a radical “generic location” language: there is no formal difference whatsoever between expressions of plain/thematic location and expressions of inverse/rhematic location (5a) or generic existence (5b).

- (4) Swahili G42d (Marten 2013: 61 for (4c), Bernander et al. Forthcoming(a) for (4a) and (4b))
- a. *ki-tabu ki-po meza=ni*
 7-book SM7-COP 9.table=LOC
 ‘the book is on the table’
- b. *meza=ni ki-po ki-tabu*
 9.table=LOC SM7-COP 7-book
 ‘there is a book on the table’

³Note that the selection of *βi* ‘be’, rather than *li* ‘be’, in the plain/thematic locational should be ascribed to the fact that *li* is a defective verb which cannot be used in the subjunctive mood (see also §2.2).

- c. *wa-po pia wa-chunguzi binafsi*
 SM2-COP also 2-investigator private
 ‘there are also private investigators’

(5) Lingala C30B (Michael Meeuwis, p.c.)

- a. *búku e-zal-í na mésá*
 book SM.3SG.INAM-COP-PRS on 9.table
 ‘the book is on the table/there is a book on the table’
- b. *bi-lamba pé e-zal-í*
 8-clothes also SM.3SG.INAM-COP-PRS
 ‘are there also clothes?’

In sum, even if additional data are certainly needed, Bantu languages can be said to show a tendency for joint constructionalisation of inverse/rhematic location and generic existence. We will refer to these shared constructions as “existential locationals” (ELs) (cf. Koch 2012: 591) but will continue to make a distinction between (rhematic/) “inverse locationals” (ILs) and “generic existentials” (GEs) where needed.

Still, some languages diverge between ILs and GEs in terms of agreement pattern and/or predicate. In Beo C45A, for example, GEs are characterised by expletive subject marking. The copula accompanied by the comitative marker *na* takes an invariable third person singular (class 1) subject marker (6a–6b). An additional locative proform figures in ILs. Following Gérard (1924: 69), it is triggered by the presence of a nominal ground.⁴

(6) Beo C45A (Gérard 1924: 69)

- a. generic existence
a-na ba-to ba-nyenye
 SM1_{EXPL}-COM 2-person 2-mean
 ‘*il y a des gens méchants*’ [‘there are mean people’]
- b. generic existence
a-li na kumu
 SM1_{EXPL}.PST-COP COM 1a.chief
 ‘*il y avait un chef*’ [‘there was a chief’]

⁴Note that it is not clear from the data what the morphosyntactic status of the locative proform is.

c. inverse location

a-li huna faranka gi-bale ka lekete
SM1_{EXPL}.PST-COP LOC.COM 10.franc 10-two in pocket
'il y avait deux francs dans la poche' ['there were two francs in the
pocket']

In Ndengeleko P11, verb agreement is governed by the postverbal figure in GEs (7a) and by the nominal ground in ILs (7b). Moreover, the postverbal figure is introduced by a comitative marker in ILs (7b) but not so in GEs (7a).

(7) Ndengeleko P11 (Ström 2013: 253, 283)

a. generic existence

ga-b-ii ma-bago ma-bili
SM6-COP-PFV 6-axe 6-two
'there are two axes'

b. inverse location

ku-b-íí ni múu-ndu ku-yéeto
SM17-COP-PFV COM 1-person 17-9.toilet
'there's someone in the toilet'

The available data suggests that the absence of the nominal ground in some expressions of GE correlates with a reduction in locative morphology. This implies that languages for which we only have GEs could end up being characterised as showing agreement with the figure whereas the actual situation might well be more diversified. Related to this, it should be noted that many languages have more than one EL and thus that the absence of data might at times lead to classifications which, upon more thorough research, will turn out to be too rigid (see also §2.5.1).

1.3 On presentationals and negative existentials

Our sample includes a fair number of presentational clefts and other presentationals (cf. Table 1) which we most typically encountered at the beginning of a narrative and whose main function is to introduce new entities into a discourse (Lambrecht 2001). They are a common usage extension of ILs and this is reflected by languages like Nyamwezi where the two constructions (2b and 2e) pattern alike. However, when taking a closer look at the languages for which we have presentationals as well as ILs, we find that divergences regarding the agreement

pattern and/or the predicate occur rather frequently, i.e. in almost half of the cases.

Shangaji P312 narratives habitually begin with the formulaic expression *khaz-aari toówo* ‘it wasn’t like this’. The narrator then introduces the story’s main character(s) or event (8a). This presentational is similar in structure to the ELs in (8b–8c): the entity new to the discourse (8a) and the figure (8b–8c) both occur in postverbal position and both trigger subject agreement on the verb. However, Shangaji ELs obligatorily include a locative enclitic which agrees with the nominal ground in ILs (8b) and with an exophoric ground in GEs (8c). They thus show double agreement. There is also an information-structural difference between presentationals and ELs. In presentationals, the postverbal subject receives contrastive focus marked by initial high tone insertion (8a, cf. *máúulu* & *úswáaiibu*),⁵ which is not the case in ELs (8b–8c).

(8) Shangaji P312 (Maud Devos, field notes)

- a. *kha-zaa-ri toówo yaa-ri má-úulu na n-khíira*
 NEG-SM10.PST-COP thus SM6.PST-COP 6-leg and 3-tail
waa-r’ ú-swáaiibu wa ngúukhu na xaága
 SM14.PST-COP 14-friendship 14.CONN 1a.chicken and 1a.eagle
 ‘It wasn’t like this, there were legs and a tail, there was a friendship between a chicken and an eagle.’
- b. *zaa-rií=vo khuúnttí z-iínkéénye z’ aá-tthu*
 SM10.PST-COP=LOC16 10.group 10-many 10.conn 2-person
va-páraaza
 16-9.terrace
 ‘there were many groups of people in front of the house’
- c. *waa-rí=wó uúca mwiínkéénye*
 SM14.PST-COP=LOC17 14.rice 14.many
 ‘there was a lot of rice’

In Malila M24, both presentationals and ELs include a locative proform. However, their morphosyntactic status differs. In the presentational, the postverbal discourse-new entity triggers subject agreement on the verb, which takes an additional locative enclitic (9a). In the EL, the preverbal ground is subject-marked on the verb and there is no agreement with the postverbal figure (9b) – see Bloom Ström (2020) for a similar pattern in Xhosa S41.

⁵See Devos (2017) for more on focus marking in Shangaji.

(9) Malila M24 (Eaton 2015: 5, 14)

- a. *á-lii=po* *u-mu-ntu* *ɯmo*
 SM1.PST-COP=LOC16 AUG-1-person 1.one
 ‘there was a certain person’
- b. *muula* *mwá-li* *i-tata*
 18.DEM_DIST SM18.PST-COP AUG-9.bush
 ‘in there was bush’

Presentationals and ELs also sometimes differ as to the choice of verb. Although presentationals often take *be* type or *have* type verbs just like ELs (cf. §2.2), they sometimes use a one-place predicate with a more specific meaning, like ‘go’ in (10), ‘do’ in (11) or ‘be at, exist’ in (12).

(10) Lega D25 (Meeussen 1962: 76)

- kwênd-ilé* *mu-ntu* *gu-mozi*
 SM17.go-PFV 1-person 1-one
 ‘there once went/was a man’

(11) Digo E73 (Nicolle 2013: 26)

- hipho kare kpwa-hend-a* *mu-tu* *na m-che-we*
 long ago SM17.PST-do-FV 1-person and 1-wife-POSS1
 ‘long ago, there was a person and his wife’

(12) Makwe P231 (Devos 2008: 449)

- á-ní-pwááw-a* *mwáali wá-ku-ít-á* *wá-lúúme*
 SM1-PST.PFV.DJ-exist-FV 1.girl 1.CONN-15-refuse-FV 2-man
 ‘there once was a girl who refused men’

The Makwe verb *pwawa* ‘be at, exist’ used in (12) also occurs in expressions including an explicit ground (13a). However, the verbs *wa* ‘be’ or *li* ‘be’ are preferred in ELs (13b).

(13) Makwe P231 (Devos 2008: 374)

- a. *n-kaátíi=mu* *mu-pwaw-a* *cíi-nu*
 18-inside=18.DEMI SM18-exist-IPFV.CJ 7-thing
 ‘is there anything inside?’
- b. *pa-méeza pa-w-ele* *kí-táabu ~ pa-li* *kí-táabu*
 16-6.table SM16-COP-PFV.CJ 7-book ~ SM16-COP.PFV 7-book
 ‘on the table there is a book’

Our data shows that although presentationals and ELs frequently pattern alike, the former often show a reduction in locative morphology or a demotion thereof from locative subject marker (9b) to locative enclitic (9a), and sometimes a different verb. This implies that languages for which we only have presentative expressions cannot be included in our typology, especially if they display agreement with the figure and absence of locative morphology.

Finally, our sample includes four languages for which we only have negative locational existentials. As shown by Bernander et al. (Forthcoming[b]), Bantu negative ELs may consist simply of standard negation applied to the corresponding affirmative construction, as in (14).

- (14) Shangaji P312 (Maud Devos, field notes)
kha-zaa-ri=wo *pwílimwiithi o-muú-ti*
 NEG-SM10.PST-COP=LOC17 10.mosquito 17-3-town
 ‘there were no mosquitos in town’

However, they often involve specialised morphosyntax. Nyamwezi is a case in point. It makes use of the adjective *dshw(ó)* ‘empty’. As can be seen in (15), the adjective agrees with the nominal ground and there is no agreement with the following figure. As dedicated negative existentials tend to be formally very divergent vis-à-vis their affirmative counterparts we therefore only consider non-dedicated negative constructions for the purposes of this chapter.

- (15) Nyamwezi F22 (Maganga & Schadeberg 1992: 226–227)
kú-weeleelo kú-dshw βó-soóndo
 17-5a.world 17-empty 14-goodness
 ‘there is no (real) goodness in the world’

In sum, our research focuses on rhematic locationals or, as Creissels (2019c) refers to them, inverse locational predications (ILs). Our data suggests that they show joint constructionalisation with generic existentials for which we use the joint term existential locationals (ELs). Bantu ELs are identical to plain/thematic locationals (PLs) or show non-canonical word order and/or specialised morphosyntax (often a locative proform). Presentational constructions often pattern with ELs. However, they show a tendency towards agreement with the figure rather than with the (implicit) ground and they sometimes select predicates different from the *be* and *have* type verbs found in ELs. Languages for which we only have presentational constructions are not further considered in this chapter which leaves us – after the additional subtraction of four languages with only inconclusive data – with ELs from 157 languages.

In the remainder of this chapter, we first look at the building blocks of Bantu ELs and morphosyntactic variation to develop a detailed typology (§2). We then take a closer look at the different types of ELs and their distribution within the Bantu domain (§3). Before suggesting actual Proto-Bantu (PB) reconstructions for ELs (§5) we investigate the non-inverted strategies in the north-western part of the Bantu area (§4). The last section (§6), finally, presents our conclusions.

2 Morphosyntactic variation in existential locationals

Recent (typological) studies on existential constructions like Bentley et al. (2013) and Bentley (2017) give the template in (16) for the typical components of existential constructions. The “pivot” is the only cross-linguistically obligatory element in this template. Given our focus on ELs and more specifically ILs, we will use the terms “figure” and “ground” rather than “pivot” and “coda” respectively, as the former are essential categories of semantic events of location (Talmy 1975).

- (16) Morphosyntactical template for existential constructions (Bentley et al. 2013)
 (EXPLETIVE) (PROFORM) (COPULA) PIVOT/FIGURE (CODA/GROUND)

The French example in (17) illustrates an existential construction including all typical components.

- (17) French (own knowledge)
il y a des livres sur la table
 EXPLETIVE PROFORM COPULA FIGURE GROUND
 ‘there are books on the table’

Let us now reconsider (2b), (7b) and (5a) to identify the relevant components in Bantu ELs. The Nyamwezi EL from (2b) has the components lined up in (18). The copula agrees with the inverted figure through the subject marker and with the (implicit) ground through a locative enclitic in the post-final slot.

- (18) Nyamwezi F22 (cf. (2b) above)
aa-li=mo m̄mnh’ ʔ́w-ŋw-ilaálé
 SM_{FIG}-COPULA-PFIN_{GROUND} FIGURE GROUND
 ‘there is a person on the farm’

The components of the Ndengeleko EL from (7b) are given in (19). The copula agrees with the ground. The figure is introduced by a comitative marker and followed by the ground.

- (19) Ndengeleko P11 (cf. (7b) above)
ku-bíí ni múundu kuyéeto
 SM_{GROUND}-COPULA COM FIGURE GROUND
 ‘there’s someone in the toilet’

The Lingala EL from (5a), finally, shows a non-inverted word order. The copula agrees with the preverbal figure. Bantu languages with this type of EL often have heavily reduced agreement systems.

- (20) Lingala C30B (cf. (5a) above)
búku e-zalí na_mésá
 FIGURE SM_{FIGURE}-COPULA GROUND
 ‘the book is on the table / there is a book on the table’

The figure is the central element of the templates and cannot be omitted. The nominal ground can be absent and we also found examples of copula dropping, in which a nominal ground is always present. Nominal grounds are characteristically expressed by locative nouns which in Bantu languages are generally derived through the addition of a locative nominal prefix of class 16 **pa-*, 17 **kɔ-* or 18 **mɔ-* (Meeussen 1967; Grégoire 1975). Other less widespread strategies for locative noun formation include the addition of the class 23/25 locative prefix **l-* (cf. Grégoire 1975; Maho 1999: 204–206) and the locative suffix *-(i)ni* (Schadeberg & Samsom 1994). Locative nouns are considered part of the noun class system and they can induce locative agreement within the noun phrase, and locative concords on the verb. However, in some Bantu languages locatives cannot induce locative agreement or concord and are therefore analysed as prepositional phrases rather than locative nouns (Grégoire 1975; Marten 2010; Zeller Forthcoming). This is most notably the case for the southern Bantu Nguni S40 and Sotho-Tswana S30 languages, but see §3.2.1 for additional cases in forest Bantu languages. Moreover, many north-western Bantu languages are devoid of productive locative marking and instead make use of prepositions unrelated to the reconstructed locative prefixes (Grégoire 1975; Guérois 2016; Zeller Forthcoming). Important variables in Bantu ELs are word order (§2.1), the verbal element (§2.2) and the agreement pattern (§2.3). The latter not only concerns the verb-initial subject marker, which can agree with the figure, with the ground or can be used expletively, but also secondary locative agreement markers which most frequently occupy the post-final verb slot (§2.4). We discuss them successively in the following sections which build up towards our typology of Bantu existential locationals (§2.5).

2.1 Word order and information structure

Bantu languages are said to display flexible word order associated with information structure (Bearth 2003; van der Wal 2015). The preverbal domain tends to be interpreted as non-focal if not topical, whereas the immediately-after-verb position receives a non-topical if not focal interpretation (cf. van der Wal 2015 and references therein).⁶ It thus does not come as a surprise that ELs show a change of word order with respect to PLs: The figure which is topical in PLs but not-topical in ELs moves from preverbal to postverbal position. The great majority of languages in our sample indeed show “figure inversion” with respect to PLs. However, non-inverted constructions are attested as well. They appear to be of two types. First, there are Liko- or Lingala-like cases, which show complete syntactic identity between ELs and PLs and thus ambiguous readings, as in (3a) and (5a), respectively. Koch (2012), who refers to these languages as “radical generic location” languages, suggests that the syntactic identity correlates with a rather fixed word order, which does not allow word order to reflect differences in information structure. However, it would also reflect joint constructionalisation of expressions of location and existence in these languages. Second, there are languages that do not adhere to the typical information-structural configuration sketched above in that they allow for non-topical or even focal constituents to occur in preverbal position. In Mbuun B87, for example, focused objects are moved to preverbal position and subjects are focused in situ but require movement of the object to sentence-initial position (Bostoen & Mundeke 2012). In ELs, this leads to the configuration in (21).

- (21) Mbuun B87 (Bwantsa-Kafungu & Meeussen 1970–71)
mw-e-saas mw-aa bá-nt àá-yé
18-7-shed 18-DEMI 2-people SM2.PRS-COP
'dans/sous ce hangar il ya des gens' ['in/under this shed there are people']

Western Serengeti languages show a similar configuration as they allow detopicalised constituents to occur in preverbal position (Nicolle 2015; Aunio et al. 2019; Bernander & Laine 2020). Although figure-inversion is possible in ELs in these languages (22a), the non-inverted word order is also attested (22b).

⁶Note that this also holds for Nen A44, well-known for its non-canonical OV word order, as “heavy” objects (objects carrying exclusive focus) tend to occur postverbally (Mous 2005).

(22) Ikoma JE45 (Bernander & Laine 2020: 74, 78)

- a. *n-tʃe-epi=hó* *tʃa-ŋómbɛ haase e=mo-té*
 FOC-SM10-PRS.COP=LOC16 10-cow under CONN9=3-tree
 ‘there are cows under the tree’
- b. *a-ká* *aβa-ýéni m-ba-ɸi=hó*
 23/25-home 2-guest FOC-SM2-PRS.COP=LOC16
 ‘there are visitors at home’

Lingala-type languages and Mbuun-type languages are hard to distinguish in the absence of data on language-specific information-structural characteristics. One way to distinguish between them could be the position of the ground which appears to move to sentence-initial position in the Mbuun-type languages illustrated in (21) and (22b). However, for now both types are classified as “no inversion” languages in our typology.

We now turn to the more regular pattern involving figure inversion. Figure inversion is part of a large range of related inversion constructions in Bantu languages referred to as subject inversion constructions (Demuth & Harford 1999; Marten & van der Wal 2014). For reasons explained further in §2.3, we prefer not to use the term “subject inversion” but rather use the term “figure inversion” because the inverted argument has the semantic role of figure whereas its syntactic function shows variation (logical subject, grammatical subject) and is subject to debate in Bantu theoretical linguistics (cf. Morimoto 2006; Diercks 2011; Salzmann 2011; van der Wal 2015). Figure inversion in ELs shares two constant characteristics with what Marten & van der Wal (2014: 3) refer to as core subject inversion constructions: (i) the logical subject (i.e. the figure for our purposes) follows the verb and cannot be omitted; and (ii) it is non-topical. The other two constant characteristics are less obvious in Bantu ELs, i.e. object marking appears to be marginally possible (cf. §2.4) and close bonding between the verb and the inverted figure does not appear to be necessary. The figure in ELs typically is non-topical; the information flow goes from the ground to the figure rather than the other way around. However, this does not imply that the figure is obligatorily indefinite or that it carries narrow or presentational focus. In many languages of the world, there is a restriction on definite figures in existential constructions (McNally 2011; Bentley et al. 2013), even if it is also generally acknowledged that indefiniteness is not an obligatory feature of the figure (Koch 2012; Creissels 2019c). Bloom Ström (2020) shows that although figures in Xhosa S41 existentials are typically indefinite, they are not obligatorily so (23).

- (23) Xhosa S41 (Bloom Ström 2020: 234)
ku-kho u-nyana wa-m apha
 SM17-be_present 1a-son 1-1SG.POSS here
 ‘there is my son here’

The main function of existentials is commonly said to be the introduction of a new referent into the discourse (Hengeveld 1992; McNally 2011; Koch 2012). However, as far as our data allow for generalisations on this topic, this does not seem to be reflected by narrow focus on the figure. Rather, the figure is typically underspecified for focus. As for the conjoint/disjoint alternation (cf. van der Wal & Hyman 2017), in Cuwabo P34, for example, there is a clear preference for the disjoint in ELs (Guérois 2015: 523), which implies a non-focal reading of the figure or athetic/sentence focus reading, as in (24). Data from Makwe show that the verb *pwawa* ‘be at, exist’ allows for a choice between conjoint and disjoint. The conjoint form implies narrow (exclusive) focus (13a), which is odd (25a) in expressions of bounded existence as they imply a habitual relation between the ground (here: the sky) and the figure (here: stars). The disjoint form is thus preferred (25b), except if one wants to emphasise that the presence of the figure is in some way exceptional. So, ELs allow for both the conjoint and disjoint, but the conjoint signalling exclusive focus on the figure (cf. van der Wal 2011) is the marked option.

- (24) Cuwabo P34 (Guérois 2015: 516)
o-ttólo=ni ókúle o-hi-kála fúlóóri
 17-well=LOC 17.DEMIII SM17-PFV.DJ-COP 9a.flower
 ‘at that well there is a flower’

- (25) Makwe P231 (Devos 2008: 386)
- a. ? *léelo ku-pwaw-ije jínóondwa ku-cáanya*
 today SM17-exist-PFV.CJ 10.star 17-high
 Int.: ‘today there are stars in the sky’
 - b. *léelo ku-ni-pwáaw-a jínóondwa ku-cáanya*
 today SM17-exist-PFV.DJ 10.star 17-high
 ‘today there are stars in the sky’

A similar situation holds for the so-called “augment” (cf. de Blois 1970). van der Wal & Namyalo (2016: 19) argue that the presence of an augment in Ganda JE15 results in athetic interpretation of the EL (26a), whereas absence of the augment signals exclusive focus on the figure (26b–26c).

(26) Ganda JE15 (van der Wal & Namyalo 2016: 19 for (26b–26c), Nanteza 2018: 30 for (26a))

- a. *e Kampala e-ri=yo a-ma-tooke*
 19 Kampala 19-COP=LOC19 AUG-6-banana
 ‘at Kampala there are bananas’
- b. *mu-katále mw-áá-báddé-mú báána b-okká*
 18-market SM18-PST-COP.PRF-LOC18 2.children 2-only
 ‘in the market were only children’
- c. *mu-katále mw-áá-báddé-mú baantú, si mbwa*
 18-market SM18-PST-COP.PRF-LOC18 2.people NEG.COP 10.dogs
 ‘in the market were people, not dogs’

In sum, except for its inverted position, the figure does not appear to be obligatorily specified for narrow focus identifiable by Bantu specific focus strategies such as the selection of a conjoint tense and the absence of an augment (cf. also the absence of a focal initial high tone in the Shangaji EL (8b–8c), which rather are the marked options in (25) and (26), respectively. We therefore adhere to Creissels’ (2019c: 10) analysis who, following Partee & Borschev (2004; 2007) and Borschev & Partee (2002), argues that “the difference between plain locational predication and inverse-locational predication is only indirectly related to information structure, and basically reflects the ‘perspectivization’ of the figure-ground relationships”. In ILs (and by extension ELs) the relationship is from the ground to the figure, whereas it is from the figure to the ground in PLs.

2.2 The verbal element

The verbal elements occurring in Bantu ELs are essentially of two types; they are related to the verbal element attested in: (i) plain locative predications (PLs); or (ii) possessive predications. In the following we discuss each type in turn before pointing out some interesting cases of merger between the two types and some rare instances of lexical specialisation in ELs.

The verb figuring in Bantu ELs is often identical to the one found in PLs, as illustrated in (2a) vs. (2b–2d), (4a) vs. (4b) and (5) above. We refer to this verbal element as a locative copula based on its function in PLs where it combines with a locative nonverbal predicate to form a verbal predicate (Dryer 2007). Different types of locative copula are attested in our sample: (i) defective ‘be’ verbs which, depending on language-specific characteristics, display more or less restricted verbal inflection; (ii) full-fledged ‘be’ verbs which do not show such a restriction;

and (iii) verbs with more specific meanings like ‘sit’ or ‘be at, exist’. Reflexes of the defective verb **di* ‘be’ are particularly common in Bantu ELs, as exemplified in (2), (8), (9), and (13). Swahili uses the defective verbs *po* (4), *ko* and *mo*, which are derived from locative enclitics, probably through the deletion of a preceding copula. Full-fledged ‘be’ verbs which do not show restricted verbal inflection are also attested in ELs where they often are in a more or less complementary distribution with a defective verb. In Shangaji, *ri* ‘be’ (< **di*) has a relatively wide usage range covering all present and past perfective verb forms (8b–8c). Other tense/aspect forms use the full-fledged verb *iya* ‘be’ (27).

- (27) Shangaji P312 (Maud Devos, field notes)
raangu zawiiy-ánk-á=vo *suphúuru*
 9.past SM10.PST.COP-PLUR-FV=LOC16 10.mat
 ‘in the past there used to be mats’

In Makwe defective *li* ‘be’ is much more restricted in use. It only occurs in present tense contexts, where it is in free variation with the regular verb *wa* ‘be’ (13b). Elsewhere only *wa* can be used. Finally, ELs with ‘be at, exist’ or ‘be, live, sit’ verbs are attested in some languages, such as Makwe (13a), which also uses *pwawa* ‘exist’ in PLs (28), but prefers *wa* ‘be’ in both ELs and PLs.

- (28) Makwe P231 (Devos 2008: 375)
kolóosho ji-pwaw-á *kwáaci? ji-pwaw-áa=pa*
 10.cashew SM10-exist-PRS.IPFV where 10-exist-PRS.IPFV=16.DEMI
 ‘where are the cashew nuts? they are here’

Cuwabo displays a different distribution: PLs typically make use of the defective verb *li* (29a) which is also attested in negative ELs (29b). Affirmative ELs (29c), however, consistently use *kala* ‘live, be, remain’, which can also be used in PLs (29d) and can thus be considered a locative copula as defined in this chapter. Still, Cuwabo shows clear signs of a split lexicalisation between PLs and negative ELs (*li*) on the one hand, and affirmative ELs on the other hand (*kala*).

- (29) Cuwabo P34 (Guérois 2015: 169, 397, 295, 371)
- a. *o-lí* *o-mabásâ=ni*
 SM1-COP 17-6.work=LOC
 ‘he is in his house’
 - b. *va-céla=ní=va* *ka-va-á-lí* *maanjé*
 16-well=LOC=16.DEF NEG-SM16-PST-COP 6.water.PL
 ‘at the well, there was no water’

14 Proto-Bantu existential locational construction(s)

- c. *mu-náá-vég-e* *o-íko* *o-hi-kála* *anyákôko*
 SM2PL-FUT-play-PROH 17-river SM17-PFV.DJ-COP 2.crocodile
 ‘do not play at the river, there are crocodiles’
- d. *bābááni* *o-ni-kál-êc-a* *va-tákûlu*
 1a.my_father 2-IPFV.CJ-COP-DUR-FV 16-9a.home
 ‘my father spends the day at home’

Apart from locative copulas, Bantu ELs also often make use of a verb identical to the one found in possessive constructions. The latter typically make use of a defective or fully-fledged ‘be’ verb in combination with a comitative marker introducing the possessee, i.e. the so-called “conjunctive” or “with-possessives” (Stassen 2013). The subject takes the role of possessor, as in (30a) from Gyeli A801 and (31a) from Cuwabo. In present tense contexts, a process reminiscent of what Stassen (2013) refers to as “HAVE-drift” often takes place: the ‘be’ verb is omitted and the comitative marker is inflected for person (30b). In some languages the comitative marker can also take restricted TAM marking (31b).⁷ The result is not a transitive HAVE-possessive (or transpossessive) construction as the possessee does not behave like an object and cannot be object-marked on the comitative. Another process reminiscent of HAVE-drift is the merger between the ‘be’ verb and comitative marker. In Cuwabo, for example, *káana* ‘have’ probably originates in *kála na* ‘be with’ (Guérois 2015: 445).

(30) Gyeli A801 (Grimm 2015: 357)⁸

- a. *mé* *bé* *nà* *nkwàno*
 1SG.PRS COP.R COM 3.honey
 ‘I have honey’
- b. *mé* *nà* *nkwàno*
 1SG.PRS COM 3.honey
 ‘I have honey’

⁷In Cuwabo, inflected *na* and *káana* are more regularly used to express ‘have’ than *li* in combination with *na*. The latter has the locative or stative meaning ‘be with’ (Guérois 2015: 444).

⁸It should be noted that we adapted the original glossing to our working definition of copulas thereby oversimplifying the Gyeli data. Gyeli has both verbal copulas like *bé* in (30a) to which a realis marking H tone may attach, and non-verbal copula like the ones in (103). For more on Gyeli copula types, see Grimm (2015: 346–378).

(31) Cuwabo P34 (Guérois 2015: 444, 345, 396)

- a. *míyó ddi-lí na ááná á-ili*
 1SG.PRO SM1SG-COP COM 2.child 2-two
 ‘I am with two children’
- b. *ka-ddi-á-ná makalra*
 NEG-SM1SG-PST-COM 6.charcoal.pl
 ‘I had no charcoal’
- c. *ba-a-kaána áyíma a-raarú ánáyánā*
 SEQ-SM2-COM 2.child 2-three 2.child.woman
 ‘they had three daughters’

We refer to (merged) combinations of ‘be’ and a comitative marker and to inflected comitatives as “comitative copulas”. ELs making use of a comitative copula were already seen in (6a–6c) and (7b). HAVE-possessives with a transitive ‘have’ verb are also used in Bantu possessive constructions, but they are rarer and often co-exist with a comitative copula, as is the case for Gyeli (32). In Rangi F33, both the comitative copula (33a) and the HAVE-possessive (33b) can be used in ELs (33a–33b).

(32) Gyeli A801 (Grimm 2015: 360)

- mè bùdé b-wánò bà-báà*
 1SG have.R 2-child 2-two
 ‘I have two children’

(33) Rangi F33 (Stegen 2011: 345, 373)

- a. *kura weerwii kwa-tiite Moosi Nkusa*
 there outside SM17.PST-have 1.sir Nkusa
 ‘there outside was Old Nkusa’
- b. *kaáyii kH-ri na isáare*
 9.home SM17-COP COM 5.matter
 ‘at home there is a matter’

Although locative copula and comitative copula or have-verbs are mostly easy to distinguish, there are some interesting cases of polysemy where the same verbal element is used to express both possession and location. Bastin (2020: 49) mentions (*i*)na, a merger of *di ‘be’ and *ná ‘with’, which in some zone H languages has acquired the meaning ‘be’. In these languages, there is no (longer) real

polysemy as the synchronic expression of possession requires the use of a comitative marker.⁹ However, in Totela K41 *ina* is polysemous between ‘be’ and ‘have’ (34a–34b). To disambiguate the two senses a comitative marker can be added in possessive constructions, but its presence is not obligatory (34c). Consequently, the EL with *ina* is relatable to both the PL and the possessive construction in Totela (34d). In addition to locative and comitative copula, we therefore distinguish a small but interesting category of locative/possessive copulas.

(34) Totela K41 (Crane 2011: 246, 107, 308)

- a. *èná lánzè èñándà*
SM1.COP 16.outside 16.9.house
‘he is outside the house’
- b. *ndin’ o-muzilili*
SM1SG.COM AUG-3.fresh_milk
‘I have fresh milk’
- c. *ndina nëñòmbè*
SM1SG.COP COM.9.cow
‘I have a cow/I am with a cow’
- d. *sùnú èchífùmò kà-kwiná* *ò-múkùlù*
today 7.morning PREHOD.IPFV-SM17.COP/COM AUG-1.elder
‘this morning there was an elder’

Bantu ELs thus typically make use of locative copula, comitative copula and less frequently of have-verbs and polysemous locative/possessive copula. Lexical specialisation is only rarely attested in (affirmative) ELs. A possible example is found in Eton A71, where ELs make use of a locative copula or of the verb ‘do’

⁹We do find interesting variation in possessive constructions suggesting that the shift from ‘have’ to ‘be’ is not completed yet in all zone H languages concerned. Dereau (1955: 30–31) gives for Central Kongo H16b a type of intransitive possessive construction which Stassen (2013) refers to as the ‘genitive possessive’ and which is generally rare in Bantu languages: *mwáana u-na yáame* (1.child SM1-COM COM.POSS1SG) ‘the child is with me/I have a child’. As Bastin (2020: 49) points out, *na* can be interpreted as ‘be’ here as it is followed by *ye* to express ‘be with/have’. Zombo H16hK has a very similar possessive construction (Araújo 2013: 178), but with a subject marker agreeing with the possessor, which suggests that the ‘have’ meaning lingers on: *a-ntú nzó é-nà záu* (2-person 10.house SM2-COP/COM 10.POSS2) ‘as pessoas têm casas’ [‘the people have houses < the people, houses they have theirs’]. Very similar examples are found in Tsootso H16hZ (Baka 1992: 87) (80a). (See <https://www.bantufirst.ugent.be/research/west-coastal-bantu-interactive-map> for more information about the referential classification of Tsootso employed here.)

extended with a valence-decreasing suffix. As Van de Velde (2008: 126) notes, this might be a semantic calque from French *se produire* ‘happen’.

- (35) Eton A71 (Van de Velde 2008: 126)
tindin à-H-kòm-bàn-H á ñ-póŋ
multiple_crash 1-PST-do-VDS-NF LOC 3-street
‘there has been a multiple crash in the street’

In sum, we identify in our sample two major types of verbal elements: locative copula and comitative copula, and three minor types: have-verbs, polysemous locative/possessive copula and specialised EL verbs. We now take a closer look at the agreement patterns attested in ELs.

2.3 Agreement patterns

Bantu inversion constructions have inspired an ongoing discussion about the status of the so-called subject (agreement) marker (Bresnan & Kanerva 1989; Demuth 1990; Bearth 2003; Morimoto 2006; Diercks 2011; Salzmann 2011; Khumalo 2012; van der Wal 2015). Let us reconsider the Malila EL in (9b). The verbal agreement marker agrees with the preverbal ground and not with the figure which in formal semantics would be referred to as the logical subject. This agreement pattern can be interpreted in two ways (Morimoto 2006: 164), either: (i) the agreement marker is a subject marker implying that the preverbal ground is the subject and that inversion is a grammatical-relation changing operation; or (ii) the agreement marker is a topic marker licensing the preverbal ground and no change in grammatical relation takes place (cf. also Bearth 2003: 141). This theoretical debate goes beyond the scope of this chapter. In this section we aim to describe the variation in agreement patterns and especially whether agreement is with the ground, the figure, both (double agreement) or none (expletive constructions). For ease of reference, we stick to the predominant Bantu tradition of referring to the verb-initial agreement marker as the subject marker.

Bantu ELs show three “single” and three “double” agreement patterns. As for the “single” ones, the subject marker agrees with the ground in most languages with figure inversion. It takes a locative subject marker which varies depending on the locative class of the ground (36a–36c). We refer to this agreement pattern as “locative inversion” (based on the terminology in Marten & van der Wal 2014).

(36) South Binja D26 (Meeussen & Sebasoni 1965)

- a. *así á-ɪ kyáta*
 16.ground SM16-COP 7.mat
 ‘*par terre est [il y a] une natte*’ [‘on the ground there is a mat’]
- b. *kʋ ndábʋ kɔ́-ɪ booba*
 17 9.house SM17-COP 8.thing
 ‘*sur la maison il y a des choses*’ [‘on the house there are things’]
- c. *mʋ ndábʋ mɔ́-ɪ booba*
 18 9.house SM18-COP 8.thing
 ‘*dans la maison il y a des choses*’ [‘in the house there are things’]

In other languages with figure inversion in ELs, such as Manda N11 in (37), the subject marker agrees with the postverbal figure. We label this agreement pattern “agreeing inversion” (cf. Marten & van der Wal 2014).

(37) Manda N11 (Bernander 2017: 250)

- pa-lóngólo y-áki, a-y-í’ mú-ndu mónɡa*
 16-front 9-POSS3SG SM1-COP-PRF 1-person 1.one
 ‘in front of it, there is a person’

Agreement with the figure is the predominant pattern in languages without figure inversion (5), although languages displaying this pattern often have reduced agreement systems.

In still other languages with ELs marked by figure inversion, the subject marker does not show agreement with the ground or with the figure but is a non-referential expletive marker. We distinguish three types of expletive markers: locative, non-locative and zero expletives. The latter concern the absence of a verb-initial agreement marker (76b). Locative expletives refer to invariable subject markers of a locative origin which do not display agreement with the ground. The Swahili example in (38) shows a mismatch between the locative class 16 of the ground and the locative class 17 of the subject marker pointing towards a non-referential expletive use of the latter.

(38) Swahili G42d (Marten 2013: 51)

- hapa ku-na kazi moja n-zuri sana ...*
 DEM16 SM17-COM 9.work 9.one 9-good very
 ‘here there is a very nice job ...’

However, mismatches in locative class agreement are not always a tell-tale sign of the expletive use of locative subject markers. Rwanda JD61, for instance, shows merger in locative class agreement: locative verb-initial agreement is always in class 16.

- (39) Rwanda JD61 (Zeller & Ngoboka 2018: 27)
mu ká-báande haa-shíze
18 12-valley SM16.PST.DJ-finish.PFV
'(the area) in the valley is finished'

In (39) the preverbal locative is clearly selected by the predicate and is therefore a thematic subject rather than an adjunct. An expletive interpretation of the class 16 subject marker is not possible in this context (Zeller & Ngoboka 2018: 27). The mismatch between class 18 of the preverbal locative and the invariable class 16 subject marker in (39) is thus not sufficient evidence for the expletive use of the latter. A similar case is found in Rundi JD62 (for which see Devos et al. 2017: 58). Unfortunately, we often do not have enough data to distinguish between a referential and an expletive use of locative subject markers. For now, we decided to categorise all inverted ELs with a subject marker of a (clear) locative origin as cases of “locative inversion”. Non-locative expletives are more easily detectable. They are invariable and do not agree with the ground or the figure (40). The agreement patterns marked by non-locative or zero expletives are referred to as “expletive inversion”.

- (40) Mboshi C25 (Prat 1917: 58)
o pu e-di la a-tsusu
17(?) village SM_{EXPL}-COP COM 2-chicken
'are there chickens in the village?'

All cases of “double” agreement involve the presence of an additional locative proform agreeing with the (implicit) ground. The most frequent pattern concerns ELs with figure inversion whereby the subject marker agrees with the postverbal figure and another secondary agreement marker agrees with the ground. In the Nyamwezi example in (2b) the subject marker agrees with the postverbal figure, whereas a locative enclitic attached to the verb agrees with the ground. We also came across an example of double agreement in a non-inverted EL in Mwera P22 in (41), where the subject marker agrees with the preverbal figure and the pre-initial locative marker agrees with the postverbal ground.

- (41) Mwera P22 (Harries 1950: 115)
mōto mu-gu-li n-nyumba
 3.fire 18-SM3-COP 18-9.house
 ‘there is fire in the house’

A less frequent pattern involves “redundant” double agreement, i.e. both the subject marker and a secondary agreement marker agree with the ground (42). In some cases, the subject marker has a locative origin which appears to be used expletively (43). A final pattern involves the combination of a non-locative expletive subject marker and a secondary locative marker agreeing with the ground (44).

- (42) Soga JE16 (Nabirye 2016: 240)
e-mmanga eyo e-li-yo aka-fo
 23-below 23.DEMII SM23-COP-LOC23 12-place
 ‘below there, there is a small place’

- (43) Nkore JE13 (Grégoire 1975: 77)
o-munju egyo ha-ri-mu a-ba-ntu
 AUG-3.house 3.DEMIII SM16-COP-LOC18 AUG-2-person
 ‘*dans cette maison il y a des gens*’ [‘in that house there are people’]

- (44) Haya JE22 (Grégoire 1975: 77)
o-musanduku egi a-li-mu e-bintu bike
 AUG-18.9.box 9.DEMI SM_{EXPL}-COP-LOC18 AUG-8.thing 8.few
 ‘*dans cette boîte, il y a peu de choses*’ [‘in this box are not many things’]

In sum, we distinguish three main agreement patterns in Bantu ELs: locative inversion, agreeing inversion and expletive inversion. Each of them has a subtype including a secondary locative proform.

2.4 Locative proforms: EL markers and lexicalisation

Locative proforms are an important element of Bantu ELs. Moreover, they sometimes constitute the basic difference between PLs and ELs, as in (2a) vs. (2b–2d).¹⁰ The locative proform can thus function as a dedicated EL marker. However, this is not always the case as a locative proform is sometimes attested in the PL as

¹⁰Recall that we argued, in line with Koch (2012) and Creissels (2019c), that a change in word order is not a sufficient characteristic to consider an expression of EL as a dedicated EL construction (cf. §1).

well as in the EL. In Nyakyusa M31, a locative enclitic agreeing with the ground appears to be obligatory in ELs (45a), but is optional in PLs (45b–45c). A similar pattern is attested in Western Serengeti languages (Bernander & Laine 2020).

(45) Nyakyusa M31 (Persohn 2017: 299, 315, 307)

- a. *lɪŋga fy-a-li=po i-fi-ndɔ paa-meesa*
if/when SM8-PST-COP=LOC16 AUG-8-food 16-table
'if there had been food on the table'
- b. *a-li=mo n-nyumba*
SM1-COP=LOC18 18-9.house
'he's in the house'
- c. *ɔ-mw-ana a-li mu-m-piki*
AUG-1-child SM1-COP 18-3-tree
'the child is in the tree'

As shown in §2.3, locative proforms mark agreement with the ground or are used expletively and most frequently occur in the subject marker and/or the post-final enclitic slot of the verb. In a few languages, such as Mwera in (41), they occupy the pre-initial proclitic slot. As indicated in §2.1, Marten & van der Wal (2014) claim that object marking is not possible in core subject inversion constructions. However, Yao P21 data from the 1920s, shown in (46), suggests that a locative proform can (or once could) occupy the object marker slot in ELs with figure inversion.

(46) Yao P21 (Sanderson 1922: 150)

- a. *wa-pa-li wa-ndu wa-jinji*
SM2-OM16-COP 2-person 2-many
'there were many people'
- b. *si-mu-li ng'ombe*
SM10-OM18-COP 10.cow
'there is cattle (in there)'

More recent Yao data from Whiteley (1966) do not show inclusion of one of three locative prefixes. Rather, the class 16 locative prefix is used irrespective of the locative class of the (implicit) ground (47a). The class 16 locative marker *pa-* and the defective verb *li* thus appear to have merged with subsequent lexicalisation giving rise to a lexical verb of existence. This *palí* verb is not preferred in ELs, which rather select the comitative copula, cf. (47b) and also Taji (2017: 77, 100).

(47) Yao P21 (Whiteley 1966: 173)

- a. *a-palí vá-ndu mú-mseu*
 SM2-exist 2-person 18-9.road
 ‘there are people in the road’
- b. *mwa-ná vá-ndu mú-mseu*
 SM18.PRS-COM 2-person 18-9.road
 ‘there are people in the road’

Similar forms, historically probably likewise including a class 16 locative marker and a ‘be’ verb, are encountered in several Makonde P23 varieties, e.g. *pawa* in Chinnima Makonde (Kraal 2005: 384) and *pagwa* in Plateau Makonde (Leach 2010: 368), as well as in Mwera P22, i.e. *pawa* and *pali* (Harries 1950: 115),¹¹ Mabiha P25, i.e. *pawa* (Harries 1940: 138), and Makwe P231, i.e. *pwawa* as in (12) and (13). In these languages too, lexicalisation has given rise to a lexical verb of ‘existence’ which is not the preferred choice in ELs. Locative proforms occupying the post-final slot also sometimes lexicalise into verbs of ‘existence’ rather than grammaticalise into specialised EL verbs. Rundi has a verb *riho* ‘exist’ resulting from the merger of *ri* ‘be’ and the class 16 locative enclitic, which is most frequently used in presentational clefts. Rundi ELs use either the form without the locative enclitic or including an enclitic agreeing with the ground (Devos et al. 2017: 77). In Xhosa, merger of the locative proform *kho* and the comitative marker *na* (cf. Bloom Ström 2020: 220) has given rise to a specialised EL verb (23) with a usage extension towards presentational expressions.

In sum, locative proforms are recurrent in Bantu ELs and sometimes function as dedicated EL markers. They most frequently occur in the subject marker or post-final slot. Locative proforms often participate in the lexicalisation of ‘existence’ verbs which tend to not be the preferred choice in ELs.

2.5 A typology of locative existential constructions

After discussing the variable features of Bantu ELs, we now combine them into two sets of features in Table 2. Vertically, variation as to type of verb in ELs is plotted (cf. §2.2): 1) locative copula, 2) comitative copula, 3) have-verbs, 4) locative/possessive copula, and 5) specialised EL verbs. The horizontal axis represents the four types pertaining to variable word order and agreement pattern, as outlined in §2.1 and §2.3: A) locative inversion, B) expletive inversion, C) agreeing

¹¹Mwera manifests a curious variation between the position of locative prefixes of classes 17 and 18 and that of class 16. While the former occur pre-initially (41), the latter occupies the object prefix slot (Harries 1950: 115).

inversion, and D) no (figure) inversion. Each of these four columns is further divided into two to indicate whether there is a secondary locative proform in addition to the primary agreement pattern. We label these subdivisions “single” (i) and “double” (ii) agreement patterns. Languages for which the EL in question is dedicated, i.e. differing from the PL in more than word order alone, are bolded in Table 2.

2.5.1 Intralingual variation

Languages for which we have diversified data often show the availability of more than one way of expressing existential location. Unfortunately, it is often not clear whether the different expressions are in free variation or not. Marten (2013), who gives a detailed account of the two ELs attested in Swahili G42d concludes that they differ in syntactic structure and usage range. The non-dedicated strategy with a locative copula and agreeing inversion in (4b), i.e. 1.C.ii in Table 2, has a less rigid word order (non-inverted constructions are possible) and wider usage range than the strategy with the comitative copula and locative inversion (38), i.e. 2.A.i in Table 2. In Manda N11, the strategy with agreeing inversion in (37), i.e. 1.C.i in Table 2, occurs more frequently than the one with locative inversion (48), i.e. 1.A.i in Table 2.

(48) Manda N11 (Rasmus Bernander, field notes)

apa pa-y-í fĩindu
PROX.DEM16 SM16-be-PRF 8.thing
'here there are things'

Shangaji P312 shows a similar difference in frequency between the strategy with agreeing inversion, as in (8b–8c), (14) and (27), i.e. 1.C.ii in Table 2, and the one with locative inversion, as in (49), i.e. 2.A.ii in Table 2. The presence of the latter strategy in Shangaji could be due to Swahili influence. Moreover, there appears to be an information-structural difference between the examples with agreeing inversion (8b–8c), (14) and (27) and locative inversion (49): the latter put focus (indicating surprise) on the figure.

(49) Shangaji P312 (Maud Devos, field notes)

o-na júguú=wó leélo
SM17-COM 1a.game=LOC17 today
'there's a game today!'

Information structure and language contact are also put forward as possible factors behind the remarkable plurality of strategies in the Western Serengeti languages (Bernander & Laine 2020). Ishenyi JE45 has up to four different strategies inventoried in Table 2: 1.C.ii (50a), 1.A.ii (50b), 2.A.i (50c) and 1.D.ii (50d). The strategy with the comitative copula could be due to Swahili influence, and Western Serengeti languages permit detopicalised constituents in preverbal position which explains the availability of both inverted and non-inverted constructions.

(50) Ishenyi JE45 (Bernander & Laine 2020: 76, 61, 78)

- a. *η-ko-réŋge=hó e-γi-táβo mu-méét̃fa*
 FOC-SM17-PST.COP=LOC16 AUG-7-book 18-table
 ‘there is a book on the table’
- b. *nu=hó t̃fé-re t̃fin-téépi t̃fen-kóro na=t̃fen-súúhu*
 FOC=LOC16 SM10-PRS.COP 10-animals 10-big COM=10-small
 ‘there are big and small animals’
- c. *haa-re βoosé mw-i-mótoka haa-na in-t̃f̃óka*
 16-DEM_{DIST} under 18-5-car SM16-COM 9-snake
 ‘there under the car there is a snake’
- d. *umw-éja o-ra-βa=hó*
 3-opportunity SM3-SIT-COP=LOC16
 ‘if there is time’

The most frequently attested interlingual variation in ELs is the choice between a locative (51a) and comitative (51b) copula (1 and 2 in Table 2), which co-occur in some languages, such as Ombo C76.

(51) Ombo C76 (Meeussen 1952: 31)

- a. *kó-lindí antu ikí*
 SM17-COP.IPFV 2.person 2.many
 ‘*il y a beaucoup de gens*’ [‘there are many people’]
- b. *ká-ík-í la=nguó*
 SM17.PST-COP.PFV COM=2.hippo
 ‘*il y avait beaucoup de hippopotames*’ [‘there were many hippos’]

Next, some languages have a strategy involving locative inversion as well as a strategy involving agreeing inversion (A & C in Table 2). In some languages, such as Malila in (9) and Xhosa (Bloom Ström 2020), this correlates to a difference in

Table 2: A typology of Bantu existential locational constructions

		A. LOCATIVE INVERSION		B. EXPLETIVE INVERSION		C. AGREEING INVERSION		D. NO INVERSION	
		i. single	ii. double	i. single	ii. double	i. single	ii. double	i. single	ii. double
1. LOCATIVE COPULA	A22, A72a,	F23	B11b,	JE22,	B87	JE42	A43a,	A62B	
	A74a	JD61,	B81	JE24	D25,	F22	A43b,	C71,	
	B31, B52,	JD62,	C32,		D311,	G11,	A44,	C81	
	B73d	JE13,	C35b,		D32	G40C,	A45,	D332	
	C76	JE15,	C502,		G63,	G42d,	A53,	F34	
	D25, D26,	JE16,	C53,		G67	G63	A622,	H16hK	
	D28, D43,	JE32b,	C73		JE431	JE24,	A71,	JE401	
	D55	JE402,			K332,	JE25,	A801,	JE45	
	E51, E55	JE45			K41	JE251,	A842,	(Ishenyi)	
	F12, F31,	(Ishenyi)			N11,	JE401,	A91	P22	
	F33				N121,	JE402,	B87		
	G52, G62,				N13	JE431,	C104,		
	G63, G67				P31	JE45	C14,		
	H16b, H16c,					(Ishenyi)	C30B,		
	H16hK,					K332	C411,		
	H21, H31,					L41	C55,		
	H41					M31	C61,		
	JD42, JD53,					N31a	C71,		
	JD61, JD62,					P31,	C75		
	JD63					P312,	D201,		
	K14, K21,					P312,	D311,		
	K33, K402					P34	D32		
	L23, L31a,						JE45		
	L32, L33,						K332		
	L52, L53								
	M15, M24,								
	M301, M41,								
	M42, M54,								
	M62, M64								
	N11, N15,								
	N21, N31a,								
	N31b, N41,								
	N43								
	P22, P23,								
P231, P25,									
P34									
R11									

14 *Proto-Bantu existential locational construction(s)*

A. LOCATIVE INVERSION		B. EXPLETIVE INVERSION		C. AGREEING INVERSION		D. NO INVERSION	
i. single	ii. double	i. single	ii. double	i. single	ii. double	i. single	ii. double
2. COMITATIVE COPULA	A34	P312	A91	C45A	F22?	A45	
	B73d		B25, B72a			F34	
	C101,		C15, C25,				
	C76		C301,				
	E51,		C61J				
	E622C,						
	E73						
	F33						
	G12, G22,						
	G23, G32,						
	G35, G38,						
	G42d,						
	G63, G65						
	H16hK,						
	H31, H41						
	JE402,						
	JE45						
	(Ishenyi)						
	K21, K33						
	L13						
	M13						
	N12, N44						
	P11, P13,						
	P14, P21,						
	P231						
	R21, R22,						
	R30						
	S10, S13,						
S21, S31,							
S33, S407,							
S43, S51,							
S53, S54							

	A. LOCATIVE INVERSION		B. EXPLETIVE INVERSION		C. AGREEING INVERSION		D. NO INVERSION	
	i. single	ii. double	i. single	ii. double	i. single	ii. double	i. single	ii. double
3. HAVE- VERBS	E622C, E65, E74 F33							
4. LOCATIVE/ POSSES- SIVE COPULA	H32 K41 R41						B865	H16hZ
5. SPECIAL- ISED EL VERBS	P21 S41, S42		H21	H21	P21		A71	

usage, i.e. existential location vs. presentative. In other languages, such as Ishenyi in (50a, 50c) vs. (50b), both are attested in ELs.

The last recurrent pattern concerns the variation between the presence and the absence of double agreement (i & ii) in the same language. In Rundi JD62, the locative copula can take both a locative subject marker and a locative enclitic as in (52a) (1.A.ii) or only a locative subject marker as in (52b) (1.A.i).

(52) Rundi JD62 (Devos et al. 2017: 72; Manoah-Joël Misago, p.c.)

- a. *mu bu-úuki ha-ri=mwó i-súkáari*
 18 14-honey SM16-COP=LOC18 AUG-9.sugar
 ‘in honey(, there) is sugar’
- b. *ha-ri i-gi-tabu ku méézá*
 SM16-COP AUG-7-book 17 9.book
 ‘there is a book on the table’

In sum, more research is needed to account for the plurality of ELs in some Bantu languages. Possible motivating factors include usage range, language contact and information structure. Moreover, seeing that most of our data on ELs in Bantu languages is limited, further research might show that intralingual variation is a more general feature of Bantu ELs.

2.5.2 Some typological generalisations

Before turning to a historical-comparative account, we check our Bantu EL typology against existing typologies of existential constructions, more specifically those by Koch (2012) and Creissels (2019c).

We find that Bantu languages overwhelmingly display split constructionalisation between expressions of thematic location on the one hand and expressions of rhematic location and existence on the other hand. The bolded languages in Table 2 all show this distinction. The unbolded ones show joint constructionalisation of thematic and rhematic location as well as existence. Some languages merely show a word order permutation (unbolded C type languages), whereas others do not even show this minimal difference (unbolded D type languages). Many of the latter languages belong to North-Western Bantu (NWB) or Central-Western Bantu (CWB) branches (cf. Grollemund et al. 2015). This allows for at least two hypotheses: (i) figure inversion emerged after the NWB and CWB branches had split off; or (ii) figure inversion in ELs became obsolete in NWB and CWB and was replaced by a non-dedicated, non-inverted construction which could be interpreted as an areal feature which these Bantu languages share with the so-called Macro-Sudan Belt linguistic area (Clements & Rialland 2008; Güldemann 2008). In fact, non-inverted ELs have been put forward as a shared feature of the latter linguistic area (Creissels 2019a,b). We take a closer look at non-inverted ELs in §4. There is no clear evidence for split constructionalisation between expressions of location (whether thematic or rhematic) and expressions of existence in Bantu. However, more diversified data is needed to ascertain this claim (cf. also the divergences between expressions of rhematic location and generic existence described in §1). Koch (2012: 582–583, fn. 24) mentions that Zulu shows evidence for both joint constructionalisation between rhematic location and existence through the use of the comitative copula *na* (53a–53b) and split constructionalisation between rhematic location and existence through the use of the specialised verb *khona* (53c) in existentials (but not in rhematic/thematic locationals). However, additional evidence shows that *khona* can be used in expressions of rhematic location displaying locative inversion (54). Notice that in generic existentials and presentationals *khona* shows a preference for agreeing inversion (53c) (cf. also Bloom Ström 2020 on Xhosa).

(53) Zulu S42 (Koch 2012: 570, 573)

- a. *ku-ne-bhuku e-tafuleni*
 SM17-COM.5-book LOC-table.LOC
 ‘there is a book on the table’

- b. *ku-na aba-ntu aba-hlupheka-yo*
17-COM 2-person 2-be_unhappy-REL
'there are unhappy people'
- c. *ba-khona aba-ntu aba-hlupheka-yo*
SM2-LOC.PRED 2-person 2-be_uhappy-REL
'there are unhappy people'
- (54) Zulu S42 (Buell & de Dreu 2013: 462)
ku-khona aba-fundi ku-lesi si-kole
SM17-LOC.PRED 2-student PREP-7.DEM 7-school
'there are students at this school'

Following Creissels (2019c), we find that dedicated Bantu ELs are overwhelmingly of the types 'there-be' (1A) and '(there-)be-with' (2A). Whereas the use of an expletive subject in impersonal constructions appears to be cross-linguistically predominant (Creissels 2019b), Bantu languages allow for a referential locative subject marker which agrees with the ground. Still, locative subject markers can be used expletively and non-locative expletive subject markers are attested as well. Both the use of referential locative subject markers and the use of the comitative copula in ELs seem to be typical Bantu features (Creissels 2019c: 26, 33).

3 Main types and variation

In this section we take a detailed look at Table 2, which clearly highlights two major EL types in Bantu: 1.A.i and 2.A.i (cf. also Creissels 2019a). Both are frequent in our sample and show a Bantu-wide distribution covering, if not all zones, all phylogenetic groups in Grollemund et al. (2015), i.e. North-Western Bantu (NWB), Central-Western Bantu (CWB), West-Western Bantu (WWB), South-Western Bantu (SWB) and Eastern Bantu (EB). Type 1.A.i is characterised by the use of a locative subject marker and a locative copula. Type 2.A.i likewise involves locative subject marking but makes use of a comitative copula. Whereas the use of a comitative copula overwhelmingly correlates with locative subject marking and a postverbal figure, locative copulas display more variation as to agreement and word order. We first discuss the verbal elements (§3.1) making a main distinction between locative (§3.1.1) and comitative (§3.1.2) copula and relating the remaining types of verbal elements to these two main types (§3.1.3). We then take a closer look at the agreement patterns (§3.2), starting with locative subject markers (A) and related expletive subject markers (B) (§3.2.1), before turning to agreement with the inverted (C) or non-inverted (D) figure (§3.2.2).

3.1 Verbal elements

One hundred and five sample languages make use of a locative copula in ELs. They are spread over the whole Bantu domain, except for zone S. This may be an accidental gap, but it ties in with the predominance of comitative copula in zone S. Fifty-nine sample languages make use of a comitative copula in ELs. These languages are also spread over the whole Bantu domain, but this time with the exception of zone D (including JD-languages, viz. zone D languages reclassified into zone J). The three other types of verbal elements (i.e. have-verbs, locative/possessive copula and specialised EL verbs) are attested in 14 languages only.

3.1.1 Locative copula

In this section we concentrate on locative copulas found in ELs of the type 1.A.i in Table 2. The variation in the choice of the locative copula in the 63 languages concerned reflects the overall variation. By reducing the number of languages to look at, we allow for a more detailed discussion. In 41 languages, listed in Table 3, ELs include a reflex of the defective verb **di* (Bastin et al. 2002).

As mentioned in §2.2, we refer to **di* as a locative copula because it consistently introduces locative predicates in PLs. It typically shows more or less restricted verbal inflection and is often found in a complementary distribution with a regular ‘be’ verb in both PLs and ELs.

In three languages, the locative copula appears to consist of the reflex of **di* and an extra element.

- (55) Kpe A22 *wéli* (Tanda & Neba 2005: 210)
 Nzebi B52 *lííd* (Marchal-Nasse 1989: 532)
 Ombo C76 *li-ndi* (Meeussen 1952: 30)

In Nzebi B52, C(V) roots are regularly extended with *-ad*, for instance *b* ‘be’ becoming *báád* (Marchal-Nasse 1989: 440, 533); *li* is only used in the perfect, i.e. *liidi*, comparable to *beedi*, the perfect of *báád*. In Ombo C76, the locative copula almost always takes the imperfective suffix *-ndi* (Meeussen 1952: 23–24). Only for Kpe A22, do we not have enough data to ascertain whether *wéli* includes a reflex of **di*.

In 63 sample languages, we identified a reflex of the full-fledged verb **bá* ‘dwell, be, become’ (Bastin et al. 2002). In six of them, listed in (56), it is the only verb attested in ELs. Admittedly, for Tsogo B31 and Holoholo D28, we only have past and negative ELs in which **di* might well be regularly replaced by **bá*. However, the other four languages in (56) appear to have lost **di*. In Makonde and Mabiha,

**di* is either entirely absent or a trace is found in the lexicalised verb of existence *pali* (cf. §2.4). In Makwe, *li* is still used in the present tense, but even there *wa* ‘be’ is preferred (13b).

- (56) Tsogo B31 *ba* (Marchal-Nasse 1979: 51)
 Holoholo D28 *ba* (Grégoire 1975: 32)
 Ndamba G52 *va* (Edelsten & Lijongwa 2010: 116)
 Manda N11 *ya* (Bernander 2017: 258-259)
 Makonde P23 *va* (Kraal 2005: 323)
 Mabiha P25 *wâ* (Grégoire 1975: 43)

Table 3: Locative copula which are reflexes of **di*

Lega	<i>lr</i>	Yombe	<i>dr</i>	Ruund	<i>d</i>	Tumbuka	<i>li</i>
D25		H16c		L53		N21	
South Binja	<i>r</i>	Shi	<i>li</i>	Mambwe	<i>lr</i>	Chewa	<i>li</i>
D26		JD53		M15		N31b	
Nyanga	<i>rr</i>	Rwanda	<i>li</i>	Malila	<i>lr</i>	Nyanja	<i>li</i>
D43		JD61		M24		N31a	
Buyu	<i>r</i>	Rundi	<i>ri</i>	Ndali	<i>li</i>	Nsenga	<i>li</i>
D55		JD62		M301		N41	
Kuyu	<i>rr</i>	Fuliiru	<i>ri</i>	Nyakyusa	<i>lr</i>	Nyungwe	<i>li/ri</i>
E51		JD63		M31		N43	
Kamba	<i>r</i>	Luvale	<i>li</i>	Taabwa	<i>lr</i>	Mwera	<i>li</i>
E55		K14		M41		P22	
Bende	<i>li</i>	Songye	<i>i</i>	Bemba	<i>lr</i>	Makwe	<i>li</i>
F12		L23		M42		P231	
Nilamba	<i>lr</i>	Luba-Kasai	<i>di</i>	Lamba	<i>li</i>	Umbundu	<i>li</i>
F31		L31a		M54		R11	
Rangi	<i>rr</i>	Kanyok	<i>dy</i>	Soli	<i>li</i>		
F33		L32		M62			
Hehe	<i>li</i>	Luba-	<i>di</i>	Tonga	<i>li</i>		
G62		Katanga		M64			
		L33					
Bena	<i>li</i>	Lunda	<i>di</i>	Tonga	<i>i</i>		
G63		L52		N15			

ní which is widespread in Eastern Bantu (Meeussen 1967: 115; Wald 1973; Gibson et al. 2019) and known to expand its usage range at the expense of **di* (Wald 1973: 248–249).

- (59) Nande JD42 (Grégoire 1975: 76)
o-mo-ba-ndw abá mu-ny ó-mwibi
AUG-18-2-person 2.DEMI 18-COP AUG-1.thief
'*parmi ces hommes-là, il y a un voleur*' ['among those people, there is a thief']

Six languages have a locative copula relatable to comitative *na*. As explained in §2.2, Bastin (2020: 49) argues that (*i*)*na* has acquired the meaning 'be' in some zone H languages and can thus be found in ELs and PLs alike. As shown in (60), this change is also attested in zones A and K.

- (60) Ewondo A72a *nə* (Grégoire 1975: 123)
Bulu A74a *ne* (Grégoire 1975: 123)
Manyanga H16b *ina* (Laman 1912: 240)
Yaka H31 *ina* (Bwendelele s.d.)
Mbala H41 *ina* (Moyo-Kayita 1981: 120)
Fwe K402 *ina* (Gunnink 2018: 84)

In Eton, closely related to Ewondo and Bulu, *ne* can also be used in copular clauses, where it sometimes optionally (61) combines with the comitative marker *èèy*. This optionality of *èèy* points towards an origin as a comitative copula (61b).

- (61) Eton A71 (Van de Velde 2005: 405, 202)
a. *à-nè èèy lè-bùm*
SM1-COP COM 5-belly
'she is pregnant < she is with belly'
b. *à-nè lè-bùm*
SM1-COP 5-belly
'she is pregnant < she is with/has belly'

In languages where the comitative copula acquired the meaning 'be' and is used as a locative copula, 'have' is expressed either through the combination of the former comitative copula and a (new) comitative (62) or through a 'have' verb typically derived from a verb meaning 'seize, grasp' (63).

- (62) Mbala H41 (Moyo-Kayita 1981: 71)
wéna i ngangú
 SM1.COP COM 9.intelligence
 ‘*il est intelligent*’ [‘he is intelligent < he is with intelligence’]
- (63) Fwe K402 (Gunnink 2018: 108)
ndi-kwèsí a-bá-mbwa
 SM1SG-have AUG-2-dog
 ‘I have dogs’

In sum, locative copulas usually are or include a reflex of **di* which originally was in complementary distribution with a full-fledged ‘be’ verb, most often a reflex of **bá*. In some languages, the latter eventually replaced the locative copula. In a small set of languages, the comitative copula has undergone a semantic shift towards the expression of location. The specialised EL verb *pali* in Yao is a variation on the main locative copula type as it probably originates in the merger of the class 16 object marker *pa-* and *li* as mentioned in §2.4. However, as explained there, it is not the preferred verb in Yao ELs.

3.1.2 Comitative copula

In this section we focus on comitative copulas found in ELs of the type 2.A.i in Table 2. This type is attested in 46 of the 58 languages using a comitative copula in ELs. As was noted in §2.2, Bantu ELs often take a possessive predicator which typically consists of a comitative copula, i.e. a locative copula followed by a comitative marker or a comitative marker inflected for subject marking. Below we first look at the full comitative copula before considering the eroded form, i.e. the form without the locative copula. We then take a look at a special comitative copula consisting of what looks like an inflected comitative marker itself followed by an invariable comitative marker.

In 15 languages, the comitative copula is a locative copula followed by a comitative marker. The locative copula is either a reflex of **di* (ten languages), **bá* (five languages), or **(j)ikad* (one language). Makwe can choose between *li* or *wa* in present tense contexts.

- | | | | |
|------|-------------------|----------------|--------------------------------|
| (64) | Teke Tyee B73d | <i>li ya</i> | (Ruth Raharimanantsoa, p.c.) |
| | Babole C101 | <i>i na</i> | (Leitch 2003) |
| | Ombo C76 | <i>lmdi la</i> | (Meeussen 1952; Grégoire 1975) |
| | Kuyu E51 | <i>ri na</i> | (Englebretson et al. 2015) |
| | Vunjo-Chaga E622C | <i>i na</i> | (Moshi 1995) |

Rangi F33	<i>ri na</i>	(Stegen 2011)
Kinga G65	<i>le na</i>	(Enock Mbilung’i, p.c.)
Kwezo L13	<i>di nɔ</i>	(Grégoire 1975)
Fipa M13	<i>li na</i>	(Struck 1911)
Makwe P231	<i>li na/wa na</i>	(Devos 2008)
Ngoni of Tanz. N12	<i>vi na</i>	(Gastor Mapunda, p.c.)
Ndengeleko P11	<i>ba ni</i>	(Ström 2013)
Matuumbi P13	<i>ba na</i>	(Odden 1996)
Ngindo P14	<i>ba na</i>	(Gromova & Uрманchieva 2005)
Kwangali K33	<i>kara na</i>	(Dammann 1957)

The comitative marker is mostly a reflex of **nà* ‘with, also, and’ (Meeussen 1967: 115; Bastin et al. 2002) (ten languages) or its variants **dà* (Ombo) or **jà* (Teke Tyee B73d) (Bastin et al. 2002). In two languages we find a comitative marker with a vowel different from *a*. The vocalic change can be explained in different ways. The use of *ni* rather than *na* in Ndengeleko could indicate that the current comitative marker in these languages is a reflex of the copula **ní* rather than of **nà*, as comitative *ni* can indeed be used as a copula, as shown in (65). In other Eastern Bantu languages, such as Shangaji in (66), the reflexes of **ní* (i.e. *ti*) and **nà* are also in free variation in at least some contexts.

(65) Ndengeleko P11 (Ström 2013: 280)

ywéembe ni ηjóoi
 PRON.1 COM 1.old_person
 ‘he is an old person’

(66) Shangaji P312 (Maud Devos, field notes)

oolay-iw-a na=siimba oolay-iw-a ti siimba
 SM1.kill-PASS-PFV COM=1a.lion SM1.kill-PASS-PFV COP 1a.lion
 ‘he was killed by a lion’

The proclitic use of the comitative marker might also be a trigger of vocalic change. In Kinga G65, *na* merges with the augment of the noun referring to the figure (67). If a specific vowel sequence is particularly frequent this could cause the vowel of the comitative marker to change.

(67) Kinga G65 (Enock Mbilung’i, p.c.)

kho-le n=u-mu-nu / kho-le n=a-va-nu
 SM17-COP COM=AUG-1-person / SM17-COP COM=AUG-2-person
 ‘there is a person’/‘there are persons’

Next, the comitative marker often has a short personal pronoun cliticised to it (Dammann 1977), which could also trigger vocalic change after intervocalic consonant loss and/or merger. In Pare G22, for example, the comitative marker has two allomorphs, i.e. *na/ne*, of which the second could be a merger of *na* and the class 1 short personal pronoun *-ye* (Mous & Mreta 2004: 225).

In 24 languages, listed in (68), the comitative marker *na* itself functions as the verbal element taking (locative) subject marking.

- | | | | |
|------|--------------|------------|------------------------------------|
| (68) | Benga A34 | <i>na</i> | (Nassau 1892) |
| | Digo E73 | <i>na</i> | (Nicolle 2013) |
| | Pare G22 | <i>na</i> | (Mous & Mreta 2004) |
| | Shambaa G23 | <i>na</i> | (Besha 1989) |
| | Ng’hwele G32 | <i>na</i> | (Legère 2010) |
| | Luguru G35 | <i>na</i> | (Mkude 1974) |
| | Vidunda G38 | <i>na</i> | (Legère 2010) |
| | Swahili G42d | <i>na</i> | (Marten 2013) |
| | Bena G63 | <i>na</i> | (Morrison 2011) |
| | Kizu JE402 | <i>na</i> | (Gray 2013) |
| | Ishenyi JE45 | <i>na</i> | (Bernander & Laine 2020) |
| | Sena N44 | <i>na</i> | (Grégoire 1975) |
| | Yao P21 | <i>na</i> | (Whiteley 1966) |
| | Kwanyama R21 | <i>na</i> | (Halme 2004) |
| | Ndonga R22 | <i>na</i> | (Fivaz 1984) |
| | Herero R30 | <i>na</i> | (Möhlig & Kavari 2008) |
| | Venda S21 | <i>na</i> | (Ziervogel et al. 1972) |
| | Kagulu G12 | <i>ina</i> | (Petzell 2008) |
| | Shona S10 | <i>ne</i> | (Grégoire 1975) |
| | Manyika S13 | <i>ne</i> | (Stevick & Machiwana 1960) |
| | Nrebele S407 | <i>ne</i> | (Grégoire 1975) |
| | Swati S43 | <i>ne</i> | (Marten 2010) |
| | Tsonga S53 | <i>ni</i> | (Sozinho Francisco Matsinhe, p.c.) |
| | Ronga S54 | <i>ni</i> | (Dimande 2020) |

In Kagulu G12, the comitative marker is preceded by the vowel *i* which could be a trace of **di* ‘be’ or an epenthetic vowel inserted to avoid a monosyllabic stem. In seven languages in (68), the inflected comitative marker has a deviant vowel, i.e. either *e* (four languages) or *i* (three languages), for which possible explanations have already been suggested above. Further conceivable origins for

a vowel other than *a* are an added inflectional final vowel suffix and merger with an additional comitative marker *le* after intervocalic consonant loss. This brings us to the special type of comitative copula, exemplified in (69), in which an inflected form of (*i*)*na* itself is followed by a comitative marker.

- (69) Zombo H16hK *ina ye* (Araújo 2013)
 Yaka H31 *ina ye* (Bwendelele s.d.)
 Mbala H41 *ina i* (Moyo-Kayita 1981)
 Lozi K21 *nani* (O’Sullivan 1993)
 Tswana S31 *na le* (Cole 1955)
 S. Sotho S33 *na le* (Salzmann 2004)
 Tsonga S53 *na ni* (Sozinho Francisco Matsinhe, p.c.)

As mentioned in §2.1 and §2.2, some inflected comitative markers with or without a trace of **di* have acquired the sense ‘be’ and are used in PLs (70a). In order to be used as a possessive (70b, 71a, 72a) or an EL predicator (70c, 71b, 72b, 73a), they must be combined with an additional comitative marker. In the zone S languages in (69), the semantic shift from ‘be with’ to ‘be’ is less clear as we do not have evidence for the use of *na* in PLs (73b).

- (70) Zombo H16hK (Araújo 2013: 148, 198, 190)
- a. *a-ntu mu-nzó ena*
 2-person 18-9.house SM2.COP
 ‘*as pessoas estão em casa*’ [‘the people are at home’]
 - b. *á-kentó ena yé a-ngúdí a-wu*
 2-woman SM2.COP COM 2-mother 2-POSS
 ‘*as mulheres estão com as mães*’ [‘the women are with their mothers’]
 - c. *vèná yè ndíngà záyingí mù-Angola*
 SM16.COP with 10.language 10.CONN.many 18-Angola
 ‘*tem muitas línguas em Angola*’ [‘there are many languages in Angola’]
- (71) Mbala H41 (Moyo-Kayita 1981: 71)
- a. *wéna i ngangu*
 SM1.COP COM 9.intelligence
 ‘*il est intelligent*’ [‘he is intelligent < he has/is with intelligence’]
 - b. *há-mu-dú hena i mu-lédi*
 16-3-head SM16.COP COM 3-garment
 ‘*sur la tête il y a un habit*’ [‘on the head there is a garment’]

(72) Tswana S31 (Cole 1955: 330, 331)

- a. *ke-na le=bana ba-le ba-bêdi*
 SM1SG-COP COM=2.child 2-DEM 2-two
 ‘I have two children’
- b. *go-na le=ba-tho*
 SM17-COP COM=2-person
 ‘there are some people’

(73) S. Sotho S33 (Salzmann 2004: 26 for (73a), Schoeneborn 2009: 58 for (73b))

- a. *mo-tse-ng há-Masúpha hó-na-lé=libetsa*
 3-village-LOC 17.CONN-Masupha SM17-COP-COM=8.firearms
 ‘at Masupha’s village there are firearms’
- b. *ke jarete-ng*
 PRON1SG garden-LOC
 ‘I am in the garden’

In sum, two types of comitative copula can be distinguished in Bantu ELs. First, there is the full form consisting of a locative copula, usually a reflex of **dì*, followed by a comitative marker, habitually a reflex of **nà*. The full form has eroded in many languages resulting in a second type consisting of the comitative marker inflected for subject. The inflected comitative marker has undergone a semantic shift from ‘be with’ to ‘be’ in some languages giving rise to a subtype of the first type of comitative copula whereby inflected *na* itself is followed by an invariable comitative marker.

3.1.3 Variations on the comitative copula

Variations on the comitative copula type include transitive have-verbs, specialised EL verbs relatable to the comitative copula and polysemous copula formally relatable to the comitative copula.

As noted in §2.2, Bantu languages typically make use of a comitative copula in possessive constructions. Some languages (also) have a transitive have/hold-verb in ELs, for instance the four eastern Bantu languages spoken in Kenya and Tanzania (74). In Vunjo-Chaga (75b) and Rangì (33b) (in §2.2), ELs may also select the more regular comitative copula in ELs. We do not have enough data to ascertain whether this choice is also available in Gweno E65 and Taita E74.

- (74) Vunjo-Chaga E622C *wozre* ‘have, hold’ (Moshi 1995: 131)
 Gweno E65 *yír* ‘have’ (Philipson & Nurse 2000: 29–30)
 Taita E74 *erekogh* ‘get+IPFV.PASS’? (Grégoire 1975: 67)
 Rangi F33 *tetε/tiite* ‘have’ (Dunham 2005; Stegen 2011: 165)
- (75) Vunjo-Chaga E622C (Salzmann 2004: 46 for (75a), Grégoire 1975: 57 for (75b))
- a. *numbe-nyi ko Ohanyi ku-wozre singi ya ki-leghe*
 9.house-LOC 17.CONN John SM17-have 9.nest 9.CONN 7-bird
 ‘On John’s house is a bird nest’
- b. *ku-lja Tšomba kw-i na ndža*
 17-DEM Tshomba SM17-COP COM 9.hunger
 ‘*au Tshomba il y a la famine*’ [‘at Tshomba there is famine’]

Mbundu H21 uses invariable *sai* in possessive constructions and ELs. Possessive constructions can also make use of a comitative copula consisting of the locative copula *ala* ‘be’ (from *(j)ikad) followed by comitative *ni*, whereas ELs may also select the locative copula with a locative subject concord referring to the ground.

- (76) Mbundu H21 (Chatelain 1888–99: 12)
- a. *eye sai jingombe*
 PRON.2SG have 10.cattle
 ‘*tu tens gado*’ [‘you have cattle’]
- b. *sai jisanji*
 have 10.chicken
 ‘*há galinhas*’ [‘there are chickens’]

Xhosa and Zulu make use of the specialised EL verb *khona* (77).

- (77) Xhosa S41 (Bloom Ström 2020: 226)
- kú-khóna úm-phánda om-khúlu ke phaya é:ntla*
 SM17-be_present 3-barrel 3-big then there inside
 ‘there is a big barrel there inside’

As argued by Bloom Ström (2020: 219–220), *khona* may be a merger between a class 17 locative marker *kho-* and the inflected comitative marker *na*, originally expressing something like ‘there be with’ (but see Louw & Jubase 1963: 123, du Plessis & Visser 1992: 239 for a different analysis).

In five sample languages the verbal element in ELs is polysemous between ‘be’ and ‘be with/have’. This polysemy likely reflects an ongoing semantic shift from ‘be with, have’ to ‘be’. Whereas in some Bantu languages this shift has been accomplished (see §3.1.1), it is ongoing in the five languages in (78), which all show some traces of the original possessive/comitative meaning.

- (78) Tsootso H16hZ *ina* ‘be with’ (Baka 1992)
 Suku H32 *ina* ‘be with’ (Piper 1977: 380–381)
 Totela K41 *ina* ‘be with’ (Crane 2011: 34)
 Yeyi R41 *na* ‘be with’ (Araújo 2013)
 Nzadi B865 *māṅ* ‘have, be’ (Crane et al. 2011)

In Tsootso H16hZ, Suku H32, and Totela K41 *ina* is used in PLs (79a) ((34a) in §2.2), which suggests that the shift to ‘be’ has been accomplished. Moreover, Suku possessive constructions require the use of the comitative marker *ye*. However, traces of the original comitative meaning are attested in Suku ELs (79b) and Tsootso and Totela possessive constructions (80a) and (34a). The Suku EL is exceptional in that the figure rather than the ground displays locative marking (79b). Our hypothesis is that the class 18 locative marker attaches to the verbal element, as we think is the case in Tsootso (80b), rather than to the figure and that the sentence can be translated as ‘the iron has/is with inside the hammer’. The optionality of the comitative marker in Totela possessive constructions suggests that the comitative meaning persists in some contexts (34b–34c). The Tsootso possessive construction (80a) appears to be of the ‘genitive possessive’ type (Stassen 2013) expressing something like ‘the person how many necks are his?’, in which case *ina* would unambiguously express ‘be’. However, it takes a subject marker referring to the possessor (‘the person’) rather than to the possessee (‘how many necks’) implying the translation ‘the person how many necks he has his?’. In Tsootso, Suku, and Totela *ina* thus mainly expresses ‘be’, but in some particularities in use the original comitative meaning persists.

- (79) Suku H32 (Piper 1977: 381)
- a. *yéna ha-máamba*
 SM1SG.COP 16-6.water
 ‘*ich bin am Wasser*’ [‘I am at the water’]
- b. *ki-séngú kyéná mu-nzúundu*
 7-iron SM7.COP/COM 18-9.hammer
 ‘*auf Eisen ist der Hammer*’ [‘on the iron is the hammer’]

- c. `pfúmú kèna ya bahika
 1a.chief SM1.COP COM 2-slave
 ‘der Häuptling hat Sklaven’ [‘the chief has slaves’]

(80) Tsootso H16hZ (Baka 1992: 87)

- a. è-mù:-nthù nsí:ngú kwá kéná záu
 AUG-1-person 10.neck how_many SM1.COP/COM 10.poss2
 ‘combien de cous l’homme a-t-il ?’ [‘how many necks does the person have’]
- b. mù-tótóphóló wú-ná mò mwà-wóóso
 3-ashes SM3-COP LOC18 18.CONN-all
 ‘il y a du cendre partout’ [‘there are ashes everywhere’]

Nzadi B865 is like Totela (34b–34c) in that the persistence of the possessive meaning is reflected by the optionality of the comitative marker in possessive constructions (81a).

(81) Nzadi B865 (Crane et al. 2011: 145, 240, 210)

- a. mi a máŋ (yε) bään
 PRON.1SG PRS COP/COM COM children
 ‘I have children’
- b. mwään a máŋ kó ndzɔ
 child PRS COP/COM in house
 ‘there is a child in the house’
- c. a máŋ kó ndzɔ
 PRS COP in house
 ‘he is in the house’

Data from Yeyi R41 suggest that na is fully polysemous in this language. It is used in PLs, possessive constructions and ELs alike.

(82) Yeyi R41 (Seidel 2008: 421, 423, 422)

- a. ka-na=po
 SM12-COP=LOC16
 ‘it (the axe) is (over) there’
- b. mu-ti wu-na ma-papa
 3-tree SM3-COM 6-leaf
 ‘the/a tree has leaves’

- c. *mu-na* *u-ndavu mu-mu-tara*
 SM18-COP/COM 1a-lion 18-3-courtyard
 ‘there is a lion in the courtyard’

3.2 Agreement patterns

One hundred and eleven sample languages display locative subject marking in ELs, which is clearly predominant when the copula is comitative. Locative copulas allow for more variation in agreement. Below we first look at locative and related expletive subject markers in ELs (§3.2.1) before turning to agreement with an inverted or non-inverted figure (§3.2.2). Many languages with non-inverted ELs have (severely) reduced agreement systems. The verbal element is often exempt of agreement markers.

3.2.1 Locative and expletive subject markers

As Grégoire (1975; 1983; 2003) points out, most forest Bantu languages (zone A, B10-70, C10-70 & D10-40) do not have agreement triggering locative classes, except for southern zone D (i.e. Mituku D13, Lega D25, South Binja D26, Holoholo D28, Nyanga D43 and Buyu D55) (Grégoire 2003: 358). Nonetheless, several of them do have ELs with locative subject marking. The class 17 subject markers in Kpe A22, Benga A34, Ewondo A72a, Teke Tyee B73d and Ombo C76 and the class 16 subject marker in Bulu A74a and Babole C101 are traces of a former locative system, as these languages only have locative prepositions (Grégoire 1975; 1983). Their synchronic use is expletive and not referential.

- (83) Kpe A22 *o* (Tanda & Neba 2005: 210)
 Benga A34 *o* (Nassau 1892)
 Ewondo A72a *o* (Grégoire 1975: 123)
 Ombo C76 *kv* (Meeussen 1952: 31)
 Bulu A74a *a* (Grégoire 1975: 123)
 Babole C101 *ha* (Leitch 2003)

Grégoire (2003: 359) notes that Tsogo B31 has two locative nouns *gòmá* (class 17) and *vòmá* (class 16) ‘place’ of which the second one can determine agreement. Tsogo ELs show that grounds of both class 17 (84a) and class 16 (84b) can determine agreement on the verb.

(84) Tsogo B31 (Marchal-Nasse 1979: 51)

- a. *go-sá-ba pógó go mó-dono*
 SM17-NEG-COP 9.rat 17 3-roof
 ‘*il n’y a pas de rat sur le toit*’ [‘there is no rat on the roof’]
- b. *va-sí-báká mó-yakó vané*
 16-NEG-COP.PST 3-food 16.DEM
 ‘*il n’y avait pas de nourriture là*’ [‘there was no food there’]

In Nzebi B52, which has locative prepositions clearly relatable to PB **pa-* (16), **kv-* (17) and **mɔ-* (18), ELs exclusively use the class 17 subject marker (85a–85c). One presentational construction shows the expletive use of a class 16 subject marker (85d).

(85) Nzebi B52 (Marchal-Nasse 1989: 530)

- a. *vaanâvá gu-líidi baatə bá-kúnu*
 here.16 SM17-COP.PRF 2.person 2-many
 ‘*ici, il y a beaucoup de gens*’ [‘here there are a lot of people’]
- b. *gú tsó nzeli gu-líidi bá-tfwí bá-kunu*
 17 inside 9.river SM17-COP.PRF 2-fish 2-many
 ‘*dans l’eau il y a beaucoup de poisons*’ [‘in the river there are a lot of fish’]
- c. *mu yul’ á maambə gu-líidi ma-mbúngu mósólo*
 18 9.top 9.CONN 6.water SM17-COP.prf 6-canoe 2.two
 ‘*sur l’eau il y a deux pirogues*’ [‘on the water there are two canoes’]
- d. *va-líidi lə-sógá lə-kíma ló ...*
 SM16-COP.PRF 11-way 11-other REL.11
 ‘*y a-t-il un autre moyen ...*’ [‘is there another way that ...’]

Southern Bantu languages of zone S also did not retain the PB locative nominal prefixes, except with some inherently locative nouns (Grégoire 1975; Marten 2010). Locative agreement is heavily reduced and typically selects the class 17 prefix (Grégoire 1975). Except for Shona S10, all zone S languages in our sample have class 17 subject marking in ELs, as illustrated in (86) with Ronga S54.

(86) Ronga S54 (Dimande 2020: 112)

- henhla ka n-sinya ku-ni nyoka*
 16.TOP 17.CONN 3-tree SM17-COM 9.snake
 ‘*em cima da árvore há cobra*’ [‘on top of the tree there is a snake’]

Except for these zone A, B, C and S languages, we find that the locative subject marker is mainly used referentially, i.e. agreeing with the locative class of the ground, as described in §2.3, in particular (36). Some exceptions do occur ranging from the loss of class 18 agreement in Kamba E55, Vunjo-Chaga E622C, Ishenyi JE45 and Tanzanian Ngoni N12, to agreement merger in favour of class 16 in Rwanda JD61 and Rundi JD62, and class 17 in Lozi K21 and Kwangali K33. Subject agreement with the ground is sometimes possible, but not obligatory, as in Swahili (38) (cf. §2.3). Similarly, the Tonga M64 expressions of generic existence show that the class of the locative subject marker may change depending on the semantics of the implicit ground. However, the IL in (87c) shows a mismatch between the locative subject marker and the locative class of the ground.

(87) Tonga M64 (Collins 1962: 110)

- a. *ku-li uu-zya*
SM17-COP SM1.REL-come
'there is someone coming'
- b. *mu-li uu-yimba*
SM18-COP SM1.REL-sing
'there is someone inside singing'
- c. *ku-li nhombe zyosanwe mu-zi-bili mu-muunda*
SM17-COP 10.cow 10.five 18-10-two 18-3.field
'there are seven cows in the field'

This might point towards an ongoing change favouring the expletive use of one of the locative classes in ELs, which would be in line with the cross-linguistic tendency for ELs to be non-referential (Koch 2012; Creissels 2019a). Some ELs have non-locative expletive subject markers. They mainly occur in forest Bantu languages (Kwakum A91, Orungu B11b, Kota B25, Ngungwel B72a, Tiene B81, Bongili C15, Mboshi C25, Doko C301, Bangi C32, Bolia C35b, Linga C502, Gesogo C53, Ntomba C61J, Nkucu C73), which lack locative classes and agreement. Unlike Tsogo (84) and Ombo (83), these forest languages do not display traces of locative agreement in the subject marker slot. Instead, they use an invariable subject marker of a non-locative class, as shown in (40) and (88).

(88) Doko C301 (Twilingiyimana 1984: 131)

- ánê, é-dí n' òmòtò*
here SM5/7?_EXPL-COP COM 1.person
'*ici, il y a une personne*' ['here, there is a person']

“Double” agreement marking which combines a locative or expletive subject marker with a locative enclitic appears to be a unique feature of interlacustrine Bantu languages.¹³ Apart from zone J and Sumbwa F23, it is only attested in Shangaji and possibly also in Beo (6b). Shangaji has several ELs, of which the most frequently used ones are of the agreeing-inversion type (8b–8c, 14). They always include a locative enclitic referring to the ground. As will become clear in §3.2.2, the presence of a locative proform is a recurrent characteristic of agreeing-inversion type ELs. The presence of a locative enclitic in (89) could therefore be attributed to analogy with the more frequently occurring existential construction in which the subject marker agrees with the figure. Note that the locative enclitic does not attach to the comitative copula but rather to the figure.

- (89) Shangaji P312 (Maud Devos, field notes)
okhúúle o-na n-názií=wo na n-ráráanja
17.DEMIII SM17-COM 3-coconut_tree=LOC17 and 3-orange_tree
‘over there is a coconut tree and an orange tree’

Otherwise, double agreement including a locative or expletive subject marker seems an innovation of Great Lakes Bantu. Some languages, such as Soga JE16 and Tsofso JE32b in (90), have redundant double agreement: both the subject marker and the locative enclitic are referential with the ground.

- (90) Tsofso JE32b (Dalgish 1976: 141)
xu-mu-saala xu-li-xwo aBa-saatsa
17-3-tree SM17-COP-LOC17 2-man
‘on the tree are the men’

In other languages, such as Sumbwa in (91), Rwanda, Rundi, and Nkore, subject agreement is restricted to a single locative class (typically class 16), whereas the locative enclitic is referential with the ground and can secure the semantics of a nominal ground in its absence.

- (91) Sumbwa F23 (Grégoire 1975: 50)
mu-numba ha-ta-li=mo shi-ntu
18-9.house SM16-NEG-COP=LOC18 7-thing
‘*dans la maison, il n’y a rien*’ [‘in the house there is nothing’]

¹³Sumbwa F23 shares several features with zone J languages, which is either due to contact or suggests that genealogically speaking Sumbwa rather belongs to zone J (Bastin 2003: 521).

In still other languages, such as Ganda in (92), Kizu in (93) and Ishenyi in (94), the situation is less straightforward as both the subject marker and the locative enclitic display restricted locative agreement, but merger in locative class agreement appears to happen at different paces in both positions. In Ganda, merger in locative class agreement is more advanced in the subject marker slot than in the enclitic slot. A class 16 subject marker is often selected but classes 17 and 23/25 occur sporadically. The locative enclitic shows regular agreement with the ground but in the case of a class 17 or 18 nominal ground mismatches do occur, leading to configurations whereby neither the subject marker nor the enclitic are referential with the ground (92). Similar cases occur in Kizu and Ishenyi.

(92) Ganda JE15 (Nanteza 2018: 36)

wa-li=yo a-ba-ana ba-na mu ki-zimbe
 SM16-COP=LOC23 aug-2-child 2-four 18 7-building
 ‘there are four children in the building’

(93) Kizu JE402 (Gray 2013: 44)

mu-char# mu-yo kw-a-re=ho m#-kari w#mwi
 18-7.village 18-DEMII SM17-PST-COP=LOC16 1-woman 1.one
 ‘in that village, there was a certain woman’

(94) Ishenyi JE45 (Bernander & Laine 2020: 76)

ŋ-ko-réŋge=hó eyi-tábo mu-méétʃa
 FOC-SM17-PST.COP=LOC16 7-book 18-table
 ‘there is a book on the table’

We also find double agreement involving a non-locative expletive subject marker in Great Lakes Bantu languages, such as Haya in (95) and Kerebe JE24. The invariable class 1 subject marker *a-* is in these languages accompanied by a locative enclitic referential with the ground.¹⁴ In Kerebe, there is a choice between agreeing and expletive inversion (Thornell 2004).

(95) Haya JE22 (Grégoire 1975: 77)

a-ha-iguru a-li=ho enyanyinyi
 AUG-16-9.sky SM1_{EXPL}-COP=LOC16 AUG.10.star
 ‘*au ciel, il y a des étoiles*’ [‘in the sky, there are stars’]

¹⁴Grégoire (1983: 152) suggests that an expletive subject marker of class 16 became reanalysed as a class 1 subject marker in some zone A languages, because of their formal similarity. It is unlikely that a similar process took place in Haya and Kerebe as they have class 16 locative prefixes of the shape *ha-*.

3.2.2 Agreement with the figure

In this section we take a closer look at ELs in which the subject marker agrees with the figure. A few counterexamples notwithstanding, this agreement pattern is restricted to ELs selecting a locative copula. This suggests that locative or related expletive agreement is a fundamental characteristic of ELs with a comitative copula. The ground and the figure function as the possessor and the possessee, respectively, and the verbal element agrees with the ground or takes an expletive subject marker.

3.2.2.1 Agreement with inverted figure

ELs with agreeing inversion occur less frequently and are less widespread than ELs with locative or expletive inversion. They are largely restricted to eastern Bantu. In some languages the locative copula agrees only with the figure and the construction does not include a locative proform. In Matengo N13, this appears to be the only way of expressing existential location. Manda has two types of ELs, a non-dedicated one characterised by agreeing inversion and the absence of a locative proform, and a dedicated one involving locative inversion.

Still, most languages displaying agreeing inversion do include a locative proform in ELs. With respect to PLs this locative proform may be non-dedicated (i.e. obligatory in ELs and PLs alike), conventionalised (i.e. obligatory in ELs and optional in PLs) or dedicated (i.e. obligatory in ELs and absent in PLs). The Mbukushu K333 example in (96a) illustrates the inclusion of a non-dedicated (pre-initial) locative marker in ELs. As seen in (96b), it is also present in PLs. In Nyakyusa, a locative enclitic is required in ELs (97a) but optional in PLs (97b–97c). Dedicated locative enclitics are found, among others, in a number of interlacustrine languages. Kerebe ELs combine a dedicated locative enclitic (98a–98c) with either agreeing or expletive inversion (98a–98b).

(96) Mbukushu K333 (Fisch 1977; 1998: 118)

- a. *mu-vinyu mo ghu di ghu-semwa*
18-wine LOC18 SM14 COP 14-truth
'in wine, there is truth'
- b. *ha-nuke po ha di pa-mbongi*
2-child LOC17 SM2 COP 16-9.mission
'the children are at the mission'

(97) Nyakyusa M31 (Persohn 2017: 310, 307, 315)

- a. *n-k-iis̩ kɪ-mo, a-a-li=ko* *ɔ-malafyale j̩-mo*
 18-7-land 7-one SM1-PST-COP=LOC17 AUG-1.chief 1-one
 ‘in some land, there was a chief’
- b. *ɔ-mw-ana a-li mu-m-piki*
 AUG-1-CHILD 1-COP 18-3-tree
 ‘the child is in a/the tree’
- c. *a-li=mo n-nyumba*
 SM1-COP=LOC18 18-9.house
 ‘(s)he is in the house’

(98) Kerebe JE24 (Thornell 2004: 25)

- a. *βa-li-ho: a-βa-ntu*
 SM2-COP-LOC16 AUG-2-person
 ‘there are people’
- b. *a-li-ho: a-βa-ntu*
 SM1-COP-LOC16 AUG-2-person
 ‘there are people’
- c. *a-n-te zi-li mu ki-βuga*
 2-10-cow SM10-COP LOC18 7-shed
 ‘the cows are in the cow shed’

Nyamwezi ELs regularly take a locative copula displaying double agreement: once with the figure through the subject marker and once with the ground through an obligatory locative enclitic (2b). We found one example where the locative copula combines with a comitative marker thus apparently constituting a comitative copula exceptionally agreeing with the figure rather than taking a locative subject marker. It could be that the comitative marker has a different function here. In Nyakyusa, we also found examples of a locative copula seemingly combining with a comitative marker in ELs characterised by agreeing inversion (100). As it turns out, the comitative marker is used as an additive focus marker, expressing ‘also, too’. For a similar use of the comitative marker in Pare, see Mous & Mreta (2004: 221). Maybe the Nyamwezi example in (99) likewise expresses that there are also snakes inside of the beehive, but this is not reflected in the (free) translation.

- (99) Nyamwezi F22 (Maganga & Schadeberg 1992: 218, 222)
nshikú zíingí gúsBaági ga-lí=mó ná=ma-yoká
 10.day 10.many SM6.COP.HAB SM6-COP=LOC18 COM=6-snake
 ‘frequently, there are snakes inside’
- (100) Nyakyusa M31 (Persohn 2017: 316)
ky-a-li=po n=I-kI-piki
 SM7-PST-COP=LOC16 COM=AUG-7-stump
 ‘there was also a wood’

ELs with agreeing inversion probably are an innovation motivated by a dis-preference for locative subject marking rather than by a loss of it. Reference to the ground tends to be demoted to the post-final slot. In languages where ELs with locative inversion and ELs with agreeing inversion co-occur, the latter could involve a usage extension of the presentational construction, which often displays a preference for agreeing inversion. More fine-grained data are needed to confirm this hypothesis.

3.2.2.2 Agreement with non-inverted figure

ELs without figure inversion occur in 23 NWB and CWB languages and in seven scattered languages spoken elsewhere. As mentioned in §2.1, they are of two types: (i) “radical generic location” languages (Koch 2012) like Liko (3) and Lingala (5) with complete syntactic identity between ELs and PLs and thus ambiguous readings; and (ii) languages like Mbuun (21) allowing non-topical or even focal constituents in preverbal position. In the absence of information-structural analyses, the distinction is not always an easy one to make. Languages for which we have good indications that the preverbal, non-inverted position of the figure is due to a non-canonical word order include Mbuun (Bostoen & Mundeke 2012) (21), Mbugwe F34 (Vera Wilhelmsen, p.c.), Zombo (Araújo 2013) and Western Serengeti languages (Nicolle 2015; Aunio et al. 2019; Bernander & Laine 2020) (22). In Mbugwe and Zombo, ELs and PLs are not syntactically identical. Whereas the figure is preverbal in ELs, as in (101a) and (102a), the ground is preverbal in PLs, as in (101b) and (102b), suggesting that non-topical/focal constituents occur in preverbal position.

- (101) Mbugwe F34 (Vera Wilhelmsen, p.c.)
 a. *kaái ve-ényi vá-re=kó*
 9.house 2-guest SM2-COP=LOC17
 ‘there are guests at home’

- b. *Ally geri á-re*
 Ally 9.car SM1-COP
 ‘Ally is in the car’

(102) Zombo H16hK (Araújo 2013: 164, 148)

- a. *mùnà dínà kàfi sukádi zénà mó*
 18.DEM 5.DEM 5.coffee 10.sugar SM10.COP LOC18
 ‘*naquele café tem açúcar*’ [‘in that coffee there is sugar’]
- b. *à-ntù mù-nzó ènà*
 2-person 18-9.house SM2.COP
 ‘*as pessoas estão em casa*’ [‘the people are in the house’]

In all these languages, the ground precedes the figure, itself preceding the verbal element, which may or may not have a locative proform added to it. Similar word orders are attested in six other languages with non-inverted ELs: Bakoko A43b, Mmala A62B, Gyeli A801, Leke C14, Tetela C71, Budu D332. In Bakoko, Gyeli, Tetela and Budu both Ground-Figure-Copula-[Locative] (103a) and Figure-Copula-Ground (103b) word orders are possible. For Mmala and Leke (104a), we only have examples with a sentence-initial ground. Still, the PLs do not display non-canonical word orders (103c, 104b). It thus remains unclear whether these languages are of the Mbuun- or the Lingala-type.

(103) Gyeli A801 (Nadine Grimm, p.c.)

- a. *kwádò dé tù m-ùdâ m-vúdú¹⁵ nùù*
 7.village LOC inside 1-woman 1-one 1.COP
 ‘in the village there is a woman’
- b. *m-ùdâ m-vúdú àà kwádò dé tù*
 1-woman 1-one 1.COP 7.village LOC inside
 ‘there is a woman in the village’
- c. *Ada àà ndáwò dé tù*
 Ada 1.COP 9.house LOC inside
 ‘Ada is in the house’

¹⁵Note that the numeral ‘one’ marks indefiniteness in this context (Nadine Grimm, p.c.).

(104) Leke C14 (Vanhoudt 1987: 131)

- a. *wó-ndákwe éde móryō ð-zi*
 17-9.house 9.DEM 3.fire SM3-COP
 ‘dans cette maison il y a du feu’ [‘in that house there is fire’]
- b. *ma-mvã ã-zi mba*
 6-dog SM6-COP where
 ‘où sont les chiens?’ [‘where are the dogs?’]

The remaining languages are of the Lingala “radical generic location” type (see also Liko in (3a). Their ELs and PLs are morphosyntactically identical. In Nyokon A45, this clearly correlates with a fixed word order. Nyokon may select a locative (105a) or a comitative copula (105b) in ELs. In both cases the figure is preverbal and the ground follows the copula. The copula in ELs is the same as in PLs (105c).

(105) Nyokon A45 (Mous 2005: 7, 8; Barreteau s.d.)

- a. *átán nà kīnōŋ*
 6.stones COP 7.road
 ‘there are stones on the road’
- b. *mánóŋ nà àŋgá nyá níkùŋ*
 6.blood COP COM POSS.2SG 5.spear
 ‘there is blood on your spear < blood is with your spear’
- c. *ù nà mi:mi nà ùkùs*
 PRON.3SG COP near COM 3.fire
 ‘il est près du feu’ [‘he is close to the fire’]

4 Non-inverted existential constructions: archaism or innovation?

Bantu existential constructions overwhelmingly display figure inversion with non-inverted constructions being largely restricted to northern Bantu borderland languages. In historical terms, this allows for at least two hypotheses.

First, seeing that non-inverted ELs are (i) cross-linguistically rare (Creissels 2013; 2015; 2019a,c), and (ii) within the Bantu domain mainly found in the area closest to the Bantu homeland, more specifically in languages belonging to the NWB and CWB branches, an obvious inference would be that PB only had non-dedicated, non-inverted ELs and that the cross-linguistically more common in-

verted ELs were innovated after these first branches split off. Interpreting non-inverted ELs as an archaism questions the PB reconstruction of “anastasis” (“*renversement*”) or subject inversion (Meeussen 1959: 215; 1967: 120). Recent studies on subject inversion claim that there is an implicational hierarchy following which there is no inversion with full lexical verbs in a language without inversion with copula (Marten & van der Wal 2014: 59). If PB did not have (locative, expletive or agreeing) inversion in ELs then it most probably did not have it in other constructions either. §4.1 further considers the hypothesis of non-inverted ELs being a PB feature.

Second, if we assume that PB had ELs with figure inversion then we need to account for the non-inverted constructions in the NWB and CWB languages. They could be interpreted as an areal feature. As suggested by Creissels (2019a), and also taken up by Güldemann (2018), exactly this type of non-dedicated and non-permuted existential construction might well be one of the defining features of a linguistic area known as the “Sudanic Belt” (Clements & Rialland 2008) or “Macro-Sudan Belt” (Güldemann 2008). Following Güldemann (2008: 152), the Macro-Sudan Belt covers an area in Northern sub-Saharan Africa “sandwiched between the Atlantic Ocean and the Congo Basin in the south and the Sahara and Sahel in the north, and spans the continent from the Atlantic Ocean in the west to the escarpment of the Ethiopian Plateau in the East”. Some features of peripheral northern Bantu have non-Bantu donors belonging to the Macro-Sudan Belt (Güldemann 2018: 456). Although some such shared features, such as base-4 numeral systems, seem confined to the eastern parts of the northern Bantu borderland (see Hammarström 2010), several other features such as labial-velar stops and cross-height ATR vowel harmony have affected languages “from the Atlantic in the west to Lake Albert in the east” (Clements & Rialland 2008: 43). The question now is whether non-inverted ELs can likewise be the result of areal diffusion of a Sudanic Belt feature. In §4.2 we take a closer look at the pros and cons of the areal innovation hypothesis.

4.1 Non-inverted ELs as an archaic feature

As was mentioned in §3.2.2.2, non-inverted ELs are mainly found in NWB and CWB languages. Table 4 shows the distribution of non-inverted ELs over the different phylogenetic groups of Grollemund et al. (2015).

Twenty-four languages with non-inverted ELs belong to the NWB and CWB branches. The remaining nine languages are scattered across the other branches. Their non-inverted ELs could be interpreted as cases of archaic persistence but it seems that at least four of them, i.e. Mbuun (21), Zombo (102), Mbugwe (101)

Table 4: Phylogenetic distribution of non-inverted ELs

NWB	Basaa A43a, Bakoko A43b, Nen A44, Nyokon A45, Kpa A53, Mmala A62B, Gunu A622, Eton A71, Gyeli A801, Koonzime A842, Kwakum A91
CWB	Aka C104, Leke C14, Lingala C30B, Bomboma C411, Kele C55, Mongo C61, Tetela C71, Kela C75, Ndengese C81, Liko D201, Bila D311, Bira D32, Budu D332
WWB	Nzadi B865, Mbuun B87, Zombo H16hK, Tsootso H16hZ
SWB	Mbukushu K333
EB	Mbugwe F34, Ngoreme JE401, Nata JE45, Mwera P22

and the Western Serengeti languages JE45, have non-canonical information structural characteristics which allow or even require (Mbugwe) the non-topical/focused figure to occur in preverbal position. Moreover, the focalisation of the figure often triggers the ground to move to clause-initial position. The Nata JE45 EL in (106) has a special word order with the ground preceding the figure, itself preceding the copula. It also includes a dedicated locative proform. The non-inverted ELs in these languages can thus be attributed to language-specific characteristics, which at least in the Western Serengeti languages could have been triggered by language contact.

(106) Nata JE45 (Bernander & Laine 2020: 61, 71)

- a. *mo-mo-súko e-βi-yérɔ m-be-epi=mú*
 18-3-bag AUG-8-thing FOC-SM8-PRS.COP=LOC18
 ‘there is a thing in the bag’
- b. *a-βá-áto βá-áru m-ba-apí mw-i-sókɔ*
 AUG-2-person 2-many FOC-SM2-PRS.COP 18-5-market
 ‘many people are at the market’

Mbukushu (107) and Mwera (41), have non-inverted ELs featuring a pre-initial locative marker. Whereas in Mwera the pre-initial locative marker is obligatory

in ELs and optional in PLs (108a–108b), Mbukushu displays the opposite pattern with a dedicated locative marker in PLs (96b) and an optional one in ELs. Tsootso includes a locative enclitic in ELs (80b) which is not present in PLs.

(107) Mbukushu K333 (Fisch 1998: 119)
ha-genda (ko) ha di ku-di-ghumbo
 2-guest LOC17 SM2 cop 17-5-village
 ‘there are guests in the village’

(108) Mwera P22 (Harries 1950: 114, 115)

a. *nguku i-li n-nyumba*
 9.chicken SM9-COP 18-9.house
 ‘the chicken is in the house’

b. *mu-tu-li muno*
 LOC18-SM1PL-COP 18.DEM
 ‘we are in here’

Nzadi, finally, appears to be a radical generic location language and should thus be interpreted as a case of archaic persistence in light of the present hypothesis.

It should be noted that most languages, except Nzadi, Tsootso and Mbugwe, also have inverted ELs. Mwera (109) and Mbuun (110), for example, have alternative ELs characterised by locative and agreeing inversion, respectively.

(109) Mwera P22 (Harries 1950: 115)
mu-li wa-ndu amula
 SM18-COP 2-person 18.DEM
 ‘there are people inside’

(110) Mbuun B87 (Léon Mundeke, p.c.)
wó ó-nkáán ká-nyéng a méés
 COP(=PP3) 3-book LOC-on_top CONN 6.table
 ‘there is a book on the table’

The NWB and CWB languages mostly do not show intralingual variation although we must admit that data are often limited. Bila D311 and Bira D32, however, do have alternative inverted constructions, both involving agreeing inversion. Kwakum has an alternative EL involving expletive inversion. Moreover, for

Table 5: Inverted ELs in NWB and CWB

	locative inversion	expletive inversion	agreeing inversion
NWB	Kpe A22, Benga A34, Ewondo A72a, Bulu A74a Tsogo B31	Kwakum A91 Orungu B11b, Kota B25	
CWB	Babole C101, Ombo C76	Bongili C15, Mboshi C25, Doko C301, Bangi C32, Linga C502, Gesogo C53, Ntomba C61J, Nkucu C73	Bila D311, Bira D32

a number of NWB and CWB languages in our database we only have ELs characterised by figure inversion and a (locative) expletive subject marker. Table 5 categorises all the languages involved.

If we consider the non-inverted ELs as archaic, then the inverted ones should be interpreted as innovations, which is not unlikely seeing that inverted ELs are much more common cross-linguistically. However, when taking a closer look at the inverted constructions in question, they rather seem to be archaisms. First, all the NWB and CWB ELs with locative inversion have expletive locative subject markers, which are interpreted as traces of a former locative system (Grégoire 1975; 1983; 2003, and §3.2.1), an interpretation which is not consistent with the supposed innovative nature of the construction. In the Ewondo example in (111) the ground is introduced by the preposition *a*, a trace of the class 16 nominal (pre-)prefix, and the copula takes a class 17 expletive subject marker.

- (111) Ewondo A72a (Grégoire 1975: 123)
á-ndá ó-nə díbi
 16-house SM17-COP darkness
 ‘*dans la maison, il fait noir*’ [‘in the house, there is darkness’]

In the corresponding Bulu utterance, the copula is co-referential with the ground. Note, however, that the class 16 subject marker has formally merged with the class 1 subject marker.

(112) Bulu A74a (Grégoire 1975: 123)

á-ndá a-nε díbi

16-house SM16-COP darkness

‘*dans la maison, il fait noir*’ [‘in the house, there is darkness’]

Babole has traces of a class 16 locative subject marker in existential constructions (113).

(113) Babole C101 (Leitch 2003: 405)

hé na múmgwà

SM16.COP COM 3.salt

‘there is salt’

Next, non-locative expletive subject marking can be analysed as the result of the total disappearance of the locative system. (Non-locative) Expletive inversion is almost entirely restricted to forest Bantu languages, which are known to have lost locative agreement. As pointed out by Grégoire (1983: 152; see also note 14) the class 16 expletive subject marker became reanalysed as a class 1 subject marker in some zone A languages, because of their formal similarity. This might well have happened in Kwakum, which has inverted ELs with an expletive subject marker of class 1/3SG. It should be noted that Kwakum has a heavily reduced concord system (Hare 2018; Njantcho Kouagang 2018). Subject markers, for example, are either 3SG (*a*) or 3PL (*je*). In (114) the 3SG subject marker is used expletively as it does not agree in number with the inverted figure.

(114) Kwakum A91 (Hare 2018: 213)

a bε me tɛfi ne akan i-dzambu

SM1/3SG COP PST4 also COM warriors CONN-war

‘there were also a lot of warriors (in Til)’

As was mentioned before, Kwakum also has non-inverted ELs. Bila and Bira similarly display variation between non-inverted (115a) and inverted (115b) ELs.

(115) Bila D311 (Brisson 1965: 66, 109)

a. *ba-bí ba-kibóko subá libo*

SM2-COP.PRF 2-7.hippo in 5.water

‘*il y avait des hippopotames dans la rivière*’ [‘there were hippos in the river’]

- b. *nyodwa ndi suba ngoli*
 9.knot COP in 9.rope
 ‘*il y a un nœud dans la corde*’ [‘there’s a knot in the rope’]

Bila and Bira also have severely reduced concord systems. Their inverted ELs are characterised as “agreeing inversion” because the subject marker agrees in number with the inverted figure (115a). However, exceptions do occur, especially in Bira, where the subject marker tends to be 3SG, irrespective of the number value of the lexical subject (Meinhof 1939: 253). As pointed out by Meinhof (1939: 284–285), the severe reduction of the concord system triggers a more rigid SVO word order. Still following Meinhof (1939: 285) ELs can constitute an exception to the SVO word order (116a). However, several examples suggest that ELs too “succumb” to word order restrictions triggered by the reduced agreement system (116b).

(116) Bira D32 (Meinhof 1939: 278, 285)

- a. *na karai a-bi-kau¹⁶ gani*
 and beginning SM3SG-8-COP.PRF 5.word
 ‘*und im Anfang war das Wort*’ [‘and in the beginning was the word’]
- b. *na mbili a tali madia a-bi-kau kube*
 and 10.pitcher CONN 6.stone six SM3SG-?-COP.PRF there
 ‘*und es waren dort sechs Krüge von Stein*’ [‘and there where six stone pitchers there’]

In sum, even though the genealogical/phylogenetic distribution of the non-inverted ELs suggests they are an archaic feature, the inverted ELs attested in the NWB and CWB branches cannot straightforwardly be analysed as innovations but rather point towards the reduction of the concord system as a possible trigger of a more rigid word order resulting in non-inverted ELs. Note that the hypothesised link between a reduced concord system and non-inverted ELs needs further research as not all NWB and CWB languages in Table 4 show heavily restricted subject agreement. Kela is a case in point. The subject marker of the copula varies in accordance with the preverbal figure (see also 126a).

¹⁶Meinhof (1939: 276) suggests that *kau* could be an old perfect form of a verb ‘to be’. It is used to express ‘to occur, be there/somewhere’ and combines with the class 8 prefix *bi-*. Note that the class 8 demonstrative *bindo* is used as a locative particle (Meinhof 1939: 253).

(117) Kela C75 (Forges 1977: 78)

- a. *ĩy a-yadí nd âtény a:nd ânt a:iko*
 1.thief SM1-COP.PRS in 2.inside 2.CONN 2.people 2.DEM
 ‘*il y a un voleur parmi nous*’ [‘there is a thief among us’]
- b. *mpw é-yadí nd ôtém o:nda mpoke*
 9.mouse SM9-COP.PRS in 3.heart 3.CONN 9.POT
 ‘*il y a une souris à l’intérieur du pot*’ [‘there’s a mouse inside of the pot’]

4.2 The areal innovation hypothesis

ELs not showing morphosyntactic differences from plain locational constructions constitute the dominant type in the Macro-Sudan Belt (Creissels 2019a,c). Furthermore, they are especially prominent in its core area where the Benue-Congo, Adamawa-Ubangi and Central Sudanic languages border on the Bantu domain. Indeed, more than 80% of the languages of the core area sample have ELs characterised by word order “rigidity” and absence of morphological specialisation in relation to PLs. Below we give examples of non-inverted ELs from Benue-Congo, Adamawa-Ubangi and Central Sudanic languages. The Benue-Congo languages are Mungbam (118a), Tiv (119a) and Mundabli (120a) (see also Creissels 2019b).¹⁷ PLs are given as well for the sake of comparison.

(118) Mungbam [Benue-Congo > Southern Bantoid] (Lovegren 2013: 441)

- a. *ā-dzāŋ ì-fē’ ì-kòŋ á m̂*
 12-fly 5-head 5-funnel PREP LOC.at
 ‘there’s a fly on the rim of the funnel’
- b. *ī-tī jī k̄ā-kp̄ē k̄ā á su*
 5-stone 5.DET 12-shoe 12.DET PREP LOC.face
 ‘the stone is in front of the shoe’

(119) Tiv [Benue-Congo > Bantoid > Tivoid] (Abraham 1940: 24, 68)

- a. *kwaghyan ŋgu*
 food COP.CL1
 ‘there is food’

¹⁷Benue-Congo languages not mentioned by Creissels (2019a,c) also possibly conflating location and existence include Kwanja (Thwing 2006), Tikar (Stanley 1991: 303), Kemezung (Smoes 2010: 35), Esimbi (Coleman et al. 2004: 58), Yemba (Bamileke) (Haynes 1996), Limbum (Fransen 1995: 316) and Kom (Shultz 1997: 40).

- b. *iyɔ ŋgi shin nya*
snake COP.CL2 in ground
'the snake is on the ground'

(120) Mundabli [Benue-Congo > Bantoid > Yemne-Kimbu] (Voll 2017: 304, 303)

- a. *mbĩ dʒ wú gbà*
6.wine COP POSS1 house.LOC
'there is wine in his house'
- b. *wù dʒ (i) fĩ mĩ*
PRON1 COP LOC 9.market in
'she is at the market'

For Adamawa-Ubangi and Central Sudanic, examples from Samba Leko and Ngambay are given (see also Creissels 2019a). As can be gathered from the translation equivalents in (121) and (122), the lack of differentiation with PLs leads to ambiguity.

(121) Samba Leko [Adamawa-Ubangi] (Fabre 2002: 297)

- wēl tǎ wūurú*
water COP backwater.at
'*il y a de l'eau dans le marigot/l'eau est dans le marigot*' ['there is water in the backwater/water is in the backwater']

(122) Ngambay [Central Sudanic] (Ndjerareou et al. 2010: 22)

- dǎu àr kái*
person 3SG.stand house
'there is someone at home/someone is at the house'

Out of the 33 sample languages with non-inverted ELs, 21 are spoken in an area more or less bordering the Macro-Sudan Belt, thus allowing for an explanation in terms of areal diffusion. The languages in question belong to zones A, C and D: Basaa A43a, Bakoko A43b, Nen A44, Nyokon A45, Kpa A53, Mmala A62B, Gunu A622, Eton A71, Gyeli A801, Koonzime A842 and Kwakum A91, Aka C104, Leke C14, Lingala C30B, Bomboma C411, Kele C55, Mongo C61, Liko D201, Bila D311, Bira D32 and Budu D332. Examples from Nyokon (105a), Gyeli (103b), Lingala (5), Bila (115a) and Bira (116b) have already been given. Additional examples follow. PLs are provided for the sake of comparison.

- (123) Gunu A622 (Rekanga 1989: 172)
nɛfɛ́bɛ́ nɛ́-lé gu tsi
 5.paper SM5-COP 17 ground
 ‘*par terre il y a un papier/il y a un papier par terre*’ [‘there is a paper on the ground/the paper is on the ground’]
- (124) Aka C104 (Thomas & Bahuchet 1991: 134)
 a. *mòbòdì ndé vê*
 mushroom COP there
 ‘*il y a des champignons par là-bas*’ [‘there are mushrooms over there’]
 b. *àmɛ ndé ngô mbúsà*
 I COP far 1.last
 ‘*moi, je suis là-bas, loin derrière*’ [‘I am there, far behind’]
- (125) Budu D332 (Asangama 1983: 400, 174)
 a. *akuu bá-noi bá=o*¹⁸
 high 2-bird SM2.COP=LOC
 ‘*au-dessus, il y a des oiseaux*’ [‘there are birds up there’]
 b. *mo-kósa u-á aká*
 3-corn SM3-COP here
 ‘*le maïs est ici*’ [‘the corn is here’]

In Budu, the ground is right-dislocated which could be suggestive of a non-canonical word order as attested in Mbuun and Western Serengeti languages. In Mbuun, subjects “are focused in situ but their focalisation triggers movement of the object to clause-initial position” (Bostoen & Mundeke 2012: 139). It could thus be the case that the non-topical nature of the figure in (125a) causes the ground to move to clause-initial position. If so, ELs in Budu do not display complete syntactic identity to PLs. Nevertheless, the non-inverted ELs in the other languages of the northern Bantu borderland are very similar to the ELs found in the core area of the Macro-Sudan Belt and could thus be the result of areal diffusion.

Still, there are eight western and four¹⁹ eastern Bantu languages in our sample which like the ones above have ELs characterised by a non-inverted word order.

¹⁸Note that the locative enclitic is not dedicated to the expression of inverse location. It also occurs in plain locational clauses, cf. a-á-o ‘he is there’ (Asangama 1983: 166).

¹⁹In Table 1 the JE45 Western Serengeti languages are counted as a single language.

However, their geographical distribution makes the areal diffusion hypothesis doubtful or even completely unlikely. With reference to the phylogeny in Grolle-mund et al. (2015), the western Bantu languages in question belong to three distinct branches: CWB, i.e. Tetela C71, Kela C75 and Ndengese C81; WWB, i.e. Nzadi B865, Mbuun B87, Zombo H16hK and Tsootso H16hZ; and SWB, i.e. Mbukushu K333. The EB languages include Mbugwe F34, the Western Serengeti languages Nata JE45, Ikoma JE45, Ishenyi JE45 and Ngoreme JE401, and Mwera P22.

Only Tetela, Kela (126), Ndengese and Nzadi (81b–81c) behave like the majority of northern languages (and Macro-Sudan Belt languages) discussed above in that they lack morphosyntactic differentiation between ELs and PLs.

(126) Kela C75 (Forges 1977: 78)

- a. *nnyàmà ì-yàdí ndá bòkòndà*
10.animal SM10-COP.PRS in 3.forest
'*il y a des bêtes dans la forêt*' ['there are animals in the forest']
- b. *njàdí ndá Bònómbà*
SM1SG.COP.PRS in Bonomba
'*j'habite à Bonomba*' ['I live/am in Bonomba']

(127) Ndengese C81 (Goemaere 1980: 42; Galerne 2001: 90)

- a. *bonto a-le=ko*
1.person SM1-COP=LOC
'*daar is iemand*' ['there is someone']
- b. *bo-sóngo bǎ-lé=kǎ lé njvúfulu*
3-tree SM3-COP=LOC behind 9.house
'*l'arbre est derrière la maison*' ['the tree is behind the house']

Mbukushu allows both ELs with agreeing inversion (96a) and non-inverted ELs (107) which are morphosyntactically similar to PLs (96b). The ELs of the other languages either display morphological specialisation in relation to their PLs or no complete syntactic identity. Tsootso includes a locative enclitic in ELs (80b) which is not present in PLs. Mwera has both an EL characterised by locative inversion (109) and a non-inverted EL (41) which takes a pre-initial locative marker which is optionally present in PLs (108a-b). Mbuun (21), Zombo (102), Mbugwe (101) and the Western Serengeti languages (22b), (50b), (106) all have non-canonical information structural characteristics which allow or even require the non-topical/focused figure to occur in preverbal position.

In sum, the majority (21) of languages with non-inverted ELs are spoken in an area compatible with the areal diffusion hypothesis. Moreover, most of the non-inverted ELs in languages spoken further away from the Macro-Sudan Belt differ from the northern non-inverted languages in that they display some morphosyntactic particularities in comparison to PLs.²⁰ This does not apply to Tetela, Kela, Ndengese and Nzadi. The first three are CWB languages which adds some weight to the archaic feature hypothesis. However, the WWB language Nzadi does not fit either the archaic feature or the areal innovation hypothesis. Interestingly, it has an extremely reduced concord system probably due to contact with non-Bantu languages (Crane et al. 2011: 4) which again points towards a link between reduced concord systems and non-inverted ELs.

5 The Proto-Bantu existential locational construction(s)

The frequencies and geographical spread of the following two types of ELs attested in the Bantu languages of our convenience sample straightforwardly suggest their reconstruction to at least node 5 in the phylogenetic tree of Grollemund et al. (2015):

- A. EL featuring a locative copula and (formal) locative inversion (1.A.i in Table 2)
- B. EL featuring a comitative copula with a locative subject marker (2.A.i in Table 2)

Both strategies are widely and frequently attested in our sample (see Table 2). This, together with the fact that the most frequently attested intralingual variation in ELs concerns the choice between a locative and a comitative copula, makes the reconstruction of two existential strategies plausible. Their reconstruction all the way up to node 1 is less straightforward because of the scarcity of both types in the NWB and CWB languages. Instead, languages belonging to these branches often have non-inverted ELs which are rare outside of these branches and cross-linguistically. This allows for at least two possible scenarios.

First, PB had non-dedicated, non-inverted ELs and the inverted constructions (A & B) were innovated after the NWB and CWB branches had split off. The rare non-inverted constructions in other branches can then be considered archaic

²⁰At least for the Western Serengeti languages and Mbugwe, it cannot be excluded that their irregular existential constructions may be connected to contacts with languages of other families, such as Nilotic and Cushitic.

heterogenities. However, several (21 languages in our sample vs. 24 with non-inverted ELs) NWB and CWB languages have inverted constructions which are not easily interpreted as independent innovations but rather seem to involve traces of a former full-fledged concord system with locative agreement.

This leads us to the second scenario, which rather argues for the presence in PB of dedicated inverted ELs (A & B). The 21 NWB and CWB languages with inverted ELs can then be considered (adapted) retentions of the original structure. However, we then still need to explain the innovation of the cross-linguistically rare non-inverted EL. We suggest that the reduction of the concord system (and more specifically the loss of locative agreement, an essential characteristic of Bantu ELs), witnessed across the north-western periphery of Bantu languages and possibly an effect of contact with non-Bantu languages (cf. e.g. Maho 1999; Good 2018; Verkerk & Garbo 2022) was an important trigger of this innovation. It prompted speakers to use alternative constructions or adapt the existing ones. Our data suggests that they had recourse to either an adapted or an alternative construction. The adapted construction is the EL with expletive inversion, which is especially frequent in zone C languages, but also occurs in other forest Bantu languages. The locative subject marker probably first became expletive, and merged at a later stage with a non-locative class (cf. Grégoire 1975; 1983: 124).

The alternative construction is the non-inverted EL, which in most cases shows no morphosyntactic differences from the plain locational construction (§3.2.2.2 & §4.2). In certain languages with dedicated ELs, PLs occasionally have existential readings. The Nata example in (128) is a case in point. In Nata, the existential reading is facilitated by the fact that the language allows non-topical constituents in preverbal position. In other languages, the existential reading of a PL is mostly observed in the absence of an explicit ground. The preverbal figure in (129) from Swahili, for example, can have a topical or non-topical interpretation resulting in ambiguous locational/existential readings.

- (128) Nata JE45 (Gambarage 2019: 54)
o-mu-sekeḗnya wojí mu-umwẹ́rǐ
AUG-3-sand SM3.PRS.COP 18-3.moon
‘there is sand on the moon/sand is on the moon’

- (129) Swahili G42d (Marten 2013: 47)
wa-tu wa-po
2-person SM2-COP
‘there are people/people are there/people are available’

If the reduction or loss of the locative agreement system triggers the loss of ELs characterised by locative agreement, this alternative reading of a PL (with a preverbal figure) may become the preferred way of expressing an existential locational meaning. The predominance of non-inverted existential constructions in the languages of the Macro-Sudan Belt might well have been an important factor in the consolidation of the non-inverted strategy. Twenty-one languages with non-inverted ELs are spoken in an area compatible with the hypothesis that the absence of figure inversion in Bantu languages of the northern borderland is an areal feature originating from non-Bantu donors from the Macro-Sudan Belt. However, a few languages (Tetela, Kela, Ndengese and Nzadi) of the radical generic location type are spoken too far away from the Macro Sudan Belt to be consistent with the areal diffusion hypothesis. In sum, we suggest that contact-induced noun class reduction and ensuing loss of locative agreement are the main explanatory factors for the innovation of non-inverted ELs in the northern Bantu borderland rather than the areal diffusion of a radical generic location type. Interestingly, counterexamples to both the archaic feature hypothesis and the areal diffusion hypothesis point towards the severe reduction of the concord system as a possible trigger for a more rigid word order and consequently non-inverted ELs. However, languages like Kela, which have non-inverted ELs and do not display heavily reduced concord systems, suggest that the latter hypothesis is in need of further research. For now, we reconstruct types A and B to node 5 and suggest that their reconstruction to PB is plausible.

The reconstruction of the first strategy implies the reconstruction of locative inversion. Meeussen (1967: 120) reconstructs “anastasis” or subject-object inversion for PB and considers locative inversion as a special case of subject-object (patient) inversion or “*renversement*” (Meeussen 1959: 215). Moreover, there is an implicational hierarchy that there is no inversion with full lexical verbs in a language without inversion with copula (Marten & van der Wal 2014: 59). Therefore, if subject-object inversion can be reconstructed for PB, then inversion as found in Bantu ELs predominantly involving locative copula can also be reconstructed. The predicator slot was most probably filled with defective **di* ‘be’ (Bastin et al. 2002), at least in present tense contexts. In sum, we suggest the following morphosyntactic pattern for the EL featuring a locative copula and locative inversion:

A. ***[(LOC.NP #) LOC.SM-dì # NP (# LOC.NP)] (# = word boundary)**

In non-present contexts, **bá* ‘be, dwell, become’ (Bastin et al. 2002) was probably used as copula.

The second EL strategy does not involve locative inversion, but still takes a locative subject marker, as the ground is considered the possessor of the figure. The comitative copula most probably consisted of **dì* or **bá* immediately followed by **nà* ‘with, also, and’ (Meeussen 1967: 115; Bastin et al. 2002). Although the inflected comitative copula is also frequent in our sample, it is less widespread and thus probably a later development. We therefore propose the following morphosyntactic pattern for the EL featuring a comitative copula:

B. *[(LOC.NP #) LOC.SM-dì (#) na (#) NP (# LOC.NP)]

Throughout the Bantu area locative morphology plays an important role in ELs and this is also true for the suggested reconstructions which both involve locative subject marking. Locative noun classes and their agreement sets have been reconstructed for PB (Meeussen 1967; Grégoire 1975).²¹

The proposed reconstructions assume a level of fusion of the verbal form which, following Güldemann (2003; 2011; 2022), did not exist in PB. Following this hypothesis, PB was characterised by a “split predicate” with a self-standing subject pronoun and verb (stem), which only at a later stage came to fuse into the synthetic verbal “template” characteristic for Bantu languages. The question of how agglutinative PB was again ties in with the larger debate on whether features witnessed in north-western Bantu which match with those of the Macro-Sudan area are to be considered retentions of the original structure or rather as representing later instances of (contact-induced) loss (see Good & Güldemann 2006; Hyman 2007; Nurse 2007; Güldemann 2008; Nurse 2008: 62–72; Güldemann 2011; Hyman 2011). However, it should be noted that Güldemann (2011) himself claims that subject pronouns or other class indexing markers – such as locative class markers – fused earlier in “simple” verb forms, i.e. predicate constructions without any intervening TAM marking, as is precisely the case with the copula in our suggested reconstructions. So, even if we accept the “split predicate” hypothesis, the suggested reconstructions could still be valid for PB and certainly for a reconstruction to node 5. However, we cannot preclude that the locative subject marker and the copula formed two disparate words and hence that the hyphen between LOC.SM and COP should rather be a <#>, marking word boundary (or a clitic <=> representing some in-between state of fusion).

²¹See also Good (2018: 33) for further Bantu-external evidence showing that the locative classes are at least as old as PB.

6 Conclusions

The main goal of this chapter was to reconstruct the morphosyntactic pattern of PB existential constructions. To be able to do so, we investigated the synchronic variation in existential strategies in 157 Bantu languages. It would have been nice to be able to include more data and especially more fine-grained data (cf. all the test sentences in (1)) in order to avoid the risk of comparing apples and oranges, as we might have done now and again when comparing inverse/rhematic locationals with instances of generic and bounded existence. Also, in order to establish whether a language has expletive or referential locative agreement, we ideally should have equivalents of utterances like ‘on the table, there is a cat’, ‘at the market, there are fruits’ and ‘in the house, there are rats’. This way we can ascertain whether the locative marker, if present, shows agreement with the locative ground or is of an expletive nature. We hope that this chapter may trigger researchers to include all these types of sentences in their elicitation lists so the current dataset can be expanded and improved.

Based on the present sample, we were still able to come up with two sets of variables regarding “existential locationals” (ELs) in Bantu languages. The first set pertains to word order and agreement patterns and distinguishes a non-inverted type and three types involving figure inversion: locative inversion, expletive inversion, and agreeing inversion. The second set concerns the type of verbal element, typically a locative or comitative copula. have-verbs, polysemous locative/possessive copulas and specialised EL verbs also occur in our sample but much less frequently.

Bantu languages often have more than one existential strategy. The most recurrent and most widely spread strategies are the ones involving figure inversion, a locative subject marker, and either a locative or a comitative copula. The use of a comitative copula and referential locative subject markers are typical features of Bantu ELs (see also Creissels 2019a: 26–33). Locative and comitative copulas are almost equally frequent and widespread. We therefore put forward two existential locational strategies as the best candidates for reconstruction to at least node 5 of the phylogenetic tree of Grollemund et al. (2015) and possibly to PB:

- A. *[(LOC.NP #) LOC.SM-dì # NP (# LOC.NP)]
- B. *[(LOC.NP #) LOC.SM-dì (#) na (#) NP (# LOC.NP)]

The reason for not straightforwardly reconstructing these morphosyntactic patterns to PB lies in the fact that they are only scarcely attested in forest Bantu

languages, which rather have non-inverted ELs or ELs characterised by expletive inversion. We suggested that the almost (!) complete absence of the A and B patterns in the NWB and CWB branches can be explained in two ways: (i) PB had a non-dedicated, non-inverted EL and the present-day non-inverted ELs are retentions; (ii) PB had the ELs in A and B and the non-inverted ELs are (contact-induced) innovations. Although further research and more data are needed, our preference goes to the second explanation which assumes that the innovation was triggered by the severe reduction or even complete loss of (locative) noun classes and the ensuing (locative) agreement system in the concerned languages. The reduced concord system resulted in ELs with expletive inversion and exempt of locative marking, or in a more rigid word order and consequently non-inverted ELs.

Acknowledgements

We wish to thank the editors as well as the reviewers whose comments and insights helped us improve this chapter. Several people helped us updating and supplementing our database with unpublished data. We thankfully recognise the help of Yuko Abe (Bende), Sebastian Dom (Kongo), Helen Eaton (Malila), Nadine Grimm (Gyeli), Deo Kawalya (Ganda), Elisabeth Jane Kerr (Nen), Antti Laine (Nata), Johnson Malema (Jita), Gastor Mapunda (Tanzanian Ngoni), Sozinho Francisco Matsinhe (Tsonga), Enock Mbiling'i (Kinga), Michael Meeuwis (Lingala), Manoah-Joël Misago (Rundi), Godian Moses (Luguru), Léon Mundeke (Mbuun), André Ndagba (Liko), Lengson Ngwasi (Hehe), Ruth Raharimanantsoa (Teke Tyee & Ngungwel), and Vera Wilhelmsen (Mbugwe). Rasmus Bernander gratefully acknowledges the support of the Kone Foundation for his part of the research.

Abbreviations

1, 2, 3 ...	noun class 1, 2, 3	COP	copula
SG/PL	person singular/plural	CWB	Central-Western Bantu
AUG	augment	DEF	definite
CJ	conjoint	DEM	demonstrative
CL	class	DET	determiner
COM	comitative	DIST	distal
CONJ	conjunction	DJ	disjoint
CONN	connective	DUR	durative

EB	Eastern Bantu	PLUR	pluractional
EXIST	existential	POSS	possessive
EXPL	expletive	PRED	predicative
FOC	focus	PREHOD	prehodiernal
FUT	future	PREP	preposition
FV	final vowel	PRF	perfect
HAB	habitual	PROH	prohibitive
INAM	inanimate	PRON	personal pronoun
IPFV	imperfective	PRS	present
LOC	locative	PST	past
NEG	negation	R	realis
NF	non-final form of the hesternal and the hodiernal past perfective	REL	relative
NPST	non-past	SEQ	sequential
NUM	numeral	SIT	situative
NWB	North-Western Bantu	SM	subject marker
OM	object marker	STAT	stative
PASS	passive	SWB	South-Western Bantu
PFV	perfective	VDS	valence-decreasing suffix
		WWB	West-Western Bantu

References

- Abraham, Roy Clive. 1940. *The principles of Tiv*. London: Crown Agents for the Colonies (on behalf of the Government of Nigeria).
- Araújo, Paulo Jefferson Pilar. 2013. *Domínios conceituais das construções locativas, existenciais, comitativas e possessivas em línguas bantas*. São Paulo & Bayreuth: University of São Paulo & University of Bayreuth. (Doctoral dissertation).
- Asangama, Atisa. 1983. *Le budu, langue bantu du nord-est du Zaïre : esquisse phonologique et grammaticale*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Aunio, Lotta, Holly Robinson, Tim Roth, Oliver Stegen & John B. Walker. 2019. The Mara languages (JE40). In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 501–532. Milton Park, Abingdon: Routledge.
- Baka, Jean. 1992. *Essai de description du tso:tso : parler ko:ngo du nord de l'Angola*. Brussels: Université libre de Bruxelles. (MA thesis).
- Barreteau, Daniel. s.d. Questionnaire d'inventaire linguistique de nyoʔɔ. Unpublished ms.

- Bastin, Yvonne. 2003. The interlacustrine zone: Zone J. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 501–528. London: Routledge.
- Bastin, Yvonne. 2020. The class 18 locative prefix and the expression of the present progressive in Bantu. *Africana Linguistica* 26. 5–58.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africanmuseum.be/collections/browsecollections/humansciences/blr>.
- Bearth, Thomas. 2003. Syntax. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 121–142. London: Routledge.
- Bentley, Delia. 2017. Copular and existential constructions. In Andreas Dufter & Elisabeth Stark (eds.), *Manual of Romance morphosyntax and syntax* (Manuals of Romance Linguistics 17), 332–366. Berlin: Walter de Gruyter.
- Bentley, Delia, Francesco Maria Ciconte & Silvio Cruschina. 2013. Existential constructions in crosslinguistic perspective. *Italian Journal of Linguistics / Rivista di Linguistica* 25(1). 1–13.
- Bernander, Rasmus. 2017. *Grammar and grammaticalization in Manda: An analysis of the wider TAM domain in a Tanzanian Bantu language*. Gothenburg: University of Gothenburg. (Doctoral dissertation).
- Bernander, Rasmus, Maud Devos & Hannah C. Gibson. Forthcoming(a). Existential constructions in Bantu languages. In Eva-Marie Bloom Ström, Hannah C. Gibson, Rozenn Guérois & Lutz Marten (eds.), *Current approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.
- Bernander, Rasmus, Maud Devos & Hannah C. Gibson. Forthcoming(b). The negative existential cycle in Bantu. In Ljuba Veselinova & Arja Hamari (eds.), *The negative existential cycle* (Research on Comparative Grammar 1). Berlin: Language Science Press.
- Bernander, Rasmus & Antti Laine. 2020. The formation of existential constructions in Western Serengeti: A micro-comparative exploration of variation and change. *Africana Linguistica* 26. 59–102.
- Besha, Ruth Mfumbwa. 1989. *A study of tense and aspect in Shambala* (Language and Dialect Studies in East Africa 10). Berlin: Dietrich Reimer.
- Bloom Ström, Eva-Marie. 2020. The existential copula in Xhosa in relation to indefiniteness. *Studies in African Linguistics* 49(2). 213–240.
- Borschev, Vladimir & Barbara H. Partee. 2002. The Russian genitive of negation: Theme-rheme structure or perspective structure? *Journal of Slavic Linguistics* 10(1–2). 105–144.

- Bostoen, Koen & Léon Mundeke. 2012. Subject marking, object-verb order and focus in Mbuun (Bantu, B87). *Southern African Linguistics and Applied Language Studies* 30(2). 139–154.
- Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chicheŵa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.
- Brisson, Robert. 1965. *Vocabulaire Bira*. Unpublished ms.
- Buell, Leston C. & Merijn de Dreu. 2013. Subject raising in Zulu and the nature of PredP. *The Linguistic Review* 30(3). 423–466.
- Burger, J. P. 1960. *An English – Lozi vocabulary*. Mongu, Barotseland Protectorate: Book Depot of P.M.S. Sefula.
- Bwantsa-Kafungu, S.-Pierre & Achiel Emiel Meeussen. 1970–71. *Notes mbuun*. Unpublished ms. Tervuren: Royal Museum for Central Africa.
- Bwendelele, A. s.d. *Notes yaka*. Unpublished ms.
- Chatelain, Héli. 1888–99. *Kimbundu grammar: Grammatica elementar do kim-bundu ou lingua de Angola*. Geneva: C. Schuchardt.
- Clements, George N. & Annie Riailand. 2008. Africa as a phonological area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 36–85. Cambridge: Cambridge University Press.
- Cole, Desmond T. 1955. *An introduction to Tswana grammar*. London; New York: Longmans; Green.
- Coleman, Arnie, Karen Coleman & Brad Koenig. 2004. *Grammar sketch of Esimbi*. Unpublished ms. Yaoundé: SIL Cameroon.
- Collins, B. 1962. *Tonga grammar*. London: Longmans.
- Crane, Thera M. 2011. *Beyond time: Temporal and extra-temporal functions of tense and aspect marking in Totela, a Bantu language of Zambia*. Berkeley, CA: University of California. (Doctoral dissertation).
- Crane, Thera M., Larry M. Hyman & Simon Nsielanga Tukumu. 2011. *A grammar of Nzadi [B.865]: A Bantu language of the Democratic Republic of Congo* (University of California Publications in Linguistics 147). Berkeley, CA: University of California Press.
- Creissels, Denis. 2013. Control and the evolution of possessive and existential constructions. In Elly van Gelderen, Jóhanna Barðdal & Michela Cennamo (eds.), *Argument structure in Flux: The Naples-Capri papers* (Studies in Language Companion Series 131), 461–476. Amsterdam: John Benjamins.
- Creissels, Denis. 2015. Existential predication and trans-possessive constructions. Paper presented at Colloque international « La prédication existentielle dans les langues naturelles : valeurs et repérages, structures et modalités », Institut National des Langues et Civilisations Orientales, Paris, 10–11 April 2015.

- Creissels, Denis. 2019a. Existential predication in the languages of the Sudanic belt. *Afrikanistik-Aegyptologie-Online*. <https://www.afrikanistik-aegyptologie-online.de/archiv/2019/4860>.
- Creissels, Denis. 2019b. Impersonal constructions in typological perspective. Paper presented at the Sixteenth Conference on Typology and Grammar for Young Scholars, Saint Petersburg, 21–23 November 2019.
- Creissels, Denis. 2019c. Inverse-locational predication in typological perspective. *Italian Journal of Linguistics / Rivista di Linguistica* 31(2). 37–106.
- Czinger, Christine. 2002. Decomposing existence: Evidence from Germanic. In Werner Abraham & C. Jan-Wouter Zwart (eds.), *Issues in formal German(ic) typology* (Linguistik Aktuell / Linguistics Today 45), 85–126. Amsterdam: John Benjamins.
- da Silva Maia, António. 1961. *Dicionário complementar português-kimbundu-kikongo: Línguas nativas do centro e norte de Angola*. Cucujães (Angola): Tipografia das Missões.
- Dalgish, Gerard M. 1976. Locative NP's, locative suffixes, and grammatical relations. In Henry Thompson, Kenneth Whistler, Vicki Edge, Jeri J. Jaeger, Ronya Javkin, Miriam Petruck, Christopher Smeall & Robert D. Van Valin Jr. (eds.), *Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society*, 139–148. Berkeley, CA: Berkeley Linguistics Society, University of California.
- Dammann, Ernst. 1957. *Studien zum Kwangali: Grammatik, Texte, Glossar* (Abhandlungen aus dem Gebiet der Auslandskunde 63). Hamburg: Cram, de Gruyter & Co.
- Dammann, Ernst. 1977. Das 'o of reference' in Bantusprachen. In Brigitta Benzinger, Otto Böcher & Günther Meyer (eds.), *Wort und Wirklichkeit: Studien zur Afrikanistik und Orientalistik, Eugen Ludwig Rapp zum 70. Geburtstag. Teil 2: Linguistik und Kulturwissenschaft*, 31–43. Meisenheim am Glan: Anton Hain.
- de Blois, Kornelis Frans. 1970. The augment in the Bantu languages. *Africana Linguistica* 4. 85–165.
- de Wit, Gerrit. 2015. *Liko phonology and grammar: A Bantu language of the Democratic Republic of the Congo*. Leiden: Leiden University. (Doctoral dissertation).
- Demuth, Katherine. 1990. Locatives, impersonals and expletives in Sesotho. *The Linguistic Review* 7(3). 233–249.
- Demuth, Katherine & Carolyn Harford. 1999. Verb raising and subject inversion in Bantu relatives. *Journal of African Languages and Linguistics* 20(1). 41–61.
- Dereau, Léon. 1955. *Cours de kikôngo*. Namur: Wesmael-Charlier.
- Devos, Maud. 2008. *A grammar of Makwe (Palma, Mozambique)* (LINCOM Studies in African Linguistics 71). Munich: LINCOM Europa.

- Devos, Maud. 2017. Shangaji paired tenses: Emergence of a CJ/DJ sytem? In Jenneke van der Wal & Larry M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu* (Trends in Linguistics. Studies and Monographs 301), 122–146. Berlin: Walter de Gruyter.
- Devos, Maud, Manoah-Joël Misago & Koen Bostoen. 2017. A corpus-based description of locative and non-locative reference in Kirundi locative enclitics. *Africana Linguistica* 23. 47–83.
- Diercks, Michael. 2011. The morphosyntax of Lubukusu locative inversion and the parameterization of Agree. *Lingua* 121(5). 702–720.
- Dimande, Ernesto. 2020. *A Concordância verbal na Língua Ronga*. Unpublished ms. Maputo: Eduardo Mondlane University.
- Dryer, Matthew S. 2007. Clause types. In Timothy Shopen (ed.), *Language typology and syntactic description, Volume 1: Clause structure*, 2nd edn., 224–275. Cambridge: Cambridge University Press.
- du Plessis, Jacobus A. & Marianna Visser. 1992. *Xhosa syntax*. Pretoria: Via Afrika.
- Dunham, Margaret. 2005. *Éléments de description du langi : langue bantu F 33 de Tanzanie : phonologie – grammaire – lexicque* (Langues et Littératures de l’Afrique Noire 13). Louvain: Peeters.
- Eaton, Helen. 2015. *Malila narrative discourse* (SIL Language and Culture Documentation and Description 27). Dallas, TX: SIL International.
- Edelsten, Peter & Chiku Lijongwa. 2010. *A grammatical sketch of Chindamba: A Bantu language (G52) of Tanzania* (East African Languages and Dialects 21). Cologne: Rüdiger Köppe.
- Englebretson, Robert, Lisa Jeon, Jessica Li, Samantha Mauney, Anai Navarro & Jonas Wittke. 2015. *A basic sketch grammar of Gikūyū* (Rice Working Papers in Linguistics 6). Houston, TX: Rice Linguistics Society, Rice University.
- Fabre, Anne Gwenaëlle. 2002. *Étude du samba leko, parler d’Allani (Cameroun du Nord, famille Adamawa)*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Fisch, Maria. 1977. *Einführung in die Sprache der Mbukushu, Ost-Kavango, Namibia* (Wissenschaftliche Forschung in Südwestafrika 15). Windhoek: S.W.A. Wissenschaftliche Gesellschaft.
- Fisch, Maria. 1998. *Thimbukushu grammar*. Windhoek: Out of Africa Publishers.
- Fivaz, Derek. 1984. Predication in Ndonga. *African Studies* 43(2). 147–160.
- Forges, Germaine. 1977. *Le kela, langue bantoue du Zaïre (zone C) : esquisse phonologique et morphologique* (Société d’études linguistiques et anthropologiques de France 59–60). Paris: SELAF.

- Fransen, Margo Astrid Eleonora. 1995. *A grammar of Limbum: A Grassfields Bantu language spoken in the North-West province of Cameroon*. Amsterdam: Vrije Universiteit Amsterdam. (Doctoral dissertation).
- Galerie, Anne. 2001. *Le ndengese : description d'une langue bantoue : étude du syntagme nominal*. Nice: University of Nice Sophia Antipolis. (Doctoral dissertation).
- Gambarage, Joash Johannes. 2019. *Belief-of-existence determiners: Evidence from the syntax and semantics of Nata augments*. Vancouver: University of British Columbia. (Doctoral dissertation).
- Gast, Volker & Florian Haas. 2011. On the distribution of subject properties in formulaic presentationals of Germanic and Romance: A diachronic-typological approach. In Andrej Malchukov & Anna Siewierska (eds.), *Impersonal constructions: A cross-linguistic perspective* (Studies in Language Companion Series 124), 127–166. Amsterdam: John Benjamins.
- Gérard, R. P. 1924. *La langue lebéo : grammaire et vocabulaire* (Bibliothèque Congo 13). Brussels: Vromant & Cie.
- Gibson, Hannah C., Rozenn Guérois & Lutz Marten. 2019. Variation in Bantu copula constructions. In María J. Arche, Antonio Fábregas & Rafael Marín (eds.), *The grammar of copulas across languages* (Oxford Studies in Theoretical Linguistics 73), 213–242. New York, NY: Oxford University Press.
- Goemaere, Alphonse. 1980. *Spraakleer van het Londengese + Woordenlijst*. Unpublished ms. Leuven.
- Good, Jeff. 2018. East Benue-Congo noun classes, with a focus on morphological behavior. In John R. Watters (ed.), *East Benue-Congo: Nouns, pronouns, and verbs* (Niger-Congo Comparative Studies 1), 27–57. Berlin: Language Science Press. DOI: [10.5281/zenodo.1314321](https://doi.org/10.5281/zenodo.1314321).
- Good, Jeff & Tom Güldemann. 2006. The Bantu verbal prefixes and S-Aux-O-V order in Benue-Congo. Paper presented at the International conference on “Bantu Grammar Description and Theory”, School of Oriental and African Studies, University of London, 20–22 April 2006.
- Gray, Hazel. 2013. *Locatives in Ikizu*. Leiden: Leiden University. (MA thesis).
- Grégoire, Claire. 1975. *Les locatifs en bantou* (Annales – Série in-8° – Sciences humaines 83). Tervuren: Royal Museum for Central Africa.
- Grégoire, Claire. 1983. Quelques hypothèses comparatives sur les locatifs dans les langues bantoues du Cameroun. *Journal of West African Languages* 13(2). 139–164.
- Grégoire, Claire. 2003. The Bantu languages of the forest. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 349–370. London: Routledge.

- Grimm, Nadine. 2015. *A grammar of Gyeli*. Berlin: Humboldt University of Berlin. (Doctoral dissertation).
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Gromova, Nelli V. & Anna Yu. Urmanchieva. 2005. Танзанийские сказки-билингвы [Tanzanian bilingual fairy tales]. In A. I. Koval (ed.), *Африканская сказка-III. К исследованию языка фольклора [The African folktale III. Study of the language of folklore]*, 215–354. Moscow: Eastern Literature.
- Guérois, Rozenn. 2015. *A grammar of Cuwabo (Mozambique, Bantu P34)*. Lyon: Université Lumière – Lyon 2. (Doctoral dissertation).
- Guérois, Rozenn. 2016. The locative system in Cuwabo and Makhuwa (P30 Bantu languages). *Linguistique et Langues Africaines* 2. 43–75.
- Güldemann, Tom. 2003. Grammaticalization. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 182–194. London: Routledge.
- Güldemann, Tom. 2008. The Macro-Sudan Belt: Towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151–185. Cambridge: Cambridge University Press.
- Güldemann, Tom. 2011. Proto-Bantu and Proto-Niger-Congo: Macro-areal typology and linguistic reconstruction. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas: With special reference to Africa* (Tokyo University of Foreign Studies, Studies in Linguistics 2), 109–141. Amsterdam: John Benjamins.
- Güldemann, Tom. 2018. Historical linguistics and genealogical language classification in Africa. In Tom Güldemann (ed.), *The languages and linguistics of Africa* (The World of Linguistics 11), 58–444. Berlin: Mouton de Gruyter.
- Güldemann, Tom. 2022. Predicate structure and argument indexing in early Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 387–421. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575831.
- Gunnink, Hilde. 2018. *A grammar of Fwe: A Bantu language of Zambia and Namibia*. Ghent: Ghent University. (Doctoral dissertation).
- Halme, Riikka. 2004. *A tonal grammar of Kwanyama* (Namibian African Studies 8). Cologne: Rüdiger Köppe.

- Hammarström, Harald. 2010. Rarities in numeral systems. In Michael Cysouw & Jan Wohlgemuth (eds.), *Rethinking universals: How rarities affect linguistic theory* (Empirical Approaches to Language Typology 45), 11–59. Berlin: Mouton de Gruyter.
- Hare, David M. 2018. *Tense in Kwakum narrative discourse*. Dallas, TX: Graduate Institute of Applied Linguistics. (MA thesis).
- Harries, Lyndon. 1940. An outline of Maŵiha grammar. *Bantu Studies, A Journal Devoted to the Scientific Study of Bantu, Hottentot and Bushmen* 14(1). 91–146.
- Harries, Lyndon. 1950. *A grammar of Mwera, a Bantu language of the eastern zone, spoken in the south-eastern area of Tanganyika Territory* (Bantu Grammatical Archives 1). Johannesburg: Witwatersrand University Press.
- Haynes, Nancy R. 1996. *Two Yémba texts interlinearised (language data)*. Yaoundé: SIL Cameroon.
- Hengeveld, Kees. 1992. *Non-verbal predication. Theory, typology, diachrony* (Functional Grammar Series 15). Berlin: Mouton de Gruyter.
- Hyman, Larry M. 2007. Reconstructing the Proto-Bantu verbal unit: Internal evidence. In Nancy C. Kula & Lutz Marten (eds.), *Bantu in Bloomsbury: Special issue on Bantu linguistics* (SOAS Working Papers in Linguistics 15), 201–211. London: Department of Linguistics, School of Oriental & African Studies.
- Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language Dynamics and Change* 1(1). 3–49.
- Khumalo, Langa. 2012. Passive, locative inversion in Ndebele and the unaccusative hypothesis. *South African Journal of African Languages* 30(1). 22–34.
- Koch, Peter. 2012. Location, existence, and possession: A constructional-typological exploration. *Linguistics* 50(3). 533–603.
- Kraal, Pieter Jacob. 2005. *A grammar of Makonde (Chinnima, Tanzania)*. Leiden: Leiden University. (Doctoral dissertation).
- Laman, Karl E. 1912. *Grammar of the Kongo language (Kikongo)*. New York, NY: Christian Alliance Publishing Company.
- Lambrecht, Knud. 1988. Presentational cleft constructions in spoken French. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse* (Typological Studies in Language 18), 135–179. Amsterdam: John Benjamins.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Cambridge Studies in Linguistics 71). Cambridge: Cambridge University Press.
- Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics, An Interdisciplinary Journal of the Language Sciences* 39(3). 463–516.

- Leach, Michael Benjamin. 2010. *Things hold together. Foundations for a systematic treatment of verbal and nominal tone in Plateau Shimakonde*. Leiden: Leiden University. (Doctoral dissertation).
- Legère, Karsten. 2010. Vidunda/Ngh'wele sentences. Unpublished field notes.
- Leitch, Myles. 2003. Babole (C101). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 392–421. London: Routledge.
- Louw, Jacobus A. & J. B. Jubase. 1963. *Handboek van Xhosa*. Johannesburg: Bonapers Beperk.
- Lovegren, Jesse Stuart James. 2013. *Mungbam grammar*. New York, NY: University at Buffalo, State University of New York. (Doctoral dissertation).
- Lyons, John. 1967. A note on possessive, existential and locative sentences. *Foundations of Language* 3(4). 390–396.
- Maganga, Clement & Thilo C. Schadeberg. 1992. *Kinyamwezi: Grammar, texts, vocabulary* (East African Languages and Dialects 1). Cologne: Rüdiger Köppe.
- Maho, Jouni F. 1999. *A comparative study of Bantu noun classes*. Gothenburg: Acta Universitatis Gothoburgensis. (Doctoral dissertation).
- Marchal-Nasse, Colette. 1979. *Esquisse de la langue tsogo: Phonologie, morphologie*. Brussels: Université libre de Bruxelles. (3rd cycle dissertation).
- Marchal-Nasse, Colette. 1989. *De la phonologie à la morphologie du nzebi, langue bantoue (B52) du Gabon*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- Marten, Lutz. 2010. The great siSwati locative shift. In Anne Breitbarth, Christopher Lucas, Sheila Watts & David Willis (eds.), *Continuity and change in grammar* (Linguistik Aktuell / Linguistics Today 159), 249–267. Amsterdam: John Benjamins.
- Marten, Lutz. 2013. Structure and interpretation in Swahili existential constructions. *Italian Journal of Linguistics / Rivista di Linguistica* 25(1). 45–73.
- Marten, Lutz, Nancy C. Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105(3). 253–338.
- Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318–368.
- McNally, Louise. 2011. Existential sentences. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 2 (Handbooks of Linguistics and Communication Science 33.2), 1829–1848. Berlin: Mouton de Gruyter.
- Meeussen, Achiel Emiel. 1952. *Esquisse de la langue ombo (Maniema – Congo belge)* (Annales – Série in-8° – Sciences de l'homme – Linguistique 4). Tervuren: Royal Museum of the Belgian Congo.

- Meeussen, Achiel Emiel. 1959. *Essai de grammaire rundi* (Annales – Série in-8° – Sciences de l’homme – Linguistique 24). Tervuren: Royal Museum of the Belgian Congo.
- Meeussen, Achiel Emiel. 1962. Lega-teksten. *Africana Linguistica* 1. 75–97.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Meeussen, Achiel Emiel & Servilien Sebasoni. 1965. *Notes binja-sud*. Unpublished ms. Tervuren: Royal Museum for Central Africa.
- Meinhof, Carl. 1939. Die Sprache der Bira. *Zeitschrift für Eingeborenen-Sprachen* 29(4). 241–287.
- Mkude, Daniel J. 1974. *A study of Kiluguru syntax with special reference to the transformational history of sentences with permuted subject and object*. London: University of London. (Doctoral dissertation).
- Möhlig, Wilhelm J. G. & Jekura U. Kavari. 2008. *Reference grammar of Herero (Otjiherero): Bantu language of Namibia, with a glossary Otjiherero – English – Otjiherero* (Southern African Languages and Cultures 3). Cologne: Rüdiger Köppe.
- Morimoto, Yukiko. 2006. Agreement properties and word order in comparative Bantu. *ZAS Papers in Linguistics* 43. 161–187.
- Morrison, Michelle Elizabeth. 2011. *A reference grammar of Bena*. Houston, TX: Rice University. (Doctoral dissertation).
- Moshi, Lioba. 1995. Locatives in KiVunjo-Chaga. In Akinbiyi Akinlabi (ed.), *Theoretical approaches to African linguistics* (Trends in African Linguistics 1), 129–145. Trenton, NJ: Africa World Press.
- Mous, Maarten. 2005. The innovative character of object-verb word order in Nen (Bantu A44, Cameroon). In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics, with special focus on Bantu and Mande. Essays in honour of Yvonne Bastin and Claire Grégoire* (Collection Sciences Humaines 169), 411–424. Tervuren: Royal Museum for Central Africa.
- Mous, Maarten & Abel Mreta. 2004. The comitative/conjunctive preposition na in Chathu (Pare), a Bantu language of Tanzania. *South African Journal of African Languages* 24(4). 219–229.
- Moyo-Kayita, Makila. 1981. *Esquisse de grammaire mbala : Morphologie et syntaxe*. Kananga: Institut Supérieur Pédagogique. (MA thesis).
- Nabirye, Minah. 2016. *A corpus-based grammar of Lusoga*. Ghent: Ghent University. (Doctoral dissertation).
- Nanteza, Moureen. 2018. *Locative enclitics in Luganda. Form and meaning*. Kampala: Makerere University. (MA thesis).

- Nassau, Robert H. 1892. *Mackey's grammar of the Benga-Bantu language*. New York, NY: American Tract Society.
- Ndjerareou, Mekoulnodji, Christy Melick & Sarah R. Moeller. 2010. A brief grammatical sketch of Ngambay. *GIAL Electronic Notes Series* 4(2). 1–39.
- Nicolle, Steve. 2013. *A grammar of Digo: A Bantu language of Kenya and Tanzania* (SIL International Publications in Linguistics 149). Dallas, TX: SIL International.
- Nicolle, Steve. 2015. *Comparative study of Eastern Bantu narrative texts* (SIL Electronic Working Papers 2015-003). Dallas, TX: SIL International.
- Njantcho Kouagang, Elisabeth. 2018. *A grammar of Kwakum*. Paris: Institut National des Langues et Civilisations Orientales. (Doctoral dissertation).
- Nurse, Derek. 2007. The emergence of tense in early Bantu. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 164–179. Somerville, MA: Cascadilla Proceedings Project.
- Nurse, Derek. 2008. *Tense and aspect in Bantu* (Oxford Linguistics). New York, NY: Oxford University Press.
- O'Sullivan, Owen. 1993. *English – Silozi dictionary*. Lusaka: Zambia Educational Publishing House.
- Odden, David. 1996. *The phonology and morphology of Kimatumbi* (The Phonology of the World's Languages). Oxford: Clarendon Press.
- Partee, Barbara H. & Vladimir Borschev. 2004. The semantics of Russian genitive of negation: The nature and role of perspectival structure. In Robert B. Young (ed.), *Proceedings of the 14th Semantics and Linguistic Theory Conference, held May 14–16, 2004, at Northwestern University*, 212–234. Ithaca, NY: Cornell University.
- Partee, Barbara H. & Vladimir Borschev. 2007. Existential sentences, BE, and the genitive of negation in Russian. In Ileana Comorovski & Klaus von Heusinger (eds.), *Existence: Semantics and syntax* (Studies in Linguistics and Philosophy 84), 147–190. Dordrecht: Springer.
- Persohn, Bastian. 2017. *The verb in Nyakyusa: A focus on tense, aspect, and modality* (Contemporary African Linguistics 2). Berlin: Language Science Press. DOI: 10.5281/zenodo.926408.
- Petzell, Malin. 2008. *The Kagulu language of Tanzania: Grammar, texts and vocabulary* (East African Languages and Dialects 19). Cologne: Rüdiger Köppe.
- Philippon, Gérard & Derek Nurse. 2000. Gweno, a little known language of Northern Tanzania. In Kulikoyela Kahigi, Yared M. Kihore & Maarten Mous (eds.), *Lugha za Tanzania / Languages of Tanzania: Studies dedicated to the memory of Prof. Clement Maganga* (CNWS Publications 89), 233–284. Leiden:

- Research School for Asian, African, & Amerindian Studies, University of Leiden.
- Piper, Klaus. 1977. *Elemente des Suku: Zur Phonologie und Morphologie einer Bantusprache*. Leiden: Rijksuniversiteit te Leiden. (Doctoral dissertation).
- Prat, Jean. 1917. *Petite grammaire mbochie et dictionnaire (rivière Alima, Congo français)*. Brazzaville: Mission Catholique.
- Rekanga, Jean-Paul. 1989. *Essai de grammaire gunu (langue bantoue du Cameroun A 62)*. Brussels: Université libre de Bruxelles. (MA thesis).
- Salzmann, Martin. 2011. Towards a typology of locative inversion – Bantu, perhaps Chinese and English – but beyond? *Language and Linguistics Compass* 5(4). 169–189.
- Salzmann, Martin David. 2004. *Theoretical approaches to locative inversion*. Zurich: University of Zurich. (MA thesis).
- Sanderson, Meredith. 1922. *A Yao grammar*. 2nd revised edition. London: Society for Promoting Christian Knowledge.
- Schadeberg, Thilo C. & Ridder Samsom. 1994. Kiinimacho cha mahali: Kiambishi tamati cha mahali -ni. *Afrikanistische Arbeitspapiere* 37. 127–138.
- Schmitz, Robert. 1912. *Les baholoholo (Congo belge)* (Collection de monographies ethnographiques : sociologie descriptive 9). Brussels: Librairie Albert Dewit (pour l'Institut International de Bibliographie).
- Schoeneborn, Anne. 2009. *The Sesotho book: A language manual developed by a Peace Corps volunteer during her service in Lesotho*. Mazenod.
- Seidel, Frank. 2008. *A grammar of Yeyi: A Bantu language of southern Africa* (Grammatical Analyses of African Languages 33). Cologne: Rüdiger Köppe.
- Shultz, George. 1997. *Kom language grammar sketch, Part 1*. Yaoundé: SIL Cameroon.
- Sitali, Georgina Nandila. 2008. *Some aspects of the semantics and syntax of adverbial clauses in Lozi: A corpus-based study*. Lusaka: University of Zambia. (MA thesis).
- Smoes, Christopher L. 2010. *A sketch grammar of the Kemezung language*. Yaoundé: SIL Cameroon.
- Stanley, Carol. 1991. *Description phonologique et morpho-syntaxique de la langue tikar (parlée au Cameroun)*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Stassen, Leon. 2013. Predicative possession. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/117>.

- Stegen, Oliver. 2011. *In quest of a vernacular writing style for the Rangi of Tanzania: Assumptions, processes, challenges*. Edinburgh: University of Edinburgh. (Doctoral dissertation).
- Stevick, Earl W. & Kingston Machiwana. 1960. *Manyika step-by-step*. Cleveland (South Africa): Central Mission Press.
- Ström, Eva-Marie. 2013. *The Ndengeleko language of Tanzania*. Gothenburg: University of Gothenburg. (Doctoral dissertation).
- Struck, Bernhard. 1911. Die Fipasprache (Deutsch-Ostafrika). *Anthropos* 6(5). 951–993.
- Taji, Julius John. 2017. *The morphosyntactic and semantic properties of grammatical relations in Chiyao*. Dar es Salaam: University of Dar es Salaam. (Doctoral dissertation).
- Talmy, Leonard. 1975. Figure and ground in complex sentences. In Cathy Cogen, Henry Thompson, Graham Thurgood, Kenneth Whistler & James Wright (eds.), *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, University of California. DOI: 10.3765/bls.v1i0.2322.
- Tanda, Vincent Ambe & Ayu'nwi N. Neba. 2005. Negation in Mokpe and two related coastal Bantu languages of Cameroon. *African Study Monographs* 26(4). 201–219.
- Thomas, Jacqueline M. C. & Serge Bahuchet (eds.). 1991. *Encyclopédie des pygmées aka : Techniques, langage et société des chasseurs-cueilleurs de la forêt centrafricaine (Sud-Centrafrique et Nord-Congo)*. Tome I, les pygmées aka. Fasc. 4, *La langue* (Société d'études linguistiques et anthropologiques de France 331). Louvain: Peeters.
- Thornell, Christina. 2004. Studies in the Kerebe language. Unpublished ms.
- Thwing, Rhonda. 2006. *Verb extensions in Vute*. Yaoundé: SIL International.
- Twilingiyimana, Chrysogone. 1984. *Éléments de description du doko* (Annales - Série in-8° - Sciences humaines 116). Tervuren: Royal Museum for Central Africa.
- Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics, with special focus on Bantu and Mande. Essays in honour of Yvonne Bastin and Claire Grégoire* (Collection Sciences Humaines 169), 425–441. Tervuren: Royal Museum for Central Africa.
- Van de Velde, Mark. 2008. *A grammar of Eton* (Mouton Grammar Library 46). Berlin: Mouton de Gruyter.
- van der Wal, Jenneke. 2011. Focus excluding alternatives: Conjoint/disjoint marking in Makhuwa. *Lingua* 121(11). 1734–1750.

- van der Wal, Jenneke. 2015. Bantu syntax. In *Oxford handbooks online. Scholarly research reviews*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199935345.013.50.
- van der Wal, Jenneke & Larry M. Hyman (eds.). 2017. *The conjoint/disjoint alternation in Bantu* (Trends in Linguistics Studies and Monographs 301). Berlin: De Gruyter Mouton.
- van der Wal, Jenneke & Saudah Namyalo. 2016. The interaction of two focus marking strategies in Luganda. In Doris L. Payne, Sara Pacchiarotti & Mokaya Bosire (eds.), *Diversity in African languages. Selected papers from the 46th Annual Conference on African Linguistics* (Contemporary African Linguistics 1), 355–377. Berlin: Language Science Press. DOI: 10.17169/langsci.b121.490.
- Vanhoudt, Bettie. 1987. *Éléments de description du leke, language bantoue de zone C* (Annales – Série in-8° – Sciences humaines 125). Tervuren: Royal Museum for Central Africa.
- Verkerk, Annemarie & Francesca Di Garbo. 2022. Sociogeographic correlates of typological variation in northwestern Bantu gender systems. *Language Dynamics and Change* 12(2). 155–223.
- Voll, Rebecca Maria. 2017. *A grammar of Mundabli: A Bantoid (Yemne-Kimbi) language of Cameroon*. Leiden: Leiden University. (Doctoral dissertation).
- Wald, Benji. 1973. Syntactic change in the lake languages of northeast Bantu. *Studies in African Linguistics* 4(3). 237–268.
- Whiteley, Wilfred H. 1966. *A study of Yao sentences*. Oxford: Clarendon Press.
- Zeller, Jochen. Forthcoming. Locatives in Bantu. In Lutz Marten, Nancy C. Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages* (Oxford Guides to the World's Languages 2). Oxford: Oxford University Press.
- Zeller, Jochen & Jean Paul Ngoboka. 2018. Agreement with locatives in Kinyarwanda: A comparative analysis. *Journal of African Languages and Linguistics* 39(1). 65–106.
- Ziervogel, Dirk, Petrus J. Wentzel & Tshikhuse N. Makuya. 1972. *A handbook of the Venda language*. Pretoria: University of South Africa.

Chapter 15

Diachronic typology and the reconstruction of non-selective interrogative pronominals in Proto-Bantu

Dmitry Idiatov

LLACAN - Langage, Langues et Cultures d'Afrique (CNRS, INaLCO, EPHE)

In this chapter, I propose a typologically informed reconstruction of the Bantu non-selective interrogative pronominals (NSIPs). Bantu NSIPs are characterised by a bewildering degree of formal variation, which makes their reconstruction particularly difficult. Therefore, I begin with a more general methodological discussion of the issue of variation in functional elements and the possible ways of dealing with it in reconstruction and by an overview of the diachronic typology of NSIPs. The most important results of the proposed reconstruction of Bantu NSIPs are that no human NSIP stem 'who?' can be reconstructed for Proto-Bantu (PB), while the morphological status of the non-human NSIP form 'what?' is ambiguous, and that the NSIP forms that can be reconstructed to PB were emerging out of complex interrogative constructions (viz. a clause-level cleft construction and a nominalisation construction) retained from some pre-PB stage within Southern Bantoid. Because of their complex constructional origin and the typical pathways of formal and semantic evolution, the reconstruction of interrogative pronominals bears significant relevance to the reconstruction of many other parts of Bantu morphosyntax, such as deictics (both spatial and discourse ones), the so-called augment and more generally referential status marking, nominalisation, noun classes, subject indexation, copulas, cleft constructions, relative clause constructions, constituent order, and root phonotactics (the question of vowel-initial roots and the identity of PB *j).



1 Introduction

In this chapter, I propose a typologically informed reconstruction of the Bantu non-selective interrogative pronominals (NSIPs), such as ‘who?’ and ‘what?’, which are used in non-selective contexts where the speaker perceives the choice as free (see Idiatov 2007 for a more detailed definition).¹ This reconstruction is intended as a major revision of the reconstructions proposed by Meeussen (1967) in *Bantu Grammatical Reconstructions* (BGR) and their minor updates in *Bantu Lexical Reconstructions 3* (BLR3) by Bastin et al. (2002), viz. **n(d)ái* ‘who?’ of the so-called class 1a and the interrogative stem **-í* ‘what?’ used with the prefix of class 7 as ‘what?’ and in the locative classes 16, 17 and 18 as ‘where?’. For purposes of reconstruction in this chapter, I take Proto-Bantu (PB) as the latest common stage that can be reconstructed using the data of the languages traditionally classified as Narrow Bantu [narr1281].²

Bantu NSIPs are characterised by a bewildering degree of formal variation (§2.1), which makes their reconstruction particularly difficult. Traditionally, Bantu historical linguistics dealt with such a high degree of variation in one of three ways, viz. by reconstructing a number of more common formal “types”, by reconstructing the simplest possible form or by reconstructing a kind of common denominator of most of the attested reflexes (§2.2). I believe that we can enhance the reconstruction of highly variable functional morphemes, such as NSIPs, by taking diachronic typology into account (§2.3).

The main generalisation that emerges from my research about the Bantu interrogative pronominals is that they go back to complex interrogative constructions, viz. a clause-level cleft construction in the case of the BGR form **n(d)ái* and a nominalisation construction in the case of the BGR form **-í*. These constructions are retentions from earlier, pre-PB stages. Because of their complex constructional origin and the typical pathways of formal and semantic evolution, the reconstruction of interrogative pronominals bears significant relevance to the reconstruction of many other parts of Bantu morphosyntax, such as deictics (both spatial and discourse ones), the so-called augment and more generally referential status marking, nominalisation, noun classes, subject indexation, copulas, cleft constructions, relative clause constructions, and constituent order. The chapter discusses some of these many implications for Bantu morphosyntax, but

¹Selective interrogative pronominals (SIPs), such as ‘which one?’, are used in selective contexts, where the choice is perceived by the speaker as restricted to a closed set of alternatives.

²For Narrow Bantu languages, I provide their Guthrie codes following Maho (2009). For non-Narrow Bantu languages, I provide their Glottolog identifier code of the shape [xxxx1111, name of the language group] (cf. Hammarström et al. 2021).

often in footnotes or relegated to the Appendix so as not to disrupt the main line of argument too much.

The chapter has the following structure. I set the stage by introducing the issue of variation in functional elements and the possible ways of dealing with it in §2. In §3, I provide an overview of the diachronic typology of NSIPs in the world's languages (cf. Idiatov 2007). In §4, I go through some of the typologically trivial changes that affect NSIPs in Bantu. In §5, I highlight some of the oddities of NSIPs across Bantu and their implications for the reconstruction of NSIPs in Bantu, especially the human NSIP 'who?'. In §6, I revise the reconstructions of the Bantu NSIPs. In order to both refine the reconstructions and to determine the level to which they belong, I equally take into consideration data from the wider Bantoid continuum, occasionally complemented by data from other Benue-Congo groups. §7 provides some concluding remarks.

2 Setting the stage: The variation and the ways of dealing with it

2.1 Bewildering variation

In Bantu languages, interrogatives in general and 'who?' and 'what?' in particular are characterised by a bewildering degree of formal variation, as can for example be observed in the forms of interrogatives in the data used for the lexicostatistic study by Bastin et al. (1999), which are available on the website of the Royal Museum for Central Africa³ and are cited in the remainder of this chapter without an explicit reference. On the one hand, we find a multitude of forms that do not seem to have anything in common, such as the forms for 'who?' in Basaa A43a *njé(é)*, Eton A71 *zá* with the construction-specific variant *zà* (Van de Velde 2008a: 176, p.c.), Ngombe C41 *ndá*, Liko D201 *wàní* (de Wit 2015), and Tswana S31 *máng*; or the forms for 'what?' in Basaa *kí(i)*, Eton *já* with the constructional and dialectal variant *jà* and the dialectal variant *yá* (Van de Velde 2008a: 176, p.c.), Fumu B77b *ima*, and Komo D23 *ékéndà*. On the other hand, we also encounter a multitude of forms that are clearly related but where this relationship is marred by irregular correspondences, such as the forms for 'who?' in A15 varieties reconstructed as **njá* by Hedinger (1987: 244), viz. Akoose *nzé(é)*, Myenge *nzáá*, Mwanhed *nzé*, Mbo (of Ekanang) *ndé*, Mwaneka *nzá*, Mkaa *njá*; or the forms for the general NSIP 'who?; what?' in a number of languages of zone C that I discuss in

³Cf. https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/lexicostatistic-study-bantu-languages

Idiatov (2009), such as Mboshi C25 *ndè ~ nê*, Mongo-Nkundo C61 *ná*, varieties of Tetela C70 *nâ*, and Ntomba-Inongo C35a *ínò*. The forms of interrogative pronominals may range from very short, lacking any internal morphological structure, such as Mongo-Nkundo *é* ‘what?; where?’, to relatively long. Such longer forms may be nominals including a class prefix, such as Mwani G403 *kì-nání* ‘what?’ with the nominal prefix of class 7. They may be nominal expressions, such as Enya Kibombo D14 *kì-úmà nàání* ‘what?’ literally meaning ‘what thing?’, with the class 7 noun *kì-úmà* ‘thing’ modified by the interrogative *nàání* that on its own means ‘who?’. They can even be clause-level constructions used as nominal expressions, such as Kagulu G12 (*i*)*yehoki* ‘who?’ and Mbula [mbul1261, Jarawan Bantu] *yá ‘n(á)* ‘who is it?; who?’. As discussed in §4.1.4.2, Kagulu (*i*)*yehoki* ‘who?’ is structurally (*i*)-*y-e-hoki*, a nominalised predication that literally means something like ‘the one that s/he is the one where?’ [(NMLS-)-1-be:NMLS-where?] (leaving the source locative interrogative *hoki* ‘where?’ unanalysed). Mbula *yá ‘ná* can be construed both as a predication [^H-*yà ná ~ V̇-yà ná*] [NMLS-which? COP.PRES] meaning ‘who is it?’, also as a base of the cleft construction, as in (1a–1c), and as a nominal expression meaning ‘who?’, as in (1d–1f). It can even take the regular nominal plural marker *à^H-*, as *à-yá/ná*, while preserving the same structural ambiguity between the predication ‘who are these?’ and the nominal expression ‘who? (PL)’.

(1) Mbula [mbul1261, Jarawan Bantu]⁴

- a. *yá* *‘ná*
 H-*yà* *ná*
 NMLS-which? COP.PRES
 ‘Who is it?’
- b. *ndà* *yá/ná*
 [3SG]COP.EQ who?
 ‘Who is he?’ (3SG subject index has no overt marker)
- c. *mà: yá* *‘ná?*
mà V̇-yà *ná*
 POSS NMLS-which? COP.PRES
 ‘Whose is it?’ (lit.: ‘It is the one of who?’)
- d. *yá* *n* *ndà* *m̀bwà:má ‘má:n*
 H-*yà* *ná* *ndà* *m̀bwá:má mǎ:n*
 NMLS-which? COP.PRES [3SG]COP.EQ woman this
 ‘Who is this woman?’ (lit.: ‘Who is it (that) she is this woman?’)

⁴The Mbula data cited in this chapter come from my joint research with Mark Van de Velde.

- e. *yá* *ʼná* *à-sán-i*
 H-yà ná àH-sèn-í
 NMLS-which? COP.PRES 2SG-see-3SG
 ‘Who did you see?’ (lit.: ‘Who is it (that) you saw him/her?’)
- f. *à-sán-i* *yá/ʼná*
 2SG-see-3SG who?
 ‘You saw who?’ (an echo-question)

2.2 Traditional ways of dealing with high variation in Bantu historical linguistics

Both in their bewildering degree of formal variation and in the way that this formal variation is structured, Bantu NSIPs resemble various deictic forms, such as personal indexes (substitutives and possessives)⁵ and demonstratives, rather than nouns and verbs, following a common cross-linguistic pattern (cf. Diessel 2003; Idiatov 2007: 564–566). Three different approaches have been applied in Bantu historical linguistics to go about the reconstruction of forms that are characterised by a high degree of variation.

The first approach is to reduce the synchronic variation by reconstructing a smaller range of more common formal “types”. Methodologically, this is a rather conservative approach that discards only irregularities that are deemed to be minor, while preserving the more radical formal variation. This approach can be illustrated with the way in which Malcolm Guthrie deals with the reconstruction of interrogative pronominals in Volume 2 of *Comparative Bantu* (Guthrie 1971), as summarised in (2). I also include the SIP ‘which?’ as it is a frequent source of NSIPs cross-linguistically (cf. §3.2).

- (2) Guthrie’s (1971: 125, 156) reconstruction of Bantu interrogative pronominals⁶

- a. ‘what?’: (*-ní C.S. 1354), *-yàní C.S. 1926
 b. ‘who?’: *náà C.S. 1337, *nání C.S. 1343, *-yàní C.S. 1925
 c. ‘which?’: (*-ká C.S. 1046), *-kí C.S. 1046, *-ní C.S. 1354, *-pí C.S. 1498, *-tí C.S. 1728

⁵The term *substitutives* in Bantu linguistics refers to free personal indexes (pronominals) and their stems, while *possessives* refers to free possessive personal indexes (pronominals) and their stems.

⁶The C.S. codes refer to the specific “comparative series” established by Guthrie.

This approach is often adopted in reconstructions of the Bantu lexicon, such as in BLR3 (Bastin et al. 2002), resulting in the so-called *osculance* in reconstructions (cf. Bostoen 2001; Ricquier & Bostoen 2008; Bostoen & Bastin 2016).

The second approach, largely adopted in BGR, radically reduces the observed variation by reconstructing the simplest possible form, which is usually short, and by discarding as much as possible any deviations from the presumed system-general regularities. In practice, this usually implies picking out one formal “type” to the expense of the others. Meeussen’s (1967) reconstruction of Bantu interrogatives is summarised in (3).

- (3) Meeussen’s (1967: 103, 107) reconstructions of Bantu interrogatives
- a. ‘what?’: “a set which looks like a fragmentary system of interrogative nouns with stem *-í* : 7 *ki-í* ‘what’, 16 *pà-í* (17 *kù-í*, 18 *mù-í*) ‘where’”
 - b. “[class] 1a *n(d)ái* ‘who’, if it belongs here [= the set based on the stem *-í*], shows an element *n(d)á-* which is not attested otherwise (also *n(d)ání*).”
 - c. ‘which?’: pronominal prefix + *-ní*

The most drastic reduction of variation in BGR with respect to interrogative pronominals concerns ‘who?’, a nominal stem with an unusual form and extremely high formal variability. In this respect, note that later treatments of this interrogative, such as Doneux & Grégoire (1977: 193) and Schadeberg (2003: 163), concede that it is not reconstructable with certainty to PB. In Idiatov (2009), I argue that no interrogative pronominal meaning ‘who?’ can be reconstructed for PB. In choosing **nai* with a variant **ndai*, both unspecified for tone, BLR3 follows BGR, although admitting the tonal uncertainty. In BGR, ‘what?’ and the related locative interrogative forms for ‘where?’ look much less controversial. However, for a nominal stem marked by a nominal prefix, **í* ‘what?’ (also taken up in BLR3) has a very unusual vowel-initial shape. Furthermore, we cannot help but notice the striking difference between this elegant reconstruction and that of Guthrie (1971) in (2) with radically different forms.

The third approach does not reduce the variation but deals with it by reconstructing a kind of common denominator of most of the attested reflexes (while discarding some of the less common variants). This generally results in longer and structurally complex forms that may include morphemes whose function remains unidentified. I am not aware of any example involving interrogative pronominals, but the reconstruction of substitutives and possessives by Kamba Muzenga (2003) is a good illustration of such an approach. For example, the

scheme in Figure 1 describes the pathways of change of substitutives in Bantu, where all the structures below the reconstructed form represent the various reflexes in modern Bantu languages, with the original form faithfully preserved or either one, two or three of the four original morphemes lost. Of the four morphemes in the reconstructed form only the last two are assigned a meaning. Thus, *pp* stands for a *pronominal prefix*, a prefix from the paradigm of pronominal prefixes, while *e* is the substitutive stem for the first and second person and class 1, and *o* is the substitutive stem for all the other classes.

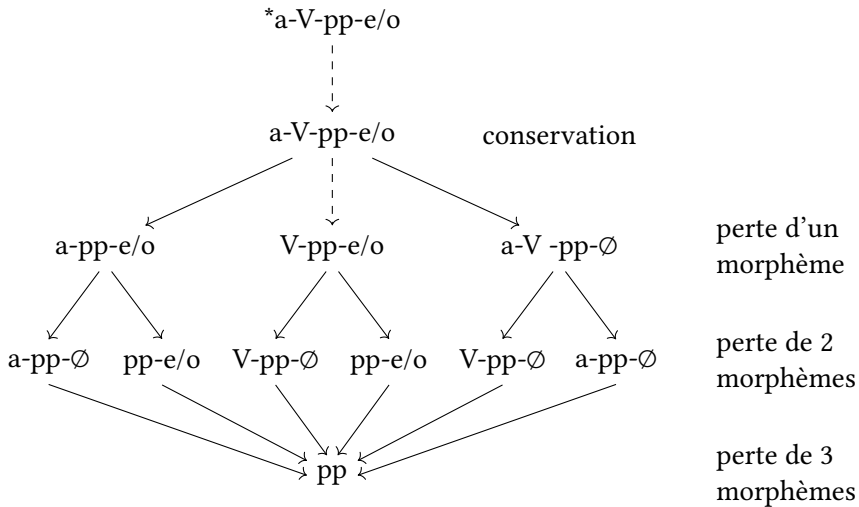


Figure 1: The pathways of change of substitutives in Bantu (Kamba Muzenga 2003: 228)

2.3 A typologically informed reconstruction

Reconstruction of functional elements, such as deictic forms and interrogative pronominals, is notoriously difficult. As we know from languages with long written traditions, the history of such functional forms tends to involve various irregular types of changes that by and large defy the rigorous application of the traditional Comparative Method. For example, it is only thanks to the older written sources that we know that Dutch *maar* ‘but; just, only’ and German *nur* ‘just, only’ are both reflexes of [not + be.OPT.PST] ‘were it not (that)’, viz. Old Dutch *ne ware* and Old High German *ni wāri*.

For languages without long written traditions, we can enhance the reconstruction of functional morphemes by taking diachronic typology into account. A

typologically informed reconstruction does this by trying to achieve the closest match between the observed variation in the presumed reflexes of a given element and the typological knowledge of common processes of change. By informing us about the typical pathways of formal and semantic change affecting a given type of linguistic items, diachronic typology provides us with the cues as to what historical sources may have produced the observed variation and thus feed back into the reconstruction by allowing us both to better account for the variation observed and to increase the plausibility of our reconstructions.

Since the tool of typologically informed reconstruction has so far been rarely applied to its full potential in Bantu historical linguistics, two methodological remarks are appropriate here. First, a typologically informed reconstruction provides best results in situations with many closely related languages characterised by fine-grained variation in the form of the functional element that we attempt to reconstruct. As in any reconstruction endeavour, the more languages, the better, as it is precisely the observed variation that should allow us to detect the pathways of change of the functional item in question. The more closely related the languages are, the better. At more shallow time depths, we can be more certain that the various formally and semantically similar items showing irregular sound correspondences indeed stem from a common source. In the case of closely related languages, independent innovations of formally and functionally closely matching forms from two completely different sources are much less likely. Bantu languages meet all these desiderata very well.

Second, the irregular changes that we may posit need to be both formally and semantically plausible. For example, the correspondence between [a:] in Dutch *maar* ‘but; just, only’ and [u:] in German *nur* ‘just, only’ is highly irregular. However, if we presume that in the source form the vowel [a:] was immediately preceded by a [w] and followed by an [i] in the next syllable, as in *ni wāri*, the proposed correspondence becomes much more phonetically plausible, despite remaining irregular. Similarly, the correspondence between *n* and *m* in the same two forms is highly irregular, but becomes less of an oddity if we presume that the change was not from *n* to *m* or vice versa but *nw* > *m*. The evidence for semantic plausibility can come from various sources. Minimally, the presumed semantic change should comply with regular mechanisms of semantic change, such as metonymically or metaphorically motivated shifts. For instance, a direct change from ‘who?’ to ‘what?’ or vice versa cannot be accounted for by any regular mechanism of semantic change, while a change from ‘where?’ to ‘which one?’ or a change from ‘which one?’ to ‘who?’ or ‘what?’ can, and in fact, is not uncommon cross-linguistically (cf. Idiatov 2007; 2009: 67–69; 2014). The presumed

semantic change may also be plausible because it would correspond to a conventionalisation of some readily available pragmatic inferences, as for instance in Italian the development from *schiaùvo* ‘slave’ > ‘(I am your) slave’ > ‘yours’ (as a farewell expression) > *ciao* ‘bye’.

3 The diachronic typology of non-selective interrogative pronominals

The diachronic typology of NSIPs in this section is largely based on Idiatov (2007).⁷ I begin by presenting a typology of the evolution of forms of NSIPs in §3.1.⁸ I highlight the fact that the generally held assumption of the universal high stability of NSIPs is not borne out by the cross-linguistic data, as in fact they prove to be highly unstable (§3.1.1). This instability is created by an interaction between two strong diachronic tendencies that work in opposite directions, viz. a strong predilection of interrogatives for substance accretion (§3.1.2) and probably an even stronger predilection for substance reduction (§3.1.3). Like many other functional elements, NSIPs are usually difficult to reconstruct. The apparent degree of difficulty of their reconstruction depends on the exact way the accretion and reduction of substance interact (§3.1.4). In §3.2, I briefly present a semantic diachronic typology of NSIPs.

3.1 Formal evolution

3.1.1 Formal (in)stability

Traditionally, NSIPs are considered to be among the most change-proof elements in any language. They are believed to be highly resistant to both replacement through borrowing (Haspelmath & Tadmor 2009, Matras 2009: 199) and language-internal renewal (Haspelmath 1997: 176). In this respect, they are believed to be similar to personal pronominals. The two kinds of pronominals are therefore often perceived as good indicators of (long-range) genetic relationships and are regularly included in basic vocabulary lists.

However, this assumption of the universal high stability of NSIPs is not borne out by the facts. In very many language families, NSIPs, like other interrogatives, turn out to be diachronically unstable and structurally complex polymorphemic

⁷Readers interested in knowing more about the typology of interrogatives and questions in general could consult the rather comprehensive overview of recent typological studies by Hölzl (2017: 55). Köhler (2016) is a dedicated study of questions in a number of African languages.

⁸This formal typology is largely applicable to many other types of interrogatives in general.

constructions (cf. Idiatov 2007; Cysouw & Hackstein 2011; Idiatov 2011; Ratliff 2011). Thus, even in Indo-European, the alleged textbook example of the stability of interrogative pronominals, we actually cannot reconstruct full NSIPs, only their initial segment **k*^w, which may or may not have had any morphological status of its own (Cysouw & Hackstein 2011).

The diachronic instability and the structural complexity of interrogatives are driven by an interaction between two strong diachronic tendencies that work in opposite directions, viz. a strong predilection of interrogatives for substance accretion and probably an even stronger predilection for substance reduction. Interrogatives share their strong predilection for substance reduction, often highly irregular and radical, with other function words due to frequency effects. The predilection for substance accretion is primarily due to two factors. The first one is the prominent information structural status of interrogatives. For instance, cross-linguistically this prominent status is manifested in the extremely recurrent use of focus constructions with interrogatives, such as French *qu'est-ce que [Marie a fait]?* 'what [did Mary do]?', which is literally a cleft construction 'what is it that [Mary did]?'. Due to frequency effects, the various elements that mark this prominent information structural status tend to become reanalysed as part of the interrogative itself. The second factor is the very strong tendency for continuity in the evolution of interrogatives: a given interrogative is almost always based on another interrogative (cf. Diessel 2003; Idiatov 2007; Cysouw & Hackstein 2011).⁹ In this respect, interrogatives are similar to other deictic forms, such as personal indexes and especially demonstratives, and differ from many other functional elements, such as conjunctions, tense and aspect markers or number markers, with which such lack of continuity is commonplace.¹⁰ A possible way to circumvent this strong continuity tendency is to use elements with cataphoric, or more precisely, suspended referential specification (on suspensive pronominals see van den Eynde & Mertens 2003: 70; Idiatov 2007: 3). In an appropriate discourse situation, this can be achieved by using constructions with demonstratives ('This one who did it [is]?...' > 'Who is the one who did it?'), nouns with

⁹See also Bostoen & Guérois (2022 [this volume]), on a somewhat similar, albeit less strong, tendency with certain verbal suffixes in Bantu.

¹⁰Thus, it is commonplace that a future marker develops from a form of a motion verb, such as 'go' or 'come'. It is equally trivial that a plural marker would evolve from a singular noun meaning 'group' or that a conjunction 'but', such as Dutch *maar*, would develop from a clause meaning 'were it not (that ...)'. The main difference here is that with these other kinds of functional elements the source form need not already contain the semantics of the target form. Thus, a motion verb that develops into a future marker need not itself be in the future tense, and moreover, its evolution into a future marker may happen in a language that did not have the category of tense before at all.

generic semantics or comparable indefinite pronominals ('A person / Somebody did it?...?' > 'Who is the one who did it?'), the word 'name' or 'call (a name)' ('The name of the one who did it [is]?...' or 'The one who did it is called (a certain name)?...' > 'Who is the one who did it?') – for some examples, see Idiatov (2007; 2014).

3.1.2 Substance accretion

Substance accretion often begins with the inclusion in the interrogative construction of free morphemes, that later become bound. The original morphological boundaries may subsequently become erased in the process of univerbation. In (4), I provide an overview of the common types of elements accreted in the diachrony of NSIPs. Often, several such elements are combined together. The same element can often also be construed with different functions at the same time, as when the same marker functions as a copula and as a focus marker or a deictic functions as a relativiser and a nominaliser, and so on.

- (4) Common types of accreted material for NSIPs
- a. various types of deictics: nominal, adverbial or modifying demonstratives, personal pronominals
 - b. focus markers
 - c. copulas
 - d. relativisers
 - e. nouns with generic semantics ('thing', 'person', 'place', 'name', etc.)
 - f. gender, number, noun class markers, classifiers
 - g. nominalisers (such as *one* in *which one?*, the augment in Bantu)

3.1.3 Substance reduction

Like with other functional elements, substance reduction in NSIPs due to frequency effects is often highly irregular and radical, and specific to particular word forms. The reduction may affect just one segment (syllable, morpheme, word) or several segments (syllables, morphemes, words) at one site or here and there. The only major cross-linguistic generalisation that can be made with respect to substance reduction in both functional and lexical forms is that it is more likely to affect the parts of a given form that are prosodically less prominent, such as the segments in non-stressed syllables, and the parts that contribute less to the

lexical meaning of the form as a whole. The possibility of the latter semantic conditioning is expected to weaken along with the gradual loss of transparency in the morphosyntactic structure of a NSIP and the contribution of each element to the lexical meaning of the form as a whole. Note, however, that before any accreted morphosyntactic material can undergo reduction, it must become integral part of the NSIP construction and its relation with the source construction must be weakened. Typically, such attrition would then only affect the morphosyntactic material in question within the NSIP and not in the source construction.

One particularly typical kind of substance reduction for NSIPs is what I have called *cosa*-type reduction in Idiatov (2007), based on the Italian example *che cosa* ‘what thing?’ > *cosa* ‘what?’. This variety of endocentric compound reduction is comparable to the use of *unions* for *trade unions* in English, with the head of the compound used to stand for the compound as a whole after the modifier is deleted.

A good example of a highly irregular and radical substance reduction, involving the loss of both the original NSIP stem and the morphosyntactic material accreted earlier, is the evolution of the NSIP ‘who?; what?’ in the French-based Louisiana Creole, as presented in Rottet (2004) and further discussed in Idiatov (2007: 253). The Louisiana Creole NSIP ‘who?; what?’ has the variants (*ki*) *sa ki* for questions about subjects and (*ki*) *sa* for questions about objects which all result from the evolution of the constructions that in standard French would be rendered as *c’est qui ça qui?* [DEM.M.SG:COP.PRS.3SG who? that.N.SG REL.SUBJ] ‘it is who that one who [did this]?’ for questions about subjects and *c’est qui ça que?* [DEM.M.SG:COP.PRS.3SG who? that.N.SG REL.OBJ] ‘it is who that one that [you saw]?’ for questions about objects.¹¹

3.1.4 Reconstructing interrogatives: The interplay between accretion and reduction

As with many other functional elements, the reconstruction of interrogatives is usually difficult, but may seem relatively easy depending on the exact way the accretion and reduction of substance interact. Three types of such interaction are possible. In the first type, accretion and reduction occur at the same side of

¹¹Dictionaries of French prescribe the spelling *ça* of the distal neuter pronominal demonstrative ‘that one’ for the element used to mark insistence with interrogatives, as in *qui ça?* ‘who? (tell me!)’, *où ça?* ‘where? (tell me!)’, *comment ça?* ‘how? (tell me!)’ (e.g. Rey-Debove 1996; cf. also <https://www.cnrtl.fr/etymologie/ça>). However, the Louisiana Creole *sa* could also reflect the homonymous proximal adverbial demonstrative *çà* ‘here’. Both options are plausible semantically and typologically.

an interrogative, as schematically illustrated in (5). Each original morpheme is represented by a succession of three identical letters, such as *aaa* and *bbb*, and as the morpheme becomes reduced the number of letters also reduces, viz. *aaa* → *aa* → *a*. The segments belonging to the original interrogative *aaa* are highlighted in bold.

- (5) Accretion and reduction occur at the same side of an interrogative
aaa → **aaa** *bbb* → **aabb** → *abb* *ccc* → *abcc*

An evolution as in (5) may remain detectable for a long period of time and create an illusion that its reconstruction is easy. In reality, we can only reconstruct a small part of the original interrogative. According to Cysouw & Hackstein (2011), this is the situation with the Proto-Indo-European NSIPs where we can only reconstruct the initial segment **k^w* of the original interrogative stem.

In a second scenario, accretion and reduction of substance occur at the opposite sides of an interrogative, as schematically illustrated in (6), where the substance is accreted on the right and reduced on the left end of the original interrogative *aaa*.

- (6) Accretion and reduction occur at the opposite sides of an interrogative
aaa → *aaa* **bbb** → *aabbb* → **abb** *ccc* → *bcc*

Evolving as in (6), the original interrogative may very quickly vanish without traces, which makes reconstruction really difficult. Only daughter languages with enough fine-grained variation involving reflexes of the original interrogative may allow us to reconstruct the latter. An example of reconstruction in such a situation is provided in Idiatov (2011) for the interrogative pronominals of Eastern Mayan languages, a very shallow linguistic group with extremely diverse forms of interrogative pronominals.

In a third and final scenario, accretion happens on the sides (left, right or both at the same time), while reduction takes place inside an interrogative, as in (7).

- (7) Accretion occurs on the sides and reduction inside an interrogative
aaa → **aaa** *bbb* → **aabb** → *ccc* **abb** → *ccabb* → *cbb*

In the scenario in (7), the original interrogative may vanish when trapped inside, which also complicates reconstruction. Like in the preceding scenario (6), this difficulty may be mitigated by the availability of many daughter languages with enough fine-grained variation in the form of reflexes of the original interrogative.

3.2 Semantic evolution

In (8), I provide an overview of the common pathways of semantic change of interrogative pronominals with particularly non-selective ones as the endpoints. Some of these pathways may lead to the emergence of a lack of differentiation between ‘who?’ and ‘what?’. Importantly, none allows for a direct change from ‘who?’ to ‘what?’ or the other way around, which is in line with the fact that neither change can be accounted for by any regular mechanism of semantic change. Note that I do not take into consideration here any possible interaction with gender-number marking or other nominalising elements.¹²

- (8) Common pathways of semantic change of interrogative pronominals
- a. ‘which one?’ > usually ‘who?’, occasionally ‘who?; what?’
 - b. ‘who?; what?’ > ‘who?’ or ‘what?’ when a new dedicated form of either NSIP emerges
 - c. ‘(be) where?’ > ‘(be) which one?’ > ‘which one?’
 - d. ‘(be) where?’ > ‘which [N]?’, ‘what (kind of) [N]?’
 - e. ‘(be) how?’ > ‘what (kind of) [N]?’
 - f. ‘(do) how?’ > ‘(do) what?’ (questions about actions)
 - g. ‘what (kind of) [N]?’ <> ‘which/what [N]?’
 - h. ‘which/what [N]?’ <> ‘which one?’
 - i. ‘what (kind of) [N]?’ > ‘(be) what?’, ‘(be) who? (classification, rather than identification)’
 - j. constructions based on a noun meaning ‘name’ or verbs meaning ‘do; say; be’, ‘name’, ‘call’ > ‘who?’, ‘what?’, ‘who?; what?’

¹²For example, when there is no interaction with gender marking, a selective interrogative pronominal ‘which one?’ tends to develop into a non-selective ‘who?’, not ‘what?’. In a sex-based gender system, the outcome depends on the semantics of the gender categories. Thus, ‘which one?’ marked for masculine gender often evolves in ‘who?’, while the same stem marked for neuter gender normally evolves into ‘what?’. Similarly, a NSIP specified for gender may become specialised as ‘who?’, ‘what?’ or both, depending on the organisation of the gender system. In a combination of an interrogative modifier ‘which [N]?’ or ‘what [N]?’ with a classifier, a deictic element such as a demonstrative pronominal or an article, a generic nominal, such as ‘person’, ‘thing’, ‘one’, the latter element not only nominalises the interrogative modifier but also contributes its own semantics which affects the possible pathways of semantic change.

4 Bantu NSIPs: Typological commonalities

In this section, I go through the formal (accretion in §4.1 and reduction in §4.2) and semantic (§4.3) changes affecting Bantu NSIPs and related interrogatives that are cross-linguistically common.

4.1 Formal evolution: Accretion

4.1.1 Overview

All the typologically common types of elements accreted in the diachrony of NSIPs cited in (4) in §3.1.2 are also attested in Bantu. In this section, I particularly focus on some of the regularities of accretion that have not enjoyed much attention in the literature so far. I do not elaborate much on the well-known accretion of class markers, such as the locative class markers with ‘where?’ subsequently inherited in any derived ‘which (one)?’ and ‘who?’ interrogatives (but see some examples in §4.1.5) and the class 7 marker with ‘what?’, as in Basaa A43a *kí(i)* ‘what?’ and Enya Kibombo D14 *kíikú* ‘what?’ (next to *ki-úmà nàání* ‘what?’, lit. ‘what thing?’).

Questions in Bantu, especially those about subjects, are often constructed as clefts, as in ‘it is who that did P?’ for ‘who did P?’, e.g. (9) for Makhuwa P31, or pseudo-clefts, as in ‘the one that did P is who?’ for ‘who did P?’, e.g. (10) for Mongo C61, with the notional predicate being topicalised and construed as a relative clause and the interrogative being focalised and construed as a nominal predicate.

(9) Makhuwa P31 (van der Wal 2009: 171, 172)

a. *ti paní o-tthik-ale erráncá?*
COP 1.who? 1-throw-PFV.REL 10.oranges

‘Who has thrown oranges?’ (lit. ‘It is who that has thrown oranges?’)

b. **paní o-n-áápéya nramá?*
1.who? 1-PRS.CNJ-cook 3.rice

‘Who cooks the rice?’

(10) Mongo C61 (Hulstaert 1965: 144)

ö-kelaki ná?
1.REL-did NSIP

‘Who has done it?’ (lit. ‘The one who has done (is) who?’)

For this reason, the accretion of substance in interrogatives often proceeds within (pseudo-)cleft structures.¹³ Related to this is the tendency to source accreted substance from various deictic forms and other forms that themselves are typically sourced from deictics, such as focus markers, copulas and relativisers. However, the exact (original) morphosyntactic function of the accreted deictic material is often difficult to establish with certainty (§4.1.2). An interesting detail is that accreted deictics in Bantu (and more broadly in Benue-Congo) appear to be preferentially sourced from deictics that are not distal, but rather intermediate or used discourse-referentially or intersubjectively (§4.1.3). Among the recurrent formal types, which I note with capital letters,¹⁴ we find N(D)I, I-, -TE, -O and -E (§4.1.4). Since SIPs and NSIPs, especially ‘who?’, in Bantu often develop from locative interrogatives, some of the accreted substance is inherited from such locative interrogatives, with the two most common types being PA- and (N)KA- (§4.1.5). Another common Bantu NSIP source is the Interrogative Modifier construction, combining a generic noun, such as ‘thing’, ‘person’, ‘place’, with an interrogative modifier, as in ‘what thing?’ for ‘what?’ or ‘which place?’ for ‘where?’ (§4.1.6). Such interrogatives may later undergo a *cosa*-type reduction. Finally, a few languages of zones D and J provide an example of the use of reduplication for accretion of interrogatives, which is cross-linguistically rather uncommon, but semantically transparent (§4.1.7).

4.1.2 Morphosyntactic function of the accreted deictic material

The exact (original) morphosyntactic function of the accreted deictic material is often difficult to establish with certainty. For example, in Liko D201, NSIPs are typically (and for questions about subjects, obligatorily) clause-initial and in that position they are “always followed by a demonstrative [of] type I” that agrees in class with the NSIP (de Wit 2015: 434: 434), as with *i-kí y-ǰ* [7-what? 7-DEM_I] and *wàní n-ǰ* [1a.who? 1-DEM_I] illustrated in (11).

¹³The use of pseudo-cleft structures may result in interrogative constructions where interrogatives are sentence-final, which typologically is particularly unusual (cf. Dryer 2013). This situation appears to be restricted to zone C (cf. Bokamba 1976; Idiatov 2009: 63–65). Due to frequency effects, such (pseudo-)cleft structures tend to become reduced to various extents (cf. §5.3).

¹⁴A formal type represented in capital letters is a schematic representation of a range of similar forms that recurrently appear as parts of interrogative pronominals, without necessarily being synchronically analysable as separate morphemes, and that are likely to (partially or fully) go back to the same morphosyntactic material diachronically.

(11) Liko D201 (de Wit 2015: 258)

wàní nǎ á-ly-á ndì nyàmá ní-nǎ?
 1a.who? 1.DEM_I 3SG.PST:1.OBJ-eat-FV.PST PST.REM 1a.animal COP-1.DEM_I
 ‘Who ate this animal?’

The question in (11) looks like a cleft construction with the subordinate clause introduced by a demonstrative used as a relativiser. However, the regular relativiser in Liko has the structure [ní COP + CL- DEM_I], as in *ní-nǎ* for class 1 and *ní-yó* for class 7, which is identical to the modifying use of the demonstrative of type I, also illustrated in (11), with the only difference that the copula is optional in the modifying demonstrative and obligatory in the relativiser. As explicitly noted by de Wit (2015: 434) the copula is not allowed in clause-initial NSIPs which precludes the synchronic analysis of the demonstrative in the NSIP construction as a relativiser. It also does not make sense to analyse it as a modifier of the NSIP. Morphosyntactically, this demonstrative is best analysed as a pronominal in apposition with the NSIP, with which it also agrees in class. This obligatory use of the demonstrative of type I with clause-initial NSIPs resembles the information-structural use of *ça* ‘that one’, typically a discourse-referential deictic, as a kind of insistence marker in French interrogatives, such as *qui ça?* ‘who? (tell me!)’, *où ça?* ‘where? (tell me!)’, *comment ça?* ‘how? (tell me!)’, mentioned regarding NSIPs in Louisiana Creole in §3.1.3 above.

4.1.3 Endophoric deictics and distance distinctions in exophoric deictics

Accreted deictics in Bantu (and more broadly in Benue-Congo) appear to be preferentially not distal, or at least not the most distal within the deixis system of a given language. In systems with more than two distance distinctions, it is often the intermediate distance deictic or the deictic pointing to a location closer to the interlocutor that is recruited for NSIPs. For example, compare Bira D32 *éké* ‘what?’ and Komo D23 *ékéndò* ‘what?’, where the Komo form has accreted the intermediate distance demonstrative *ndò* (Constance Kutsch Lojenga, p.c.). In Basaa, the SIP ‘which one?’ may be constructed with the near-addressee demonstrative pronominal (instead of the regular noun class marker), as in *hû-mbéé* [19.this_one_closer_to_you-which] ‘which one? (class 19)’ (Bôt 1986: 68). The accreted deictic often also has some endophoric or discourse-referential uses, and more broadly intersubjective uses in coordinating the attention of the speaker and addressee to objects and places (cf. Evans et al. 2018: 123–134 on the relevance of the intersubjective use in the typology of demonstrative systems). The Liko demonstrative of type I accreted with the clause-initial NSIPs as mentioned

in §4.1.2 can be used as an example here. It is different from both the proximal and distal demonstrative, although used as a building block for the latter. In origin, it seems to be an intermediate distance deictic or deictic pointing to an object closer to the interlocutor. Importantly, demonstratives of type I are “often used for text-internal reference or for the activation of a participant in a text”, or when “it is not relevant to indicate whether the referent is present or not [at the] site of the speech act” but just to draw the attention of the interlocutor to it (de Wit 2015: 256–257).

A similar formal link between (non-selective) interrogatives, on the one hand, and non-distal, discourse-referential and intersubjective demonstratives, on the other hand, is found in the wider Bantoid domain (see for some examples the Appendix A). In fact, this link may just reflect a general cross-linguistic tendency.¹⁵

4.1.4 Some recurrent formal types of accreted deictic material

Across Bantu, there are a number of recurrent formal types that appear as accreted material on non-selective interrogatives and that are likely to have a deictic origin, such as N(D)I (§4.1.4.1), I- (§4.1.4.2), -TE, -O and -E (§4.1.4.3). However, in practice, it is often difficult to decide which of the possible deictic sources was involved in each particular case and whether it was accreted as a deictic form or as one of the forms typically derived from deictics, such as copulas, relativisers and focus markers.

4.1.4.1 Type N(D)I

Type N(D)I is attested throughout Bantu and probably inspired Guthrie’s (1971) reconstructions of ‘which?’ and ‘what?’ cited in (2) and Meeussen’s (1967) reconstruction of ‘which?’ cited in (3). See Idiatov (2009: 70) on forms such as Duala A24 *we(ni)* ‘where?’ or Punu B43 *ave(ni)* ‘where?’, where I draw attention to the fact that the demonstrative stem *ni* is well-attested across Bantu.¹⁶ Also recall the Liko copula *ní*, suspiciously similar to the second syllable of *wàní* ‘who?’

¹⁵For instance, in Russian the NSIPs often combine with the neuter proximal demonstrative *èto*. Also recall the Louisiana Creole *sa* (§3.1.3) which reflects either the French neuter demonstrative *ça* or the homonymous proximal adverbial demonstrative *çà* ‘here’. Etymologically, the former French neuter demonstrative *ça* is a distal demonstrative, but synchronically it is non-specified for distance and is rather used discourse-referentially as anaphor or intersubjectively to attract attention to something.

¹⁶It is particularly widespread in zone C (Claire Grégoire, p.c.). We find similar forms as far as Jarawan Bantu, such as the Mbula anaphoric determiner *nì ~ ì*.

and *yáni* ‘where?’.¹⁷ Across Bantu, similar accreted material is found on the NSIPs for ‘who?’ and ‘what?’ (see Appendix B) and on substitutives (cf. Kamba Muzenga 2003). Besides the demonstrative stem *ni* mentioned above, the two major sources of type N(D)I are reflexes of other deictic(s) and (identificational, presentative, ascriptive) copulas (aka nominal predicative markers). In the interrogatives with N(D)I accreted on the left side, most likely N(D)I- functioned as a copula (cf. the evidence provided by Givón 1974), reflecting the copulas **ní* ~ **ní* and/or **ndí* ~ **ndí*.¹⁸ On the right side of the interrogatives, the range of possible source functions of -N(D)I is more diverse. The more common sources in this case are likely to be a pronominal or demonstrative used for information-structural purposes or as a relativiser (cf. §4.1.2), although a copula function is also possible if the presentative copula construction in a given language used to be [N COP] ‘N it is’ rather than [COP N] ‘it is N’. Besides the demonstrative stem *ni*, two other promising etymons are the (possessive) pronominal stem of class 1 **ndi* ~ **ndi* (cf. Kamba Muzenga 2003: 48) and BGR’s pronominal and verbal prefix of class 5 **dí*. In this respect, note that the latter prefix may need to be reconstructed with a nasal-consonant cluster as **ndí*, given that class 5 prefixes may have the shape *nV-* in languages as diverse as Orungu B11b, Nen A44 and Kenyang [keny1279, Mamfe].

4.1.4.2 Type I-

Type I- is relatively well attested throughout Bantu. It is mostly found with ‘who?’ interrogatives. The form is typically *í-*, as in Tsogo B31 *índa*, Pinji B304 *inde*, Kagulu G12 CL-(*i*)*hoki* ‘which (one)?’ and (*i*)*yehoki* ‘who?’ (in addition to a presumed Swahili G42d loan *nani* ‘who?’) both derived from *hoki* ‘where?’, Nande JD42 (*i*)*ndi*, Hunde JD51 *inde* (class 1) / *bānde* ~ *bēnde* (class 2), Rwanda JD61 ‘who?’ (*i*)*ndé*, or just a H tone on the initial nasal *n̄-* as in Ntomba-Inongo C35a *n̄nò* ‘who?; what?’ and Salampasu L51 *n̄ny* ‘who?’.

¹⁷Tonally, *yáni* ‘where?’ (*yá* ‘towards, in the direction of’ + *àní*) matches perfectly with *kéki* ‘why?’ which combines the preposition *ká* and *i-kí* ‘what?’. The interrogative stem *àní* itself must originate in a locative interrogative ‘where?’, similar to Kagulu CL-*ani* ‘where?’ (Petzell 2008: 89–92, 177), with -N(D)I accreted on the right like in the Duala and Punu forms mentioned above. The Liko NSIP ‘who?’ also provides an example of a regular semantic evolution of ‘where?’ to a SIP and subsequently a NSIP. Compare also Ndaka D301 *àní* ‘who?’ and *imáni* ‘what?’, where **imá* reflects BLR3 **jómà* ‘thing’ (cf. §4.1.6) and *àní* ‘who?’ is from earlier ‘which (one)?’. It is most likely that such ANI interrogatives result from some earlier substance reduction on the left. Thus, in Idiatov (2009), I suggest that the left-sided material minimally included the locative marker of class 16 **pa*. One other well-attested option is the type (N)KA- that I discuss in §4.1.5, which is sourced from **ka* ‘be at (X’s place)’ (cf. also Appendix F).

¹⁸See also Appendix C on the possible copula stacking in the forms with the initial *nd-* cluster.

The brackets around the type I- accreted material highlight the fact that it is often subsequently reduced. It is best preserved in two environments. First, it may merge with a class prefix by changing the vowel quality of the latter, as in Kagulu (i)yehoki (class 1) and Hunde bēnde (class 2). Second, the vowel of type I- runs less risk of being elided and its H tone of being delinked and deleted in utterance-initial position, which may lead to a divergent evolution of the interrogative form in different positions. For example, in Rwanda ‘who?’ is usually *ndé*, however some speakers also have the form *indé* but only in the beginning of an utterance (Jacob 1984–87, via Bastin et al. 1999). A somewhat different example is provided by Mongo, where interrogatives are never sentence-initial. In questions about subjects (and optionally in those about objects), they are sentence-final (Idiatov 2009: 63–65). Interestingly, in the Nkundo variety of Mongo, the reference dialect of Hulstaert (1957), the sentence-initial polar question marker *nà* is very similar to the non-selective interrogative pronominal *ná* ‘who?; what?’. As I suggest in Idiatov (2009: 71), the polar question marker and the interrogative pronominal likely result from a divergent evolution of the same interrogative in different constructions (compare A70 ‘who?’ interrogatives discussed in §5.3.2). The HL tone must reflect the older tone pattern of this interrogative in Mongo, where the initial H tone is a type I- accretion. In this respect, compare Doko C301 *ndâ* ‘who?’ and *-ndâ* ‘which/what N?’. Interestingly, in a number of other Mongo varieties, the sentence-initial polar question marker has the form *yà*, which is similar to the dialectal variant CL-*yá* of the interrogative modifier ‘what/which [N]?’. The two forms differ in exactly the same way as *nà* and *ná* in Mongo-Nkundo, as well as *ndâ* and *-ndâ* in Doko. With respect to *yà* and CL-*yá*, note also the human interrogative pronominals of Bwamba C10 *yá* ‘who?’, Libobi C412 *ya* ‘who?’, Chokwe K11 *i-ya ~ a-ya* ‘who?’, Mbunda K15 *íyà* ‘who?’ (cf. §5.2.4).

As can be observed above, type I- often accretes on forms already containing other accreted material, especially that of type N(D)I-. While N(D)I- accreted on the left side most likely functioned as a copula, Type I- is more likely to derive from a nominaliser, such as the element often referred to as ‘augment’ in Bantu linguistics (cf. de Blois 1970, and Van de Velde 2017 on the augment in A70 languages). This nominaliser augment origin is most clear in forms such as CL-(i)hoki ‘which (one)?’ and (i)yehoki ‘who?’ in Kagulu, where *i-* is synchronically the augment (‘initial vowel’) of classes 1, 4, 5, 7, 8, and 9 (Petzell 2008: 49), while it is absent from the source locative interrogative *hoki* ‘where?’. In fact, in (i)yehoki ‘who?’, the nominaliser augment is present twice, viz. the optional initial *i* and merged with the vowel *a* of the class 1 subject prefix as *e*. It should be mentioned that type I- also bears strong formal resemblance to the variants **í ~ *i* and **^H* (a floating high tone) of the so-called nominal predicative marker

‘it is [N]’, whose other variants are **ní* ~ **ní* and **ndí* ~ **ndí* (cf. Givón 1974; Grégoire 1975: 125; Coupez 1977).¹⁹ However, the data from languages such as Kagulu make a copula origin of type I- less plausible. In Idiatov (2009: 71–72), I also hypothesised that type I- could go back to the pronominal or subject prefix of class 9 *(*j*)*l-*, (locative) class 24 **l-* or class 7 **kl-*. Here, I propose a different source, the pre-PB determiner **yé*, discussed in §6.1.5.3, which better matches tonally and semantically and has more coherent cognates beyond Narrow Bantu.

The rising LH tone pattern in forms such as Hunde *inde* suggests that besides the type I- accreted material represented by the H tone, such interrogatives also contain some additional accreted material on their left edge whose trace is the initial L tone. It is again the Kagulu data that provides clear indications on the origin of this L-toned element as the class 1 subject marker and ultimately a form of the verb ‘be’ (a locative copula) fused with the class 1 agreement prefix. Thus, in Kagulu (*i*)*yehoki* ‘who?’ the class 1 prefix is the “non-past and non-perfective subject marker” *ya-* (and not the pronominal prefix *yu-*) (cf. Petzell 2008: 90, 101), which further accounts for the addition of yet another nominaliser augment, viz. the optional initial *i*. As illustrated in (12) with the class 10 agreement, *ya-* derives from an inflected form of the locative copula *-a* fused with the class 1 agreement prefix.²⁰ That is, the interrogative (*i*)*yehoki* ‘who?’ is structurally a nominalised predication, as overtly marked by the augment, literally meaning something like ‘the one that s/he is the one where?’ (*i*)*y-e-hoki* [(NMLS)-1-be:NMLS-where].

(12) Kagulu G12 (Petzell 2008: 178)

sa hoki

si-a hoki

10-be where?

‘Where are they (class 10)?’

As discussed in §5.3.2, the earlier presence of the H tone of the augment, as a nominaliser or a construct form marker, can also account for the generalisation of the H tone forms *zá* ‘who?’ and *já* ‘what?’ in Eton A71, and similar H-toned forms in other A70 varieties, as well as in A15 and A40 languages.

¹⁹Although the copula type N(D)I, viz. **ní* ~ **ní* and **ndí* ~ **ndí*, and the copula type I, **i* ~ **i* and *^H, traditionally seem to be considered allomorphs, I believe that historically they are not related (see Appendix C).

²⁰In fact, subject prefixes originating in the inflected forms of the copula *-a* appear to be rather common in Eastern Bantu. See Appendix D for various examples from zones D, E, G, N and P.

4.1.4.3 Types -TE, -O, and -E

Another relatively recurrent type of accreted material of deictic origin is -TE, which seems to be absent from Eastern and South-Western Bantu and functionally restricted to ‘what?’ interrogatives, as in Nen A44 *yătè* ‘what?’, *yătè* N ‘what kind of N?’, Maande A46 *àáté* ‘what?’, Shake B251 *índè ~ íntè* ‘what?’, Boa C44 *tě* ‘what?’. The type -TE may have been sourced from an anaphoric demonstrative, such as Eton CL-*tə* (Van de Velde 2008a: 146–148), Ewondo A72a CL-*tă* (Abessolo Nnomo & Etogo Mbezele 1982: 185) and Aghem [aghe1239, West Ring Grassfields] CL-^L*té^H* CL-*ś* ‘the one in question, the one you and I know about or have been talking about’ (Hyman 1979: 40).

It is possible that the same type occurs in forms such as Nkucu Wela C73 *nàtó* ‘what?’ (one of the forms) and Kele Yawembe C55 -*tò* ‘what?’. Their final *o* may be due to the further accretion of the common Bantu ‘*o* of reference’ (Dammann 1977), i.e. BGR’s substitutive stem *-*o*, to which I refer as type -O.²¹ Other possible examples of the type -O accretion include ‘who?; what?’ in zone C, such as Ntomba-Inongo C35a *ńńɔ* and Bolia C35b *ńɔ* (cf. Idiatov 2009: 72).

Finally, type -E is manifested in final vowels of mostly ‘what?’ and sometimes ‘who?’ interrogatives in zones A, B and C, such as in Duala A24 *njé* (vs. *njá* ‘who?’), Noho A32a *njáe* (vs. *njani* ‘who?’), Basaa A43a *njé(é)* ‘who?’ (but see §5.2.2 on its use as a stem for ‘what?’), Tsogo B31 *inde* (vs. *índa* ‘who?’), Koyo Ehamba C24 *nde* (vs. *nda* ‘who?’), Balobo C314 *ndé* (vs. *ndá* ‘who?’), Motembo C371 *nde* (vs. *nda* ‘who?’), and Bwamba C10 *yé* ‘what?’ (vs. *yá* ‘who?’). Recall the forms for ‘who?’ in A15 cited in §2.1, such as Akoose *nzé(é)* and Mwahed *nzé* (vs. Mwaneka *nzá*, Mkaa *njá*). Also compare C30 varieties Gyando (Ngiri) *ye* ‘what?’ and Doko *yó* ‘what?’, where type -E seems to alternate with type -O. Type -E may originate in BGR’s substitutive stem *-*e*, another deictic stem, or be a reduction of type -TE.

4.1.5 Accreted material inherited from locative interrogatives

In Bantu, following a common cross-linguistic path of change, locative interrogatives often develop into SIPs and subsequently into NSIPs, especially ‘who?’. For this reason, accreted substance is often inherited from such locative interrogatives. The most common and transparent type here is PA-, which is sourced from the class 16 marker **pa-*, in forms such as Makhuwa Nampula P31 *páni* ‘who?’ and Kagulu CL-(*i*)*hoki* ‘which (one)?’ and *yehoki* ‘who?’ (see Appendix B).

²¹Alternatively, the forms with the back rounded vowel may reflect **ntw* ‘some (entity), any’ (BLR 4807), which has reflexes meaning ‘thing’. In this respect, see §4.1.6.

Doneux (1971: 134–135) reports that in languages of zone J locative interrogatives meaning ‘where?’ are often accreted with *nka-* and *-na*.²² While there are no such clear examples of NSIPs involving *-na*,²³ both NSIPs and SIPs that are likely to contain (N)KA- are more widespread.²⁴ Some particularly clear examples are found in zones C and L: Babanda C44 *kàní* ‘who?’, Boa Buta C44 *kàné* ‘who?’, Luba-Kasai L31a *ɲanyì* ‘who?’ and *ci-ɲanyì* ‘what?’ (class 7). Another possible example is Lower Pokomo E71B *ga ~ gá* ‘who?’, also used as an interrogative modifier of ‘thing’ in the construction for ‘what?’ (see §4.1.6).²⁵

The use as ‘where?’ as in zone J is clearly the source of the NSIPs and SIPs with reflexes of (N)KA-. To begin with, this is suggested by the typical paths of semantic change of interrogatives from ‘where?’ to ‘which one?’ and further to ‘who?’ and ‘what?’ rather than the other way around (see §3.2). Furthermore, if we assume a locative source, we can propose a coherent etymology for (N)KA-, plausible both semantically and formally. Thus, (N)KA- was transparently sourced from **ka* ‘be at (X’s place)’.²⁶

4.1.6 NSIPs instantiating the Interrogative Modifier construction

It is common cross-linguistically for NSIPs to be construed as nominal expressions based on generic nouns, such as ‘thing’, ‘person’, ‘place’, and an interrogative modifier, as in ‘what thing?’ for ‘what?’ or ‘which place?’ for ‘where?’. With other nouns the same interrogative modifier may have primarily selective semantics (‘which [N]?’), a variety of non-selective semantics (‘what [N]?’, ‘what kind of [N]?’), or be largely indifferent to this distinction (such as French *quel [N]?*).²⁷

²²Doneux (1971) does not discuss any possible sources. I argue that *-na* most likely goes back to the intermediate deictic stem **ná* (see Appendix E) and *nka-* to a form of **ka* ‘be at (X’s place)’ (see Appendix F).

²³However, see §4.1.6 on a number of ‘what?’ interrogatives ending in *-(i)na*, where a different etymology is more likely.

²⁴This also applies to ‘where?’ interrogatives, such as Northern Sotho S32 *kae* (Poulos & Louwrens 1994) and Mongo C61 *nkó* (Nkundo), *ńkó*, *ńkò*, *nká* and *nké* (other varieties) (Hulstaert 1957; 2007: 290).

²⁵As highlighted in footnote 17, (N)KA- plausibly also formed the initial part of some *ani*-like forms for ‘where?’, ‘which (one)?’, ‘who?’ and ‘what?’ in Eastern Bantu (with the locative class 16 **pa* being another plausible candidate).

²⁶See Appendix F for a discussion of the etymology of the accreted element (N)KA-.

²⁷Descriptions of individual Bantu languages often remain vague with respect to the semantics of the interrogative modifier and rely exclusively on the translational equivalent. Thus, descriptions in French often use the translation *quel*, which is indifferent to the distinction between selective and non-selective semantics, while descriptions in English often use *which*, which is selective by default, but may also be used non-selectively. I suspect that in many cases we indeed deal with real semantic ambiguity, as may be confirmed by (contextualised) sentential examples when they are provided.

NSIPs instantiating the Interrogative Modifier construction are attested throughout Bantu. However, they remain relatively infrequent. To some extent, this is likely to be due to the presence of a rich noun class system in which noun class markers can function similarly to nouns with generic semantics. In Bantu, such NSIPs mostly mean ‘what?’ and somewhat less frequently ‘where?’. I was able to identify two to three nominal stems on which such interrogatives are based.

The first stem is **ntò* ‘some (entity), any’ (BLR 4807) which has reflexes meaning ‘thing’ in class 7, ‘place, somewhere’ in the locative classes 16, 17 and 18, and ‘person, somebody’ in class 1 (cf. Grégoire 1975: 137–138). Despite this range of possible meanings, I found it only as a part of ‘what?’, such as in Kwange D102 *ki-ntù nàání* ‘what?’ and Lower Pokomo E71B *kinthu ga ~ ki-ntú-gá* ‘what?’ (see also §4.1.4 on some less clear examples in the east of zone C).²⁸

The second well-attested nominal stem is **jómà* ‘thing; bead; iron’ (BLR 3619), whose reflexes can mean ‘thing’, ‘place’ or ‘person’ depending on the noun class, and convey a number of more specific meanings, such as ‘bead’, ‘iron’, ‘belongings’ (cf. Grégoire 1975: 139–142). This stem is found in ‘what?’, ‘where?’ and ‘who?’ interrogatives. Thus, we find Babole Bakolu C101A *zumba nza*, Enya Kibombo D14 *ki-úmà nàání* ‘what?’, Enya Manda D14 *ki-úmà nàání* ‘what?’, Ndaka D301 *ímání* ‘what?’, Lunda L52 *yumanyi* ‘what?’.²⁹ Other such ‘what’ constructions have undergone the *cosa*-type reduction (see §3.1.3), as in Bodo D308 *è má*, Kukuya B77a *ki-má*, Fumu B77b *ima*, Teke Laali B73b *ímá ~ kii-ma*. ‘Where?’ interrogatives involving **jómà* occur in Fang A75 *vom ave*, Lundu A11 *oe oma*, and Kele C55 *ánima*, where *vom*, *oma* and *áma* respectively mean ‘place’ (Grégoire 1975). Finally, in B10 and Eastern and South-Western Bantu languages, we also find ‘who?’ interrogatives involving **jómà*. In B10, they are transparently based on class 1 reflexes of **jómà* meaning ‘person’, as in Mpongwe B11a (Raponda-Walker 1934) *o-ma* ‘person’, *mande* ‘who?’ next to *oma ande* ‘what person?; who?’, *ande* ‘what?; what (kind of) [N]?’ and Orungu B11b (Ambouroué 2007) *ò-má* ‘person’, *méndè* (after a low tone: *mèndè*), *ò-má ándè* ‘what person?; who?’, *ándè* ‘what?;

²⁸Both *nàání* in Kwange *ki-ntù nàání* ‘what?’ and *ga ~ gá* in Lower Pokomo *kinthu ga ~ ki-ntú-gá* ‘what?’ mean ‘who?’ on their own. These two uses illustrate the typical evolution from the SIP ‘which (one)?’ (person or thing) to the NSIP ‘who?’ (cf. §4.3, §5.2.4, §4.1.5).

²⁹Note that just like in the Kwange and Lower Pokomo forms above, the interrogative elements in Babole Bakolu, Enya and Ndaka ‘what?’ also mean ‘who?’ when used nominally on their own, viz. Babole Bakolu *nza* ‘who?’, Enya Kibombo *nàání* ‘who?’, Enya Manda *ní-nàání* ‘who?’, Ndaka *ání* ‘who?’, and illustrate the same type of evolution.

what (kind of) [N]?’.³⁰ According to BLR3, reflexes of **jómà* meaning ‘person’ are restricted to zones A and B. In contrast, the comparable Eastern and South-Western Bantu ‘who?’ forms, such as Tswana S31 *máng*,³¹ must be derived from an earlier selective ‘which one?’ indifferent to the distinction between persons and things, and ultimately from a locative ‘where?’, based on reflexes of **jómà* meaning ‘place’, not ‘person’ or ‘thing’. This is suggested by the possibility to use such ‘who?’ interrogatives in questions about non-personal proper names, such as toponyms or names of species of flora and fauna, and as interrogative modifiers ‘what kind of, what [N]?’ equally indifferent to the distinction between persons and things (cf. Idiatov 2009: 66–67, 69). Such uses cannot be accounted for if we take the original meaning of these interrogatives to be ‘who?’.³²

Finally, a number of ‘what?’ interrogatives in zones C, D and J may instantiate the Interrogative Modifier construction involving *(*j*)*íná* ‘thing’ attested in zones B, D and R (cf. Meeussen 1967: 103; Grégoire 1975: 142), which subsequently underwent the *cosa*-type reduction. Such ‘what?’ interrogatives are Beo C45A and Ngelema C45 *etina*, Komo D23 *sínà*, Bukusu JE31c *sínà*, Kisa JE32D *sina* ~ *shina*, Isukha JE412 *shiina*, Samia JE34 *sina*. Alternatively, the (*i*)*na* part may also represent a reduction of the same SIP that resulted in the ‘who?; what?’ NSIP in some languages of zone C, such as Mongo C61 *ná* (see Idiatov 2009), and the (modifying) interrogative stems that can have both a selective and non-selective reading, such as Doko C301 *-ndá* ‘which/what N?’.

4.1.7 Reduplication

Doneux (1971: 134–135) reports that in a number of languages of zone J locative interrogatives ‘where?’ have been accreted through reduplication. I found only a few examples of accretion through reduplication with interrogative pronominals in zones D and J, such as Enya Kibombo D14 *kiíkíí* ‘what?’ (next to *kì-úmà nàání* ‘what?’, lit. ‘what thing?’) and Ziba JE22D *-kí(kí)* ‘what?’. Reduplication for accretion of interrogatives is somewhat unusual typologically, but it is easy to account

³⁰Interestingly, at least Mpongwe must have had another reflex of **jómà* in class 5 that meant ‘thing’. This is suggested by the fact that Mpongwe also has a placeholder word of class 5 *mande* ~ *mamande* ~ *mande-mande* ‘whatchamacallit’ that is used exclusively to refer to things or places whose name escapes one’s mind at the moment of speaking (Raponda-Walker 1934).

³¹Other such ‘who?’ forms including Bhele D31 *mání*, Luguru G35 *mani*, Pende L11 *mapi*, Tswana S31 *máng*, Southern Sotho S33 *máng*, Nkuna S53D and Luleke S53A *maní*, Tswa S51, Tsonga S53 and Ronga S54 *mání*, Konde S54 *má(ni)* are examples of *cosa*-type reduction.

³²Following the same line of reasoning, we can equally exclude the possibility that the initial *m*- in these ‘who?’ interrogatives results from a reduction of the class 1 prefix *mɔ-* > *mw-* > *m-*.

for by information-structural uses of reduplication for meanings such as ‘really X’, ‘exactly X’, ‘X and nothing else’ (where X is the reduplicated element).³³

4.2 Formal evolution: Reduction

As is common cross-linguistically, reduction of substance with interrogatives in Bantu is largely irregular. Recall the forms for ‘who?’ in the A15 varieties cited in §2.1, which illustrate this point well. The *cosa*-type reduction (cf. §3.1.3) is also attested in Bantu (see §4.1.6 for some examples). Regarding the interaction between reduction and accretion of substance in interrogatives presented in §3.1.4, my impression is that overall the evolution of NSIPs in Bantu is best represented by the scenario schematised in (7) above, with accretion on the sides and reduction inside, although with a certain preference for accretion on the right. The right side of interrogatives appears to be generally more stable in non-North-Western Bantu, especially in Eastern Bantu, which matches more general morphological and phonological patterns in that North-Western Bantu languages often have maximality constraints on stems. These are generally absent elsewhere, while in Eastern Bantu we sometimes observe the opposite situation with minimality constraints on stems.

4.3 Semantic evolution

The two most common semantic pathways of change at the origin of interrogative pronominals in Bantu are: (i) ‘(be) where?’ > ‘(be) which one?’, ‘which [N]?’ , ‘what (kind of) [N]?’ > ‘which one?’ resulting in SIPs; and (ii) ‘which one?’ > ‘who?’ resulting in human NSIPs. Both are also very common cross-linguistically (cf. §3.2).

The change ‘which one?’ > ‘who?’ is for example reported for the languages of zone J by Doneux (1971). The same evolution must have taken place in those numerous cases where ‘who?’ corresponds to an interrogative modifier ‘which/what [N]?’ indifferent to the distinction between persons and things. Some examples are provided in §4.1.6. Compare also Libobi C412 *ya* ‘who?’, Chokwe K11 *i-ya* ~

³³Cross-linguistically, reduplication of interrogatives seems to be more typical for echo-questions. Thus, in Russian we can have an echo-question with the reduplication of *čego*, the genitive form of *čto* ‘what?’, that expresses a nuance of disbelief *Čego-čego on skazal?* ‘What did he say exactly? (Have I really heard what you say he said?)’, while a more neutral echo-question would either use the non-reduplicated genitive form or the non-reduplicated accusative form *čto* ‘what?’. The latter accusative form is also the normal form in regular, non-echo-questions about objects.

a-ya ‘who?’, Mbunda K15 *íyà* ‘who?’ and the dialectal variant CL-*yá* of the interrogative modifier ‘what/which [N]?’ in Mongo. The change from ‘(be) where?’ to ‘which one?’ or ‘which/what [N]?’ can be illustrated with Akoose A15C *héé* ‘where?’ and CL-*héé* ‘which (one)?’, Mongo C61 *nkó* ‘where?’ and CL-*lé nkó* [CL-COP where?] ‘which (one)?’ (lit. ‘the one that is where?’), Kagulu G12 *hoki* ‘where?’ and CL-(*i*)*hoki* ‘which (one)?’ (see also Doneux & Grégoire 1977: 191–192).

Since these two common pathways of change share the selective interrogative step, the output of (i) can obviously be the input for (ii), resulting in an evolution from ‘where?’ to ‘who?’. A particularly transparent example is provided by Kagulu (Petzell 2008: 89–92, 177), where we have both *hoki* ‘where?’ and CL-(*i*)*hoki* ‘which (one)?’, *yehoki* ‘who? (class 1)’, *wehoki* ‘who? (class 2)’. A more common situation is where the original locative origin of ‘who?’ has been masked by subsequent changes, but can be traced back thanks to both language-internal and comparative evidence, as illustrated in §4.1.5 and §4.1.6. Thus, besides formal evidence, such as the frozen locative class 16 prefix in Makhuwa Ile P31 *pání* ‘who?’ and Giryama E72a *hani* ‘who?’, the reconstruction of the locative or selective origin of ‘who?’ is facilitated by the fact that often the same interrogative concurrently evolves into an interrogative modifier ‘which/what [N]?’ indifferent to the distinction between persons and things, or into the stem of the non-human interrogative ‘what?’, two uses that cannot be accounted for if we take the original meaning to be ‘who?’.

5 Bantu NSIPs: Typological oddities

5.1 Overview

The oddities of a system, such as unnatural or lexical conditioning for allomorphs in morphology or unusual combinations of meanings for semantics, are most telling for the purposes of internal reconstruction. In this section, I highlight two of the major types of peculiarities of NSIPs across Bantu and the implications for their reconstruction, especially ‘who?’. The first type (§5.2) pertains to the surprising patterns of colexification of ‘who?’ and various interrogatives that are either non-human, such as ‘what?’, or indifferent to the difference between humans and things, such as ‘which/what [N]?’ which imply that such ‘who?’ constructions originate in selective and locative interrogatives. The second type (§5.3) pertains to the tendency to construe interrogative pronominals, especially those questioning subjects, as nominal predicates, because they have their source in clause-level constructions of the cleft type. Given the natural correlation be-

tween subjects and agentivity, in the long run the effect of this tendency is most noticeable with the human interrogative ‘who?’.

5.2 Colexification of human and non-human interrogatives

5.2.1 Lack of differentiation between ‘who?’ and ‘what?’ in zone C

As I discuss in detail in Idiatov (2009), a number of languages in zone C have NSIPs used as both ‘who?’ and ‘what?’, such as Mboshi C25 *ndè ~ nê*, Mongo-Nkundo C61 *ná*, varieties of Tetela C70 *nâ*, Ntomba-Inongo C35a *ínò* and Bolia C35b *ńò*. At least in Mongo-Nkundo, there is also a rare dedicated non-human NSIP, viz. *é* ‘what?’.³⁴ We can multiply such examples if we take into consideration cases where one and the same form means ‘who?’ in one language, but ‘what?’ in another. For example, we have Ligendza C414 *ndá* ‘who?’ and Buja C37 *ndá* ‘what?’. All these ‘who?; what?’ interrogatives are supposed to be reflexes of the BGR form **n(d)ái* ‘who?’.

5.2.2 ‘who?’ as the stem for ‘what?’

In a number of languages, we find ‘who?’ interrogatives corresponding to the BGR form **n(d)ái* ‘who?’ used as the stem for ‘what?’ in combination with the class 7 prefix, e.g. Mwani G403 *náni* ‘who?’ and *ki-náni* ‘what?’, Luba-Kasai L31a *nyani* ‘who?’ and *ci-nyani* ‘what?’ (Kabuta 2006), Nyasa N31D *yani* ‘who?’ vs. *ci-yani* ‘what?’. A slightly more complex example is found in Basaa A43a, where we have *njé(é)* ‘who?’ vs. *kí(i)* ‘what?’, but also *kí.njé(é)* ‘what?’, additionally used as a modifier ‘what kind of [N]?’ (Moreton & Bôt Bá Njock 1975: 372, 468; Bôt 1986: 66) (see also §5.2.4 and §5.3 below). The complication here is that synchronically the class 7 prefix in Basaa is not *ki-*, but zero or *y-* as a noun prefix and *í-* or *y^H-* as an agreement marker. Finally, see §4.1.4.3 above on the type -E accretion in zones A, B and C that very often appears to derive ‘what?’ interrogatives from ‘who?’ interrogatives corresponding to the BGR form **n(d)ái* ‘who?’.

³⁴Remarkably, *é* in Mongo-Nkundo can also mean ‘where?’ with motion verbs as an equivalent of the regular locative interrogative *nkó*. Such a colexification pattern is very unusual and probably due to the accidental merger of two interrogatives based on the same interrogative stem ‘what?’: one marked by class 7 **kɪ-*, as typical for ‘what?’ interrogatives, and the other one by locative class 17 **kɔ-*. The class 7 prefix in Mongo is zero with vowel-initial nominal stems and *e-* elsewhere. There is no more class 17 in Mongo, but its reflex would be expected to be *o-* or zero with the same distribution as class 7 (Grégoire 1975: 126–128).

5.2.3 ‘who?’ as ‘(be) what?’ about a name of a person or thing

In a number of languages, ‘who?’ is used as ‘(be) what?’ in questions about both personal proper names and non-personal proper names, such as toponyms or names of species of flora and fauna. As I illustrated in Idiatov (2009: 69), this use is found with ‘who?’ in Ligendza *ndá*, which is supposed to be a reflex of the BGR form **n(d)ái* ‘who?’, and Tswana *máng*, the univerbation of an interrogative construction literally meaning ‘which/what place?’ (cf. §4.1.6 above).

5.2.4 ‘who?’ as the interrogative modifier ‘which/what [N]?’

In a number of languages, the same form is used for ‘who?’ and for the interrogative modifier ‘which/what [N]?’ with human and non-human nouns. For example, recall the ‘what?’ interrogatives instantiating the Interrogative Modifier construction with the noun ‘thing’ discussed in §4.1.6.³⁵ I do not know whether ‘who?’ in these languages can also be used in the Interrogative Modifier construction with nouns other than ‘thing’. Synchronically more productive uses can be illustrated with Tswana *máng* ‘who?’ and [N] *máng* ‘what kind of, what [N]?’ (Idiatov 2009: 66–67), Basaa *njé(é)* ‘who?’ vs. *njé(é)* [N] ‘which/what [N]?’ (Hyman 2003),³⁶ and Akoose *nzé* ‘who?’ vs. *nzé* [N]\H-é ‘what/which [N]?’³⁷ (Hedinger 2008).³⁸ In some cases, the difference may be only tonal, as in Doko C301 *ndá* ‘who?’ and *-ndá* ‘which/what [N]?’. Given that the comparative evidence clearly suggests that this particular interrogative, presumably a reflex of the BGR **n(d)ái* ‘who?’, used to have a more complex structure, the tonal difference may be due to a divergent evolution of the earlier complex tonal pattern in pronominal and modifying uses respectively. See also §5.3.2 below on the modifying use of ‘who?’ in A15, A40 and A70 languages, which simultaneously demonstrates the divergent tonal evolution and the gradual simplification of a biclausal cleft construction into a monoclausal construction. However, in some cases the tonal differences may also be due to additional nominalising morphology in the interrogative pronominal ‘who?’, as in Mbula [mbul1261, Jarawan Bantu] *yá* [N] ‘which/what [N]?’ vs. *yá ná* [^H-yà ná ~ V-yà ná] [NMLS-which? COP.PRES] ‘who is it?; who?’ (cf. §2.1), where the H tone is likely to come from a nominaliser that otherwise appears to be restricted to deictics.

³⁵See also footnotes 27 and 28 above.

³⁶See also §5.2.2 above and §5.3.2 below.

³⁷In this construction, \H marks that the tone of the noun is replaced with H, which may be considered as an instance of H tone plateauing between the H of the interrogative and that of final -é.

³⁸See also §5.3.2 below.

We can multiply similar examples if we take into consideration cases where one and the same form is used as ‘who?’ in one language, but as ‘which/what [N]?’ in another. Thus, compare the Mongo C61 dialectal variant CL-*yá* ‘what/which [N]?’ with Bwamba C10 *yá* ‘who?’, Libobi C412 *ya* ‘who?’, Chokwe K11 *i-ya* ~ *a-ya* ‘who?’, and Mbunda K15 *íyà* ‘who?’.

5.3 Interrogative pronominals as nominal predicates

In Bantu, questions (especially those about subjects) are often constructed as clefts, as in ‘it is who that did P?’ for ‘who did P?’, or pseudo-clefts, as in ‘the one that P is who?’ for ‘who did P?’, with the notional predicate being topicalised and construed as a relative clause and the interrogative being focalised and construed as a nominal predicate (cf. §4.1.1). Due to frequency effects, such (pseudo-)cleft structures tend to become reduced to various extents with univertation, formal erosion and simplification of a biclausal construction into a monoclausal one as a result (compare the case of Louisiana Creole interrogative pronominals presented in §3.1.3). Given the natural correlation between subjects and agentivity, in the long run the effect of this tendency is most noticeable with the human interrogative ‘who?’. Traces of the former cleft structure may be found both in the form of the interrogative itself (§5.3.1) and of the constituent question construction (§5.3.2).

5.3.1 Cleft traces in the form of the interrogative itself

As discussed in §4.1.2–4.1.4, since the accretion of substance in interrogatives often proceeds within cleft structures, the accreted substance is often sourced from various deictic forms and other forms that themselves are typically sourced from deictics and used as building blocks of cleft constructions, such as copulas, focus markers and relativisers. Various traces of such morphemes may remain discernible.

For example, across Bantu many NSIPs begin with a nasal-consonant cluster, such as *nd-*, *nz-*, *nj-*. Such NC clusters are particularly common in ‘who?’ interrogatives as reflected in the BGR reconstruction **n(d)ái* ‘who?’, but are also found in ‘what?’ interrogatives, since the interrogative construction reconstructed in BGR as **n(d)ái* was originally not a dedicated human interrogative (see §6.1). As discussed in Idiatov (2009: 71), the unusual shape and sound correspondences, such as *d/z* before *a*, most likely reflect the copulas **ní* ~ **ní* and **ndí* ~ **ndí* (see §4.1.4.1 on type N(D)I). Such clause-level constructions may later be overtly nominalised. Thus, as discussed in §4.1.4.2, the type I- accreted material more frequently found

with ‘who?’ interrogatives is likely to originate in a nominaliser, especially the augment, e.g. Kagulu (*i*)*yehoki* ‘who?’ < (*i*-)*y-e-hoki* [(NMLS-)1-be:NMLS-where?], a nominalised predication literally meaning something like ‘the one that s/he is the one where?’. Because such a nominaliser tends to be reduced for phonological reasons, like the prosodic weakness of a V-shaped prefix, the interrogative may end up looking like a nominalisation by conversion, i.e. a word category change that is not marked by any explicit morphology, like the verb *drink* > the noun *drink*. At the same time, there are cases where clause-level interrogative constructions were effectively nominalised by conversion, as in Mbula *yá’ ná* ‘who is it?; who?’ [^H-yà ná ~ V̇-yà ná] [NMLS-which? COP.PRES] (cf. §2.1).³⁹

The copula origin of many interrogative pronominals, especially the forms of ‘who?’, is indirectly further supported by another peculiarity of their morphosyntax. Such ‘who?’ interrogatives regularly lack any overt (human) class 1 marker, the reason for which they are typically set apart together with other prefix-less human nominals as a subclass of the human class 1, the so-called class 1a (cf. Van de Velde 2006). In this respect, they differ radically from ‘what?’ interrogatives, which are often overtly marked for noun class, typically class 7. This lack of overt class marking is expected if these ‘who?’ interrogatives come from a cleft construction with a copula. It is common for copulas to be invariable and not to be agreement targets (cf. Gibson et al. 2019 specifically on Bantu).

5.3.2 Cleft traces in the form of the constituent question construction

Often, the last (supra)segmental trace (besides word order) that remains of the former interrogative cleft construction is the use of the relative prefix on the verb or the dedicated relative verb form in constituent questions. For instance, in Orungu B11b interrogatives are normally utterance-initial and require a relative prefix on the verb suggesting an earlier cleft structure, possibly with additional prosodic traces in the case of ‘who?’ (cf. Ambouroué 2007: 141–142, 166–167). Similarly, in Ewondo A72a, the relative verb form marked by a postposed floating ^H tone is used with (sentence-initial) interrogatives, as well as focus pronominals and a number of (historically complex) clause-linkers, such as *áná* ‘like’, *ámú* and *àsú* ‘because’ (Abessolo Nnomo & Etogo Mbezele 1982: 75–76, 166). This last suprasegmental trace of the interrogative cleft construction may be partly lost in the closely related language Eton A71, which has a similar relative verb form. However, only a “limited number of verb forms have a special form in relative clauses”, viz. the present affirmative form of *nà* ‘be’, the present tense form

³⁹The nominaliser in the Mbula form nominalises the interrogative modifier *yà* [N] ‘which/what [N]?’; not the predication.

in southern dialects, the resultative verb form, and the future auxiliary (Van de Velde 2017: 54–55). This relative form is used with sentence-initial interrogatives when such a dedicated form is available, as in (13) with the copula *nà*, except in the future tense where the speakers consulted use the non-relative form of the future auxiliary, as in (14) (Mark Van de Velde, p.c.).

- (13) Eton A71 (Mark Van de Velde, p.c.)

zá nà 'vá-lá
zá à-nè-H Lvá-lá
who? 1-COP-REL ADV.DEM-NADR

‘Who is it?’ (for example, asking a person approaching in the dark about their identity) (lit.: ‘(It is) who that s/he is there near you?’)

- (14) Eton A71 (Mark Van de Velde, p.c.)

z éyèyì sò
zá èyèyì L-sò
who? FUT.AUX INF-COME

‘Who will come?’

Except in those limited cases mentioned above where the relative verb form is used, Eton interrogatives can be used in situ (15a) or sentence-initially (15b) without any further morphosyntactic changes.

- (15) Eton A71 (adapted from Van de Velde 2008a: 329)

a. *ùyéń zà á mákíd?*
ù-H-jén-H zà á mákíd
2SG-PST-see-NF who? LOC market

‘Whom did you see at the market?’

b. *zá ù-H-jén-H á mákíd*
who? 2SG-PST-see-NF LOC market

‘Whom did you see at the market?’

Comparison of Eton with Ewondo illustrates another important point. In a Bantu language, fronting of interrogatives should normally reflect an older cleft construction even in the absence of any other morphosyntactic traces, such as relative clause morphology. In fact, this finding is supported by a more general observation. Given that Bantu languages, especially in the north-west, are characterised by a rigid constituent order that is typical for languages of Northern Sub-Saharan Africa in general, it is expected that an interrogative can be used

sentence-initially only as a result of a more profound reorganisation of the morphosyntax of the utterance, as in a cleft construction. These findings also suggest that what is synchronically described in terms of fronting of an interrogative out of its in-situ position, historically represents a change in the opposite direction. An erstwhile clause-level constituent interrogative construction used sentence-initially as part of a larger cleft construction was first deranked into a nominal expression that can no longer be used in an independent declarative clause. A full predication gets stripped of its predicative properties and starts being used as a nominal expression. As a consequence, it can be used in situ in constructions restricted to nominal expressions, such as the postverbal (non-subject) argument construction.⁴⁰ In this respect, recall also the Mbula NSIP *yá' ná* 'who is it?; who?' presented in §2.1.

A particularly interesting example of cleft reduction is the colexification of 'which/what [N]?; what kind of [N]?' and 'who?' in A15, A40 and A70 languages. It not only showcases a gradual simplification of a biclausal into a monoclausal construction, but also demonstrates the possibility of a divergent tonal evolution depending on the construction in which an interrogative is used (see also §4.1.4.2 on Mongo). Ewondo, for example, has besides *zá* 'who?' also a rare interrogative modifier *zǎ* [N](-V̇) 'what kind of [N]?', where a low-toned copy vowel is added to monosyllabic nouns and the verb takes the relative form (Abessolo Nnomo & Etogo Mbezele 1982: 75–76, 166). The low-toned copy vowel is likely to have its origin in a proximal deictic stem used here as a relativiser or copula.⁴¹ Eton has, besides *zá* 'who?' with a restricted constructional variant *zà*, also a rare exclamatory *zá* [N] 'what (kind of) [N]!', both followed by a non-relative verb form (Van de Velde 2008a: 178, p.c.). The tonal difference between *zá* 'who?' and *zǎ* 'what kind of [N]?' in Ewondo has been levelled in Eton in favour of the tone of the interrogative pronominal, which is much more frequent than the modifier. The LH tone pattern of the modifier is likely to be closer to the original tone pattern. In this respect, recall the constructional variant *zà* 'who?' in Eton and compare the 'who?' interrogatives in some other A70 varieties, such as Ntumu A75A *zà* and Meke A75C *nzá*. In fact, a comparable tonal and segmental variation within A70

⁴⁰Obviously, this historical scenario does not preclude the possibility that once the in-situ use of an erstwhile sentence-initial clause-level interrogative, such as 'it is who [that P]?', has become established, the alternation between the in-situ and the sentence-initial position may have been later generalised to other interrogatives which did not originate in a sentence-initial cleft-type interrogative.

⁴¹Compare the low tone in the Ewondo relic proximal adverbial demonstrative forms of class 16 *vá* and class 18 *mú* (Grégoire 1975: 118) and the Basaa near-addressee demonstrative stem, viz. just a low tone (Hyman 2003) or a copy vowel with a low tone (Bôt 1986).

is also found with ‘what?’. Thus, in Eton, we have *já* with the constructional and dialectal variant *jà* and the dialectal variant *yá* (Van de Velde 2008a: 176), while we have *dzé* in Ewondo, *ndzè* in Ntumu and *zè* in Meke. The generalisation of the H tone forms *zá* ‘who?’ and *já* ‘what?’ in Eton and similar cases elsewhere may be accounted for by the earlier presence of the H tone of the augment, as a nominaliser and/or a construct form marker, which would represent a case of the I- type accretion (see §4.1.4.2). In this respect, note that (at least some) speakers of Eton use the L-toned forms *zà* ‘who?’ and *jà* ‘what?’ as a nominal predicate introduced by the copula *nà*, as in (16) and (17) respectively, which can be compared to (13–15) above for ‘who?’. This is exactly the context where there may be less need for these interrogatives to be overtly marked as nominals by a nominaliser augment, and where they definitely cannot be marked by the augment as the construct form marker (cf. Van de Velde 2019: 249).

- (16) Eton A71 (Mark Van de Velde, p.c.)

à-nà zà

1-COP who

‘Who is s/he?’

- (17) Eton A71 (Mark Van de Velde, p.c.)

é-nâ jà

5-COP what

‘What is it?’

In Basaa, the evolution observable in the A70 languages seems to be even more advanced than in Eton, in that *njé(é)* ‘who?’ and the interrogative modifier *njé(é)* [N] are identical in form and neither requires the use of a relative clause (Hyman 2003; Van de Velde 2017: 64). In Akoose A15C, the situation is intermediate between Eton and Basaa in that *nzé* ‘who?’ and *nzé* [N]\H-é ‘what/which [N]?’ are identical in form and neither requires the use of a relative clause (Hedinger 2008). However, in a question, the verb used with the interrogative pronominal or the phrase with the interrogative modifier takes the relative form when the question is not about a subject, a property they share with the cleft construction described by Hedinger (2008) as a “topicalisation” construction. Another feature that the interrogative modifier construction *nzé* [N]\H-é ‘what/which [N]?’ shares with both relative clauses and topicalisation (clefts) is the final element -é. It is reminiscent of the ‘reduced’ forms of the relativiser [N]-é/é (the full form is [N] CL-è) and the ‘topicalisation’ marker ‘it is the [N] that...’ [N]=éé (the full form is [N] CL-à). Like in A70 languages, final -é is likely to have been sourced from a non-distal deictic stem used here as a relativiser or a copula.

6 A revision of the reconstruction of the PB NSIPs

In this section, I propose a revision of previous NSIP reconstructions which is informed by the diachronic typology of NSIPs presented in §3 and applied to Bantu in §4 and §5. I equally take into consideration data from outside of Narrow Bantu in order to refine the reconstructions and to determine the level to which they belong. In §6.1, I revise the PB reconstruction for ‘who?’ as **n(d)ái* ‘who?’. For ease of reference, I refer to the interrogatives that would formerly be considered as reflexes of PB **n(d)ái* ‘who?’ as NDAI type interrogatives. In §6.2, I propose a critical reassessment of the PB reconstruction for ‘what?’ as the interrogative stem **í*.

6.1 The human NSIP ‘who?’ and the NDAI type interrogatives

6.1.1 Overview

As I argued in Idiatov (2009) and further elaborate here, no simplex ‘who?’ interrogative can be reconstructed for PB. The only form proposed so far, viz. **n(d)ái* ‘who?’, results from univerbation and nominalisation, either by conversion or by means of an overt nominaliser, such as the augment, of a clause-level interrogative cleft construction. The latter was most likely based on an erstwhile SIP meaning ‘which one?’ indifferent to the distinction between persons and things. The primary development was from a cleft content question construction ‘it is which one [that P]?’ > ‘it is who [that P]?’ > ‘who [(that) P]?’ (sentence-initial NSIP with some traces of the former cleft, cf. §5.3.2) > ‘who?’ (NSIP usable in situ in non-sentence-initial positions, cf. §5.3.2).⁴² Furthermore, thanks to the original indifference to the distinction between persons and things, we also find interesting patterns of colexification of ‘who?’ and ‘what?’ or ‘which/what [N]?’ (cf. §5.2).

In Idiatov (2009), I proposed that the NDAI type interrogatives go back to the structure **[AG9(or AG7)-COP CL16-‘what?’]* ‘(it) is where?’, viz. something like PB **í-ndí pà-í*. I now believe that this reconstruction should be revised, except for the copula part. As discussed in Appendix C, the initial *nd-* cluster of the copula may reflect a stacking of two copulas, **ní* and **dì ~ *li*. As I show in §6.1.2, the NDAI type interrogative construction predates PB, but is probably limited to Southern Bantoid. We should therefore also consider data from outside of Narrow Bantu. The complexity of the tonal patterns and tonal correspondences of the NDAI

⁴²I do not reconstruct a pseudo-cleft, such as ‘The one that P is who?’, since the preference for construing content questions as pseudo-clefts appears to be largely restricted to zone C.

type suggests that we should reconstruct three to four tones, probably *LHL(H) (§6.1.3). I propose to reconstruct the pre-copula part of the NDAI construction as the 3SG personal index *à used as a dummy subject of the copula (§6.1.4) and the post-copula part as a nominalisation of the interrogative modifier *yà ~ *là ‘which/what [N]?’ (§6.1.5).

6.1.2 The NDAI type interrogative cleft construction in Southern Bantoid

The clause-level interrogative cleft construction that resulted in NDAI type interrogatives can be safely reconstructed well beyond Narrow Bantu, but probably limited to Southern Bantoid. Related forms are well-attested in Narrow Grassfields. For example, for the Mbam-Nkam Grassfields group, Elias et al. (1984) reconstruct two ‘who?’ stems, viz. *-gú, with a wide distribution,⁴³ and *^Hndā, as in Limbum [limb1268, Mbam-Nkam Grassfields] *ndāā* (Fransen 1995). The latter stem is limited to Nkambe [nkam1238, Mbam-Nkam Grassfields], a small group of languages in the very north of the Mbam-Nkam domain. We also find similar forms in Ring Grassfields, such as Babungo [veng1238, South Ring Grassfields] *ndā ~ ndá* (Schaub 1985), Babanki [baba1266, Centre Ring Grassfields] *ndô* (Paulin 1995), Mmen [mmen1238, Centre Ring Grassfields] *āndē* ‘who?’ (Paulin 1995), Weh [wehh1238, West Ring Grassfields] *ndēē* (from *HLH)⁴⁴ ‘who?’ (Paulin 1995), Isu [isum1240, West Ring Grassfields] *ndiā* ‘who?’ (Paulin 1995). Examples of related interrogatives in other Bantoid groups are Mundabli [mund1328, Southern Bantoid] *ndē* ‘who?’ (Voll 2017) and Esimbi [esim1238, Tivoid] *āndā* ‘who?’ (Coleman et al. 2004).

Given its complex constructional origin, the NDAI type may have been conventionalised independently in a number of Bantoid groups. Similarly, its initial univerbation and formal reduction (or its complete loss) may also have occurred at a relatively late stage, long after the diversification of Southern Bantoid. However, it must have emerged when Bantoid languages were still very closely related. We can therefore reconstruct one construction with the same slots and the same or very similar elements filling these slots for all the relevant Southern Bantoid groups.

⁴³Most likely, the stem *-gú ‘who?’ is yet another example of the typical evolution of ‘which one?’ to ‘who?’, presumably augmented with a class 1 prefix. Thus, compare Babanki [baba1266, Centre Ring Grassfields] *cl-kô^H* ‘which [N]?’ or nominalised as ‘which one?’ (cf. Hyman 1980: 241), which in principle could also come from earlier *^Hkô^H and where the two floating ^H tones could reflect the same nominalising morphology as that discussed in §6.1.5 below.

⁴⁴Davison (2009: 11) explains that in the Weh orthography “the phonetic mid-level mark [...] should probably be thought of as a lowered high tone”.

6.1.3 Revising the tonal reconstruction of the NDAI type interrogative construction

A complex constructional origin of the NDAI type interrogatives is also indirectly corroborated by the complexity of their tonal patterns and possible tonal correspondences. It is no coincidence that BLR3 adopts BGR's reconstruction but removes its tonal specification, viz. **nai* ~ **ndai*, admitting the tonal uncertainty of the reconstruction. Within Narrow Bantu, NDAI type interrogatives usually have one to two tones and all possible tone patterns are attested, viz. L, H, LH, and HL. This suggests **LHL* or **HLH*, unless we can demonstrate that all cases of HL are due to the H tone of a later type I- accretion (cf. §4.1.4.2), in which case **LH* would suffice but **LHL* would also be acceptable. However, outside of Narrow Bantu, we also find NDAI type interrogatives with three tones, such as LHL (as in Babanki *ndô*) and HLH, as in Weh *ndéé*,⁴⁵ and probably other Grassfields forms with surface M tones. This suggests **LHLH*, or less likely **HLHL*.⁴⁶ In any event, we should reconstruct three to four tones for the NDAI type. Presuming the tone-bearing unit was a syllable, the construction must have had at least three to four syllables. Furthermore, the attested segmental forms suggest that in this reconstruction one tone, most likely L, should precede the ND-cluster and two or more should follow it. Given that the morphemes involved in the NDAI type interrogative construction are most likely to have been short functional morphemes, such as a copula, a subject index, a deictic stem, an interrogative stem, a focus marker, and the like, we are dealing with at least three to four distinct morphemes.

6.1.4 The pre-copula part: the 3SG personal index **à* as a dummy subject

The initial **í-* in my earlier reconstruction (Idiatov 2009) is a later type I- accreted form (cf. §4.1.4.2). As discussed above, the element preceding the copula most likely had a L tone. From a comparative Bantoid and wider Benue-Congo (and Niger-Congo) perspective, the best candidate is the 3SG personal index **à* used as a dummy subject. Compare the floating L tone dummy subject before the copula in the cleft construction in Mundabli [^L dummy subject + *dí* 'be' + X + P] 'It is X that P' (Voll 2017: 139). This is a well-attested Niger-Congo root, with a rather stable L tone.

The pre-PB 3SG personal index **à* is the same morpheme as the BGR class 1 subject marker **á*, as I believe the H tone of this marker in BGR is due to an

⁴⁵See previous footnote 44.

⁴⁶**HLHL* is less likely because outside Narrow Bantu the tone preceding the ND cluster is hardly ever H.

overreliance on Eastern Bantu data and is a later innovation. In Bantu, the agreement of class 1 is known to be one of the possible options in constructions with enforced agreement, such as ‘It is X that P’ or ‘There is X that P’ (cf. Van de Velde 2006: 202–203), as illustrated in (18) from Mongo and (19) from Orungu.

- (18) Mongo C61 (Hulstaert 1966: 331, 618)
- a. *a-le ndé nsé*
1-COP.PRS really 9.fish
‘It’s really a fish.’
- b. *a-le ngá [áótosangelaka josó]*
1-COP.PRS like
‘It’s as if [he had already said this to us before].’
- (19) Orungu B11b (Van de Velde & Ambouroué 2011: 124)
- èpóswá sídyàbí*
à-í-póswá sídyàbí
1-IPFV-fall.PRS 10b.leaf
‘There are leaves falling.’ (lit.: ‘It falls leaves.’)

From a typological perspective, the use of the agreement pattern strongly associated with human nouns (viz. of class 1) as the enforced agreement pattern in Bantu is a perplexing choice (cf. Corbett 1991: 208, as discussed by Van de Velde 2006: 202–203). However, this synchronic oddity can be straightforwardly accounted for as a trace of the original indifference of the 3SG personal index **à* to the human semantics typically associated with the nouns of class 1 in modern Bantu languages.

6.1.5 The post-copula part: a nominalised interrogative modifier

In Idiatov (2009), I proposed that the post-copula part of the NDAI type interrogatives should be reconstructed as **pà-í* ‘where?’ [CL16-‘what?’]. This reconstruction is semantically plausible and matches the Bantu data relatively well formally, but as discussed in §6.1.5.1 below, it also has a number of problematic aspects. From a Bantu-internal perspective, none of the issues is crucial but taken all together and given that the NDAI type interrogative construction predates PB, I believe a different reconstruction provides a better account of the data. In particular, I propose to reconstruct the post-copula part as a nominalisation of the interrogative modifier **yà ~ *là* ‘which/what [N]?’ that functioned as the SIP ‘which one?’. This interrogative modifier is comparable to *yà* [N] ‘which/what

[N]?’ in Mbula and the dialectal variant [N] CL-*yá* ‘what/which [N]?’ in Mongo C61.⁴⁷ A possibility that an earlier form of this interrogative may have been **lá* is suggested by the existence of such interrogatives as Noone [noon1243, Beboid] CL-*lá* ‘which [N]?; which one?’, *lá* ‘what?’ (Hyman 1981: 25, 119).⁴⁸ For ease of reference, in the rest of the chapter I use only the form **yá*.

Nominalisation of modifiers is typically achieved in Bantoid by means of noun class affixes or deictics (cf. on the nominaliser augment §4.1.4.2). While in Narrow Bantu such markers are typically prefixes, in Bantoid we also find suffixes and combinations of prefixes and suffixes. Therefore, the post-copula part of the NDAI type interrogative construction may have had one of the following structures, **[NMLS-which?]*, **[which?-NMLS]* or **[NMLS-which?-NMLS]*. To make a choice between these options and to identify the nominaliser(s) involved, I present in §6.1.5.2 some interesting data on the different ways of nominalising the interrogative modifier *yá* in Mbula. In §6.1.5.3, adducing data from Bantoid and wider Benue-Congo, I reconstruct the pre-PB determiner **yé* that gave origin (among other things) to the markers used to nominalise the interrogative modifier in the post-copula part of the NDAI type interrogative construction. Finally, in §6.1.5.4 I propose to reconstruct two variants of the pre-PB (Southern Bantoid) NDAI construction **à ndé yé-yà* (~ *yé-là*) [3SG COP NMLS₁-which?] ‘it is which one?’ and **à ndé yé-yà-yé* (~ *yé-là-yé*) [3SG COP NMLS₁-which?-NMLS₂] ‘it is which one exactly?’.

6.1.5.1 Issues with reconstructing the post-copula part as **pà-í* ‘where?’ [CL16-what?]

Locative interrogatives of the PAI type appear to be largely restricted to (Mbam-Nkam) Grassfields and Narrow Bantu and are likely to be more recent.⁴⁹ Additionally, reflexes of **p* of the presumed **pà-í* part are often irregular, even though this could be due to the irregularity of the reduction following the construction’s

⁴⁷See also §4.1.4.2, §4.1.4.3 and §5.2.4 for some examples of ‘who?’ and ‘what?’ as possible reflexes of this interrogative stem in constructions other than the NDAI type.

⁴⁸Within the Noone tonology, the H tone of this interrogative may also come from an earlier **^Hlá^H* (cf. Hyman 1981: 10–11), where the two floating ^H tones could reflect the same nominalising morphology as that discussed later in this section. The NSIP *lá* ‘what?’ looks like a noun of class 5, while its plural form *mù-là* is class 12 in Noone, which corresponds to the Mbam Bantu plural class *mɔ-*, also identified in the literature as class 18 or 6 (cf. Boyd 2015: 19).

⁴⁹We may find locative interrogatives containing cognates of the PB locative class 16 **pa* beyond these groups. However, they reflect different interrogative constructions and different interrogative stems, such as the Tikar [tika1246, Northern Bantoid] interrogative *fɛn* ‘where?’ (cf. Appendix A).

univerbation. There are also considerably less traces of the labial articulation reflecting **p* in NDAI type interrogatives across Bantu than may have been expected. Many instances of labialisation or round vowels in NDAI type interrogatives may also be accounted for by the -O type accretion (cf. §4.1.4.3, Idiatov 2009: 72). Yet some other instances may be due to the accretion of the class 1 subject **ù*-, as probably in Liko D201 *wàní* ‘who?’ (de Wit 2015) (cf. §4.1.4.2 concerning Kagulu; also see Appendix D). Another important issue concerns the problematic status of the interrogative stem **í*, especially for any level beyond Narrow Bantu (cf. §6.2).

6.1.5.2 Different ways of nominalising the interrogative modifier *yà* in Mbula

In Mbula [mbul1261, Jarawan Bantu], the interrogative modifier *yà* [N] ‘which/what [N]?’ can be nominalised in two ways, viz. like classifying modifiers or like identifying modifiers resulting in the (human) NSIP *yá* ‘ná and the SIP *mà-yè: ná* respectively.

The interrogative modifier *yà* is nominalised as the (human) NSIP *yá* ‘ná ‘who is it?; who?’, structurally |^H-yà ná ~ *V*-yà ná| [NMLS-which? COP.PRES], by means of a prefixed floating H ~ an underspecified vowel with a H tone (cf. §2.1 on the accretion of the copula *ná*). This nominaliser, which appears to be restricted to demonstratives in certain contexts, looks like a former class prefix or an element similar to the nominaliser augment (§4.1.4.2). This nominalisation construction can be compared to the productive construction [*mà*- + X] used with other types of stems; *mà*- is a nominal derivational prefix that can roughly be glossed as ‘the one with’. The construction [*mà*- + X] functions as a noun, where X can be a noun itself, as in *mà-là* ‘village head’ (*là* ‘village’) and *mà-ntà* ‘hunter, archer’ (*ntà* ‘bow’ itself a frozen nominalisation of the verb *tà(:)* ‘shoot with a bow’), a verb, as in *mà-bà* ‘builder (of buildings); potter’ (*bà(w)* ‘build; mould, make (a pot)’), or an adjective, as in *mà-gùlà* ‘elder sibling’ ([N] *gùlà* ‘big [N]’, *gùló* ‘it/s/he is big’).

The interrogative modifier *yà* [N] ‘which/what [N]?’ is nominalised as the SIP *mà-yè: ná* |*mà*-yà-yí ná| ‘which one is it?; which one?’ by the construction [*mà*- + X + -yí], where -yí is sourced from the 3SG non-subject person index used as a nominaliser (*ná* is the copula like in *yá* ‘ná ‘who is it?; who?’).⁵⁰ The construction [*mà*- + X + -yí] is primarily used to create adnominal modifiers that can also be used independently as nouns without any additional marking. Although synchronically, [*mà*- + X + -yí] may often be the only way to use a given element X as adnominal modifier, the original use of this construction must have

⁵⁰Like the 3SG non-subject person index, the nominaliser -yí has an allomorph -i which fuses with the preceding *a* into *e*.

been to form localising (anchoring, identifying) modifiers in terms of Rijkhoff (2008). Thus, compare *ɲɔ̀ʒàr gùlà* ‘road’ (lit.: ‘big path’), where *gùlà* ‘big’ is a classifying (or perhaps just qualifying) modifier, and *ɲɔ̀ʒàr m̀-ɔ̀gùlé* ‘big path, the path that is big (as opposed to paths with other properties)’, where *m̀-ɔ̀gùlé* is an identifying modifier ‘the one that is big’. Given that *-yí* in [*m̀- + X + -yí*] is sourced from the 3SG non-subject person index, the construction is likely to have originally been in appositional relation with the preceding noun, i.e. *ɲɔ̀ʒàr m̀-ɔ̀gùlé* literally meant something like ‘path, the big one’.

Thus, we have an interesting parallel between the nominalisation of a classifying modifier (as in *m̀-ɔ̀gùlà* ‘elder sibling’) and the nominalisation used to derive a NSIP from an interrogative modifier ($|\text{H-}yà \sim \acute{V}\text{-}yà|$ in *yá ‘ná* ‘who is it?; who?’), on the one hand, and the nominalisation of an identifying modifier (as in *m̀-ɔ̀gùlé* ‘the one that is big’) and the nominalisation used to derive a SIP from an interrogative modifier ($|\text{m̀-}yà\text{-}yí|$ in *m̀-ɔ̀yè: ná* ‘which one is it?; which one?’), on the other hand.

6.1.5.3 The pre-PB determiner **yé* as the nominaliser of the interrogative modifier

I argue that the nominaliser prefix $|\text{H-} \sim \acute{V}\text{-}|$ of the interrogative modifier *yá* in *yá ‘ná* ‘who is it?; who?’ in Mbula is sourced from the same referential element as the nominaliser augment of type I- in Bantu, such as the construct form markers *í-* and *é-* in A70 (Van de Velde 2017) and the type I- accretion in interrogatives (cf. §4.1.4.2), and the Mbula 3SG non-subject person index and identifying nominaliser *-yí*. The referential element in question is the pre-PB determiner **yé* (where *e* is the front vowel of a second degree of aperture), corresponding to PB **yí*. This pre-PB determiner had two major functions.⁵¹ First, within the noun class system, **yé* was a determiner of class 5, as reflected in the PB class 5 nominal prefix **i-*.⁵² Second, outside of the noun class system, **yé* was a selective ‘this/that very (one from a range of possible referents, from a set, a mass, etc.)’ or restrictive determiner ‘this/that very (one and not another one)’ that did not

⁵¹The two functions result from a divergent evolution of a single noun, most likely meaning ‘seed, grain, kernel’. Its reconstruction goes beyond the scope of the present chapter.

⁵²One way to account for the L tone of this class prefix in PB is analogical levelling, as all other nominal prefixes are reconstructed with L tone (in this respect, see an interesting discussion on the tones of PB class prefixes in Hyman 2005: 338–340). Another possibility is the merger with some **à* morpheme, such as the 3SG personal index **à* (see below on the L tone in person indexes sourced from **yé*). In this respect, note for example that in zone A the nominal prefix of class 5 is sometimes *à-* as in Ewondo A72a or *è-* as in Eton A71. The same variants *à-* and *è-*, as well as one case of *i-*, are found in A15 varieties (Hedinger 1987: 94–96).

agree in noun class with the noun whose reference it determined. It restricted the reference of a given referential element to one particular referent to the exclusion of any other possible referents, or in the case of collective and mass referents, to the exclusion of the rest of the group or a mass. In this sense, it can also be referred to as strongly or exclusively identifying.

Various traces of this double, agreeing and non-agreeing, usage of the determiner *yé can be found across Bantoid and beyond.⁵³ The selective or restrictive usage is reflected in the recurrent use of class 5 in Bantu for singulative or partitive derivation, as in Eton A71 *mà-ndím* ‘water’ (class 6) > *è-ndím* ‘drop of water’ (class 5) > *mà-ndím* ‘drops of water’ (class 6), *mà-kálá* ‘doughnut batter’ (class 6) > *è-kálá* ‘doughnut’ (class 5) > *mà-kálá* ‘doughnuts’ (class 6), *mà-njáj* ‘xylophone’ (class 6) > *è-njáj* ‘bar, wooden piece of a xylophone’ (class 5) > *mà-njáj* ‘xylophone bars’ (class 6) (Van de Velde 2008a: 97–98). Another interesting reflex of this selective or restrictive usage is the Liko D201 type III demonstrative stem *-í* indicating the “exclusiveness of the referent” (de Wit 2015: 260). Beyond Narrow Bantu, particularly telling evidence is provided by Babungo [veng1238, South Ring Grassfields] (Schaub 1985), where the pairing class 5 *yí-* / class 6 *má-* “includes only objects and body parts which occur in groups or pairs (the singular referring to one of the pair or group)” (Schaub 1985: 177). Furthermore, the anaphoric demonstrative modifier of class 5 *yɔ* can be used with a few nouns that are not in class 5 in the locative construction to focus on “certain one out of a group” (Schaub 1985: 70). Finally, Babungo has an identical prefix *yí-* that can be added to an ‘emphatic’ demonstrative modifier of any class and “again has ‘selective’ function (‘that one, not the other one’), as in *bú yí-njii* ‘that dog (not the other one)’ (Schaub 1985: 205), and which appears on the restrictive anaphoric locative adverbial demonstrative *yí-fi* ‘there (the particular place mentioned, not any other place)’ (Schaub 1985: 98).

Another class of elements that is likely to have been sourced from the restrictive or selective usage of the determiner *yé is represented by person indexes, such as the Mbula 3SG non-subject person index *-yí*, Kenyang [keny1279, Mamfe] class 1 (3SG human) person index *yí ~ yĩ* (Ittmann 1935–36; Voorhoeve 1980; Mbuagbaw 2000), and probably the ‘preprefixal’ morpheme reconstructed for the PB substitutives and possessives by Kamba Muzenga (2003) as **i-* in 1PL and 2PL and as **i-* ~ **i-* in class 1. Meeussen (1967) reconstructs this preprefix only in substitutives as **i-* in 1SG and as **i-* in 1PL and 2PL.⁵⁴ The person indexes in

⁵³Beyond Bantoid, a particularly interesting set of forms sourced from the determiner *yé can be found in Bena-Yungur [bena1260, Buto]. See Appendix G for more details.

⁵⁴The use of a restrictive or selective element on person indexes, which are inherently identifying anyway, may have an intensifying origin, something like ‘I myself’ > ‘I’.

question may be restricted to logophoric use, such as Babungu *yì* SG.LOG (Schaub 1985) and Nizaa [suga1248, Mambiloid] *yí* SG.LOG (Kjelsvik 2002: 18). The L tone that occasionally shows up on the person indexes reflecting the determiner **yé* is likely to come from a fusion with another morpheme, such as the 3SG personal index **à* (cf. §6.1.4 and Appendix G on Bena-Yungur 3SG.ANIM free pronominal).

In Bantu, we find yet another morpheme that in all probability is part of the same cluster of reflexes of the determiner **yé* as person indexes. The morpheme in question is the reflexive prefix ('infix' in the traditional Bantu terminology). The reflexive prefix is reconstructed in BGR as **í-*. However, as suggested by the data on Bantu reflexives discussed in Polak (1983), most likely BGR's reconstruction represents just one member of the paradigm of reflexive markers, presumably agreeing in noun class with the subject. Reflexive use is similar to logophoric in that both uses mark co-reference between two arguments. Cross-linguistically, it is not uncommon that in languages lacking dedicated logophoric person indexes, reflexive person indexes are used in logophoric contexts or that in languages with dedicated logophoric person indexes, the latter can be used in reflexive contexts or at least show strong formal similarity with the reflexive person indexes.

Diachronically, it is clear that the human reference and personal pronominal uses of the reflexes of **yé* cited above have evolved out of their selective/restrictive reference uses. In this respect, note that the evolution from selective/restrictive reference to human reference is very similar to the evolution from a selective interrogative pronominal 'which one?' to a human non-selective interrogative pronominal 'who?', which is typologically common. Both evolutions reflect the typical tendency for the feature [+human] to correlate with various features restricting the reference, such as [+unique], [+specific], [+definite], [+identification], as reflected in the various versions of the so-called Animacy (or Referential) Hierarchy (cf. Croft 2002: 130, among others, see also various chapters in Cristofaro & Zúñiga 2018).

6.1.5.4 [NMLS-which?] and [NMLS-which?-NMLS]

I propose that, like in Mbula, the interrogative modifier **yà* (**lâ*) 'which/what [N]?' could be nominalised in two different ways, viz. as **yé-yà* [NMLS-which?] and as **yé-yà-yé* [NMLS-which?-NMLS], both originally indifferent to the distinction between humans and things. It is actually likely that initially both interrogatives were selective and the distinction was rather between 'which one?' and something like 'which one exactly?'. Hence, originally there also existed two variants of the NDAI construction **à ndé yé-yà* [3SG COP NMLS₁-which?] and **à*

ndé yé-yà-yé [3SG COP NMLS₁-which?-NMLS₂] with a similar semantic distinction. Given the common pathways of semantic change of interrogative pronominals (§3.2), both constructions are most likely to ultimately evolve in non-selective ‘who?’, but they can also remain selective or become non-selective ‘who?; what?’ or ‘what?’. At the same time, it is clear that this semantic evolution happened long after PB. On the formal side, as soon as the original semantic distinction between the two constructions became blurred, either of the two constructions may have outcompeted the other, probably after a long period of co-existence as free variants.

6.2 The non-human NSIP ‘what?’

6.2.1 Overview

Meeussen (1967) reconstructs the interrogative stem **í* used in combination with nominal class prefixes, viz. 7 **kì-í* ‘what?’, 16 **pà-í* (17 **kù-í*, 18 **mù-í*) ‘where?’. That Meeussen (1967) does not provide any English gloss for this stem is because he hypothesises that it may also be part of **ndá-í* ‘who?’. In BLR3, Bastin et al. (2002) take the basic meaning of **í* to be non-human ‘what?’ in class 7, with a derived use as ‘where?; which?’ in class 16.

As briefly mentioned in §2.2, there are a number of seemingly minor formal issues with the reconstruction **í*. To begin with, **í* ‘what?’ is supposed to be a nominal stem since it is reconstructed with a nominal prefix. For a nominal stem, however, its vowel-initial shape is exceptional in PB. For all other nominal (and verbal) stems whose stem-initial consonant tends to be zero in modern Bantu languages, BGR and BLR3 consistently reconstruct a stem-initial **j*. Although I do not agree with the choice of **j*, I do agree that such stems did have an initial consonant – contra Bulkens (2009), and contra Wills (2022 [this volume]); see Appendix H for some evidence. In particular, I believe that BLR’s **j* minimally confounds PB **s*, **z*, **ʃ*, **y* and **g*. In the case of BLR’s **í* ‘what?’, I believe that the stem-initial consonant was a palatal glide **y* as it never has “strong” reflexes as a stop or a fricative. Furthermore, this stem was in all probability a heavy monosyllable with a long vowel or it was disyllabic (§6.2.3). In either case, the vowels must have had the quality *i* or *ɪ*. I discuss supporting data that come from reflexes of class 7 **kì-í* ‘what?’ and class 16 **pà-í* ‘where?’ in §6.2.2 and §6.2.3 respectively. Finally, in §6.2.4, I consider the implications of these findings for the reconstruction of the PB stem ‘what?’ within a wider Bantoid perspective. By comparing them with the reconstruction of the NDAI type in §6.1, I propose to reconstruct PB ‘what?’ as something like **yíi* or **yíí*, probably going to the

pre-PB structure *[NMLS-which?-NMLS] as reconstructed in §6.1.5.4 as part of the NDAI construction.

6.2.2 The class 7 form **kì-í* ‘what?’

In the case of the class 7 form **kì-í* ‘what?’, corroborating evidence for the reconstruction of the stem-initial **y* comes from languages such as Basaa A43a, which has *kí(i)* ‘what?’ in addition to *kí.njé(é)* (cf. §5.2.2). As pointed out in §4.1.1, in *kí(i)* the class 7 prefix has been integrated in the stem and *k-* is the stem-initial consonant, not a prefix consonant anymore. Synchronically, the class 7 prefix in Basaa is not *ki-*, but zero before a consonant or *y-* before a vowel as a nominal prefix and *i-* or *y^H-* respectively as an agreement marker. Although synchronically Basaa has many VV sequences in stems, they all result from the loss of an intervocalic consonant. All sequences of identical vowels and the PB sequence **ai* have been reduced to a short vowel (cf. Teil-Dautrey 1991). Given that *kí(i)* is a stem and not a combination of a prefix and a stem, its allomorph *kíí* with a long vowel points to an earlier presence of an intervocalic consonant, just like the long vowel in *njéé* ‘who?’, a reflex of the NDAI type interrogative construction. In this respect, compare Basaa **gíjì* (BLR 1386) ~ **gíjé* (BLR 1385) > *y-ìì* / *gw-ìì* ‘(hatched) egg’ (7/8), *lí-ŋéé* / *mà-ŋéé* ‘egg’ (5/6) (cf. Teil-Dautrey 1991: 53, 73–74).⁵⁵

Outside of Narrow Bantu, a very similar example is provided by Limbum [limb1268, Mbam-Nkam Grassfields] (Fransen 1995). Thus, Limbum has *kēē* ‘what?’ in class 7 with no prefix, which can be pluralised as *b-kēē* with the prefix resulting from a merger of the classes 2, 8 and 14 (Fransen 1995: 101), and which therefore is a stem and not a combination of a prefix and a stem. Like in Basaa, the vowel length in *kēē* ‘what?’ suggests the loss of an intervocalic consonant. Again like in Basaa, the length of the vowel in *kēē* ‘what?’ is comparable to the length of the vowel in *ndāā* ‘who?’, a reflex of the NDAI type interrogative construction.

6.2.3 The class 16 form **pà-í* ‘where?’

The class 16 form **pà-í* ‘where?’ contains the vowel sequence **ai*. According to Doneux & Grégoire (1977), besides **pà-í* ‘where?’ this vowel sequence is found in a limited number of PB stems, viz. the adjective **dài* ‘long, tall, high’ (BLR 3705), the derived verb **dài-p* ‘be(come) long, tall, high’ (BLR 784), the nouns **tái* ‘saliva’

⁵⁵For ‘egg’, compare also the relevant forms in Mbam Bantu languages, such as Baca A621 *n-hègé*, Yangben A62A *ni-kèé* and Mbule A623 *ki-ŋéé* (cf. Boyd 2015: 190), that both confirm the loss of the intervocalic consonant in this stem and suggest it was **g* rather than **j* as in BLR, viz. **gígí*.

(BLR 6231), **jài* ‘outside’ (BLR 8928), the numeral **nài*^H ‘four’ (BLR 3683) and the interrogative **ndai* ‘who?’ (BLR 8161). The typical reflexes of **ai* are *-i*, *-e*, and *-a*, although in a limited number of languages we also find *-ai*, *-ei*, *-ei*, *-ayi*, *-azi*, *-aci*. Interestingly, Doneux & Grégoire (1977: 194, 196–197) observe that the reflexes of the sequence **ai* in the two interrogatives, **pà-í* ‘where?’ (with its derivate ‘which (one)?’) and **ndai* ‘who?’, are typically the same in languages that have reflexes of both forms. At the same time, the reflexes of **ai* in the interrogatives tend to differ from the reflexes of **ai* in the other stems.⁵⁶ The reflexes of **ai* in **pà-í* tend to match those in the other stems only in zones D, J, E, F, H, and less consistently across different stems in zones B and C. This makes Doneux & Grégoire (1977: 197) wonder why the interrogatives have evolved differently from other stems.

I believe that the answer is that in **pà-í* ‘where?’, the sequence **ai* should be reconstructed differently from the other, non-interrogative stems. More specifically, since the vowel *a* in the prefix **pà-* is uncontroversial, it is the stem **i* that should be reconstructed differently. In the data of Doneux & Grégoire (1977: 190, 192), the most common reflex of **ai* in **pà-í* ‘where?’ and its derivate ‘which (one)?’ is by far *i*. Interestingly, the reflex *i* is rare in the other stems.⁵⁷ This suggests that the form that resulted in *i* in **pà-í* was in some way more prominent than *i* in the other stems reconstructed with **ai*. For example, it could have had a CV or CVV shape, such as **yí(i)* or **yí(i)*, or a CVCV shape, such as **yíyí*, **yíyí* or **yíyí*. Although we could have hypothesised that the divergent behaviour of the reflexes of **pà-í* ‘where?’ is due to the fact that *i* there is a prosodically strong stem-initial vowel preceded by a prosodically weak vowel of the noun class prefix, this account is invalidated by the fact mentioned above that in the languages that have both a reflex of **pà-í* ‘where?’ and **ndai* ‘who?’ in Doneux & Grégoire’s (1977) data, the two tend to pattern together despite the fact that **pà-* is a noun class prefix and **nda-* is not. Furthermore, this alternative hypothesis is weakened by the fact that the class 16 prefix **pà-* tends to become part of the stem in reflexes of **pà-í*, just like the class 7 prefix **ki-* tends to become part of the stem in reflexes of **ki-í* (cf. §6.2.2).

6.2.4 PB ‘what?’ and its pre-PB source

The observations in §6.2.2–6.2.3 suggest that PB ‘what?’ reconstructed in BGR as **i* should be reconstructed as **yíí* or **yíí*, or perhaps even as disyllabic as **yíyí*,

⁵⁶Here, we could also add **pái* ‘new’ (BLR 3281) discussed by Baka (2005).

⁵⁷For example, in Tswana S31 **ai* can result in *ɪ*, *e* or *ɛ*, but the most closed reflex *ɪ* is found only for **pà-í* ‘where?’ giving *-fí* ‘which [N]?; which one?’ and for **tái* ‘saliva’ giving *-tʰí* (cf. Creissels 2005: 195–196).

*yíyí or *yíyí. The only matching interrogative ‘what?’ that I have been able to identify in Bantoid is the Kenyang form *yì*. However, it has a low tone. It is therefore possible that the PB form should rather be reconstructed with a LH tone as *yìi or *yìi. A possible reflex of the earlier LH tone pattern within Bantu may be provided by the Eton sentence-initial polar question marker *yì ~ yí* (cf. Van de Velde 2008a). In this respect, recall the case of the Mongo varieties discussed in §4.1.4.2 where the sentence-initial polar question markers *nà* (Nkundo) and *yà* (some other varieties) reflect the older tone pattern that was simplified in *ná* ‘who?; what?’ and the (dialectal) variant of ‘what/which [N]?’ CL-*yá*. Typologically, the evolution from ‘what?’ to a polar question marker is also commonplace (cf. some examples in Hölzl 2017: 73).

That *yìi ~ *yìi ‘what?’ could be used as a free nominal form suggests that it already contained some kind of nominalising morphology and that its combination with the noun class prefix of class 7 (as well as that of class 16) became conventionalised at a later stage. In this respect, compare the situation in Mundabli where the interrogative *mān* ‘what?’ is not marked for noun class but can take the prefix *kì-* of class 7 when “the speaker already has a referent in mind, i.e. it implies a certain degree of definiteness” (Voll 2017: 141). That is, the marked form *kì-mān* means something like ‘what exactly?’ or ‘which one (a thing)?’.

If we now compare *yìi ~ *yìi ‘what?’ with the results of the reconstruction of the NDAI type interrogative construction in §6.1, which was indifferent to the distinction between persons and things, it becomes likely that *yìi ~ *yìi ‘what?’ also goes back to a pre-PB nominalisation of the interrogative modifier *yà (*lā) with the structure *[NMLS-which?-NMLS]. One possibility would be that this pre-PB nominalisation had the same structure *yé-yà-yé as the variant reconstructed in §6.1.5.4 for the NDAI type interrogative construction. Alternatively, the first nominaliser in this ‘what?’ interrogative could be related to the PB pronominal prefix of class 9 *jì.

7 Conclusions

In this chapter, I proposed a typologically informed reconstruction of the Bantu NSIPs ‘who?’ and ‘what?’ that was introduced by a more general discussion of the issue of variation in functional elements and the possible ways of dealing with it in reconstruction in §2, and by an overview of the diachronic typology of NSIPs in §3. The most important findings are that no ‘who?’ stem can be reconstructed for PB, while the morphological status of the non-human form ‘what?’ is ambiguous and that the NSIPs that can be reconstructed to PB were

emerging out of complex interrogative constructions retained from some pre-PB stage within Southern Bantoid. Thus, we can reconstruct two variants of the pre-PB NDAI interrogative construction *à ndé yé-yà (~ *à ndé yé-là) [3SG COP NMLS₁-which?] ‘it is which one?’ and *à ndé yé-yà-yé (~ *à ndé yé-là-yé) [3SG COP NMLS₁-which?-NMLS₂] ‘it is which one exactly?’, that often gave rise to the NSIPs meaning ‘who?’, as reflected by the BGR reconstruction *n(d)ái, but that also have many reflexes meaning ‘what?’ or both ‘who?’ and ‘what?’. The initial *nd-* cluster of the copula part of the construction may reflect a stacking of two copulas (see Appendix C). Furthermore, I propose to reconstruct PB ‘what?’ as something like *yíí or *yíí, probably going to the same pre-PB structure *yé-yà-yé (~ *yé-là-yé) [NMLS₁-which?-NMLS₂].

Given that the NSIPs that can be reconstructed to PB were at this stage emerging out of complex interrogative constructions retained from some pre-PB stage within Southern Bantoid, the proposed reconstructions cannot help us much in locating PB within Southern Bantoid on the phylogenetic tree of Grollemund et al. (2015). Thus, as discussed in §6.1.2, within Southern Bantoid the pre-PB NDAI interrogative construction should be minimally reconstructed to the most recent common ancestor of Narrow Bantu, Narrow Grassfields, Ring Grassfields, Mundabli [mund1328, Southern Bantoid], and Tivoid.⁵⁸ At the same time, as discussed in §6.2.4, the pre-PB construction that resulted in the PB interrogative stem *yíí or *yíí ‘what?’ is likely to have already achieved this degree of fusion minimally on the stage of the most recent common ancestor of Narrow Bantu and Mamfe.

I discussed various formal and semantic changes affecting Bantu NSIPs, some of which are typologically rather trivial, while others are more peculiar. I particularly highlighted two such peculiarities, viz. the surprising patterns of colexification of the human interrogative ‘who?’ and various interrogatives that are either non-human, such as ‘what?’, or indifferent to the difference between humans and things, such as ‘which/what [N]?’ (§5.2) and the tendency to construe interrogative pronominals, especially those questioning subjects, as nominal predicates (§5.3).

The evolution of the Bantu NSIPs discussed in this chapter contains a number of seemingly minor details that however have a more general relevance beyond Bantu linguistics. For example, methodologically, ‘who?’ interrogatives instantiating the Interrogative Modifier construction based on the same BLR3 stem *júmà ‘thing; bead; iron’ in B10 and Eastern and South-Western Bantu discussed in §4.1.6 illustrate the importance for reconstruction of paying attention to the whole range of uses of a given form and not to write them off as insignificant

⁵⁸“Minimally” means “given my current knowledge and understanding of the data”.

quirks. From the perspective of semantic typology, an interesting detail is that the deictics accreted in NSIPs in Bantu (and more broadly in Benue-Congo) (cf. §4.1.3) are preferentially not distal, but intermediate and near-addressee, and that the accreted deictics often also have some endophoric (discourse-referential) and more broadly intersubjective uses. With respect to syntax, Bantu languages provide an example of sentence-final position of interrogatives, which is typologically surprising but natural within the morphosyntax of the respective languages (cf. §4.1.1). Another aspect of the syntax of content questions where Bantu languages provide a particularly interesting theoretical contribution is the finding that the constituent order alternation that synchronically is typically described in terms of fronting of an interrogative out of its in-situ position, historically represents a change in the opposite direction (§5.3.2), which is deeply problematic for any syntactic framework generating the surface constituent order with a sentence-initial interrogative from some underlying syntactic structure where the interrogative is in a different position. The evolution of Bantu NSIPs also highlights the relevance in the morphosyntax of Benue-Congo languages of the distinction between identifying and non-identifying modification (and their respective nominalisations), as well the interesting parallel with the distinction between SIPs and NSIPs respectively (§6.1.5.2–6.1.5.4).

Last but not least, because of their complex constructional origin and the typical pathways of formal and semantic evolution, the reconstruction of interrogative pronominals bears significant relevance to the reconstruction of many other parts of Bantu morphosyntax. Some of the topics of historical Bantu morphosyntax where this chapter made a contribution include:

- The reconstruction of class 5 (§4.1.4.1, §4.1.4.2, §6.1.5.3), with respect to both its markers (BGR forms **i-* and **di-*) and a number of forms historically related to them (augment, copulas, person index morphology, reflexive marker, determiners), as well as its semantics (singulative, selective or restrictive reference, relation to the features [+human] and [+animate] and class 1);
- The reconstruction of the pre-PB predicate **ka* ‘be at (X’s place)’ as the source of the similarly shaped prepositions, infinitive prefixes, the so-called amplexive morphemes in the connective construction, the NKA- type accretion in locative interrogatives and interrogative pronominals, possessive nominalisers, the diminutive class 12 marker (§4.1.5, Appendix F);
- The reconstruction of the pre-PB intermediate distance demonstrative **ná* (§4.1.5, Appendix E);

- The reconstruction of subject prefixes in Eastern Bantu (§4.1.4.2, Appendix D);
- The reconstruction of pre-PB cleft construction (§6.1, §6.1.5.4);
- The reconstruction of the PB class 1 subject marker *à as the continuation of the pre-PB 3SG personal index *à (§6.1.4);
- Refining the reconstruction of the vowel sequence traditionally reconstructed as *ai (§6.2.3);
- Clarifying PB root phonotactic patterns with respect to the absence of vowel-initial roots – contra Wills (2022 [this volume]) – and the identity of PB *j
(§6.2.2, Appendix H);
- Refining the reconstruction of the PB augments (pronominal prefixes) of classes 1, 9 and 10 (Appendix H).

Acknowledgements

This work is part of the project GL7 “Reconstruction, genealogy, typology and grammatical description in the world’s two biggest phyla: Niger-Congo and Austronesian” of the Labex EFL, supported by a public grant overseen by the French National Research Agency (ANR) as part of the program “Investissements d’Avenir” (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris (reference: ANR-18-IDEX-0001). I also gratefully acknowledge the support of the AdaGram project (<http://llacan.cnrs.fr/AdaGram/index.html>), funded by the “Emergence(s)” program of the City of Paris. Special thanks with respect to the present chapter are due to Mark Van de Velde. Last but not least, I am grateful to the referees and the editors for their helpful comments.

Abbreviations

[xxxx1111,...]	glottocode (https://glottolog.org), and language family	HOD	hodiernal (past)
		INF	infinitive
		IPFV	imperfective
		LOC	locative
		LOG	logophoric
1, 2 ...	noun class, unless combined with SG and PL, in which case it refers to person	M	masculine
		N	(i) noun; (ii) neuter
		NADR	near-addressee (demonstrative)
		NF	non-final form in Eton A71
ADV	adverbial	NMLS	nominaliser
AG	agreement pattern	NSIP	non-selective interrogative pronominal
ANIM	animate		
AUG	augment	OBJ	object
AUX	auxiliary	OPT	optative
CL	noun class	P	predication, predicate
CNJ	conjoined	PFV	perfective
COP	copula	PL	plural
DEM	demonstrative	POSS	possessive
DEM _I	demonstrative of type I in Liko D201	PRES	presentative
		PRS	present tense
		PST	past
DEM _{II}	demonstrative of type II in Liko D201	REL	relativiser, relative form
		REM	remote (past)
		SG	singular
EQ	equational	SIP	selective interrogative pronominal
EXCL	exclusive		
FUT	future	SUBJ	subject
FV	final vowel	V	(i) verb; (ii) vowel

Appendix A Some examples of a formal link between interrogatives and demonstratives in the wider Bantoid domain

A formal link between (non-selective) interrogatives, on the one hand, and non-distal, discourse-referential and intersubjective demonstratives, on the other hand, similar to the one highlighted in §4.1.3 for Bantu, is equally found in the wider Bantoid domain.

For example, in Mundabli [mund1238, Southern Bantoid] *nā* ‘where?’ is occasionally accreted with the postposed ‘locative modifier’ *f-á*, which is believed to be a proximal deictic in origin (Voll 2017: 256, 333). In Tikar [tika1246, Northern Bantoid] (Stanley 1991), the demonstrative system is fundamentally based on two stems, the proximal *-ε* and the distal *-i*, as in the pairs of locative adverbials *f.ε* or *c.ε̃* ‘here’ and *f.i* or *c.ĩ* ‘there’, presentative demonstratives marked for class *n-ε* ‘this one, here it is’ and *n-i* ‘that one, there it is’ (class 1),⁵⁹ or manner adverbials *l.ε* ‘so, like this’ and *l.i* ‘so, like that’.⁶⁰ When a demonstrative can be used for discourse-referential purposes, the distal forms with *i* are used anaphorically, while the proximal forms with *ε* are used cataphorically (cf. Stanley 1991: 295). NSIPs all have the same proximal/cataphoric *ε*-vocalism, viz. *w.ε.n* ‘who?’ (class 1), *y.ε.n* ‘what?’ (class 3), *f.ε.n* ‘where?’.⁶¹ Furthermore, NSIPs are used in a cleft

⁵⁹The class numbering in Tikar does not follow the Bantu system, except for the use of odd numbers for singular classes and even numbers for plural classes and the use of class 1 for the class for nouns with mostly human referents. In the Tikar spelling used in Stanley (1991), the vowels unmarked for tone have high tone in classes 1 and 6 and mid tone elsewhere.

⁶⁰Strictly speaking, synchronically the two deictic stems can be analysed as morphemes only in the presentative demonstratives that agree in class. However, the submorphemic structure that reflects the past morphological borders is sufficiently transparent in the remaining forms and I indicate it with dots instead of hyphens. Thus, the initial *f* in the locative adverbials is recurrent in such forms in Bantoid and is a cognate of the PB locative class 16 **pa*. The initial *c* in the variant forms of the locative adverbials is suggested by Stanley (1991: 297) to come from *ci* ‘place’ and the adverbials from *ci s-ε* [3.place 3-this] and *ci s-i* [3.place 3-that].

⁶¹The submorphemic structure of these non-selective interrogatives is sufficiently transparent, even though the etymological source of some of the submorphemic elements may be debatable. Like with the locative adverbials above, the initial *f* in ‘where?’ is a trivial cognate of the PB locative class 16 **pa*. The initial *w* in ‘who?’ can be a class 1 prefix, which often has this shape in Bantoid, perhaps the same as the PB class 1 subject prefix **ɔ-*. The initial *y-* in ‘what?’ is also in all probability a reflex of a class prefix, such as class 5 or 7. The final *n* in all these forms may have a variety of sources, a copula, a relativiser, a focus marker, a demonstrative, but most likely it is a reflex of the older Benue-Congo interrogative stem **nà* ‘where?’, the same stem as reflected in Mundabli [mund1238, Southern Bantoid] *nā* ‘where?’ mentioned above and in Bena-Yungur [bena1260, Buto] (cf. Appendix G) *nā* ‘where?’.

construction where the interrogative is followed by the proximal presentative demonstrative based on the stem *-ε* agreeing in class with the interrogative and identical to the relativiser, viz. *wεn n-ε* ‘who is it that [P]?’ (class 1) and *yεn s-ε* ‘what is it that [P]?’ (class 3). Interestingly, while for regular nominals focalised by means of a cleft construction the distal form of the presentative demonstrative may also be used contributing some (unclear) additional deictic meaning (cf. Stanley 1991: 496), this option does not seem to be available for the NSIPs.

In Kenyang [keny1279, Mamfe] (Voorhoeve 1980: 280–282), the same deictic stem *né* (after a V- or N- class prefix, which is deleted) $\sim \acute{e}n$ (after a CV- class prefix whose vowel is dropped) is used to form presentative demonstratives (no distance distinctions, but necessarily visible), anaphoric demonstratives, relativisers and the selective interrogative ‘which [N]?’. The latter selective interrogative has the structure [^L + class agreement + *né* \sim *én*], as in ^L*né* (class 1) and ^L*b-én* (class 2). The floating ^L tone may be the same morpheme as the nominal marker (basically, a nominaliser) *à-* and *è-* (depending on the noun class) found in independent (presentative) demonstratives and relative pronouns, as in *à-b-én* ‘these ones / those ones that (class 2)’ and *è-n-én* ‘this one / the one that (class 5)’.

Ngwo [ngwo1241, Momo Grassfields] (Eyoh 2011) shows an intriguing parallelism between the stems of its NSIPs and intermediate demonstratives (close to the addressee) on the one hand and its SIPs and distal demonstratives (far from both the speaker and the addressee) on the other. Thus, in Ngwo we find (*à*)*wé* ‘who?’ (class 1) and (*à*)*yé* ‘what?’ (presumably, class 7), both bearing a resemblance to the Tikar non-selective interrogatives and Kenyang selective interrogatives cited above, vs. *w-ē* ‘be there (close to the addressee)’ (class 1) with the stem *-ε* on the one hand, and N *w-ē* ‘which [N]?’ (class 1) vs. *w-ē* ‘be there (far from both the speaker and the addressee)’ (class 1) with the stem *-e*, on the other hand.

Appendix B Type N(D)I: Further examples

Across Bantu, accreted material of type N(D)I is often found on the NSIPs ‘who?’ and ‘what?’. For example, compare Batanga A32C *njani* ‘who?’ vs. *njaε* ‘what?’ with Duala A24 *njá* ‘who?’ vs. *njé* ‘what?’; Ntomba-Bikoro C35a *nòní* ‘who?’ vs. Ntomba-Njale C35a *no* ‘who?’; Songola Kasenga D24 *nàíndí* ‘who?’ vs. Enya Kibombo D14 *kì-úmà nàání* ‘what?’ and Enya Manda D14 *kì-úmà nàání* ‘what?’. In Kagulu G12, next to the older forms *cl-ani* ‘where?’ and *=ki* ‘what?’, you also find *=ni* ‘what?’, *nhani* ‘how?; why?’, *choni* ‘what?’ (default), *dyoni* ‘what?’ (something said), *hoki* ‘where?’ (Petzell 2008: 89–92, 177). The latter three forms are

analysed in the source as [class marker + “reference marker” -o + =ni or =ki ‘what?’], so that these forms may represent yet another cycle of substance accretion, this time with pronominal forms, viz. *ch-o* of class 7, *dy-o* of class 5, *h-o* of class 16, based on the substitutive stem *-o.

Appendix C The copula I is not related to the copula N(D)I

Although traditionally the copula type I, viz. **i* ~ **i* and *^H, and the copula type N(D)I, viz. **ni* ~ **ni* and **ndi* ~ **ndi*, are considered allomorphs, I believe that they have different origins. The copula I has a deictic origin and most likely derives from the same deictic source as the nominaliser augment, viz. the pre-PB determiner **yé*, discussed in §6.1.5.3. The copula N(D)I may have a number of origins, which probably are all ultimately deictic as well. For the moment, I find the hypothesis proposed by Givón (1974) most appealing. Givón was focusing on data for some Eastern Bantu languages, for which the initial *nd-* cluster of the copula may reflect a stacking of two copulas, viz. the pre-PB (Niger-Congo) copula **ni* ‘be at, be with’ and the copula **li* (corresponding to the form **di* in BGR). Although not further discussed by Givón (1974) the latter copula **li* is also likely to be much older than PB.

Appendix D Subject prefixes originating in the inflected forms of the locative copula in Eastern Bantu

Subject prefixes originating in the inflected forms of a copula, which typically has the form -*a*, appear to be rather common in Eastern Bantu. So far, this evolution has been attested in zones D, E, G, N and P. For example, Bernander (2017: 82) reports the use of the 1st and 2nd person copula that “consists of the subject marker and a particle -*a*” in a number of Tanzanian Eastern Bantu languages of zones G, N and P. In this respect, note that Petzell’s (2008) synchronic analysis of the Kagulu G12 verb *kuwa* ‘be’ as *k-uw-a* [15-be-fv] with the stem -*uw* and the final vowel -*a* is mostly likely inadequate from a historical perspective. The copula stem -*a* may be lost without traces resulting in subject markers that appear to be used on their own as copulas, as Gibson et al. (2019: 219) report for Digo E73 and Swahili G42d. Much further away, we find a very similar situation in Liko D201, where the present form of the verb ‘be’, that exists only for persons and class 1, is *nà* ‘1sg.be’, *wà* ‘2sg.be’, *à* ‘3sg.be’ and it is formally identical to the respective subject prefixes (de Wit 2015: 395). Although de Wit (2015) does not

analyse these inflected forms further, at least the 1SG and 2SG forms are clearly analysable as a subject index prefix and the stem *-á*.

Appendix E The intermediate deictic **ná* as the source of the element -NA accreted in locative interrogatives

The element *-na* reported by Doneux (1971: 134–135) to be often accreted on ‘where?’ interrogatives in languages of zone J is likely to originate in some kind of deictic element or an information-structural element, somewhat like in French *où ça?* ‘where? (tell me!)’, lit. ‘where that one?’ (or ‘where here?’) next to the neutral *où* ‘where?’ (cf. §4.1.2). This is suggested by its position and shape. A particularly plausible source is the deictic stem **ná* which in all likelihood originally functioned as an intermediate deictic. Thanks to its wide distribution, we can safely reconstruct it to PB as a retention from an earlier stage. For example, compare the Bangi C32 intermediate demonstrative stem *-ná* ‘that [N] (visible)’ (Whitehead 1899: 21; MacBeath 1940: 14), the Leke C14 distal demonstrative stem *-ná* (Vanhoudt 1987), and the demonstrative stem *-ná* in Ewondo A72a, which has the proximal meaning ‘here’ when used to build modifying demonstratives, as in *é-m-óngó jó-ná* [AUG-1-child 1.PRES-here] ‘this child (here)’ (Abessolo Nnomo & Etogo Mbezele 1982: 190), and the intermediate meaning ‘there (intermediate)’ in the relic adverbial demonstrative forms of class 16 *vá-ná* and class 18 *mú-ná* (Grégoire 1975: 118). In Liko D201, the “connecting clitic” *-ná* is “often present [after] a type II demonstrative” (i.e. a proximal demonstrative), when it modifies a noun in the construction [N + *ní* COP + CL-DEM_{II}] and is not “at the end of a clause” (de Wit 2015: 259). Limbum [limb1268, Mbam-Nkam Grassfields] has *ná* ‘here’ (Fransen 1995). Finally, Mbula [mbul1261, Jarawan Bantu] has *ná* as a (presentative) copula and a kind of focus marker, which is also an integral part of the Mbula NSIPs.

Appendix F **ka* ‘be at (X’s place)’ as the source of the accreted element (N)KA- in interrogatives and of a number of other KA elements in Bantu

The interrogatives accreted with (N)KA- are originally locative interrogatives ‘where?’ which in some languages, following the usual paths of semantic change

of interrogatives, evolved to the selective interrogative ‘which one?’ and ultimately to the non-selective interrogatives ‘who?’ and ‘what?’. The etymology of the element (N)KA- which matches the locative origin of these interrogatives particularly well, both semantically and formally, is **ka* ‘be at (X’s place)’.

In fact, there is a whole range of functional elements in Bantu (and far beyond) that can be argued to have been ultimately sourced from the locative predicate **ka* ‘be at (X’s place)’. For example, in Liko D201, we find *ká* the general preposition ‘to, at, in, on, for’, *ká-* the infinitive prefix of class 9b, *kà-* the possessive relator in the “genitival” construction (different from the “associative” construction) and (with an allomorph *kǎ-*) the possessive nominaliser prefix with person indexes (‘the one of X’ as in ‘the one of me, mine’) (de Wit 2015). In Mongo C61, we find *ěkà* the preposition ‘at somebody’s place’ (Hulstaert 1957), with dialectal variants *kà ~ ká* (cf. Hulstaert 2007: 294, 296), and the *kà ~ ká* part of several of the connective stems (cf. Van de Velde 2013: 231). The Mongo forms for ‘where?’ are particularly relevant: *nkó* (Nkundo), *ńkó*, *ńkò*, *nká* and *nké* (some other varieties) (Hulstaert 1957; 2007: 290). In Konda C61E, the “locative possessive” construction uses the “old locative *-(n)ka*”, as in the preposition *è-kà ~ é-kà ~ è-nká* ‘at (somebody’s place)’ (Motingea Mangulu 2018: 53–54). The initial *e-* in these forms is either the old (locative) class 24 (as suggested by Motingea Mangulu 2018: 54) or the class 9 verbal prefix used for the enforced agreement with locative and temporal predicates (cf. Motingea Mangulu 2018: 44). Note that in Mongo, the rising tone of *ě-* in the preposition *ěkà* suggests that it was a relative (verbal) prefix and that *kà* has a predicative origin. In Mbula [mbul1261, Jarawan Bantu], *kà* is one of the possible possessive relators and a nominaliser ‘one of, from, among X’, as well as a deictic element that usually expands other demonstratives and closes some types of dependent clauses. In the latter two functions, it sometimes appears preceded by a nasal.

Reflexes of the same element **ka* throughout Bantu have also been described as the so-called amplexive morpheme in the connective construction (cf. Van de Velde 2013: 229–230). See Van de Velde (2013: 230) for an overview of various hypotheses on the origin of *ka* in the Bantu connective construction. One such hypothesis suggests that *ka* in the Bantu connective construction is a reflex of the often diminutive class 12 prefix *ka-*. As a side comment, I argue that there is indeed a relation between the two forms but that this relation is indirect and that the two elements both ultimately go back to the predicate ‘be at (X’s place)’. The diminutive use of *ka-* is likely to have evolved from its use as a nominaliser ‘one of, from, among X’, as in Liko, where it has a primarily possessive meaning ‘the one of X’, and Mbula, where the meaning is broader ‘one of, from, among X’ (e.g. *pwàrì kà à-nlé:rú nà* [sun one.of PL-star COP.PRES] ‘The sun is a star’, lit.

‘The sun, it is one of, one among the stars’). A diminutive would be a natural evolution for a form with a partitive meaning ‘one of, from, among X’ > ‘just a part of X’ > ‘small part of X’ > ‘small X’.

As already suggested by Welmers (1963) with respect to the uses of **ka* in the connective construction, in Bantu **ka* is clearly a retention from a much older stage. Its reflexes are well-attested not only throughout the Bantu domain, but well beyond it, both within Benue-Congo and in other Niger-Congo groups. Given that in Bantu reflexes of **ka* occasionally show up with (relative) verbal prefixes, as in the Mongo and Konda examples mentioned above, or in Zulu S42 (cf. Van de Velde 2013: 229), **ka* is likely to ultimately have a verbal origin, as something like ‘be at (X’s place)’, ‘be near (X’s place)’ or ‘be in contact, relation with (X)’. Its uses as a preposition ‘at (X’s place)’ or a connective relator are clearly later evolutions.⁶² Such verbal sources of prepositions are not uncommon in Niger-Congo.

The nasal part in (N)KA- may have at least two origins. First, it may originate in a (presentative, identificational) copula (aka nominal predicative marker), which is sometimes assumed to have had a variant **n* in addition to **n(d)í ~ *n(d)í, *í ~ *í*, and a purely tonal **H* variant (cf. Grégoire 1975: 125; Coupez 1977). Second, the nasal may reflect the class 9 nominal prefix **n-* used to nominalise the locative predicate ‘be at (X’s place)’ (or the preposition ‘at (X’s place)’) into a relational noun ‘the one at (X’s place)’. In this respect, compare the class 9 noun *pǎ* ‘place’ in Liko which is “similar to **pa-*, the reconstructed Proto-Bantu noun-class prefix of class 16” (de Wit 2015: 175). Another interesting example in this respect is provided by Konda, where the preposition ‘at (somebody’s place)’ appears to be used without the nasal when combined with a bound personal index, as in *ékǎsó* ‘at our place’ containing the bound 1PL index *-ísó*, but with the nasal when combined with a free personal pronominal, as in *ènká nísó* ‘at our place’ containing the free 1PL pronominal *nísó* (Motingea Mangulu 2018: 53–54).

⁶²Motingea Mangulu (2018: 53) hypothesises an evolution in the opposite direction suggesting that we may be dealing with a locative form that became an auxiliary (“locatif auxiliarisé”). However, such an evolution is unlikely given both the usual directionality of change known for such elements across Niger-Congo and cross-linguistically and the fact that the predicative properties of reflexes of **ka* are well-attested beyond zone C languages. Motingea Mangulu (2018: 53) cites further predicative uses of this element, such as Bangi C32 defective verb *kà* ‘be(come) (with a certain quality, e.g. blindness)’ used only in the present tense (Whitehead 1899: 32; MacBeath 1940: 28) or a similar verb *kà* in Yasanyama (a language from the Upper Tshuapa, presumably zone D or C), as in *linà lí-k’èé lí-kà nání* [5.name 5-COP-2SG.POSS 5-COP who?] ‘What is your name?’ (lit.: ‘The name that is of you is who?’). It is likely that the same locative predicate **ka* ‘be at (X’s place)’ is reflected in various copula forms in languages of zone C, such as Bangi C32 *ngá* COP.PRS, *líki* COP.PST.HOD, *líki* COP.PST.REM (Whitehead 1899: 32; MacBeath 1940: 28) and Konda C61E *kí* COP.PST (Motingea Mangulu 2018: 53).

When used as a preposition, a connective relator or in the locative interrogative, this nominalised form may have been preceded by a copula, comparable to the Mongo connective relator *˘-lěkà* expressing ownership and contrastive focus on the owner and based on the relative form of the copula *lè* and the preposition *ěkà* ‘at (somebody’s place)’ (cf. Van de Velde 2013: 231).

Appendix G Reflexes of the determiner **yé* in Buto (aka Bena-Mboi)

Beyond Bantoid, a particularly interesting set of forms sourced from the determiner **yé*, both as a noun class determiner and as a selective/restrictive determiner, can be found in Bena-Yungur [bena1260, Buto] and other languages of the Buto group (aka Bena-Mboi [bena1258]), a small Benue-Congo subgroup spoken in the north-east of Nigeria immediately to the north of Mbula (cf. Idiatov & Van de Velde 2019 on the classification of Buto as a Benue-Congo group).⁶³ As described in Van de Velde & Idiatov (2017), adnominal modifiers in Bena-Yungur can agree in class with the noun they modify. The three agreement classes (noun classes), viz. *WA*, *YA* and *ɓA* as referred to by the inflected determiner forms, equally used as demonstrative modifiers that do not distinguish distance in space, can each be triggered by either a singular or a plural noun. The three inflected determiners, *wā*, *yā* and *ɓā*, are based on the proximal presentative demonstrative stem *-ā* (cf. Idiatov & Van de Velde 2018). Third person indexes do not agree in class, but in animacy. The determiner **yé* is likely to have been the source of a number of elements in Bena-Yungur. Most noticeably, the determiner **yé* is particularly plausible as the source of the class *YA* morphology, such as *y-*, the agreement marker on possessive pronominals and the remnant class prefix on nouns; *-e*, the agreement marker on some modifiers and the frozen class marker on nominal stems; and *yī*, the determiner of class *ya* without the proximal presentative demonstrative stem *-ā* (cf. Idiatov & Van de Velde 2018). While in Bena-Yungur, the assignment of nouns to class *YA* does not have a clear semantic basis, in the related language Mboi [mboi1246, Buto] class *YA* is limited to human nouns and acceptable for some nouns designating animals. This is reminiscent of the Bantu reflexes of the determiner **yé* in forms of class 1 and person indexes. Like with the Babungo [veng1238, South Ring Grassfields] class 5 agreement markers and the Mbula nominaliser *-yí*, class *YA* agreement markers can also be used in Bena-Yungur for purposes other than the expression of agreement or nominalisation.

⁶³The data on Buto languages come from my joint research with Mark Van de Velde.

In particular, in certain constructions requiring the presence of a determiner, when the controller is of noun class WA, the agreeing determiner of class WA may be replaced by a non-agreeing determiner of class YA to change the interpretation of the preceding adnominal modifier licensed by the determiner from qualifying or classifying (as in ‘[I like] watery_{WA} porridge_[WA] (in general)’ to identifying (as in ‘[I like] the porridge_[WA] that is watery_{YA} (when there are several types of porridge under discussion)’ (cf. Idiatov & Van de Velde 2018). Outside of the noun class system, we find in Bena-Yungur, like in Bantoid, a singular logophoric person index, *yí ~ yá* SG.LOG.ANIM. And like in PB, in Bena-Yungur we also find reflexes of **yé* at the beginning of free personal pronominals (the Bantu substitutes) as *í-* in *ínâ* 1SG and *ítâ* 1PL.EXCL vs. *áysâ ~ áísâ* 3SG.ANIM,⁶⁴ and possibly also in the beginning of bound possessive modifiers as the vowel length of *a:* in *-a:n^M*-AG 1SG.POSS and *-a:t^M*-AG 1PL.EXCL.POSS vs. *-a:t^H*-AG 3SG.POSS and *-a:y^H*-AG SG.LOG.POSS.

Appendix H On BLR’s initial **j*: PB roots were consonant-initial

The reconstruction of PB **j* in BLR has long been known to be highly problematic. As Wills (2022 [this volume]) correctly concludes, BLR’s PB **j* is “a collection of distinct stories which require separate reconstructions, some clearer than others”. While elucidating all these distinct stories would go far beyond the scope of this chapter, I have to address one aspect of Wills’ reconstruction that is relevant for the reconstruction of PB NSIPs, viz. the reconstruction of initial \emptyset (zero) for BLR’s **j* in verb and noun roots, as well as pronominal prefixes (arguments). For brevity’s sake, I will refer to these roots as JZ-roots, short for roots with initial **j* or zero. I argue that PB roots were consonant-initial and that BLR’s initial **j* confounds several PB consonants, including minimally **s*, **z*, **ʃ*, **y*, and **g*. In what follows, the discussion will necessarily be limited to the gist of the argument. For a more detailed account, see Idiatov (In preparation).

To begin with, there is no doubt that at some pre-PB stage JZ-roots were consonant-initial. This is the canonical phonotactic pattern throughout Niger-

⁶⁴Compare the difference between the PB preprefixes **i-* in 1PL and 2PL and **i- ~ *i-* in class 1 person indexes in Kamba Muzenga’s (2003) reconstruction. In Bena-Yungur, more like in Meeussen’s (1967) PB reconstruction, the free personal pronominals of the first and second persons have a different structure from those of the third person. In the former, the person is indexed by the second morpheme, such as *n-* in *í-n-â* 1SG. In the latter, the person is indexed by the first morpheme, viz. *á-* in *áysâ ~ áísâ* 3SG.ANIM and *bá-* in *bá:bô* 3PL.ANIM (where *bá:-* < *bá-i-*).

Congo for noun and verb roots, while vowel-initial roots emerge through consonant-loss or borrowing. Numerous reliable cognates of JZ-roots can be found beyond Bantu with their initial consonant still preserved. Thus, in (20) below several PB JZ-roots are compared with their cognates in Nizaa [suga1248, Mambiloid], as well as the corresponding pre-Nizaa internal reconstructions (Endresen 1991). In (21), several PB JZ-roots are compared with their cognates in the Buto group (aka Bena-Mboi [bena1258], cf. Appendix G.) accompanied by the initial consonants reconstructed for these roots in Proto-Buto (based on the internal reconstruction by Idiatov & Van de Velde 2020).

- (20) Some BLR JZ-roots and their modern Nizaa (MN) and pre-Nizaa (PN) cognates (Endresen 1991)
- a. BLR 6142 **jíd* ‘become dark, become black’ || MN *sír* ‘black’ < PN **síd*
 - b. BLR 3616 **jóm* ‘be dry’ || MN *sóm* ‘be dry’ < PN **sóm*
 - c. BLR 1602 **jòd* ‘laugh’; BLR 1604 **jòdà* ‘laughter’ || MN *swēē* ‘laugh’ < PN **sōd-ā*; MN *sòr* ‘laughter’ < PN **sōd*
 - d. BLR 3577 **jónk* ‘suck, suckle’ || MN *swāā* ‘suck’ < PN **sOŋ-a*
 - e. BLR 3429 **jíjad* ‘be full’; BLR 3430 **jíjvd* ‘become full’ || MN *yír* ‘be full’ < PN **yíd*
- (21) The cognates of some BLR JZ-roots in the Buto group (with reconstructions of their initial consonant in Proto-Buto, based on Idiatov & Van de Velde 2020)
- a. BLR 3615 **jòm* ‘hit’ || Bena-Yungur *zəmà* (Guto), *səmà* (Pra) ‘kick’ (*z-)
 - b. BLR 1583 **jénjé* ‘cricket’ || Bena-Yungur *zèèzè* (Guto), *sèèsé* (Pra) ‘cricket’ (*z-)
 - c. BLR 3350 **jíki* ‘bee’ || Bena-Yungur *zì-ò* (Guto), *sì-ò* (Pra), Mboi *zih-ò* ‘bee’ (*z-)
 - d. BLR 3525 **jóg* ‘bathe’ || Mboi *só?* ‘bathe, take a bath’ (*s-)
 - e. BLR 3530 **jòk* (vi) ‘roast, (vi) burn’ || Bena-Yungur *yóó* ‘roast, fry’ (*y-)
 - f. BLR 1553 **jáb-ɔk* ‘cross river’; BLR 3138 **jáb-ík* ‘soak in water’; BLR 9809 **jáb-am* (vi) ‘soak’; BLR 3140 **jáb-ú* ‘crossing place, bridge’ || Bena-Yungur *yàbà* ‘bathe, take a bath’ (*j-)

Even closer to Narrow Bantu, Elias et al. (1984: 36–38) reconstruct for Proto-Mbam-Nkam Grassfields only consonant-initial verb roots and just a few vowel-initial noun roots “which have incorporated the prefix as part of the stem”. This is reminiscent of the tendency for the reflexes of the PB JZ-roots not preceded by **i* or **n* to be vowel-initial when they are nominal, and consonant-initial when they are verbal, as highlighted by Wills (2022 [this volume]): “a major difference between the vowel-initial nouns and verbs is the frequent presence of glides before the verb stems”. Wills accounts for this difference between nominal and verbal JZ-roots by assuming that a palatal glide appeared due to hiatus resolution only in verbal vowel-initial roots and then “in some languages [...] the glide variant of the verb was generalised throughout (and sometimes even strengthened)”. This is not inconceivable, but it is definitely not the most straightforward interpretation, especially given that beyond Bantu or Bantoid the roots are consonant-initial. Both for Proto-Mbam-Nkam Grassfields and PB, we can simply assume that certain kinds of root-initial consonants were lost in the relevant nouns because there they only occurred in an intervocalic environment following the same one or two (viz. singular and plural) CV- noun class prefixes, while they often happened to survive in verbs because there they appeared in a variety of contexts, most importantly word- and utterance-initially after a pause (as in the imperative construction). The inventory of the root-initial consonants in Proto-Mbam-Nkam Grassfields reconstructed by Elias et al. (1984: 39) includes both the palatal series **c* and **j* and the alveolar voiceless fricative **s*. As a side consequence, we also have to reject Wills’ suggestions of “relabelling both **c* and **j* as **s* and **z*” and “to remove the palatal series altogether”.

Another problematic aspect of the scenario proposed by Wills with the initial consonants consistently emerging in JZ-roots out of zero through epenthesis and strengthening is that often in a given language, especially in the north-west, we find a whole range of different reflexes of **Ø* whatever the environment, with no way to account in any principled fashion for why in some cases no epenthesis would take place (i.e. **Ø* would stay *Ø*), while elsewhere some glide would be epenthesised and occasionally further strengthened to a specific fricative, affricate or stop. For example, as illustrated in (22), in Eton A71, reflexes of the presumed **Ø* in verbs can be as diverse as *Ø*, *y*, *j*, *ɲ*, *c* and *s*. This comparison is limited to verbs to avoid the complication of a possible merger of the stem with a class prefix in nouns. In Eton nouns, we find an additional reflex *z*, as in (22f).

(22) The reflexes of some JZ verb roots in Eton A71 (based on Van de Velde In preparation)

a. \emptyset

- i. BLR 1602 **jòd* ‘laugh’ > *wè* ‘laugh’ from earlier **ʒl*⁶⁵
- ii. BLR 3525 **jóg* ‘bathe’ > *wágô* ~ *wógô* ‘bathe’ from earlier **ʒgá*⁶⁶

b. *y*

- i. BLR 3145 **jác* ‘open the mouth; yawn’ > *yààni* ‘yawn’, *yázi* ‘open (the door)’
- ii. BLR 3295 **jén* ‘see’ > *yén* ‘see’
- iii. BLR 3338 **jíg* ‘learn; imitate’ > *yágî* ‘learn; imitate’

c. *j*

- i. BLR 3429 **jíád* ‘be full’ > *já* ‘be(come) full’
- ii. BLR 3387 **jíb* ‘steal’ > *jíb* ‘steal’

d. *ɲ*

- i. BLR 3177, 3178 **jám(ú)* ‘suck’ > *ɲáŋ* ‘suck’
- ii. BLR 3147 **jàd* ‘spread’ > *ɲèd* ~ *sèd* ‘spread’

e. *c*

- i. BLR 3167 **jàk* ‘be lit; (vi) burn’; BLR 9595 **jàki* ‘(vt) light’ > *càk* ‘(vt) light’

f. *s*

- i. BLR 8668 **jáng* ‘say no, refuse; hate’ > *sááni* ~ *sééni* ‘quarrel, argue’, *záŋ* (9/10) ‘(n) quarrel’
- ii. BLR 5329 **jògò* ‘loud noise’; BLR 7098 **jògòd* ‘make confused noise’ > *sòg* ‘shout to scare off (e.g. a thief); boo, jeer at’
- iii. BLR 3147 **jàd* ‘spread’ > *ɲèd* ~ *sèd* ‘spread’

Finally, a scenario implying frequent consonant epenthesis, especially in verb roots, is problematic because all Bantu languages allow vowel-initial utterances and most modern Bantu languages are also perfectly fine with vowel-initial roots.

⁶⁵The form of this verb in Eton is the result of two productive morphonological processes, viz. the breaking of [ɔ] to *wa* in certain stem-initial syllables (cf. Van de Velde 2008a) and the subsequent fronting of *a* to *ɛ* due to the vocalisation of the word-final [l] to *i* followed by vowel coalescence (cf. Van de Velde 2008b: 35, 246–247).

⁶⁶The form of this verb in Eton is due to the same breaking of [ɔ] to *wa* as with *wè* ‘laugh’ in combination with the assimilation of the final vowel. In this respect, compare *dóló* ~ *dwáló* ‘5 francs’ which is a borrowing from English *dollar* (cf. Van de Velde 2008a).

It is true that many Bantu languages do not tolerate hiatus, but if hiatus is resolved through consonant epenthesis, the choice of the epenthetic consonant is determined by the vowels involved and is limited to a palatal or labial-velar glide, *y* or *w*.

The cognates from outside of Bantu cited in (20) and (21) suggest that the two JZ verbs with \emptyset reflex in Eton in (22a), viz. ‘laugh’ and ‘bathe’, should be reconstructed with initial **s* in PB. That the initial **s* had not yet lenited to \emptyset in PB is confirmed by the reflexes of the augment of class 10, reconstructed by Meeussen (1967: 97) as **ji*, but which should rather be reconstructed as **si*. Contra Wills (2022 [this volume]), the initial consonant of the augment (pronominal prefix) of class 10 is not an “Eastern innovation”, as we also find it in Grassfields, such as the Proto Grassfields connective marker of class 10 reconstructed by Hyman & Tadadjeu (1976: 76) as **sí ~ *í*, and in zone A, such as the Yangben A62A class 10 prefix allomorph *sy^(L)-* before some vowel initial-roots (cf. Boyd 2016), and zone B, such as the augment of class 10 *(s)i-* and pronominal and verbal prefixes of class 10 *s^H-* in Orungu B11b (cf. Ambouroué 2007: 60, 86; and example (19) in §6.1.4 of the present chapter). Similarly, the initial consonant of the augment of class 9 is not an “Eastern innovation” but a retention from pre-PB and should be reconstructed as PB **zi*, while the augment of class 1 should be reconstructed as PB **gɔ*.

As amply illustrated by Wills (2022 [this volume]), the initial consonants of nominal JZ-roots are best preserved when protected by a preceding nasal of the noun class prefix. In order to reconstruct the initial consonants of those nominal JZ-roots whose reflexes never happen to be preceded by the nasal of a noun class prefix in any Bantu language, such as BLR 3252 **játò* ‘canoe’ (N 14), we need to find their cognates beyond Bantu with initial consonants preserved. These consonants would probably reflect earlier **s*, **z*, **ʃ*, **y*, or **g*. Thus, for BLR 3252 **játò* external evidence suggests a velar, such as PB **g*, as the most likely candidate. We can be quite sure that the lenition of the initial consonants of these particular problematic nominal roots to \emptyset postdates the PB stage because we can demonstrate that the same initial consonants were still there in PB in other nominal and verbal roots. The most straightforward case here is obviously PB **g*, whose presence in PB has never been a matter of debate. In this respect, see also footnote 55 in the present chapter, on BLR 1386 **gìjí* ~BLR 1385 **gìjé* ‘egg’, which should be reconstructed as **gìgì*.

References

- Abessolo Nnomo, Thierry & Luc Etogo Mbezele. 1982. *Eléments de grammaire ewondo* (Langues et Littératures Nationales 10). Douala: Collège Libermann.
- Ambouroué, Odette. 2007. *Eléments de description de l'orungu : langue bantu du Gabon (B11b)*. Brussels: Université libre de Bruxelles. (Doctoral dissertation).
- Baka, Jean. 2005. Pourquoi une forme de type *-pái “nouveau” en protobantou. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics, with special focus on Bantu and Mande. Essays in honour of Yvonne Bastin and Claire Grégoire* (Collection Sciences Humaines 169), 139–146. Tervuren: Royal Museum for Central Africa.
- Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and divergence in the Bantu languages: Perspectives from a lexicostatistic study* (Annals – Human Sciences 162). Tervuren: Royal Museum for Central Africa. https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/lexicostatistic-study-bantu-languages.
- Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa. <http://www.africamuseum.be/collections/browsecollections/humansciences/blr>.
- Bernander, Rasmus. 2017. *Grammar and grammaticalization in Manda: An analysis of the wider TAM domain in a Tanzanian Bantu language*. Gothenburg: University of Gothenburg. (Doctoral dissertation).
- Bokamba, Eyamba Georges. 1976. *Question formation in some Bantu languages*. Bloomington, IN: Indiana University. (Doctoral dissertation).
- Bostoen, Koen. 2001. Osculance in Bantu reconstructions: A case study of the pair °-kádang-/°-káng- (‘fry’, ‘roast’) and its historical implications. *Studies in African Linguistics* 30(2). 121–146.
- Bostoen, Koen & Yvonne Bastin. 2016. Bantu lexical reconstruction. In *Oxford handbooks online*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199935345.013.36.
- Bostoen, Koen & Rozenn Guérois. 2022. Reconstructing suffixal phrasemes in Bantu verbal derivation. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 343–383. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575829.
- Bôt, Dieudonné Martin Luther. 1986. *Les substituts ou représentants en basaa*. Yaoundé: University of Yaoundé. (MA thesis).
- Boyd, Virginia. 2016. *Dictionnaire yangben – français*. Dallas, TX: SIL International. <https://www.webonary.org/yangben/>.

- Boyd, Virginia Lee. 2015. *The phonological systems of the Mbam languages of Cameroon with a focus on vowels and vowel harmony*. Leiden: Leiden University. (Doctoral dissertation).
- Bulkens, Annelies. 2009. Quelques thèmes pour ‘pirogue’ dans les langues bantoues. *Africana Linguistica* 15. 27–58.
- Coleman, Arnie, Karen Coleman & Brad Koenig. 2004. *Grammar sketch of Esimbi*. Unpublished ms. Yaoundé: SIL Cameroon.
- Corbett, Greville G. 1991. *Gender* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Coupez, André. 1977. Note sur les morphèmes prédicatifs en protobantou. *Africana Linguistica* 7. 89–91.
- Creissels, Denis. 2005. L’émergence de système à neuf voyelles en bantou S30. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics, with special focus on Bantu and Mande. Essays in honour of Yvonne Bastin and Claire Grégoire* (Collection Sciences Humaines 169), 191–198. Teruren: Royal Museum for Central Africa.
- Cristofaro, Sonia & Fernando Zúñiga (eds.). 2018. *Typological hierarchies in synchrony and diachrony* (Typological Studies in Language 121). Amsterdam: John Benjamins.
- Croft, William. 2002. *Typology and universals*. 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Cysouw, Michael & Olav Hackstein. 2011. The (in)stability of interrogatives: The case of Indo-European. Paper presented at the Workshop “Stability and borrowability of interrogative pronominals”, 20th International Conference on Historical Linguistics. Osaka. http://idiatov.mardi.myds.me/ICHL/ICHL_CYSOUW_HACKSTEIN_slides.pdf.
- Dammann, Ernst. 1977. Das ‘-o of reference’ in Bantusprachen. In Brigitta Benzling, Otto Böcher & Günther Meyer (eds.), *Wort und Wirklichkeit: Studien zur Afrikanistik und Orientalistik, Eugen Ludwig Rapp zum 70. Geburtstag. Teil 2: Linguistik und Kulturwissenschaft*, 31–43. Meisenheim am Glan: Anton Hain.
- Davison, Phil. 2009. *Weh orthography guide*. Yaoundé: SIL Cameroon.
- de Blois, Kornelis Frans. 1970. The augment in the Bantu languages. *Africana Linguistica* 4. 85–165.
- de Wit, Gerrit. 2015. *Liko phonology and grammar: A Bantu language of the Democratic Republic of the Congo*. Leiden: Leiden University. (Doctoral dissertation).
- Diessel, Holger. 2003. The relationship between demonstratives and interrogatives. *Studies in Language* 27(3). 635–655.
- Doneux, Jean Léonce. 1971. Deux études sur les interrogatifs en zone J du bantou. *Africana Linguistica* 5. 119–138.

- Doneux, Jean Léonce & Claire Grégoire. 1977. Une séquence vocalique à la finale de quelques thèmes en protobantou. *Africana Linguistica* 7. 159–199.
- Dryer, Matthew S. 2013. Position of interrogative phrases in content questions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/93>.
- Elias, Philip, Jacqueline Leroy & Jan Voorhoeve. 1984. Mbam-Nkam or Eastern Grassfields. *Afrika und Übersee: Sprachen, Kulturen* 67(1). 31–107.
- Endresen, Rolf Theil. 1991. Diachronic aspects of the phonology of Nizaa. *Journal of African Languages and Linguistics* 12(2). 171–194.
- Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018. The grammar of engagement I: Framework and initial exemplification. *Language and Cognition* 10(1). 110–140.
- Eyoh, Julius A. 2011. *Engwo lexicon: Engwo – English, English – Engwo* (Kay Williamson Educational Foundation – Language Monographs: Local series 3). Cologne: Rüdiger Köppe.
- Fransen, Margo Astrid Eleonora. 1995. *A grammar of Limbum: A Grassfields Bantu language spoken in the North-West province of Cameroon*. Amsterdam: Vrije Universiteit Amsterdam. (Doctoral dissertation).
- Gibson, Hannah C., Rozenn Guérois & Lutz Marten. 2019. Variation in Bantu copula constructions. In María J. Arche, Antonio Fábregas & Rafael Marín (eds.), *The grammar of copulas across languages* (Oxford Studies in Theoretical Linguistics 73), 213–242. New York, NY: Oxford University Press.
- Givón, Talmy. 1974. Syntactic change in Lake-Bantu: A rejoinder. *Studies in African Linguistics* 5(1). 117–139.
- Grégoire, Claire. 1975. *Les locatifs en bantou* (Annales – Série in-8° – Sciences humaines 83). Tervuren: Royal Museum for Central Africa.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112(43). 13296–13301.
- Guthrie, Malcolm. 1971. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages. Part I, Volume 2: Bantu prehistory, inventory, and indexes*. Farnborough: Gregg International.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank (eds.). 2021. *Glottolog 4.5*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://glottolog.org/>.
- Haspelmath, Martin. 1997. *Indefinite pronouns* (Oxford Studies in Typology and Linguistic Theory). Oxford: Clarendon Press.

- Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the world's languages: A comparative handbook*. Berlin: Mouton de Gruyter.
- Hedinger, Robert. 1987. *The Manenguba languages (Bantu A.15, Mbo cluster) of Cameroon*. London: School of Oriental & African Studies, University of London. (Doctoral dissertation).
- Hedinger, Robert. 2008. *A grammar of Akɔ̀se: A Northwest Bantu language* (SIL International Publications in Linguistics 143). Dallas, TX: SIL International.
- Hölzl, Andreas. 2017. *A typology of questions in Northeast Asia and beyond: An ecological perspective* (Studies in Diversity Linguistics 20). Berlin: Language Science Press. DOI: 10.5281/zenodo.1344467.
- Hulstaert, Gustaaf. 1957. *Dictionnaire lomóngɔ – français* (Annales – Série in-8° – Sciences de l'homme – Linguistique 16). Tervuren: Royal Museum of the Belgian Congo.
- Hulstaert, Gustaaf. 1965. *Grammaire du lomóngɔ. Deuxième partie : morphologie* (Annales – Série in-8° – Sciences humaines 57). Tervuren: Royal Museum for Central Africa.
- Hulstaert, Gustaaf. 1966. *Grammaire du lomóngɔ. Troisième partie : syntaxe* (Annales – Série in-8° – Sciences humaines 58). Tervuren: Royal Museum for Central Africa.
- Hulstaert, Gustaaf. 2007. *Eléments pour la dialectologie móngɔ, suite et fin : formes relatives et particules*. *Annales Æquatoria* 28. 201–332.
- Hyman, Larry M. 1979. Phonology and noun structure. In Larry M. Hyman (ed.), *Aghem grammatical structure* (Southern California Occasional Papers in Linguistics 7), 1–72. Los Angeles, CA: Department of Linguistics, University of Southern California.
- Hyman, Larry M. 1980. Babanki and the Ring group. In Larry M. Hyman & Jan Voorhoeve (eds.), *L'expansion bantoue : actes du colloque international du CNRS, Viviers (France) – 4-16 avril 1977, Vol. 1 : les classes nominales dans le bantou des Grassfields* (Société d'études linguistiques et anthropologiques de France, Numéro special 9), 225–258. Paris: SELAF.
- Hyman, Larry M. 1981. *Noni grammatical structure: With special reference to verb morphology* (Southern California Occasional Papers in Linguistics 9). Los Angeles, CA: Department of Linguistics, University of Southern California.
- Hyman, Larry M. 2003. Basaá (A43). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Routledge Language Family Series 4), 257–282. London: Routledge.
- Hyman, Larry M. 2005. Initial vowel and prefix tone in Kom: Related to the Bantu augment? In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics, with special focus on Bantu and Mande. Essays in honour of*

- Yvonne Bastin and Claire Grégoire (Collection Sciences Humaines 169), 313–341. Tervuren: Royal Museum for Central Africa.
- Hyman, Larry M. & Maurice Tadadjeu. 1976. Floating tones in Mbam-Nkam. In Larry M. Hyman (ed.), *Studies in Bantu tonology* (Southern California Occasional Papers in Linguistics 3), 57–111. Los Angeles, CA: Department of Linguistics, University of Southern California.
- Idiatov, Dmitry. 2007. *A typology of non-selective interrogative pronominals*. Antwerp: University of Antwerp. (Doctoral dissertation). http://idiatov.mardi.myds.me/papers/PhD_Idiatov_2007_Final_Version+ERRATA.pdf.
- Idiatov, Dmitry. 2009. A Bantu path towards lack of differentiation between ‘who?’ and ‘what?’ *Africana Linguistica* 15. 59–76.
- Idiatov, Dmitry. 2011. The instability of Eastern Mayan interrogative pronominals: Interrogative pronominals as complex constructions. Paper presented at the Workshop “Stability and borrowability of interrogative pronominals”, 20th International Conference on Historical Linguistics. Osaka. http://idiatov.mardi.myds.me/talks/2011_ICHL_Mayan_slides_Idiatov.pdf.
- Idiatov, Dmitry. 2014. Egyptian non-selective interrogative pronominals: History and typology. In Martin Haspelmath, Eitan Grossmann & Tonio Sebastian Richter (eds.), *Egyptian-Coptic linguistics in typological perspective* (Empirical Approaches to Language Typology 55), 289–322. Berlin: Mouton de Gruyter.
- Idiatov, Dmitry. In preparation. *Proto-Bantu roots were consonant-initial: The multiple identities of Proto-Bantu *j*.
- Idiatov, Dmitry & Mark Van de Velde. 2018. The relation between agreement classes and noun form (morphological) classes. Paper presented at the Workshop “Gender across Niger-Congo”. Berlin. http://idiatov.mardi.myds.me/talks/2018_BERLIN_Gender_Workshop_Idiatov_Van_de_Velde_SLIDES.pdf.
- Idiatov, Dmitry & Mark Van de Velde. 2019. Bena-Mboi is Benue-Congo. Paper presented at the Adamawa Conference. Mainz. http://idiatov.mardi.myds.me/talks/2019_MAINZ_Adamawa_BENA_Classification_Idiatov_Van_de_Velde_SLIDES.pdf.
- Idiatov, Dmitry & Mark Van de Velde. 2020. The internal reconstruction of Bena-Yungur consonants and tone patterns. *Language in Africa* 1(3). 100–124.
- Ittmann, Johannes. 1935–36. Kenyāñ, die Sprache der Nyang. *Zeitschrift für Eingeborenen-Sprachen* 26(1-4). 2–35, 97–133, 174–202, 272–300.
- Jacob, Irénée. 1984–87. *Dictionnaire rwandais – français, en 3 volumes*. Butare: Institut National de Recherche Scientifique.
- Kabuta, Ngo Semzara. 2006. *Dictionnaire cilubà – français*. <http://www.ciyem.ugent.be>.

- Kamba Muzenga, Jean-Georges. 2003. *Substitutifs et possessifs en bantou* (Société d'études linguistiques et anthropologiques de France 401). Louvain: Peeters.
- Kjelsvik, Bjørghild. 2002. *Verb chains in Nizaa*. Oslo: University of Oslo. (MA thesis).
- Köhler, Bernhard (ed.). 2016. *Interrogative and syntactic inquiries: Case studies from Africa* (Frankfurt African Studies Bulletin 23). Cologne: Rüdiger Köppe.
- MacBeath, A. G. W. 1940. *Bobangi in twenty-one lessons, with exercises & key*. Bolobo: Baptist Missionary Society.
- Maho, Jouni F. 2009. *NUGL online: The online version of the new updated Guthrie list, a referential classification of the Bantu languages*. https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf.
- Matras, Yaron. 2009. *Language contact* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Mbuagbaw, Tanyi Eyong. 2000. *Kenyang segmental phonology*. Yaoundé: Cameroon Association for Bible Translation & Literacy.
- Meeussen, Achiel Emiel. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Moreton, Rebecca L. & Henry Marcel Bôt Bá Njock. 1975. *Je parle basaá : manuel d'initiation au basaá en 265 cycles*. Nouvelle edition. Douala: Centre Culturel Africain, Collège Libermann.
- Motingea Mangulu, André. 2018. *Le lokonda des Ilângá la Bakonda : un échantillon des langues inclassables du Bassin central congolais* (Grammatical Analyses of African Languages 56). Cologne: Rüdiger Köppe.
- Paulin, Pascale. 1995. *Etude comparative des langues du groupe Ring – langues Grassfields de l'ouest, Cameroun*. Lyon: Université Lumière – Lyon 2. (MA thesis).
- Petzell, Malin. 2008. *The Kagulu language of Tanzania: Grammar, texts and vocabulary* (East African Languages and Dialects 19). Cologne: Rüdiger Köppe.
- Polak, Louise. 1983. Le réfléchi en bantou. *Africana Linguistica* 9. 271–304.
- Poulos, George & Louis J. Louwrens. 1994. *A linguistic analysis of Northern Sotho*. Hatfield, Pretoria: Via Afrika.
- Raponda-Walker, André. 1934. *Dictionnaire mpongwé – français, suivi d'éléments de grammaire*. Metz: La Libre Lorraine.
- Ratliff, Martha. 2011. Interrogative pronouns in isolating languages. Paper presented at the Workshop “Stability and borrowability of interrogative pronominals”, 20th International Conference on Historical Linguistics. Osaka. http://idiatov.mardi.myds.me/ICHL/ICHL_RATLIFF_slides_notes.pdf.
- Rey-Debove, Josette. 1996. *Le nouveau Petit Robert : dictionnaire alphabétique et analogique de la langue française*. Paris: Dictionnaires Le Robert.

- Ricquier, Birgit & Koen Bostoen. 2008. Resolving phonological variability in Bantu lexical reconstructions: The case of 'to bake in ashes'. *Africana Linguistica* 14. 109–149.
- Rijkhoff, Jan. 2008. Descriptive and discourse-referential modifiers in a layered model of the noun phrase. *Linguistics* 46(4). 789–829.
- Rottet, Kevin J. 2004. Inanimate interrogatives and settlement patterns in Francophone Louisiana. *Journal of French Language Studies* 14(2). 169–188.
- Schadeberg, Thilo C. 2003. Historical linguistics. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages* (Language Family Series 4), 143–163. London: Routledge.
- Schaub, Willi. 1985. *Babungo* (Croom Helm Descriptive Grammars). London: Croom Helm.
- Stanley, Carol. 1991. *Description phonologique et morpho-syntaxique de la langue tika* (parlée au Cameroun). Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Teil-Dautrey, Gisèle. 1991. *Evolution phonétique d'une langue bantoue du nord-ouest : le bàsàà A.43a*. Lyon: Université Lumière – Lyon 2. (MA thesis).
- Van de Velde, Mark. 2006. Multifunctional agreement patterns in Bantu and the possibility of genderless nouns. *Linguistic Typology* 10(2). 183–221.
- Van de Velde, Mark. 2008a. *A grammar of Eton* (Mouton Grammar Library 46). Berlin: Mouton de Gruyter.
- Van de Velde, Mark. 2008b. Un cas de changement phonologique par réanalyse morphologique en éton. *Africana Linguistica* 14. 177–184.
- Van de Velde, Mark. 2013. The Bantu connective construction. In Anne Carlier & Jean-Christophe Verstraete (eds.), *The genitive* (Case and Grammatical Relations Across Languages 5), 217–252. Amsterdam: John Benjamins.
- Van de Velde, Mark. 2017. The augment as a construct form marker in Eton relative clause constructions. In Gratien Gualbert Atindogbé & Rebecca Grollemund (eds.), *Relative clauses in Cameroonian languages: Structure, function and semantics* (Empirical Approaches to Language Typology 58), 47–66. Berlin: De Gruyter Mouton.
- Van de Velde, Mark. 2019. Nominal morphology and syntax. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn. (Routledge Language Family Series), 237–269. Milton Park, Abingdon: Routledge.
- Van de Velde, Mark. In preparation. *Dictionnaire eton – français*.
- Van de Velde, Mark & Odette Ambouroué. 2011. The grammar of Orungu proper names. *Journal of African Languages and Linguistics* 32(1). 113–141.

- Van de Velde, Mark & Dmitry Idiatov. 2017. Morphological classes and gender in Bóná-Yungur. In Shigeki Kaji (ed.), *Proceedings of the 8th World Congress of African Linguistics, Kyoto 2015*, 53–65. Tokyo: Research Institute for Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.
- van den Eynde, Karel & Piet Mertens. 2003. La valence : l'approche pronominale et son application au lexique verbal. *Journal of French Language Studies* 13(1). 63–104.
- van der Wal, Guenever Johanna. 2009. *Word order and information structure in Makhuwa-Enahara*. Leiden: Leiden University. (Doctoral dissertation).
- Vanhoudt, Bettie. 1987. *Éléments de description du leke, langue bantoue de zone C* (Annales – Série in-8° – Sciences humaines 125). Tervuren: Royal Museum for Central Africa.
- Voll, Rebecca Maria. 2017. *A grammar of Mundabli: A Bantoid (Yemne-Kimbi) language of Cameroon*. Leiden: Leiden University. (Doctoral dissertation).
- Voorhoeve, Jan. 1980. Kenyang. In Larry M. Hyman & Jan Voorhoeve (eds.), *L'expansion bantoue : actes du colloque international du CNRS, Viviers (France) – 4-16 avril 1977, Vol. 1 : les classes nominales dans le bantou des Grassfields* (Société d'études linguistiques et anthropologiques de France, Numéro special 9), 276–285. Paris: SELAF.
- Welmers, William E. 1963. Associative *a* and *ka* in Niger-Congo. *Language* 39(3). 432–447.
- Whitehead, John. 1899. *Grammar and dictionary of the Bobangi language; as spoken over a part of the Upper Congo, West Central Africa. Compiled and prepared for the Baptist Missionary Society's Mission in the Congo Independent State by John Whitehead*. London: Baptist Missionary Society & Kegan Paul, Trench, Trübner & Co.
- Wills, Jeffrey. 2022. Sorting out Proto-Bantu *j. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.), *On reconstructing Proto-Bantu grammar*, 59–101. Berlin: Language Science Press. DOI: 10.5281/zenodo.7575817.

Name index

- Aaron, Uche E., 140
Abad, Isidoro, 445–447
Abasheikh, Mohammad I., 291
Abe, Yuko, 652
Abéga, Prosper, 43
Abels, Klaus, 557, 573
Abessolo Nnomo, Thierry, 688, 697,
699, 721
Abraham, Roy Clive, 643
Achiri-Taboh, Blasius, 510
Adam, Jean-Jérôme, 224
Adams, Gustaf A., 29, 42
Akumbu, Pius W., 156
Alexandre, Pierre, 43, 364
Allan, Edward Jay, 510
Allen, James P., vi
Alsina, Alex, vii
Ambouroue, Odette, 361, 444, 485,
690, 697, 704, 729
Andeme Allogo, Marie-France, 43
Anderson, Gregory D. S., 389, 395,
414
Anderson, Stephen C., 43, 132, 133,
141, 144, 151, 156, 249, 262,
271
Angenot, Jean-Pierre, 43
Apuge, Michael E., 370
Araújo, Paulo Jefferson Pilar, 601,
617, 622, 625, 634, 635
Ardener, Edwin, 42
Arnott, D. W., 249
Asangama, Atisa, 645
Ashton, Ethel O., 302, 349, 549
Asiimwe, Allen, 427, 551, 563, 574
Asoshi, Melvice, 156
Atindogbé, Gratien Gualbert, 42, 190,
499, 507, 510
Atta, Samuel Ebongkome, 42, 79
Aunio, Lotta, viii, xi, 594, 634
Ayotte, Charlene, 249
Ayotte, Michael, 249
Ayuninjam, Funwi F., 249
Ba, Ibrahima, 510
Babaev, Kirill V., 395–399, 403–409,
412
Bachmann, Armin R., 38
Bahuchet, Serge, 645
Baka, Jean, 601, 625, 626, 712
Baker, Gary K., 67
Baker, Mark C., vii, 346
Baldi, Philip, xiv, xv
Bámgbóšé, Ayò, 249, 260
Bancel, Pierre, 4, 11
Barreteau, Daniel, 636
Bassong, Paul R., 547, 573
Bastin, Yvonne, xvi, xviii–xx, xxvi, 8,
14, 17, 29, 39, 41, 60–64, 107,
111, 113, 120, 174, 180, 205,
244, 254, 284, 286, 293, 298,
315, 320, 326, 329, 344, 355,
356, 362–365, 367–371, 404,
450, 600, 601, 615, 617, 618,

- 620, 630, 649, 650, 668, 669,
672, 686, 710
- Bates, George L., 43
- Baucom, Kenneth L., 74
- Baumann, Oskar, 42
- Bearth, Thomas, 424, 495, 507, 514,
522, 586, 594, 602
- Beaudoin-Lietz, Christa, 444, 448,
517
- Beavon, Keith H., 43, 44
- Beavon, Mary, 43
- Bébiné, Adriel J., 43
- Beck, David, 344, 345, 348–350
- Belliard, François, 44
- Beltzung, Jean-Marc, 520, 521
- Bendor-Samuel, John T., 123
- Bentley, Delia, 582, 592, 595
- Bernander, Rasmus, xii, 586, 591, 594,
595, 603, 606, 608, 609, 616,
617, 621, 631, 634, 638, 652,
720
- Besha, Ruth Mfumbwa, 621
- Bickmore, Lee S., 182, 286, 287, 295,
296, 390, 440
- Biloa, Edmond, 43, 364, 518, 548, 573
- Bitjaa Kody, Zachée, 298
- Blackings, Mairi John, 401
- Blanchon, Jean A., 11, 480
- Blasi, Damián E., 407
- Bleek, Wilhelm H. I., ix, xiv, 239
- Blench, Roger M., v, xi, xxviii, 42, 43,
111, 121, 123, 236, 241, 244,
251, 254, 260, 262, 263, 310,
316, 332, 394, 456
- Bloom Ström, Eva-Marie, 589, 595,
596, 607, 609, 613, 624
- Bokamba, Eyamba Georges, 427, 682
- Bolekia Boleká, Justo, 42, 324, 361,
447
- Bolioki, Léonard-Albert, 43
- Bonneau, Joseph, 215–217
- Bontinck, François, vii
- Borschev, Vladimir, 597
- Bostoën, Koen, v, ix–xi, xiii–xv, xix,
4, 7, 17, 40, 41, 62, 64, 73, 75,
84, 88, 95, 114, 152, 177, 180–
182, 201, 216, 227, 238, 241,
251, 267, 268, 271, 298, 312,
318, 320, 323–329, 332–334,
344, 345, 347, 353–355, 359,
360, 362, 364, 368, 369, 392,
406, 411, 414, 431, 456, 490,
496, 526, 538, 549, 551, 573,
594, 634, 645, 672, 676
- Bôt Bá Njock, Henry Marcel, 694
- Bôt, Dieudonné Martin Luther, 42,
191–194, 683, 694, 699
- Botne, Robert, 4, 12, 13, 85, 116, 137,
152, 156
- Bouka, Léonce-Yembi, 44, 552
- Bourquin, Walther, 60, 61, 70
- Boutwell, Richard Lee, 156, 249, 258
- Bowden, John, 326
- Boyd, Raymond, 242, 249, 251
- Boyd, Virginia, 43, 705, 711, 729
- Breedveld, Anneke, 42, 43, 79, 364
- Breen, John G., 400
- Bresnan, Joan, vii, 312, 602
- Breton, Roland, 249
- Brisson, Robert, 641
- Bubenik, Vit, 148
- Buell, Leston C., 614
- Bufe, E., 42
- Bulkens, Annelies, 62, 69, 710
- Burger, J. P., 617
- Bwantsa-Kafungu, S.-Pierre, 594
- Bwendelele, A., 618, 622
- Bybee, Joan L., xxiii, 67, 556

Name index

- Bynoe-Andriolo, Esla Y., 546, 557,
573
- Campbell, Lyle, 325, 498, 499, 515
Cann, Ronnie, 312, 318
Cardoso, Mattheus, vi
Carstens, Vicki M., 501
Chatelain, Heli, 624
Chatzikyriakidis, Stergios, vii
Cheng, Lisa L.-S., 500, 501
Cheucle, Marion, 26, 43, 44
Chia, Agnes F. S., 42
Chibaka, Evelyn Fogwe, 156
Childs, Tucker G., 367
Clark, David J., 410
Clarke, John, 519
Clements, George N., vii, xxv, 613,
637
Cole, Desmond T., 299, 432, 433, 622,
623
Coleman, Arnie, 643, 702
Collins, B., 629
Collins, Frank, 526
Connell, Bruce, 95, 126, 127, 148, 156
Contini-Morava, Ellen, 436
Corbett, Greville G., 704
Coupez, André, xxvi, 60, 61, 284, 285,
687, 723
Crane, Thera M., xii, 73, 75, 152, 180,
406, 559, 572, 574, 601, 625,
626, 647
Creissels, Denis, 62, 70, 82, 92, 311,
312, 315, 320, 322, 324, 327,
328, 389, 433, 441, 484, 582,
583, 591, 595, 597, 605, 613,
614, 629, 636, 637, 643, 644,
651, 712
Cristofaro, Sonia, 709
Croft, William, 709
- Cysouw, Michael, 676, 679
Czinger, Christine, 582
- da Silva Maia, António, 617
Daeleman, Jan, 359
Dahl, Östen, 556
Dalby, David, xviii
Dalgish, Gerard M., 630
Dammann, Ernst, 332, 344, 345, 351,
353, 617, 620, 621, 688
Davison, Phil, 702
de Blois, Kornelis Frans, 471, 596, 686
de Dreu, Merijn, 614
de Gastines, François, 298
De Kind, Jasper, 120, 318, 327, 496,
551–554, 556, 562, 563, 570,
573
de Schryver, Gilles-Maurice, ix, xi,
xiii, xix, 40, 83, 186, 271, 354,
431, 551, 552
de Vos, Mark, vii
de Wit, Gerrit, 584, 669, 682–684,
706, 708, 720–723
de Wolf, Paul P., 240
Demuth, Katherine, vii, 502, 504,
506, 512, 595, 602
den Besten, Margaret G. G., 79
Dereau, Léon, 601
Devos, Maud, viii, ix, xi–xiii, xxiii,
406, 589–591, 596, 598, 604,
607, 608, 612, 620, 630
Di Carlo, Pierpaolo, 246
Diercks, Michael, 595, 602
Diessel, Holger, 671, 676
Dieu, Michel, 246, 264, 364
Dik, Simon C., 539
Dimande, Ernesto, 621, 628
Dimmendaal, Gerrit J., vii, 130, 148,
352, 498

- Dinkelacker, Ernst, 42
Djiafeua, Prosper, 262
Doke, Clement M., 359, 392, 569, 574
Dom, Sebastian, viii, ix, xi, xii, 192,
345, 353–355, 373, 572, 652
Doneux, Jean Léonce, ix, 39, 361, 672,
689, 691–693, 711, 712, 721
Dorsch, Heinrich, 42
Downing, Laura J., vii, xiii, 13, 189,
292, 295, 390, 499, 502, 507
Dryer, Matthew S., 597, 682
du Plessis, Jacobus A., 312, 624
Dugast, Idelette, 43, 195, 196
Dunham, Margaret, 624
Duranti, Alessandro, 436
Dzameshie, Alex K., 510

Easterday, Shelece, 67
Eaton, Helen, 590, 652
Ebarb, Kristopher J., 113, 282, 284,
287
Eberhard, David M., v, 241, 498
Ebobissé, Carl, 4, 42
Edelsten, Peter, 616
Edika, E. Solange F., 43
Egbokhare, Francis O., 136
Ehret, Christopher, 4, 524
Elias, Philip, 29, 32, 35, 84, 702, 727
Elimelech, Baruch, 42
Ellington, John Ernest, 69, 73, 182
Endemann, Karl, 317, 318
Endresen, Rolf Theil, 319, 726
Englebretson, Robert, 439, 619
Epps, Patience, 326
Ernst, Urs, 44, 207–209, 370
Essono, Jean-J. Marie, 43, 80, 364
Etogo Mbezele, Luc, 688, 697, 699,
721
Evans, Nicholas, 683

Ewane Etame, Jean, 42
Eyoh, Julius A., 719

Fabb, Nigel, 401
Fabre, Anne Gwenaëlle, 644
Faltz, Leonard M., 313
Faytak, Matthew, 95
Fehn, Anne-Maria, 355, 356
Fernandez, Galilea L., 42
Festen, Bradley, 44
Fiedler, Ines, 523, 567
Fisch, Maria, 350, 351, 559, 560, 574,
632, 639
Fivaz, Derek, 74, 621
Fleisch, Axel, viii
Fontaney, V. Louise, 215–217
Forges, Germaine, 354, 359, 643, 646
Fortune, George, 392
François, Alexandre, 4
Fransen, Margo Astrid Eleonora, 156,
643, 702, 711, 721
Friesen, Dan T., 30, 36, 42
Friesen, Lisa, 266, 268, 448, 453

Galerie, Anne, 646
Galley, Samuel, 27, 43
Gambarage, Joash Johannes, 648
Garbo, Francesca Di, 648
Gary, Judith O., vii
Gast, Volker, 583, 584
Gautier, Jean-Marie, 143, 211, 212,
361, 370
Gérard, R. P., 587
Gerhardt, Ludwig, 91, 177, 260, 265,
266, 326
Gibson, Hannah C., vii, xii, 558, 573,
574, 618, 697, 720
Gildea, Spike, xxii, xxiii
Gilman, Charles, 546, 566

Name index

- Givón, Talmy, viii, xxv, 299, 319, 327,
332, 401, 402, 405, 497, 502,
503, 512, 685, 687, 720
- Goemaere, Alphonse, 646
- Goes, Heidi, 355
- Goldsmith, John A., vii
- Good, Jeff, xi, xxvii–xxix, 123, 126,
176, 183, 223, 242, 246, 256,
344–347, 365, 390, 429, 648,
650
- Goodman, Morris F., 546
- Goody, Jack, vi
- Gowlett, Derek F., 36, 299
- Gray, Hazel, 621, 631
- Grebe, Karl H., 249, 262, 264
- Green, Christopher, 284
- Green, Margaret M., 405
- Green, Melanie, 544
- Greenberg, Joseph H., viii, 18, 61, 69,
240, 242
- Gregersen, Edgar A., 516
- Grégoire, Claire, 39, 72, 113, 114, 177,
182–185, 196, 290, 314, 324,
364, 443, 471, 554, 593, 605,
616, 618–621, 624, 627, 628,
630, 631, 640, 641, 648, 650,
672, 684, 687, 690, 691, 693,
694, 699, 711, 712, 721, 723
- Grimm, Nadine, 18, 43, 186, 202–204,
599, 600, 635, 652
- Grollemund, Rebecca, x, xviii–xx,
xxviii, xxxi, xxxiii, 4, 5, 7,
40, 63–65, 68, 71, 106–109,
120, 121, 123–125, 129, 177,
186, 222, 226, 242, 244, 298,
311, 324, 332, 352, 356, 358,
360, 361, 363, 364, 366, 370,
372, 397, 410, 430–432, 438,
443, 448, 467, 488, 490, 495,
498, 499, 507, 508, 510, 511,
513, 524, 526, 528, 570, 581,
613, 614, 637, 646, 647, 651,
714
- Gromova, Nelli V., 620
- Guarisma, Gladys, 43, 196–198, 254,
370
- Gueche Fotso, Hugues Carlos, 249
- Guérois, Rozenn, viii–xi, 114, 177,
227, 251, 267, 268, 326, 333,
368, 369, 484, 490, 538, 572,
593, 596, 598–600, 617, 676
- Güldemann, Tom, viii, xi, xiii, xxiii,
xxv, xxxi, xxxv, 175, 176,
180, 189, 226, 267, 387, 388,
390, 391, 394, 395, 400, 401,
403–406, 409–411, 413, 449,
450, 488, 499, 514, 523, 537–
541, 546, 550–552, 555, 556,
558, 568–571, 574, 613, 637,
650
- Gunnink, Hilde, ix, xii, 152, 314, 559,
572, 574, 618, 619
- Guthrie, Malcolm, xv, xvii, xix, 4, 6,
8, 9, 14, 25, 27, 29, 33, 34, 37–
43, 60, 61, 63, 69, 70, 74, 76,
77, 84, 88, 91, 107, 110, 112,
152, 177, 180, 218, 219, 235,
236, 239, 282, 315, 344, 361,
362, 368, 369, 388, 395, 397,
398, 430, 431, 499, 539, 671,
672, 684
- Haas, Florian, 583, 584
- Hackstein, Olav, 676, 679
- Hadermann, Pascale, 538, 551–557,
573
- Hagège, Claude, 42, 254
- Hall, T. Alan, 67, 95

- Halle, Morris, 4
Halme, Riikka, 621
Hamlaoui, Fatima, xii, 466, 479, 496,
507, 512, 523
Hamm, Cameron, 256
Hammarström, Harald, 4, 241, 317,
637, 668
Han, Xu, vi
Hannan, Michael, 82
Hare, David M., 501, 508, 521, 522,
641
Harford, Carolyn, vii, 311, 502, 504,
506, 512, 595
Harries, Lyndon, 473, 605, 607, 639
Harris, Alice C., 499, 515
Harro, Gretchen, 133, 156, 161, 262
Hartell, Rhonda L., 123
Hasheela, Paavo, 432
Haspelmath, Martin, 675
Hawkinson, Ann Katherine, 42, 190,
311, 448, 453
Haynes, Nancy R., 133, 156, 161, 643
Heath, Daniel, 43, 205, 206
Heath, Teresa, 43, 406
Hedinger, Robert, 5, 11, 42, 70, 113,
123, 151, 187, 188, 240, 246,
249, 266, 267, 271, 669, 695,
700, 707
Hedinger, Sylvia, 42
Heine, Bernd, xxii, xxiii, 120, 318, 319,
361
Hellwig, Birgit, 269
Helmke, Christophe, vi
Helmlinger, Paul, xvi, 42
Henderson, Brent M., 499
Hendery, Rachel, 487
Hengeveld, Kees, 582, 596
Henrici, Alick, 397
Henson, Bonnie J., 44, 370
Hernández-Green, Néstor, 328
Hetherwick, Alexander, 574
Hetzron, Robert, xxiii
Hewson, John, 148
Hinnebusch, Thomas J., 83, 94, 355
Hock, Hans H., xiv, xvii, xxii, 11
Hoenigswald, Henry M., xxii
Hölzl, Andreas, 675, 713
Hombert, Jean-Marie, 36, 43
Homburger, Liliás, 60, 61, 69
Hon, Luther, 271
Honeybone, Patrick, 11
Horton, Alonzo E., 559, 560, 574
Hualde, José Ignacio, 67
Hulstaert, Gustaaf, 325, 357, 681, 686,
689, 704, 722
Hyman, Larry M., vii–ix, xi, xxv,
xxix, xxxiii, xxxv, 7, 23, 29,
41–43, 77, 117, 118, 121, 133,
151, 156, 176, 177, 179, 183,
188, 189, 197, 202, 223, 226,
227, 236–239, 248, 249, 258,
260, 267, 269, 283, 284, 286,
288–290, 292, 294, 295, 298,
299, 301, 303, 310, 311, 314,
316–318, 323, 324, 326, 334,
344–347, 355, 357, 359–362,
364, 367, 369, 372, 373, 389,
390, 392, 402, 429, 444, 451,
457, 487, 488, 497, 499, 506,
508, 514, 518, 524, 525, 539,
571, 596, 650, 688, 695, 699,
700, 702, 705, 707, 729
Ibrahim, Njoya, 547, 573
Idiata-Mayombo, Daniel-Franck, 89
Idiatov, Dmitry, vi, xiii, xv, xvi, xxv,
xxxv, 227, 242, 249, 265,
266, 371, 472, 490, 668–672,

Name index

- 674–679, 682, 684–688, 691,
694–696, 701, 703, 704, 706,
724–726
- Igwe, G. Egemba, 405
Isaac, Kendall M., 43
Ittmann, Johannes, 42, 358, 708
- Jackson, Ellen W., 254, 256
Jacob, Irénée, 686
Jacob, Peggy, 543, 564, 574
Jacquot, André, 44, 239
Jaggar, Philip J., 543
Janson, Tore, 355, 356
Janssens, Baudouin, xv, 4, 11, 12, 42,
471
Jenks, Peter, 501, 517
Jerro, Kyle J., 311, 312
Johnson, Silas F., 43
Johnston, Harry H., 239
Jones, Patrick J., 288, 290, 300
Jubase, J. B., 624
Jungraithmayr, Herrmann, 239
- Kabuta, Ngo Semzara, 359, 694
Kadenge, Maxwell, vii
Kagaya, Ryohei, 42
Kahl, Jochem, vi
Kähler-Meyer, Emmi, 317, 318, 332
Kaji, Shigeki, 284, 286
Kamba Muzenga, Jean-Georges, xiii,
81, 359, 396, 472, 488, 672,
673, 685, 708, 725
Kanerva, Jonni M., vii, 295, 602
Katamba, Francis X., 77, 283, 284,
290, 298, 299, 303
Katupha, Jose Mateus Muaria, 428
Kavari, Jekura U., 296, 297, 621
Kawalya, Deo, 652
Kawasha, Boniface, 478, 505, 506
- Kaze, Bitrus, 456
Keegan, John M., 542
Keenan, Edward L., vii, 313
Kelly, John, 43
Kemmer, Suzanne, 365
Kempson, Ruth M., vii
Kenmogne, Michel, 43, 364, 370
Kerr, Elisabeth Jane, 652
Kettunen, Harri, vi
Khumalo, Langa, 602
Kießling, Roland, 249, 269, 326
Kimenyi, Alexandre, 311, 332, 443
Kisseberth, Charles W., vii, 291, 390,
506
Kjelsvik, Bjørghild, 249, 253, 709
Koch, Harold, xxii
Koch, Peter, 582, 584, 586, 587, 594–
596, 605, 613, 629, 634
Koelle, Sigismund W., 239
Köhler, Bernhard, 675
Köhler, Oswin, 296, 297
Koile, Ezequiel, xviii
Kongne Welaze, Jacques, 370
Koni Muluwa, Joseph, 73, 84, 538,
562, 572, 573
Koops, Robert, 249
Kouoh Mboundja, Christian Josué,
42
Kozhanov, Kirill, 326
Kraal, Pieter Jacob, 607, 616
Kracht, Marcus, 313
Krause, Gottlieb A., 239
Kröger, Heidrun, 347, 348
Kuijpers, Em, 359
Kula, Nancy C., x, 439
Kulikov, Leonid, xi
Kuno, Susumu, 506
Kuperus, Juliana, 42, 364
Kuteva, Tania, xxiii, 120, 319

- Kutsch Lojenga, Constance, 683
Kwenzi-Mikala, Jérôme T., 217
- Labroussi, Catherine, 355
Lafon, Michel, 563, 569, 572, 574
Laine, Antti, 594, 595, 606, 609, 617,
 621, 631, 634, 638, 652
Laman, Karl E., 359, 618
Lamberty, Melinda, 42
Lambrecht, Knud, 583, 588
Landman, Meredith, 557, 574
Leach, Michael Benjamin, 607
LeBlanc, Paul D., vi
Legère, Karsten, 621
Leitch, Myles, vii, 619, 627, 641
Lemb, Pierre, 298
Leroy, Jacqueline, 4, 121, 123, 240,
 262, 269
Letsholo, Rose, 505
Lichtenberk, Frantisek, 347
Lijongwa, Chiku, 616
Liphola, Marcelino M., 286
Livinhac, Léon, 359
Lonfo, Etienne, 133, 156, 249, 262
Louw, Jacobus A., 624
Louwrens, Louis J., 689
Lovegren, Jesse, 156, 506, 508, 510,
 522, 524, 525, 643
Lovestrand, Joseph, 43
Lubasa, N'ti Nseendi, 566, 567
Lusekelo, Amani, vii
Lyons, John, 582
- Mabugu, Patricia Ruramisai, 311–
 313, 318
MacBeath, A. G. W., 721, 723
Machiwana, Kingston, 621
Mackey, James L., 361
Maddieson, Ian, 36, 264
- Maddox, Harry E., 359
Maganga, Clement, 350, 351, 583, 591,
 634
Magba, Elizabeth Ann, 156
Maho, Jouni F., 4, 74, 110, 152, 284,
 398, 431, 593, 648, 668
Makasso, Emmanuel-Moselly, 496,
 507, 523
Malema, Johnson, 652
Manfredi, Victor, 546
Maniacky, Jacky, ix, xi, xiii, 526
Manus, Sophie, 497
Mapunda, Gastor, 620, 652
Marantz, Alec Paul, vii
Marchal-Nasse, Colette, 112, 218, 219,
 615, 616, 628
Marks, Gianna, 572
Marlo, Michael R., vii, xi, 85, 113, 189–
 191, 282, 284, 287, 288, 291,
 390, 427, 428, 435, 517
Marten, Lutz, vii, xx, 439, 496, 507,
 511–513, 586, 593, 595, 602,
 603, 606, 608, 617, 621, 628,
 637, 648, 649
Martin, Marieke, 262, 263, 510
Maslova, Elena S., viii, 353
Mathangwane, Joyce T., 84, 301
Mathaus, Njeck, 43
Matras, Yaron, 675
Matsinhe, Sozinho Francisco, 621,
 622, 652
Mba, Gabriel, 262
Mbanji, Bawe E., 262
Mbiling'i, Enock, 620, 652
Mbolifouye, François, 510
Mbuagbaw, Tanyi Eyong, 708
McClellan, Greg L., 156
McCracken, Chelsea, 510
McGill, Stuart, 200, 247

Name index

- McGinnis, Martha, vii
Mchombo, Sam A., vii, 174, 359
McLean, Greg L., 249
McNally, Louise, 582, 595, 596
Medjo Mvé, Pither, 27, 43, 44
Meeussen, Achiel Emiel, v, viii–x, xii, xiii, xv, xvii, xviii, xx, xxi, xxiv–xxxiv, 8, 25, 34, 39, 59, 61, 67, 69, 76, 84, 85, 90, 92–94, 107, 108, 111, 113, 114, 117, 151, 152, 174–177, 179, 182, 189, 191, 202, 216, 219, 226, 238, 267, 281–283, 285, 287, 288, 294, 297, 303, 304, 344–346, 356, 360, 361, 366, 368, 369, 387, 388, 390, 391, 395, 400, 406, 410, 423, 424, 430, 448, 465, 466, 468, 469, 471, 472, 476, 477, 480, 483, 485, 486, 489, 495, 496, 498, 515, 516, 524, 537, 539, 545, 546, 560, 569–571, 590, 593, 594, 603, 609, 615, 618–620, 627, 637, 649, 650, 668, 672, 684, 691, 708, 710, 725, 729
Meeuwis, Michael, 538, 562, 573, 587, 652
Meinhof, Carl, ix, xiv–xvi, 34, 42, 60, 69, 82, 85, 94, 238, 344, 361, 368, 369, 390, 642
Mékina, Émilienne-Nadège, 43
Mel'čuk, Igor, 344, 345, 348–350
Merrill, John, 7
Mertens, Piet, 676
Mfonyam, Joseph Ngwa, 156
Mickala Manfoumbi, Roger, 44
Miehe, Gudrun, 369, 406
Misago, Manoah-Joël, 612, 652
Mitchley, Hazel, vii
Mithun, Marianne, 352, 405
Mkochi, Winifred, 292, 295, 296
Mkude, Daniel J., 496, 621
Möhlig, Wilhelm J. G., 39, 239, 296, 297, 303, 559, 560, 572, 574, 621
Moise, Eyinga Essam, 79
Mongo, Raoul, 43
Monikang, Evelyn Neh, 42
Mora-Marin, David F., 328
Moreton, Rebecca L., 694
Morimoto, Yukiko, xii, 558, 561, 572, 573, 595, 602
Moroz, George, 221
Morrison, Michelle Elizabeth, 621
Morrison, William McC., 439
Mort, Katherine, 247
Mortensen, David R., 95
Moses, Godian, 652
Moshi, Lioba, vii, 312, 619, 624
Motingea Mangulu, André, 438, 722, 723
Mous, Maarten, 42, 43, 79, 195, 196, 324, 364, 445, 450, 495, 515, 518, 594, 621, 633, 636
Moyo-Kayita, Makila, 618, 619, 622
Mpiranya, Fidèle, 360
Mreta, Abel, 621, 633
Mtenje, Al, 295, 502
Mufwene, Salikoko S., 538, 546, 566, 573
Mugane, John M., vi, 353, 561
Mukash Kalel, Timothée, 359
Mundeke, Léon, 180, 312, 325, 328, 406, 496, 594, 634, 639, 645, 652
Muriungi, Peter, 557, 573
Murrell, Paul, 44

- Mutaka, Nguessimo M., 284, 287–292
Mutaka, Philippe, 292
Mwenegoha, Hamza A. K., 346
Myers, Scott, vii
- Nabirye, Minah, 359, 360, 550, 551,
572, 574, 605
Nagano-Madsen, Yasuko, 44
Nam, Seungho, 313
Namyalo, Saudah, 549, 550, 574, 596,
597
Nanteza, Moureen, 597, 631
Nassau, Robert H., 42, 621, 627
Nchare, Abdoulaye Laziz, 156
Ndagba, André, 585, 652
Ndamsah, Gratiana, 547, 548, 573
Ndedje, René, 156, 249
Ndembe Nsasi, D., vii
Ndjerareou, Mekoulnodji, 644
Neba, Ayu'nwi N., 370, 615, 627
Nganganu, Kenfac Lucy, 249
Ngo-Ngijol Banoum, Bertrade B.,
viii
Ngoboka, Jean Paul, xiii, 434, 435,
443, 551, 572, 574, 604
Ngoran, Loveline Lenaka, 249
Ngoyani, Deo, 501
Ngunga, Armino S. A., 74, 286, 359
Ngwasi, Lengson, 652
Nichols, Johanna, 407
Nicolle, Steve, 590, 594, 621, 634
Nida, Eugene A., 36
Njantcho Kouagang, Elisabeth, 44,
401, 641
Njock, Pierre E., 42
Nkemnji, Michael, 547, 573
Nkiko, Munya Rugero, 359
Noonan, Michael, 514
Norde, Muriel, xxv
- Nouguier Voisin, Sylvie, 328
Nsuka-Nkutsi, François, xx, xxiv,
xxxii, 11, 217, 470, 471, 474,
477, 478, 480, 482, 485, 488,
495, 498, 505–508, 510, 514,
524–526
Nunn, Nathan, vi
Nurse, Derek, ix, xii, xiv, xv, xviii,
xxiii, xxvii, xxxiii, 4, 64,
83, 94, 107–114, 116–118, 120,
125, 127, 140, 143, 145, 147,
151, 152, 154, 174–176, 178,
179, 181, 182, 185, 188, 189,
196, 208, 226, 355, 356, 362,
388–390, 394, 395, 441, 488,
498, 499, 510, 514, 518, 524,
624, 650
Nzang-Bie, Yolande, 362, 364, 445
- O'Sullivan, Owen, 622
Odden, David, vii, viii, 74, 182, 189–
191, 286, 287, 390, 428, 620
Ohala, John J., 363
Ollomo Ella, Régis, 43
Orwig, Carol, 198, 199, 364, 519
- Pacchiarotti, Sara, viii, ix, xi, xii, xv,
xix, xxix, xxxv, 4, 40, 41, 73,
180, 181, 186, 227, 309, 311,
313, 317, 320, 325, 332, 335,
348, 354, 365, 373, 411, 431,
456, 490, 555
Pae, Hye K., vi
Paluku, André Mbula, 359
Parker, Elizabeth Ann, 156
Partee, Barbara H., 597
Pasch, Helma, 510
Paterson, Rebecca, 135, 247
Patman, Keith E., 43

Name index

- Paulian, Christiane, 41–43
Paulin, Pascale, 702
Perlmutter, David M., vii
Perrin, Mona, 43, 156
Persohn, Bastian, xii, 472, 606, 617,
633, 634
Peterson, David A., 313, 407
Petzell, Malin, 621, 685–687, 693, 719,
720
Philippon, Gérard, ix, x, xv, xviii,
xxvii, 4, 6, 7, 29, 40, 41, 64,
95, 107, 109–111, 116, 143, 175,
176, 320, 388, 394, 624
Piper, Klaus, 516, 625
Piron, Pascale, xviii, xix, 123, 129,
212–214, 242, 244, 363, 364
Pohlig, James N., 260, 262
Polak, Louise, ix, xx, xxxi, 284, 402,
423–426, 428, 430, 439, 443,
447–449, 451, 453, 467, 516–
518, 709
Postal, Paul M., vii
Poulos, George, viii, 689
Pozdniakov, Konstantin, 179
Prat, Jean, 604
Prince, Alan S., vii
Prittie, Rebecca, 43
Puech, Gilbert, 43, 44
Pullum, Geoffrey K., 33
Pylkkänen, Liina, vii

Qian, Nancy, vi

Racine, Odile, 360
Raharimanantsoa, Ruth, 619, 652
Ranero, Rodrigo, 433, 557, 574
Rapold, Christian, 312, 313, 325, 328
Raponda-Walker, André, 361, 690,
691

Ratliff, Martha, 676
Redden, James E., 389, 450
Reh, Mechthild, xxiii
Rekanga, Jean-Paul, 72, 214, 215, 324,
645
Renaud, Patrick, 246, 264, 364
Renaudier, Marie, 322
Rey-Debove, Josette, 678
Rialland, Annie, xxv, 613, 637
Richardson, Irvine, 43, 239
Richter, Doris, 364
Ricquier, Birgit, 17, 672
Riedel, Kristina, 428
Rijkhoff, Jan, 707
Robinson, Clinton D. W., 43
Rose, Françoise, 328
Rose, Sharon, 510, 517
Rosen, Carol G., vii
Roth, Tim, 558, 573, 574
Rottet, Kevin J., 678
Rottland, Franz, 180
Rubanza, Yunus Ismail, 436
Rubongoya, L. T., 359
Rueck, Michael J., 264, 271
Rugemalira, Josephat M., xiii, 311,
312, 320, 332, 359, 437
Rurangwa, Innocent M., 42

Sa'ad, Isa, 249, 251
Sacleux, Charles, xvi
Salzmann, Martin, 595, 602, 617,
622–624
Samsom, Ridder, 593
Sanderson, Meredith, 475, 574, 606
Sands, Bonny, 152
Satre, Scott A., 156
Schadeberg, Thilo C., ix, x, xviii, 7,
60, 182, 201, 216, 226, 236,
238, 267, 309, 310, 317, 318,

- 320–322, 328, 329, 332–334,
344, 349–351, 360–364, 366,
390, 395, 396, 405, 410, 456,
490, 496, 514, 549, 583, 591,
593, 634, 672
- Schaefer, Ronald P., 136, 312
- Schaub, Willi, 156, 702, 708, 709
- Schiffer, Maya, 359
- Schmitz, Robert, 617
- Schoeneborn, Anne, 623
- Schwarz, Anne, 510
- Schwarz, Florian, 558
- Sebasoni, Servilien, xxvii, 108, 109,
188, 351, 359, 603
- Segerer, Guillaume, 178, 367, 489
- Seidel, Frank, 312, 626
- Sharman, John Campton, 315
- Shinagawa, Daisuke, 456
- Shultz, George, 156, 643
- Sibanda, Galen, 312
- Sihler, Andrew L., 67
- Siiyaatan, Patrick, 249, 262, 264
- Sillery, Anthony, 556, 558, 574
- Simango, Silvester R., vii
- Sitali, Georgina Nandila, 617
- Smits, Helena Johanna, 510
- Smoes, Christopher L., 643
- Smolensky, Paul, vii
- Snoxall, Ronald A., 298, 302
- Song, Sanghoun, 506
- Sonkoue, Eliane Kamdem, 135, 144,
156
- Spreda, Klaus W., 249
- Stanford, Ronald, 249
- Stanley, Carol, 156, 249, 254, 256, 410,
643, 718, 719
- Stapleton, Walter Henry, 312
- Stappers, Leo, xx, xxx, 69, 315, 344,
359–363, 396, 472, 476
- Stassen, Leon, 599, 601, 625
- Stegen, Oliver, 600, 620, 624
- Stevick, Earl W., 621
- Stewart, John M., 4, 8, 10–13, 28, 29,
32, 33, 35, 39, 397
- Stilo, Donald L., 223
- Ström, Eva-Marie, 359, 588, 620
- Struck, Bernhard, 620
- Stucky, Susan U., 428
- Tabe, Florence A. E., 510
- Tadadjeu, Maurice, 729
- Tadmor, Uri, 675
- Taji, Julius John, 606
- Takizala, Alexis, 503, 504
- Talmy, Leonard, 592
- Tamanji, Pius N., 487, 510
- Tanda, Vincent Ambe, 615, 627
- Taylor, Carrie, 364
- Taylor, Charles, 475, 476, 551, 574
- Teil-Dautrey, Gisèle, 4, 12, 16, 17, 42,
62, 711
- Thomas, Elaine, 405
- Thomas, Jacqueline M. C., 645
- Thornell, Christina, 44, 631, 633
- Thwala, Nhlanhla, 312
- Thwing, Rhonda, 130, 148, 156, 249,
254, 255, 510, 643
- Torrend, Julius, 361
- Traugott, Elizabeth C., xxii
- Trithart, Mary Lee, 236, 310, 311, 313,
314, 316–321, 324, 325, 328,
332, 333, 348, 369
- Tsala, Théodore, 43
- Tucker, Archibald N., 61, 67, 76, 84
- Turvey, B. H. C., 74
- Twilingiyimana, Chrysogone, 629
- Urmanchieva, Anna Yu., 620

Name index

- Valente, José Francisco, 437
Valinande, Nzama, 288, 289, 301
van Coillie, Gustaaf, 479
Van de Velde, Mark, v, vi, viii, ix, xii,
 xxv, xxxv, 43, 73, 200, 201,
 241, 242, 249, 265, 266, 298,
 364, 444, 445, 450, 467, 471,
 473, 480, 485, 497, 500, 526,
 602, 618, 669, 670, 686, 688,
 697–700, 704, 707, 708, 713,
 716, 722–726, 728
van den Eynde, Karel, 676
van der Auwera, Johan, viii
van der Spuy, Andrew, 363
van der Wal, Jenneke, ix, xii, xx, 428,
 436, 437, 440, 443, 453, 456,
 507, 511–513, 549–551, 563,
 571, 572, 574, 586, 594–597,
 602, 603, 606, 637, 649, 681
van Eeden, Bernardus Izak Christi-
 aan, 177, 183, 317, 318
Van Leynseele, Helene, 4, 8, 10, 13, 32,
 33
Van Olmen, Daniel, 406
Van Otterloo, Karen, 283, 284, 302
Vanhoudt, Bettie, 636, 721
Vansina, Jan, v, xix, 186, 356, 431
Velten, Carl, 346
Vennemann, Theo, 515
Verkerk, Annemarie, 648
Viana, Miguel José, 362
Visser, Marianna, 312, 624
Voeltz, Erhard F. K., 177, 236, 267,
 299, 310, 312, 316–318, 332,
 361, 367, 369
Voisin, Sylvie, 320
Voll, Rebecca M., 156, 249, 506, 508,
 510, 522, 524, 525, 644, 702,
 703, 713, 718
Voorhoeve, Jan, 708, 719
Wald, Benji, xii, xvi, xxxv, 321, 423,
 428, 442, 450, 453, 618
Wang, William S-Y, 40
Warnier, Jean-Pierre, 246
Wasow, Thomas, 515
Watters, John R., vi, xii, xxi, xxxv,
 4, 111, 120, 121, 123, 125, 126,
 130, 131, 135, 136, 140, 147,
 148, 156, 178, 226, 240, 249,
 510, 539, 549, 569
Wéga Simeu, Abraham, 19, 44
Weiss, Michael, xiv
Welmers, William E., 178, 187, 723
Werner, Alice, 361
Westermann, Diedrich, 240
Westphal, Ernst O. J., 559, 560, 574
Wetter, Andreas, 541
Whitehead, John, 451, 721, 723
Whiteley, Wilfred H., 558, 574, 606,
 607, 621
Wilhelmsen, Vera, 634, 652
Wilkendorf, Patricia, 518
Willems, Emile, 456
Williamson, Kay, 111, 123, 236, 240,
 242, 264, 394
Wills, Jeffrey, xxxv, 16, 17, 25, 242,
 710, 716, 725, 727, 729
Wilson, Janet Evelyn, 487
Wiltshire, Caroline R., 67
Woolford, Ellen, 406, 425
Wung, Bong M., 269
Wynne, R. C., 350
Yanes, Serge, 79
Yates, Anthony D., 33
Yillah, M. Sorie, 546, 557, 573
Yokoyama, Tomohiro, 434, 435

Yoneda, Nobuko, xii, 428, 456, 496,
564, 574

Yukawa, Yasutoshi, 43, 207

Zeller, Jochen, 435, 441, 500, 501, 593,
604

Ziervogel, Dirk, 621

Zimmermann, Wolfgang, 432

Zukoff, Sam, 33

Zúñiga, Fernando, 709

Bantu language codes

Close to 500 Bantu language varieties are mentioned in the present book. In the reverse index which follows, these are listed according to their Guthrie/Maho codes.

A10-20-30	Sawabantu languages	A21	Mboko
A101	Oroko	A22	Kpe (Mokpwe, Mokpe, Bakweri)
A11	Lundu (Londo)		Bubia (Bobe)
A111	Ngolo	A221	Wovia
A112	Bima	A222	Su (Isubu)
A113	Batanga	A23	Kole
A114	Lokoko	A231	Duala
A115	Londo ba Diko	A24	Bodiman
A12	Lue	A241	Oli
A121	Mbonge	A25	Pongo
A122	Kundu (Lukundu, Bakundu)	A26	Mongo (Mungo)
A123	Ekombe	A261	Limba (Malimba)
A13	Balong	A27	Bubi (Bobe)
A14	Bonkeng	A31	Northern Bubi
A141	Bafo (Lefo)	A31a	Batete
A15	Manenguba (Mbo)	A31b	South-West Bubi
A15A	Mbuu (Mboo)	A31b	South-East Bubi
A15B	Myenge	A31c	Noho (Banoho)
A15C	Akoose (Akooṣe, Akossi)	A32a	Bapuku (Poko)
A15C	Belon	A32b	Batanga
A15C	Elung	A32C	Yasa
A15C	Mkaa	A33a	Kombe
A15C	Mwahed	A33b	Benga
A15C	Mwaneka	A34	Mbam
A15C	Nnenong	A40 & A60	Lombi
A151	Nkongho	A41	Abo (Bankon)
		A42	Basaa (Basaá)
		A43a	

Bantu language codes

A43a	Dibum	A75D	Atsi
A43b	Bakoko (Koko, North Kogo)	A75F A80	Mvai Nyong-Dja group
A43c	South Kogo	A801	Gyeli (Gyele, Bagyeli)
A44	Nen (Tunen)	A803	Shiwa
A45	Nyokon (Nyo'on)	A81	Kwasio (Mvumbo)
A46	Maande (Nomaande, Mandi)	A82 A83	So Makaa (Mekaa)
A461	Tuotomb	A832	Kol
A462	Yambeta	A84	Njem (Njyem)
A50	Bafia group	A841	Bajwee
A501	Hijuk	A842	Koonzime
A51	Fa' (Zakaan, Maja, Balom)	A85b A86c	Bekwel Mpiemo
A52	Dimbong	A91	Kwakum
A53	Kpa (Kpā?, Bafia, Rikpa, Pe)	A92a A92b	Polri Pomo
A54	Bea (Ngayaba)	A93	Kako
A60	Sanaga group	B11	Myene
A60?	Leti	B11a	Mpongwe
A601	Tuki (Ki)	B11b	Orungu
A62	Yambasa (Yambassa)	B11c	Galwa
A62A	Yangben	B11e	Nkomi
A62B	Mmala	B21	Seki
A62C	Elip	B22a	West Kele
A621	Baca	B22b	Ngom
A622	Gunu (Nugunu)	B24	Wumbvu
A623	Mbule (Mbure)	B25	Kota
A63	Njowi (Mangisa, Mengisa)	B251 B302	Shake (Sake) Himba
A70	Ewondo-Fang group	B304	Pinji
A70 & A63	Beti group	B305	Vove
A71	Eton	B31	Tsogo
A72a	Ewondo	B42	Sangu
A74a	Bulu	B43	Punu (Yipunu)
A75	Fang	B44	Lumbu (Yilumbu)
A75A	Ntumu	B51	Duma
A75B	Okak	B52	Nzebi
A75C	Meke	B53	Tsaangi

Bantu language codes

B602	Kaningi Nord	C35a	Ntomba-Bikoro
B61	Mbede	C35a	Ntomba-Inongo
B63	Ndumu	C35a	Ntomba-Njale
B70	Teke group	C35b	Bolia
B701	Atsitsege	C36c	Mbudza
B71aIb	Teke d'Ibali	C36e	Boloki
B72a	Ngungwel	C37	Buja
B73b	Teke Laali	C371	Motembo
B73c	Teke Yaa	C41	Ngombe
B73d	Teke Tyee	C411	Bomboma
B75	Bali-Teke	C412	Libobi
B77a	Kukuya	C414	Ligendza
B77b	Fumu	C44	Babanda
B78	Wuumu	C44	Boa
B81	Tiene	C44	Boa Buta
B82	Boma (North Boma)	C45	Ngelema
B83	Mfinu	C45A	Beo
B85	Yanzi	C50	Soko-Kele group
B85d	Nsong	C502	Linga
B861	Ngwi	C51	Mbesa
B865	Nzadi	C52	Soko (Sogo)
B87	Mbuun	C53	Gesogo
C10	Bwamba	C55	Kele (Lokele)
C101	Babole	C55	Kele Yawembe
C101A	Babole Bakolu	C61	Mongo (Nkundo, Mongo-Nkundo)
C104	Aka		
C14	Leke	C61A	Bakutu
C15	Bongili	C61E	Konda (Lokonda)
C24	Koyo Ehamba	C61J	Ntomba
C25	Mboshi	C61L	Mongo-Liinja
C25A	Bunji	C70	Tetela group
C30	Ngiri group	C71	Tetela
C30B	Lingala	C73	Nkucu
C301	Doko	C73	Nkucu Wela
C31b	Gyando	C74	Yela
C314	Balobo	C75	Kela
C32	Bangi (Bobangi)	C76	Ombo
C322	Dzamba	C81	Ndengese
C35a	Ntomba	C83	Bushong (Bushoong)

Bantu language codes

D ~ C ?	Yasanyama	E72d	Duruma
D102	Kwange	E73	Digo
D11	Mbole	E74	Taita
D12	Lengola	E74a	Dawida
D13	Mituku	E741	Saghala
D14	Enya	F12	Bende
D14	Enya Kibombo	F21	Sukuma
D14	Enya Manda	F22	Nyamwezi
D201	Liko		(Kinyamwezi)
D23	Komo	F23	Sumbwa
D24	Songola	F31	Nilamba (Nilyamba)
D24	Songola Kasenga	F32	Rimi (Remi, Nyaturu)
D25	Lega	F33	Rangi
D26	South Binja	F34	Mbugwe
D27	Bangubangu	G11	Gogo
D28	Holoholo	G12	Kagulu
D301	Ndaka	G22	Pare
D308	Bodo	G23	Shambaa (Shambala)
D31	Bhele	G32	Ng'hwele
D311	Bila	G35	Luguru
D32	Bira	G38	Vidunda
D33	Nyali	G40C	Kisetla (Settler Swahili)
D332	Budu	G402	Makwe
D43	Nyanga	G403	Mwani
D54	Bembe	G41	Tikuu
D55	Buyu	G412	Mwiini (Chimwiini)
E51	Kuyu (Kikuyu, Gikuyu)	G42a	Amu
E54	Tharaka	G42d	Swahili (Kiswahili,
E55	Kamba		Unguja)
E60	Chaga	G44	Comorian
E60 & E74	Kilimanjaro languages	G44a	Ngazija (Ngazidja)
E621B	Mashami	G52	Ndamba (Chindamba)
E622C	Vunjo-Chaga	G61	Sango
E623B	Mashati	G62	Hehe
E65	Gweno	G63	Bena
E70 & G40	Sabaki languages	G65	Kinga
E71	Pokomo	G67	Kisi
E71B	Lower Pokomo	H10A	Kituba (Kikongo ya Leta)
E72a	Giryama	H11	Bembe (Kibembe)

Bantu language codes

H112A	Kamba (Kikamba)	JE13/14	Nkore-Kiga
H12L	Vili (Civili)		(Runyankore-Rukiga)
H131K	Sundi (Kisundi)	JE15	Ganda (Luganda)
H16	Kongo (Kikongo)	JE16	Soga (Lusoga)
H16a	Sikongo (Kisikongo)	JE17	Gwere (Lugwere)
H16aM	Solongo (Kisolongo)	JE17	Lamogi (Lulamoogi)
H16b	Manyanga	JE18	West Nyala
	(Kimanyanga)	JE20	Haya-Jita group
H16bZ	Ndibu (Kindibu)	JE21	Nyambo
H16c	Yombe (Kiyombe)	JE22	Haya
H16d	Fiote	JE22D	Ziba
H16dK	Woyo (Kiwoyo)	JE22E	Hanja
H16f	Laadi (Cilaadi)	JE24	Kerebe
H16g	Ntandu (Kintandu)	JE25	Jita
H16hK	Zombo (Kizombo)	JE25I	Kwaya
H16hZ	Tsootso (Kitsootso)	JE31	Masaba
H21	Mbundu (Kimbundu)	JE31c	Bukusu
H31	Yaka (Kiyaka)	JE32	Luyia
H32	Suku (Kisuku)	JE32a	Wanga
H321	Soonde (Kisoonde)	JE32b	Tsotso
H41	Mbala (Kimbala)	JE32C	Marama
H42	Hungan	JE32D	Kisa
J	Great Lakes Bantu,	JE32E	Kabarasi
	Interlacustrine Bantu	JE32G	Tura
JD41	Konzo	JE34	Samia
JD42	Nande (Kinande,	JE341	Khayo
	Nandi)	JE342	Marachi
JD51	Hunde	JE401	Ngoreme
JD53	Shi	JE402	Kizu (Ikizu)
JD531	Tembo	JE41	Logooli
JD61	Rwanda	JE411	Idakho (Idaxo)
	(Kinyarwanda)	JE412	Isukha (Isuxa)
JD62	Rundi (Kirundi)	JE413	Tiriki
JD63	Fuliiru	JE42	Gusii
JD66	Ha	JE43	Kuria
JE102	Talinga (Kitalinga)	JE431	Simbiti
JE11	Nyoro	JE45	Western Serengeti
JE12	Tooro (Rutooro)		languages
JE13	Nkore (Runyankore)	JE45	Ikoma

Bantu language codes

JE45	Ishenyi	L53	Ruund (Ruwund)
JE45	Nata	M13	Fipa
K11	Chokwe (Ciokwe)	M14	Lungu (Cilungu)
K13	Lucazi	M15	Mambwe
K14	Luvale	M22	Mwanga (Namwanga)
K15	Mbunda	M23	Nyiha
K21	Lozi	M24	Malila
K30, K40 & M60	Botatwe languages	M301	Ndali
K31	Luyana	M31	Nyakyusa
K33	Kwangali	M41	Taabwa
K332	Gciriku	M42	Bemba (Icibemba)
	(Dciriku, Manyo)	M54	Lamba
K333	Mbukushu	M61	Lenje
K36	Shanjo	M62	Soli
K402	Fwe	M63	Ila
K41	Totela (Totela of Zambia)	M64 N101	Tonga (Plateau Tonga) Ndendeule
K411	Totela (Totela of Namibia)	N11 N12	Manda Ngoni (Ngoni of Tanzania)
K42	Subiya		
L11	Pende	N121	Ngoni (Ngoni of Malawi)
L12a	Samba		
L13	Kwezo (Kwese)	N122	Ngoni (Ngoni of Mozambique)
L21	Kete		
L22	Mbagani (Binji)	N13	Matengo
L221	Lwalwa	N14	Mpoto
L23	Songye	N15	Tonga
L30	Luba group	N21	Tumbuka
L31a	Luba-Kasai	N31a	Nyanja
	(Ciluba, Tshiluba)	N31b	Chewa (Chichewa, Cewa)
L31b	Lulua		
L32	Kanyok	N31D	Nyasa
L33	Luba-Katanga	N31E	Ntcheu
	(Kiluba, Luba-Shaba)	N31F	Nkhotakota
L34	Luba-Hemba	N41	Nsenga
L41	Kaonde	N43	Nyungwe
L51	Salampasu	N44	Sena
L52	Lunda	P11	Ndengeleko
L53	Kanincin	P13	Matuumbi

Bantu language codes

P14	Ngindo	S16	Kalanga
P20	Ruvuma group	S21	Venda (Tshivenda)
P21	Yao (Ciyao)	S30	Sotho-Tswana group
P22	Mwera	S31	Tswana (Setswana)
P23	Makonde (Shimakonde)	S31	Hurutshe
P23	Chinnima Makonde	S31	Kgatla
P23	Plateau Makonde	S31a	Rolong
P231	Makwe	S32	Northern Sotho
P25	Mabiha		(Sesotho sa Leboa)
P31	Makhuwa (Makua, Makwa, Makwe)	S33	Southern Sotho (Sesotho)
P31	Makhuwa Ile	S40	Nguni group
P31	Makhuwa Nampula	S407	Nrebele
P312	Shangaji	S41	Xhosa (isiXhosa)
P34	Cuwabo	S42	Zulu (isiZulu)
R11	Umbundu	S43	Swati (siSwati)
R20	Wambo (Oshiwambo)	S44	Sindebele (Ndebele of Zimbabwe)
R21	Kwanyama		Tswa
R22	Ndonga	S51	Tsonga (Xitsonga, Changana)
R23	Kwambi	S53	Luleke
R24	Ngandjera	S53A	Nkuna
R30	Herero (Otjiherero)	S53D	Konde
R41	Yeyi	S54	Ronga
S10	Shona (Chishona)	S54	Copi
S12	Zezuru	S61	
S13	Manyika		

Language index

Bantu languages are listed with their Guthrie/Maho code as well as their frequently used variants between round brackets '()'. All other languages are listed with an indication of their affiliation and their variants between square brackets '[]'. As a rule, Bantu language names are shorn of their noun class prefix.

- Abanyom [Ekoid], 259
Abar [Yemne-Kimbi], 257
Abo (A42; Bankon), 6, 22, 42, 79, 86, 91, 298
Abon [Tivoid], 245, 272
Adamawa [Niger-Congo subgroup], 239, 248, 394
Adamawa-Ubangi [Niger-Congo subgroup], 643, 644
Adere [North-Eastern Grassfields], 32
African languages, viii, xxiii, 566, 571, 675
Afro-Asiatic, 541, 543, 544
Aghem [Ring Grassfields], 90, 113, 127, 128, 130, 131, 134–136, 141, 142, 145, 146, 150, 156, 158, 161, 164, 243, 261, 272, 389, 688
Ahanta [Kwa], 10
Aja [Kwa], 567
Ajumbu [Yemne-Kimbi; = Mbu'], 142, 257, 261, 272
Aka (C104), 638, 644, 645
Akamkpa-Ejagham [Ekoid], 259
Akan [Kwa], 10, 13, 39
Akkadian [Semitic, Afro-Asiatic], 319
Akɔɔse, *see* Akoose (A15C)
Akoose (A15C; Akɔɔse, Akossi), 6, 21, 22, 24, 30, 42, 90, 110, 113, 117, 185–189, 191, 220, 223, 225, 226, 243, 249, 251, 263, 266, 267, 270, 406, 669, 688, 693, 695, 700
Akossi, *see* Akoose (A15C)
Albanian [Indo-European], 319
Alege [Bendi], 243
Ambele [Bantoid], 242, 243, 249, 261, 272
Ambo [Tivoid], 245, 272
Amharic [Semitic, Afro-Asiatic], 541, 542, 546, 550
Amu (G42a), 34
Anyi-Baule [Kwa], 10
Arabic [Afro-Asiatic], vi
Aramaic [Semitic, Afro-Asiatic], 319
Arawak, *see* Mojeño Trinitario
Asante Twi [Kwa], 510
Atlantic [Niger-Congo subgroup], 111, 178, 316, 319, 320, 322, 323, 328, 361, 367, 394, 489, 510, 546, 637
Atong [South-West Grassfields], 261, 272

- Atsi (A75D), 6, 19, 31, 445, 450
 Atsitsege (B701), 91
 Austronesian, 326, 716, *see* Maori, *see*
 Rapa Nui, *see* To'aba'ita
 Awing [Ngemba Grassfields], 261,
 272
- Baba [Nun Grassfields], 261, 272
 Babanda (C44), 689
 Babanki [Ring Grassfields], 113, 128,
 135, 136, 138, 142, 145, 150,
 156, 158, 161, 164, 243, 261,
 272, 702, 703
 Babole (C101), 73, 110, 112, 619, 627,
 640, 641
 Babole Bakolu (C101A), 690
 Babungo [Ring Grassfields; =
 Vengo], 128, 130, 131, 134,
 140–142, 146, 150, 156, 158,
 161, 164, 261, 272, 702, 708,
 709, 724
- Baca (A621), 5, 13, 14, 43, 45, 79, 711
 Bafanji [Nun Grassfields], 261, 272
 Bafia, *see* Kpa (A53)
 Bafia (A50 languages), 5, 7, 19
 Bafo (A141; Lefo), 6, 30, 42, 370
 Bafut [Ngemba Grassfields], 128, 136,
 137, 140, 142, 150, 156, 158,
 161, 164, 165, 243, 261, 272,
 487, 510
- Bagirmi [Bongo-Bagirmi, Central
 Sudanic], 542, 543, 546, 549
 Bagyeli, *see* Gyeli (A801)
 Bajwee (A841), 6
 Bakoko (A43b; Koko, North Kogo), 6,
 30, 43, 72, 364, 370, 635, 638,
 644
- Bakundu, *see* Kundu (A122)
 Bakutu (C61A), 517
- Bakweri, *see* Kpe (A22)
 Balep [Ekoid], 259
 Bali-Teke (B75), 80, 91
 Balo [South-West Grassfields], 261,
 272
 Balobo (C314), 688
 Balom, *see* Fa' (A51)
 Balong (A13), 6, 21, 30, 42
 Bamali [Nun Grassfields], 261, 272
 Bambalang [Nun Grassfields], 261,
 272
 Bambili-Bambui [Ngemba Grass-
 fields], 249, 261, 272
 Bamenyam [Nun Grassfields], 261,
 272
 Bamileke [Mbam-Nkam subgroup],
 110, 116, 128, 131–133, 135,
 137, 138, 141–143, 146, 148–
 151, 156, 158, 161, 164, 243,
 249, 260–262, 272, 547, 643
 Bamukumbit [Ngemba Grassfields],
 261, 272
 Bamun [Nun Grassfields], 261, 272
 Bamunka [Ring Grassfields], 243,
 261, 272
 Bangi (C32; Bobangi), 73, 90, 91, 425,
 431, 451, 629, 640, 721, 723
 Bangolon [Nun Grassfields], 261,
 272
 Bangubangu (D27), 356
 Bankal [Jarawan], 91, 265, 272
 Bankon, *see* Abo (A42)
 Banoho, *see* Noho (A32a)
 Bantoid [Benue-Congo subgroup],
 xi, xix, xxi–xxiv, xxvii–
 xxix, 63, 66, 70, 71, 80, 89,
 92, 94, 95, 105, 106, 108–
 111, 113, 114, 117, 118, 120–
 127, 129–136, 138–143, 145–

- 151, 165, 235–242, 244–248,
251, 253, 254, 258, 262, 265,
269–272, 316, 326, 332, 364,
372, 389, 395, 396, 409, 410,
414, 445, 447, 448, 499, 508,
510, 523–525, 547, 569, 643,
644, 669, 684, 702, 703, 705,
708, 710, 713, 718, 724, 725,
727
- Bantoid-Cross [Benue-Congo], 241
- Bantu (= Narrow Bantu), iii, v–x,
xiii–xxxv, 4–8, 10, 11, 15,
18, 26, 29, 32–37, 39–41, 60,
62–67, 70, 71, 75, 78, 80–
95, 105–107, 110, 111, 118,
120–122, 124, 132, 136, 139,
142, 145–150, 152–155, 173–
186, 188, 189, 191, 196–198,
201, 202, 205, 208, 210–
212, 214–219, 223–227, 236–
244, 247–249, 253, 264–267,
269–271, 281, 283, 286, 290,
292, 294, 298, 299, 301,
303, 304, 309, 311, 313–324,
326, 327, 329, 333, 334, 343,
345–348, 350–356, 359–373,
388–415, 423–434, 436, 437,
439, 441, 443–453, 465–
468, 470–474, 476–478, 480,
481, 483, 485–490, 495–502,
505–508, 510, 511, 513–525,
538, 539, 541, 545–547, 549,
551, 552, 554–561, 564, 565,
567–571, 584–587, 591–595,
597–602, 605, 607, 612–615,
617, 619, 623, 625, 627–632,
636, 637, 641, 643, 645–651,
668–674, 676, 677, 681–690,
692, 693, 695–698, 701–711,
713–715, 718, 719, 721–729
- Bapuku (A32b; Poko), 5, 42
- Basaá, *see* Basaa (A43a)
- Basaa (A43a; Basaá), 6, 9, 10, 12–17,
19–22, 24, 25, 27, 30, 36, 38,
42, 79, 86, 90, 110, 117, 118,
150, 298, 324, 364, 444, 448,
450, 455, 496, 500, 501, 507,
513–516, 522, 523, 525, 547,
573, 638, 644, 669, 681, 683,
688, 694, 695, 699, 700, 711
- Basque [language isolate], 67, 319
- Batanga (A113), 5
- Batanga (A32C), 5, 30, 42, 79, 719
- Batete (A31b), 30, 446
- Batu [Tivoid], 240, 245, 272
- Bea (A54; Ngayaba), 5, 43
- Beba [Ngemba Grassfields], 261, 272
- Bebe [Beboid], 257, 272
- Beboid [Bantoid subgroup], xxviii,
89, 105, 122–126, 128, 129,
133, 136–139, 142, 143, 145,
147–151, 156, 158, 161, 164,
236, 242–245, 248, 249,
256–258, 269–271, 510, 705
- Befang [Menchum], 249
- Bekwara [Bendi], 243, 249
- Bekwel (A85b), 6, 26, 31, 44
- Belon (A15C), 6
- Bemba (M42; Icibemba), vii, 38, 315,
439, 440, 500, 512, 616
- Bembe (D54), 512
- Bembe (H11), 76
- Bena (G63), 361, 616, 621
- Bena-Mboi, *see* Buto [Benue-Congo]
- Bena-Yungur [Adamawa], 708, 709,
718, 724–726
- Bende (F12), 616, 652
- Bendeghe [Ekoid], 259

- Bendi [Bantoid subgroup], 122–124, 236, 241–244, 249
- Benga (A34), 5, 8, 42, 79, 92, 110, 117, 150, 361, 370, 621, 627, 640
- Benue-Bantu, 514
- Benue-Congo [Niger-Congo subgroup], vii, xii, xviii–xxi, xxiv, xxvii, xxx, 10, 36, 111, 136, 152, 200, 236–240, 242, 247, 267, 310, 316, 326, 334, 373, 394, 447, 448, 466, 486–488, 507, 513, 555, 643, 644, 669, 682, 683, 703, 705, 715, 718, 723, 724
- Benue-Kwa [Niger-Congo subgroup], xxv, xxxi, 389, 393–398, 401, 403–408, 410–414, 567
- Beo (C45A), 587, 630, 691
- Beti (A70 languages + Njowi A73), 6, 19, 20, 22, 24–26, 29, 37
- Bhele (D31), 691
- Bijogo [Atlantic], 178, 179, 227, 367, 489
- Bikya [a Furu language], 261, 272
- Bila (D311), 638–642, 644
- Bile [Jarawan], 265, 272
- Bima (A112), 5
- Bini [Edoid], 487
- Binji, *see* Mbagani (L22)
- Bira (D32), 76, 406, 409, 638–642, 644, 683
- Birom [Plateau], 121, 487
- Bishuo [Ring Grassfields], 261, 272
- Bitare [Tivoid], 240, 245, 272
- Biya [Yemne-Kimbi; = Za’], 257
- Boa (C44), 443, 688
- Boa Buta (C44), 689
- Bobangi, *see* Bangi (C32)
- Bobe, *see* Bubi (A31), *see* Bubia (A221)
- Bodic [Sino-Tibetan; = Tibeto-Kanauri], 319
- Bodiman (A241), 5, 42
- Bodo (D308), 690
- Bokyi [Bendi], 243
- Bolia (C35b), 113, 629, 688, 694
- Boloki (C36e), 80
- Boma (B82; North Boma), 73, 76, 110, 315
- Bomboma (C411), 638, 644
- Bongili (C15), 80, 629, 640
- Bongo-Bagirmi [Central Sudanic], 542, 543
- Bonkeng (A14), 6
- Border (Papua New Guinea), *see* Imonda
- Botatwe (K30, K40 & M60 languages), 74, 75
- Bu [Yemne-Kimbi], 257
- Bubi
- Northern (A31a), 5, 7, 18, 25, 27, 28, 30
- South-East (A31c), 5, 18, 27, 28, 30
- South-West (A31b), 5, 18, 25, 27, 28, 30
- Bubi (A31; Bobe), 5–7, 18, 21, 28, 30, 34, 36, 42, 84, 90, 91, 150, 325, 361, 434, 438, 445–448, 450, 452, 518, 519
- Bubia (A221; Bobe), 5, 19, 42
- Budu (D332), 635, 638, 644, 645
- Buja (C37), 694
- Bukusu (JE31c), 69, 284, 512, 691
- Bulgarian [Indo-European], 319
- Buli [Gur], 510

Language index

- Bulu (A74a), 6, 9, 27, 43, 70, 72, 73, 79, 91, 150, 370, 618, 627, 640, 641
- Bum [Ring Grassfields], 261, 272
- Bunji (C25A), 86
- Buru [Bantoid], 241, 243–245, 249, 272
- Busam [South-West Grassfields], 261, 272
- Bushong (C83; Bushoong), 73, 80, 110, 112
- Bushoong, *see* Bushong (C83)
- Busuu [Ring Grassfields], 261, 272
- Buto [Benue-Congo; = Bena-Mboi], 708, 718, 724, 726
- Buu [Yemne-Kimbi], 142
- Buyu (D55), 616, 627
- Bwamba (C10), 686, 688, 696
- C'lela [Kainji], 487
- Caka [Tivoid], 245, 272
- Cama [Kwa; = Ebrié], 10, 11, 39
- Central Nigerian, *see* Platoid
- Central Sudanic [Nilo-Saharan], 401, 407, 542, 543, 643, 644
- Central-Western Bantu (CWB), xviii, xix, xxxiv, 65, 72, 107, 109, 117, 153, 324, 325, 328, 356, 358, 360, 363, 366, 370, 373, 431, 432, 438, 443, 449, 451–454, 457, 507, 511, 518, 613, 614, 634, 636–640, 642, 646–648, 652
- Cewa, *see* Chewa (N31b)
- Chadic [Afro-Asiatic], 523, 543, 544
- Chaga (E60), vii, 83, 512
- Chamba Daka, *see* Sama Mum [Dakoid]
- Changana, *see* Tsonga (S53)
- Chewa (N31b; Chichewa, Cewa), vii, 82, 86, 95, 174, 175, 181, 295, 297, 303, 346, 347, 500–502, 512, 616
- Chibchan, *see* Rama
- Chichewa, *see* Chewa (N31b)
- Chimwiini, *see* Mwiini (G412)
- Chindamba, *see* Ndamba (G52)
- Chinese [Sino-Tibetan], vi
- Chinnima Makonde (P23), 607
- Chishona, *see* Shona (S10)
- Chokwe (K11; Ciokwe), 478, 505, 686, 692, 696
- Cicipu [Kainji], 200, 247, 271, 487
- Cilaadi, *see* Laadi (H16f)
- Ciluba, *see* Luba-Kasai (L31a)
- Cilungu, *see* Lungu (M14)
- Ciokwe, *see* Chokwe (K11)
- Civili, *see* Vili (H12L)
- Ciyao, *see* Yao (P21)
- Common Bantu, xvii, 17, 39, 63, 237
- Comorian (G44), 83
- Copi (S61), 574
- CR, *see* Cross River [Benue-Congo subgroup]
- Cross River [Benue-Congo subgroup] (CR), 118, 120, 122, 123, 139, 140, 153, 236, 240–242, 247, 258, 269
- cross-Bantu, 393, 426, 487, 546
- Cushitic [Afro-Asiatic], 647
- Cuwabo (P34), 36, 368, 369, 371, 484, 538, 596, 598–600, 617
- CWB, *see* Central-Western Bantu
- Daka [Dakoid], 243, 249, 250, 272
- Dakoid [Bantoid subgroup], xxviii, 122–124, 126, 236, 241–243, 245, 248–250, 254, 269–271

- Dawida (E74a), 37, 83
Dciriku, *see* Gciriku (K332)
Degema [Edoid], 487
Denya [Bantoid], 139, 249
Dibum (A43a), 79
Digo (E73), 442, 590, 621, 720
Dimbong (A52), 5, 43
Dogon [Niger-Congo subgroup], 118, 236
Doko (C301), 629, 640, 686, 688, 691, 695
Dong [Dakoid], 250, 272
Dravidian, *see* Tamil
Dschang, *see* Yemba [Bamileke]
Duala (A24), xv, xvi, 5, 8, 9, 18, 21, 27, 28, 30, 37, 38, 42, 72, 77, 79, 80, 110, 113, 150, 358, 364, 370, 448, 518, 684, 685, 688, 719
Duguri [Jarawan], 265, 272
Dulbu [Jarawan], 265, 272
Duma (B51), 8, 110, 112–114, 224, 520
Duruma (E72d), 442
Dutch [Indo-European], 673, 674, 676
Dzamba (C322), 427, 507, 512
Dzodinka [North-Eastern Grassfields], 261, 272

East Benue-Congo (EBC), xxi, xxii, 456
Eastern Bantu (EB), vii, xix, xxix, xxxiii, 4, 38, 65, 78, 81, 84, 86–88, 95, 107, 117, 153, 312, 320, 324, 325, 352, 356, 358–362, 365, 366, 370, 372, 373, 431–437, 439–443, 451–455, 457, 507, 511, 518, 614, 618, 620, 638, 646, 653, 687, 689, 692, 704, 716, 720
Eastern Grassfields, 15, 29, 32, 84, 106, 124, 126, 128, 131–133, 135, 137, 140–143, 146, 148–151, 156, 158, 161, 164, 261, 269, 510
EB, *see* Eastern Bantu
EBC, *see* East Benue-Congo
Ebrié, *see* Cama [Kwa]
Edoid [Niger-Congo subgroup], 136, 405, 487
Efutop [Ekoid], 259
Ejagham [Ekoid], 118, 125, 135, 136, 139, 140, 243, 249, 259, 487, 510, 549, 569
Ekajuk [Ekoid], 259
Ekoid [Bantoid subgroup], xxviii, 71, 80, 89, 122–126, 135, 139, 236, 239, 243, 244, 248, 249, 258–260, 510, 549, 569
Ekombe (A123), 5
Ekparabong [Ekoid], 259
Ekpeye [Igbooid], 410
Ekwe [Ekoid], 259
Elemé [Cross River], 487
Elip (A62C), 5, 13, 14, 21, 30, 43, 45, 77
Elung (A15C), 6, 72
Emai [Edoid], 136
Eman [Tivoid], 245, 272
Engenni [Edoid], 405, 487
English [Indo-European], xvi, 26, 313, 319, 515, 539, 544, 564, 569, 572, 582, 678, 689, 710, 728
Enya (D14), 112, 520, 690
Enya Kibombo (D14), 670, 681, 690, 691, 719
Enya Manda (D14), 690, 719

Language index

- Esimbi [Bantoid], 245, 272, 643, 702
Eton (A71), 6, 25, 27, 31, 43, 72, 73, 90,
179, 181, 185, 186, 200–202,
205, 220, 223, 224, 298, 431,
444, 445, 450, 455, 497, 601,
602, 618, 638, 644, 669, 687,
688, 697–700, 707, 708, 713,
717, 727–729
Etung [Ekoid], 243, 259
Evant [Tivoid], 245, 272
Ewe [Kwa], 319, 510
Ewondo (A72a), 6, 21, 27, 31, 43, 72,
79, 110, 112, 142, 146, 150,
370, 389, 401, 450, 451, 618,
627, 640, 688, 697–700, 707,
721
Ewondo-Fang (A70 languages), 364
Fa' (A51; Zakaan, Maja, Balom), 5, 9,
14, 20, 21, 29, 30, 36, 37, 43
Fam [Bantoid isolate?], 123, 252, 272
Fang (A75), 6, 7, 27, 34, 43, 72, 690
Fang [Yemne-Kimbi], 257, 261, 272
Fante [Kwa], 10
Fe'fe', *see* Fefe [Bamileke]
Fefe [Bamileke; = Fe'fe'], 90, 141, 261,
272
Finno-Ugric, xiv
Fiote (H16d), 552, 553
Fipa (M13), 620
French [Indo-European], 67, 109, 319,
425, 592, 602, 676, 678, 683,
684, 689, 721
Fula [Atlantic], 319, 323
Fuliiru (JD63), 283, 284, 294, 302, 616
Fum [North-Eastern Grassfields],
261, 272
Fumu (B77b), 669, 690
Fungom [Yemne-Kimbi], 243
Furu [Bantoid], 242–244, 249, 261,
272
Fwe (K402), 75, 313, 314, 559, 574, 618,
619
Fyem [Plateau], 487
Gaa [Dakoid], 250, 272
Galwa (B11c), 143
Ganda (JE15; Luganda), vii, 61, 67, 76,
77, 82, 85, 90, 284, 290, 291,
293, 294, 298–303, 359, 360,
424, 436, 455, 549, 550, 574,
596, 597, 631, 652
Gciriku (K332; Dciriku, Manyo), 512,
559, 574
German [Indo-European], 67, 95,
319, 539, 544–546, 550, 555,
565, 673, 674
Gesogo (C53), 110, 629, 640
Ghomala' [Bamileke], 142, 261, 272
Gikuyu, *see* Kuyu (E51)
Giryama (E72a), 693
Gogo (G11), 117, 361
Grassfields [Bantoid subgroup], xix,
xxviii, 7, 29, 32, 38, 40, 88–
90, 95, 105, 108, 117, 118, 120–
123, 125, 126, 128, 131, 133,
135–139, 141, 143, 145, 147–
149, 151, 152, 156, 158, 161,
164, 165, 236, 239, 240, 242–
244, 246, 248, 249, 256, 260–
262, 269, 270, 326, 389, 430,
523, 525, 526, 547, 548, 570,
574, 702, 703, 705, 711, 714,
721, 727, 729
Great Lakes Bantu, *see* zone J lan-
guages
Greek [Indo-European], vi, 67

- Gunu (A622; Nugunu), 5, 9, 13, 14,
 27, 43, 45, 110, 142, 185, 186,
 198–200, 220, 225, 364, 518,
 519, 638, 644, 645
 Gur [Niger-Congo subgroup], 111,
 316, 367, 394, 406, 510
 Gur-Adamawa [Niger-Congo sub-
 group], 316
 Gusii (JE42), 38, 361, 558, 574
 Gwa [Jarawan], 265, 272
 Gwak [Jarawan], 265, 272
 Gweno (E65), 36, 623, 624
 Gwere (JE17; Lugwere), 284
 Gyando (C31b), 688
 Gyele, *see* Gyeli (A801)
 Gyeli (A801; Gyele, Bagyeli), 6, 19, 21,
 31, 43, 179, 181, 185, 186, 202,
 203, 205, 220, 223, 224, 599,
 600, 635, 638, 644, 652

 Ha (JD66), 76
 Hanja (JE22E), 436
 Hausa [Chadic, Afro-Asiatic], 543,
 544, 546, 550
 Haya (JE22), 314, 359, 360, 369, 435,
 436, 452, 455, 605, 631
 Haya-Jita (JE20 languages), 284
 Hebrew [Semitic, Afro-Asiatic], 319
 Hehe (G62), 90, 616, 652
 Herero (R30; Otjiherero), xv, xvi, 80,
 85, 182, 284, 296, 297, 303,
 304, 354–356, 512, 621
 Hijuk (A501), 6
 Himba (B302), 110, 184, 186, 214, 215,
 221, 324
 Holoholo (D28), 69, 112, 284, 285, 356,
 615–617, 627
 Hunde (JD51), 685–687
 Hungan (H42), 503, 504

 Hup [a Nadahup language of Ama-
 zonia], 327
 Hurutshe (S31), 433, 456

 Ibibio [Lower Cross River], 456, 487
 Iceve-Maci [Tivoid], 245, 272
 Icibemba, *see* Bemba (M42)
 Idakho (JE411; Idaxo), 113, 114, 284
 Idaxo, *see* Idakho (JE411)
 Idoma [Idomoid], 406
 Idomoid [Niger-Congo subgroup],
 406
 Igbo [Igboid], vii, 405, 406
 Igboid [Niger-Congo subgroup],
 240, 406, 410
 Ijò [Ijoid, Niger-Congo], 236
 Ikizu, *see* Kizu (JE402)
 Ikoma (JE45), 595, 646
 Ila (M63), 75, 76, 91
 Imonda [Border (Papua New
 Guinea)], 319
 Indo-European, xiv, 67, 105, 319, 326,
 544, 545, 676
 Interlacustrine Bantu, *see* zone J lan-
 guages
 Ipulo [Tivoid], 245, 272
 Ishenyi (JE45), 609–612, 621, 629, 631,
 646
 isiXhosa, *see* Xhosa (S41)
 isiZulu, *see* Zulu (S42)
 Isu [Ring Grassfields], 243, 261, 272,
 702
 Isubu, *see* Su (A23)
 Isukha (JE412; Isuxa), 284, 691
 Isuxa, *see* Isukha (JE412)
 Italian [Indo-European], 67, 675, 678
 Iyive [Tivoid], 243, 245, 272
 Izere [Platoid], 406, 456

Language index

- Jaku [Jarawan], 91
Jar [Jarawan], 243
Jarawan [Bantoid subgroup], 7, 63,
65, 89–91, 122–124, 126,
177, 186, 236, 242–244, 249,
264–266, 271, 334, 430, 670,
684, 695, 706, 721, 722
Jita (JE25), 90, 652
Ju Ba [Mambiloid], 126–128, 131, 137,
142, 147–150, 156, 158, 161,
164
Jukun [Jukunoid], 118
Jukunoid [Benue-Congo subgroup],
121, 122, 240–242, 247, 487

Kabarsi (JE32E), 284
Kagulu (G12), 117, 512, 621, 670, 685–
688, 693, 697, 706, 719, 720
Kainji [Benue-Congo subgroup], 121,
135, 200, 241, 247, 267, 271,
456, 487
Kaje [Platoid], 456
Kako (A93), 6, 7, 18, 44, 86, 110, 150,
179, 181, 182, 184–186, 207–
210, 220, 223, 224, 226, 370
Kalanga (S16), 84, 505
Kalenjin [Southern Nilotic], 149
Kamba (E55), 37, 615, 629
Kamba (H112A), 552, 553, 573
Kamuku, *see* tiCind
Kanincin (L53), 354
Kaningi Nord (B602), 76, 91
Kanyok (L32), 359, 616
Kaonde (L41), 610
Katloid [Kordofanian subgroup],
267
Keaka [Ekoid], 259
Kela (C75), 110, 638, 642, 643, 646,
647, 649
Kele (C55; Lokele), 33, 72, 117, 520,
638, 644, 690
Kele Yawembe (C55), 688
Kemezung [Beboid], 248, 257, 272,
643
Kenswei Nsei [Ring Grassfields], 261,
272
Kenyang [Nyang], 243, 249, 510, 685,
708, 713, 719
Kerebe (JE24), 631–633
Kete (L21), 359
Kgatla (S31), 433
Khayo (JE341), 284
Ki, *see* Tuki (A601)
Kibembe, *see* Bembe (H11)
Kikamba, *see* Kamba (H112A)
Kikongo, *see* Kongo (H16)
Kikongo Language Cluster (KLC),
83, 84, 354, 355, 359, 552
Kikongo ya Leta, *see* Kituba (H10A)
Kikuyu, *see* Kuyu (E51)
Kilimanjaro (E60 languages and
Taita E74), 37
Kiluba, *see* Luba-Katanga (L33)
Kimanyanga, *see* Manyanga (H16b)
Kimbala, *see* Mbala (H41)
Kimbundu, *see* Mbundu (H21)
Kinande, *see* Nande (JD42)
Kinga (G65), 620, 652
Kintandu, *see* Ntandu (H16g)
Kinyamwezi, *see* Nyamwezi (F22)
Kinyarwanda, *see* Rwanda (JD61)
Kirundi, *see* Rundi (JD62)
Kisa (JE32D), 691
Kisetla (G40C; Settler Swahili), 610
Kisi (G67), 616
Kisi [Atlantic], 367
Kisikongo, *see* Sikongo (H16a)
Kisolongo, *see* Solongo (H16aM)

- Kisoonde, *see* Soonde (H321)
 Kisuku, *see* Suku (H32)
 Kisundi, *see* Sundi (H131K)
 Kiswahili, *see* Swahili (G42d)
 Kitalinga, *see* Talinga (JE102)
 Kitsootso, *see* Tsootso (H16hZ)
 Kituba (H10A; Kikongo ya Leta), 538, 547, 573
 Kiwoyo, *see* Woyo (H16dK)
 Kiyaka, *see* Yaka (H31)
 Kiyombe, *see* Yombe (H16c)
 Kizombo, *see* Zombo (H16hK)
 Kizu (JE402; Ikizu), 621, 631
 KLC, *see* Kikongo Language Cluster
 Koko, *see* Bakoko (A43b)
 Kol (A832), 6, 26, 31, 44, 370
 Kole (A231), 5, 42
 Kom [Ring Grassfields], 128, 137, 142, 150, 156, 158, 161, 164, 243, 261, 272, 643
 Kombe (A33b), 5, 42
 Komo (D23), 669, 683, 691
 Konda (C61E; Lokonda), 517, 722, 723
 Konde (S54), 691
 Kongo (H16; Kikongo), vi, xv, xvi, 38, 60, 73, 85, 90, 116, 325, 547, 551–557, 562, 563, 566, 567, 573, 601, 652
 Konzo (JD41), 517
 Koonzime (A842), 6, 31, 44, 409, 638, 644
 Kordofanian [Niger-Congo subgroup], 111, 118, 236, 269, 271, 317, 510, 517
 Koshin [Yemne-Kimbi], 142, 257, 261, 272
 Kota (B25), 69, 110, 112, 185, 186, 212–214, 221, 363, 366, 629, 640
 Koyo Ehamba (C24), 688
 Kpa (A53; Kpāʔ, Bafia, Rikpa, Pe), 5, 12, 14–17, 20, 21, 23–25, 30, 34, 36, 37, 43, 79, 110, 112, 120, 142, 150, 176, 185, 186, 196–198, 220, 224, 226, 370, 406, 638, 644
 Kpāʔ, *see* Kpa (A53)
 Kpe (A22; Mokpwe, Mokpe, Bakweri), 5, 19, 21, 30, 36, 42, 72, 86, 110, 113, 146, 150, 181, 184, 186, 189–191, 193, 211, 214, 221–223, 364, 370, 447, 448, 453, 615, 627, 640
 Kru [Niger-Congo subgroup], 111
 Kuche [Plateau], 487
 Kuk [Ring Grassfields], 261, 272
 Kukuya (B77a), 690
 Kulung [Jarawan], 90, 91, 265, 272
 Kundu (A122; Lukundu, Bakundu), 5, 30, 42, 79, 86, 358
 Kung [Ring Grassfields], 261, 272
 Kuria (JE43), 86, 433, 452, 556–558, 574
 Kuteb [Jukunoid], 487
 Kuyu (E51; Kikuyu, Gikuyu), 91, 439, 440, 513, 547, 557, 558, 561, 573, 616, 619
 Kwa [Niger-Congo subgroup], xx, 10, 111, 223, 240, 316, 319, 394, 510
 Kwa' [Bamileke], 261, 272
 Kwaja [North-Eastern Grassfields], 261, 272
 Kwakum (A91), 6, 7, 9, 14, 16–18, 21, 25, 27, 31, 34, 35, 44, 110, 142, 146, 150, 401, 402, 501, 508, 521, 522, 629, 638–641, 644
 Kwambi (R23), 89

Language index

- Kwangali (K33), 183, 559, 574, 617, 620, 629
Kwange (D102), 690
Kwanja [Mambiloid], 243, 252, 272, 643
Kwanyama (R21), 73, 74, 354–356, 425, 432, 437, 438, 452, 455, 621
Kwasio (A81; Mvumbo), 6, 9, 19, 20, 31, 43, 79, 186
Kwaya (JE251), 610
Kwese, *see* Kwezo (L13)
Kwezo (L13; Kwese), 353, 354, 359, 620
- Laadi (H16f; Cilaadi), 83
Labir [Jarawan], 265, 272
Laimbue [Ring Grassfields], 261, 272
Lamba (M54), 359, 616
Lame [Jarawan], 265, 272
Lamnso? [Ring Grassfields], 243, 249, 261–264, 266, 269–272, 372
Lamogi (JE17; Lulamoogi), 284
Latin [Indo-European], vi, 67, 106, 148, 319
Lefo, *see* Bafo (A141)
Lega (D25), 112, 282, 284, 285, 356, 409, 476, 486, 501, 502, 512, 590, 616, 627
Leke (C14), 443, 635, 636, 638, 644, 721
Lelemi [Kwa], 510
Lengola (D12), 69
Lenje (M61), 75, 90
Leti (an A60 language), 6
Lezgian [North-East Caucasian], 319
Libobi (C412), 86, 686, 692, 696
Legendza (C414), 694, 695
- Liko (D201), 584, 585, 594, 634, 636, 638, 644, 652, 669, 682–685, 706, 708, 717, 720–723
Limba (A27; Malimba), 5, 42
Limum [North-Eastern Grassfields], 32, 128, 136, 137, 140, 142, 150, 156, 158, 161, 164, 243, 261, 272, 547, 548, 553, 573, 643, 702, 711, 721
Linga (C502), 629, 640
Lingala (C30B), 112, 325, 538, 547, 562, 565, 573, 586, 587, 593–595, 634–636, 638, 644, 652
Logooli (JE41), 284
Lokele, *see* Kele (C55)
Lokoko (A114), 5
Lokonda, *see* Konda (C61E)
Lombi (A41), 6, 42
Londo, *see* Lundu (A11)
Londo ba Diko (A115), 5
Louisiana Creole [French Creole], 678, 683, 684, 696
Lower Pokomo (E71B), 37, 689, 690
Lozi (K21), 36, 513, 617, 622, 629
Luba (L30 languages), 356, 439, 454, 476, 480, 486, 489
Luba-Hemba (L34), 354–356
Luba-Kasai (L31a; Ciluba, Tshiluba), 38, 80, 84, 318, 359, 360, 431, 439, 456, 616, 689, 694
Luba-Katanga (L33; Kiluba, Luba-Shaba), 80, 89, 91, 182, 359, 616
Luba-Shaba, *see* Luba-Katanga (L33)
Lucazi (K13), 354–356
Lue (A12), 5, 113
Luganda, *see* Ganda (JE15)
Luguru (G35), 496, 621, 652, 691
Lugwere, *see* Gwere (JE17)

- Lukundu, *see* Kundu (A122) 449, 512, 547, 563, 565, 571,
574, 681
- Lulamoogi, *see* Lamogi (JE17)
- Luleke (S53A), 691
- Lulua (L31b), 438, 439, 454
- Lumbu (B44; Yilumbu), 72, 83, 91
- Lumun [Kordofanian], 510
- Lunda (L52), 616, 690
- Lundu (A11; Londo), 5, 8, 9, 21, 30, 42,
69, 72, 79, 84, 110, 113, 150,
364, 690
- Lungu (M14; Cilungu), 286, 440
- Luo [Mambiloid], 252, 272
- Luo [Western Nilotic], 149
- Lus [North-Eastern Grassfields], 32
- Lusoga, *see* Soga (JE16)
- Luvale (K14), 80, 354–356, 505, 559,
574, 616
- Luyana (K31), 81
- Luyia (JE32), 284, 291
- Lwalwa (L221), 354, 355
- Ma'di [Moru-Madi, Central Su-
danic], 401
- Maande (A46; Nomaande, Mandi), 5,
9, 12, 14, 16, 18, 21, 30, 43–
45, 84, 110, 117, 118, 142, 150,
364, 450, 518, 519, 688
- Mabiha (P25), 607, 615, 616
- Macro-Sudan Belt, xxiv, 389–391,
395, 399, 401, 546, 569, 613,
637, 643–647, 649
- Maja, *see* Fa' (A51)
- Makaa (A83; Mekaa), 6, 31, 43, 79, 112,
118, 150, 181, 184–186, 205–
207, 219, 220, 227, 406, 409,
547, 573
- Makhuwa (P31; Makua, Makwa,
Makwe), 34, 85, 428, 443,
449, 512, 547, 563, 565, 571,
574, 681
- Makhuwa Ile (P31), 693
- Makhuwa Nampula (P31), 688
- Makonde
Plateau (P23), 607
- Makonde (P23; Shimakonde), xvi, 85,
286, 497, 607, 615, 616
- Makua, *see* Makhuwa (P31)
- Makwa, *see* Makhuwa (P31)
- Makwe (G402), 512
- Makwe (P231), 590, 596, 598, 607, 615,
616, 619, 620
- Makwe, *see* Makhuwa (P31)
- Malila (M24), 589, 590, 602, 609, 616,
652
- Malimba, *see* Limba (A27)
- Mama [Jarawan], 243, 265, 272
- Mambila [Mambiloid], 150, 240, 243,
252, 253, 272
- Mambiloid [Bantoid subgroup],
xxviii, 95, 105, 113, 122–124,
126–129, 131, 139, 142, 145,
147–151, 156, 158, 161, 164,
165, 236, 241–243, 248, 249,
251, 252, 254, 258, 409, 510,
709, 726
- Mambwe (M15), 361, 616
- Mamfe, *see* Nyang [Bantoid sub-
group]
- Manda (N11), 117, 603, 608, 616, 632
- Mande [Niger-Congo subgroup], 111,
236
- Mandi, *see* Maande (A46)
- Manenguba (A15; Mbo), 5, 6, 19, 20,
22–24, 29, 37, 42, 90, 113,
150, 669
- Mangisa, *see* Njowi (A63)

Language index

- Mankon [Ngemba Grassfields], 15,
32, 243, 261, 262, 269, 272
- Manta [South-West Grassfields],
243, 249, 261, 272
- Manyanga (H16b; Kimanyanga), 80,
84, 85, 91, 359, 618
- Manyika (S13), 85, 621
- Manyo, *see* Gciriku (K332)
- Maori [Austronesian], 319
- Marachi (JE342), 284, 287, 288
- Marama (JE32C), 284
- Mari [Pama-Nyungan (Australia)],
400
- Masaba (JE31), 69
- Mashami (E621B), 37
- Mashati (E623B), 34
- Matengo (N13), 428, 496, 564, 565,
571, 574, 632
- Matuumbi (P13), 428, 512, 620
- Mayan languages, 679
- Mbagani (L22; Binji), 479, 481
- Mbala (H41; Kimbala), 4, 83, 117, 618,
619, 622
- Mbam (Western Mbam A40 &
Sanaga A60), 5–7, 9, 12, 13,
15–18, 21, 24, 27, 28, 45, 124,
364, 705, 711
- Mbam-Nkam [Eastern Grassfields
subgroup], 128, 131, 132, 137,
141–143, 150, 151, 156, 158,
161, 164, 165, 261, 547, 548,
702, 705, 711, 721
- Mbat [Jarawan], 265, 272
- Mbatto [Kwa], 10, 39
- Mbay [Bongo-Bagirmi, Central Su-
danic], 542, 546
- Mbe [Bantoid], xxviii, 136, 140, 243,
244, 248, 249, 258–260, 262
- Mbede (B61), 224, 520
- Mbembe [Jukunoid], 487
- Mbesa (C51), 443, 454
- Mbə' [Ring Grassfields], 261, 272
- Mbo, *see* Manenguba (A15)
- Mboa [Jarawan], 265, 272
- Mboi [Buto], 724, 726
- Mboko (A21), 5, 42
- Mbole (D11), 90, 110, 520
- Mbonge (A121), 5, 90, 266–268, 447,
448, 453
- Mbongno [Mambiloid], 252, 272
- Mboo, *see* Mbuu (A15A)
- Mboshi (C25), 73, 86, 110, 409, 520,
521, 525, 604, 629, 640, 670,
694
- Mbu', *see* Ajumbu [Yemne-Kimbi]
- Mbudza (C36c), 110
- Mbugwe (F34), 558, 573, 634, 637–
639, 646, 647, 652
- Mbuk [Beboid], 257, 272
- Mbukushu (K333), 350, 351, 559, 560,
574, 632, 638, 639, 646
- Mbula [Jarawan; = Mbula-Bwazza],
90, 91, 243, 265, 266, 271,
272, 670, 684, 695, 697, 699,
705–709, 721, 722, 724
- Mbula-Bwazza, *see* Mbula [Jarawan]
- Mbule (A623; Mbure), 5, 13, 14, 43, 711
- Mbunda (K15), 686, 693, 696
- Mbundu (H21; Kimbundu), 89, 117,
354, 617, 624
- Mbure, *see* Mbule (A623)
- Mbuu (A15A; Mboo), 5, 30, 91, 113
- Mbuun (B87), 110, 312, 320, 325, 406,
496, 507, 512, 594, 595, 634,
635, 637–639, 645, 646, 652
- Medumba [Bamileke], 142, 261, 272,
487, 510
- Mekaa, *see* Makaa (A83)

- Meke (A75C), 699, 700
Menchum [Bantoid], 242, 243, 249, 261, 272
Mendankwe-Nken [Ngemba Grassfields], 261, 272
Mengaka [Bamileke], 127, 128, 131, 138, 140, 141, 150, 156, 158, 161, 164, 261, 272
Mengisa, *see* Njowi (A63)
Menka [South-West Grassfields], 261, 272
Mesoamerican languages, 328
Meta' [Momo Grassfields], 249, 261, 272
Metta [Bantoid], 121
Mfinu (B83), 91
Mfumte [North-Eastern Grassfields], 127, 128, 136, 140, 142, 150, 156, 158, 161, 164, 165, 243, 249, 261, 272
Migili [Plateau], 487
Missong [Yemne-Kimbi], 257
Mituku (D13), 112, 116, 117, 472, 476, 627
Mkaa (A15C), 6, 21–23, 30, 42, 72, 86, 669, 688
Mmala (A62B), 5, 13, 14, 43, 45, 635, 638, 644
Mmen [Ring Grassfields], 142, 243, 261, 272, 702
Moghamo [Momo Grassfields], 243
Mojeño Trinitario [Arawak], 328
Mokpe, *see* Kpe (A22)
Mokpwe, *see* Kpe (A22)
Momo Grassfields, 124, 126, 128, 131, 135, 136, 142, 148–150, 156, 158, 161, 164, 242, 249, 260, 261, 272, 719
Mongo (A261; Mungo), 5, 42
Mongo (C61; Nkundo, Mongo-Nkundo), 22, 73, 86, 90, 325, 356–358, 360, 453, 638, 644, 670, 681, 686, 689, 691, 693, 694, 696, 699, 704, 705, 713, 722–724
Mongo-Linja (C61L), 438
Mongo-Nkundo, *see* Mongo (C61)
Moro [Kordofanian], 510, 517
Motembo (C371), 688
Mpiemo (A86c), 6, 31, 44
Mpongwe (B11a), 72, 85, 90, 91, 143, 361, 370, 690, 691
Mpoto (N14), 117
Mufu [Yemne-Kimbi], 257
Mundabli [Yemne-Kimbi], 126, 128, 135, 142, 150, 156, 158, 161, 164, 165, 243, 249, 257, 261, 272, 487, 506, 508, 510, 522, 524, 525, 643, 644, 702, 703, 713, 714, 718
Mundani [Momo Grassfields], 128, 131, 135, 137, 142, 150, 156, 158, 161, 164, 261, 272
Mundum [Ngemba Grassfields], 261, 272
Mungaka [Nun Grassfields], 261, 272
Mungbam [Yemne-Kimbi], 127, 128, 137, 142, 150, 156, 158, 161, 164, 261, 272, 506, 508, 510, 522, 524, 525, 643
Mungo, *see* Mongo (A261)
Mungong [Beboid], 126, 128, 142, 150, 156, 158, 161, 164, 249, 257, 258, 269, 272, 487
Munken [Yemne-Kimbi], 257
Mvai (A75F), 6, 31
Mvanip [Mambiloid], 252, 272
Mvumbo, *see* Kwasio (A81)

Language index

- Mwahed (A15C), 6, 669, 688
Mwaneka (A15C), 6, 669, 688
Mwanga (M22; Namwanga), 286
Mwani (G403), 86, 670, 694
Mwera (P22), 74, 75, 80, 473, 604–607, 616, 638, 639, 646
Mwiini (G412; Chimwiini), 291
Myene (B11), 72, 110, 112–114, 184, 186, 211–213, 221, 223, 224, 516, 520
Myenge (A15B), 6, 30, 669
- Nadahup, *see* Hup
Nagumi [Jarawan], 265, 272
Naki [Beboid], 126, 257, 272
Namwanga, *see* Mwanga (M22)
Nande (JD42; Kinande, Nandi), 80, 81, 91, 284, 287–292, 294, 300, 301, 388, 398, 512, 617, 618, 685
Nandi, *see* Nande (JD42)
Narrow Bantu, *see* Bantu
Nata (JE45), 638, 646, 648, 652
NC, *see* Niger-Congo
Ncane [Beboid], 243
Nchane [Beboid], 126–128, 142, 150, 156, 158, 161, 164, 249, 257, 258, 272
Nda'nda' [Bamileke], 141, 261, 272
Ndaka (D301), 685, 690
Ndaktup [North-Eastern Grassfields], 261, 272
Ndali (M301), 616
Ndamba (G52; Chindamba), 616
Nde [Ekoid], 259
Ndebele of Zimbabwe, *see* Sindebele (S44)
Ndemli [Wider Grassfields], 113, 124, 128, 131, 135, 139, 142, 150, 156, 158, 161, 164, 243, 249, 261, 272
Ndendeule (N101), 512, 574
Ndengeleko (P11), 359, 588, 592, 593, 620
Ndengese (C81), 638, 646, 647, 649
Ndibu (H16bZ; Kindibu), 552
Ndonga (R22), 73, 74, 81, 117, 354–356, 621
Ndoró [Mambiloid?], 240, 242, 243, 252, 272
Ndumu (B63), 110, 224, 520
Ndunda [Mambiloid], 252, 272
NECB, *see* North-East Coast Bantu
Nen (A44; Tunen), 5, 9, 10, 12–14, 18, 21, 24, 29, 30, 32, 33, 38, 43–45, 79, 110, 141, 150, 176, 182, 185, 186, 195, 196, 220, 324, 325, 364, 406, 445, 446, 495, 515, 518, 594, 638, 644, 652, 685, 688
Ng'hwele (G32), 621
Ngamambo [Momo Grassfields], 261, 272
Ngambay [Central Sudanic], 644
Ngandjera (R24), 73, 74
Ngayaba, *see* Bea (A54)
Ngazidja, *see* Ngazija (G44a)
Ngazija (G44a; Ngazidja), 36, 83
Ngelema (C45), 691
Ngemba [Mbam-Nkam subgroup], 32, 128, 132, 136, 137, 142, 150, 156, 158, 161, 164, 165, 243, 249, 261, 272
Ngie [Momo Grassfields], 113, 127, 128, 131, 135, 136, 142, 150, 156, 158, 161, 164, 261, 272
Ngiemboon [Bamileke], 128, 132, 133, 137, 141, 143–145, 150, 156,

- 158, 161, 164, 243, 249, 260–263, 269, 270, 272, 547, 573
- Ngindo (P14), 620
- Ngiri (C30 languages), 22, 688
- Ngolo (A111), 5
- Ngom (B22b), 33, 72, 79, 80, 90–92
- Ngomba [Bamileke], 128, 137, 138, 141, 150, 156, 158, 161, 164, 243, 261, 272
- Ngombale [Bamileke], 142, 261, 272
- Ngombe (C41), 92, 95, 112, 669
- Ngoni (N121; Ngoni of Malawi), 610
- Ngoni (N122; Ngoni of Mozambique), 347–351
- Ngoni (N12; Ngoni of Tanzania), 574, 620, 629, 652
- Ngoni of Malawi, *see* Ngoni (N121)
- Ngoni of Mozambique, *see* Ngoni (N122)
- Ngoni of Tanzania, *see* Ngoni (N12)
- Ngoreme (JE401), 558, 574, 638, 646
- Ngoshie [Momo Grassfields], 261, 272
- Ngun [Yemne-Kimbi], 257
- Ngungwel (B72a), 629, 652
- Nguni (S40 languages), 428, 479, 569, 593
- Ngwe [Bamileke], 142, 261, 272, 547, 549, 573
- Ngwi (B861), 180
- Ngwo [Momo Grassfields], 261, 272, 719
- Niger-Congo (NC), v, xix–xxi, xxiv, xxv, xxvii, xxix, xxx, xxxiii, 88, 109, 111, 113, 114, 117–120, 136, 147, 153, 154, 176–180, 202, 227, 236, 240, 267, 269, 310, 316, 317, 322–324, 334, 344, 345, 357, 361, 367, 369–373, 389, 393–396, 406, 410, 427, 488, 489, 498, 499, 506, 510, 524, 546, 567, 703, 716, 720, 723, 725
- Niger-Congo-Kordofanian, 317
- Niger-Kordofanian, 317, 318
- Nilamba (F31; Nilyamba), 38, 91, 361, 616
- Nilotic [Eastern Sudanic, Nilo-Saharan], 130, 148, 149, 151, 647
- Nilyamba, *see* Nilamba (F31)
- Nizaa [Mambiloid], 249, 252–254, 269, 272, 709, 726
- Njem (A84; Njyem), 6, 31, 44, 91, 110, 150
- Njen [Momo Grassfields], 261, 272
- Njerep [Mambiloid], 252, 272
- Njowi (A63; Mangisa, Mengisa), 6, 25, 27, 31, 43
- Njyem, *see* Njem (A84)
- Nkambe [Mbam-Nkam subgroup], 243, 249, 261, 272, 702
- Nkem [Ekoid], 259
- Nkhotakota (N31F), 295
- Nkomi (B11e), 72, 110
- Nkongho (A151), 5, 42
- Nkore (JE13; Runyankore), 427, 605, 630
- Nkore-Kiga (JE13/14; Runyankore-Rukiga), 284, 475, 476, 479, 551, 574
- Nkot [North-Eastern Grassfields], 32
- Nkucu (C73), 629, 640
- Nkucu Wela (C73), 688
- Nkum [Ekoid], 259
- Nkuna (S53D), 691

Language index

- Nkundo, *see* Mongo (C61)
Nnam [Ekoid], 259
Nnenong (A15C), 6
Noho (A32a; Banoho), 5, 29, 42, 364, 688
Nomaande, *see* Maande (A46)
non-Bantoid, 499
non-Bantu, xxxiv, 4, 147, 405, 410, 499, 637, 647–649
non-interlacustrine Bantu, 284
Noni, *see* Noone [Beboid]
Noone [Beboid; = Noni, Nooni], 126, 128, 131, 136, 139, 142, 145, 150, 156, 158, 161, 164, 243, 248, 249, 256–258, 266, 271, 272, 372, 487, 705
Nooni, *see* Noone [Beboid]
North Boma, *see* Boma (B82)
North Kogo, *see* Bakoko (A43b)
North-East Coast Bantu (NECB), 81, 83, 362, 441, 442, 457
North-Western Bantu (NWB), xviii–xxi, xxv, xxvi, xxviii, xxx, xxxi, xxxiii, 33, 62, 63, 65, 72, 107–109, 114, 117, 141, 153, 324, 325, 328, 358, 360–366, 370, 372, 374, 430–432, 434, 438, 443–451, 453–457, 496, 500, 507, 510, 511, 518, 524, 613, 614, 634, 636–640, 642, 647, 648, 652, 653, 692
Northern Bantoid, 123, 124, 126, 487, 705, 718
Northern Sotho (S32; Sesotho sa Leboa), xv, xvi, 85, 504, 689
Nrebele (S407), 621
Nsari [Beboid], 257, 272
Nsele [Ekoid], 259
Nsenga (N41), 512, 616
Nsong (B85d), 538, 573
Nta [Ekoid], 259
Ntandu (H16g; Kintandu), 359, 563, 565–567
Ntcheu (N31E), 295
Ntomba (C35a), 113
Ntomba (C61J), 629, 640
Ntomba-Bikoro (C35a), 719
Ntomba-Inongo (C35a), 670, 685, 688, 694
Ntomba-Njale (C35a), 719
Ntumu (A75A), 6, 31, 699, 700
Nuba Mountain [Niger-Congo subgroup], 323
Nugunu, *see* Gunu (A622)
Nun [Mbam-Nkam subgroup], 128, 132, 138, 142, 150, 156, 158, 161, 164, 261, 272
Nupe [Benue-Congo], 111
NWB, *see* North-Western Bantu
Nyakyusa (M31), 471, 472, 606, 616, 617, 632–634
Nyali (D33), 112
Nyambo (JE21), 312, 359, 437, 440
Nyamwezi (F22; Kinyamwezi), 350, 351, 583, 584, 586, 588, 591, 592, 604, 633, 634
Nyang [Bantoid subgroup; = Mamfe], 118, 122–126, 129, 139, 239, 243, 244, 249, 510, 685, 708, 714, 719
Nyanga (D43), 616, 627
Nyanja (N31a), 361, 616
Nyasa (N31D), 694
Nyaturu, *see* Rimi (F32)
Nyiha (M23), 361
Nyo'on, *see* Nyokon (A45)
Nyokon (A45; Nyo'on), 5, 43, 636, 638, 644

- Nyong-Dja (A80 languages), 6, 7, 18,
19, 22, 24–26, 34, 37
- Nyoro (JE11), 81, 359
- Nyungwe (N43), 361, 616
- Nzadi (B865), 73, 180, 406, 497, 506–
508, 510, 512, 514, 625, 626,
638, 639, 646, 647, 649
- Nzebi (B52), 8, 91, 110, 112, 117, 186,
218, 219, 221, 224, 555, 573,
615, 628
- Obang [Wider Grassfields], 127, 128,
131, 142, 150, 156, 158, 161,
164, 261, 272
- Obolo [Lower Cross River], 118, 140,
487
- Ogonoid [Cross River subgroup],
456
- Okak (A75B), 6
- Oko [Benue-Congo], 394, 487
- Oku [Ring Grassfields], 89, 142, 243,
261, 272
- Old Dutch, 673
- Old Egyptian, vi
- Old High German, 673
- Old Norse, 67
- Oli (A25), 5, 21, 30, 42
- Ombo (C76), 112, 282, 283, 609, 615,
619, 620, 627, 629, 640
- Oroko (A101), 5, 36, 42, 266–268, 270
- Orungu (B11b), 361, 431, 444, 450,
455, 485, 629, 640, 685, 690,
697, 704, 729
- Osatu [South-West Grassfields], 261,
272
- Oshiwambo, *see* Wambo (R20)
- Otank [Tivoid], 245, 272
- Otjiherero, *see* Herero (R30)
- Pama-Nyungan (Australia), *see* Mari
pan-Bantu, 346
- Pare (G22), 621, 633
- PB, *see* Proto-Bantu
- Pe, *see* Kpa (A53)
- PEG, *see* Proto-Eastern Grassfields
- Pende (L11), 69, 90, 691
- PG, *see* Proto-Grassfields
- Pinji (B304), 685
- Pinyin [Ngemba Grassfields], 261,
272
- Plateau [Benue-Congo subgroup],
121, 240, 241, 247, 269, 456,
487
- Plateau Tonga, *see* Tonga (M64)
- Platoid [= Kainji + Plateau +
Jukunoid], 241, 326, 406
- PNC, *see* Proto-Niger-Congo
- Poko, *see* Bapuku (A32b)
- Pokomo (E71), 36, 361
- Polish [Indo-European], 319
- Polri (A92a), 6, 18, 19, 31, 44
- Pomo (A92b), 6, 18
- Pongo (A26), 5, 42, 92
- post-PB, xx, 226, 447, 448, 454
- Potou-Tano [Kwa subgroup], 10
- pre-Bantoid, 89
- pre-Bantu, xxxi, 92, 94, 120, 127, 151,
486
- pre-Nizaa, 726
- pre-North-East Coast Bantu, 83
- pre-PB, xxxv, 111, 119, 226, 299, 450,
486, 488, 668, 687, 703, 705,
707, 711, 713–716, 720, 725,
729
- Proto-Afro-Asiatic, viii
- Proto-Atlantic, 361
- Proto-Bantoid, xxi, 89, 148

Language index

- Proto-Bantu (PB), iii, v, vi, viii–x, xiv–xxi, xxiii–xxxv, 4, 8, 10–13, 15, 24, 29, 32, 33, 35–40, 59–72, 74–84, 86, 88–94, 105–108, 111, 113, 114, 116–121, 125–127, 129, 130, 133, 135, 136, 139–143, 145–147, 151, 153, 174–181, 189, 192, 193, 202, 205, 211, 217, 218, 221, 222, 225–227, 236–238, 254, 258, 266–268, 270, 281–284, 286, 287, 289–291, 293–295, 297, 298, 301–304, 309–312, 315–318, 320–323, 325–328, 331–333, 344–346, 352, 353, 355, 357, 360, 361, 363, 367–374, 388, 391–393, 395–398, 400, 401, 403, 404, 407, 408, 411, 413–415, 423–425, 430, 445, 447–457, 465–473, 477, 482, 483, 485–488, 490, 491, 496, 498, 499, 510, 511, 513, 515, 524, 526, 537–539, 549, 570, 572, 592, 628, 636, 637, 647–652, 668, 672, 701, 704, 705, 707, 708, 710–714, 716, 718, 720, 721, 723, 725–727, 729
- Proto-Bantu-Potou-Tano, 39
- Proto-Beboid, 137
- Proto-Benue-Congo, 236, 241, 247, 367, 467, 488
- Proto-Buto, 726
- Proto-Cross River, 121
- Proto-East Benue-Congo, xxi
- Proto-Eastern Grassfields (PEG), 29, 32, 35, 41
- Proto-EBC, *see* Proto-East-Benue-Congo
- Proto-Ejagham, 139
- Proto-Germanic, 67
- Proto-Grassfields (PG), 29, 32, 41, 326
- Proto-Indo-European, 679
- Proto-Jukunoid, 121
- Proto-Kainji, 247
- Proto-Kikongo, 354–356
- Proto-Mambiloid, 147, 149
- Proto-Manenguba, 70
- Proto-Mayan, 328
- Proto-Mbam-Nkam, 727
- Proto-Niger-Congo (PNC), 39, 179, 180, 236, 316, 322, 323, 361, 367, 369, 392, 394, 398
- Proto-North-East Coast Bantu, 83
- Proto-Sabaki, 83, 94
- Proto-Sinaitic, vi
- Proto-SWB, 355, 356
- Proto-Wambo, 74
- pseudo-PB, 483
- Pulaar [Atlantic], 510
- Punu (B43; Yipunu), 72, 90, 183, 184, 186, 215–218, 221, 223, 224, 480, 573, 684, 685
- Rama [Chibchan], 319
- Rangi (F33), vii, 558, 573, 600, 616, 620, 623, 624
- Rapa Nui [Austronesian], 319
- Remi, *see* Rimi (F32)
- Rikpa, *see* Kpa (A53)
- Rimi (F32; Remi, Nyaturu), 36, 37, 117, 406, 425, 426, 431, 438
- Ring Grassfields, 124, 126, 128, 130, 131, 135–138, 142, 148–150, 156, 158, 161, 164, 242, 249, 261, 262, 269, 270, 272, 326, 688, 702, 708, 714, 724

- Rolong (S31a), 433, 455, 456
Romance languages [Indo-European], 67, 105, 148, 450, 523
Ronga (S54), 361, 621, 628, 691
Rundi (JD62; Kirundi), 90, 92, 284, 512, 604, 607, 612, 616, 629, 630, 652
Runyankore, *see* Nkore (JE13)
Runyankore-Rukiga, *see* Nkore-Kiga (JE13/14)
Russian [Indo-European], 684, 692
Rutooro, *see* Tooro (JE12)
Ruund (L53; Ruwund), 354, 453, 616
Ruvuma (P20 languages), 74
Ruwund, *see* Ruund (L53)
Rwanda (JD61; Kinyarwanda), 284, 425, 429, 434–436, 443, 455, 512, 551, 571, 574, 604, 616, 629, 630, 685, 686

Sabaki (E70 and G40 languages), 37, 83, 442
Saghala (E741), 29
Sake, *see* Shake (B251)
Salampasu (L51), 354, 685
Sama Mum [Dakoid; = Samba Daka, Chamba Daka], 248, 250, 251, 271, 272
Samba (L12a), 4
Samba Daka, *see* Sama Mum [Dakoid]
Samba Leko [Adamawa-Ubangi], 248, 644
Samia (JE34), 691
Sanaga (A60 languages), 5, 6
Sango (G61), 85
Sangu (B42), 89

Sawabantu (A10-20-30 languages), 4–6, 19, 22–27, 37
Seereer [Atlantic], 322
Seki (B21), 4, 6, 7, 17, 18, 27, 31, 34, 35, 38, 44
Semi-Bantu, 239
Semitic [Afro-Asiatic], vi, 319, 541
Sena (N44), 621
Senufo [Gur], 111
Sesotho, *see* Southern Sotho (S33)
Sesotho sa Leboa, *see* Northern Sotho (S32)
Setswana, *see* Tswana (S31)
Settler Swahili, *see* Kisetla (G40C)
Shake (B251; Sake), 72, 688
Shambaa (G23; Shambala), 37, 85, 361, 428, 437, 440, 443, 454, 621
Shambala, *see* Shambaa (G23)
Shangaji (P312), 589, 591, 597, 598, 608, 620, 630
Shanjo (K36), 75
Shi (JD53), 38, 89, 284, 369, 467, 490, 616
Shiki [Jarawan], 265, 272
Shimakonde, *see* Makonde (P23)
Shiwa (A803), 6, 19, 43, 186
Shona (S10; Chishona), vii, 82, 86, 301, 318, 392, 402, 403, 504, 512, 555, 556, 564, 565, 574, 621, 628
Shupamem [Nun Grassfields], 128, 138, 142, 150, 156, 158, 161, 164
Sikongo (H16a; Kisikongo), 83, 496
Simbiti (JE431), 558, 574
Sindebele (S44; Ndebele of Zimbabwe), 512
Sino-Tibetan, *see* Bodic, *see* Chinese

Language index

- siSwati, *see* Swati (S43)
So (A82), 43
Soga (JE16; Lusoga), 284, 290–294,
299, 300, 303, 359, 360, 512,
517, 550, 551, 574, 605, 630
Sogo, *see* Soko (C52)
Soko (C52; Sogo), 73, 76, 520
Soko-Kele (C50 languages), 112
Soli (M62), 75, 616
Solongo (H16aM; Kisolongo), 83,
562, 565
Somyev [Mambiloid], 252, 272
Songola (D24), 91
Songola Kasenga (D24), 719
Songye (L23), 354–356, 616
Soonde (H321; Kisoonde), 83
Sotho-Tswana (S30 languages), 36,
593
South Binja (D26), 603, 616, 627
South Kogo (A43c), 6
South-West Grassfields, 242, 260
South-Western Bantu (SWB), xviii,
xix, 65, 66, 73, 87, 107, 117,
154, 324, 325, 352, 354–356,
358–361, 366, 370, 372, 374,
431, 432, 438, 439, 452–457,
511, 518, 614, 638, 646, 653,
688, 690, 691, 714
Southern Bantoid, xxxiv, 123, 242,
487, 499, 506, 514, 643, 701,
702, 705, 714, 718
Southern Sotho (S33; Sesotho), 80,
86, 504, 513, 622, 623, 691
Spanish [Indo-European], 67, 319,
447
Su (A23; Isubu), 5, 42, 370
Subiya (K42), 75, 76
Sudanic, *see* Central Sudanic, *see*
Western Sudanic
Sudanic Belt, 637
Suku (H32; Kisuku), 516, 625
Sukuma (F21), 80, 409
Sumbwa (F23), 630
Sundi (H131K; Kisundi), 538, 551, 556,
573
Swahili (G42d; Kiswahili, Unguja),
vi, vii, xv, xvi, 34, 37, 69, 83,
85, 346, 348–351, 360, 363,
368, 369, 371, 423, 428, 441,
442, 449, 454, 497, 501–503,
505, 512, 514, 515, 525, 586,
598, 603, 608, 609, 621, 629,
648, 685, 720
Swati (S43; siSwati), vii, 512, 621
SWB, *see* South-Western Bantu
Taabwa (M41), 616
Taita (E74), 623, 624
Talinga (JE102; Kitalinga), 359
Tamil [Dravidian], 319
Taram [Dakoid], 243, 250, 272
Tarok [Plateau], 247, 248
Teke (B70 languages), 73
Teke d'Ibali (B71a1b), 91
Teke Laali (B73b), 690
Teke Tyee (B73d), 619, 620, 627, 652
Teke Yaa (B73c), 84, 110, 112
Tembo (JD531), 284, 286
Temne [Atlantic], 323
Tep [Mambiloid], 252, 272
Tetela (C70 languages), 670, 694
Tetela (C71), 91, 635, 638, 646, 647,
649
Tharaka (E54), 513, 557, 573
Tiba [Dakoid], 243
Tibeto-Kanauri, *see* Bodic [Sino-
Tibetan]
tiCind [a Kamuku language], 247

- Tiene (B81), 69, 73, 182, 359, 360, 629
 Tikar [Bantoid isolate], 105, 117, 122–126, 128–131, 134, 136, 139, 141, 142, 145–151, 156, 158, 161, 164, 165, 239, 241–243, 248, 249, 254, 256, 262, 272, 410, 487, 643, 705, 718, 719
 Tikuu (G41), 89
 Tima [Kordofanian], 271
 Tiriki (JE413), 284
 Tiv [Tivoid], 70, 71, 80, 89, 90, 92, 121, 240, 243, 245, 249, 272, 643
 Tivoid [Bantoid subgroup], xxviii, 121–126, 129, 236, 241, 243–245, 249, 643, 702, 714
 To'aba'ita [Austronesian], 319
 Tonga (M64; Plateau Tonga), vii, 75, 361, 615, 629
 Tonga (N15), 292, 295, 297, 303, 361, 615
 Tooro (JE12; Rutooro), 359
 Totela (K411; Totela of Namibia), 75, 559
 Totela (K41; Totela of Zambia), 75, 559, 574, 601, 625, 626
 Totela of Namibia, *see* Totela (K411)
 Totela of Zambia, *see* Totela (K41)
 Transeurasian, 415
 Tsaangi (B53), 8, 38
 Tshiluba, *see* Luba-Kasai (L31a)
 Tshivenda, *see* Venda (S21)
 Tsogo (B31), 79, 86, 361, 615, 616, 627–629, 640, 685, 688
 Tsonga (S53; Xitsonga, Changana), 69, 90, 621, 622, 652, 691
 Tsootso (H16hZ; Kitsootso), 556, 557, 601, 625, 626, 638, 639, 646
 Tsotso (JE32b), 630
 Tswa (S51), 82, 84, 691
 Tswana (S31; Setswana), 62, 70, 82, 85, 92, 315, 424, 425, 429, 432, 433, 437, 441, 443, 452, 455, 513, 622, 623, 669, 691, 695, 712
 Tuki (A601; Ki), 5, 13, 14, 27, 43, 45, 370, 518, 548, 553, 573
 Tumbuka (N21), 89, 512, 616
 Tunen, *see* Nen (A44)
 Tuotomb (A461), 5, 43
 Tura (JE32G), 284
 Twendi [Mambiloid], 252, 272
 Tyap [Plateau], 487
 Ubangi [Niger-Congo subgroup], 111, 239, 394, 510
 Ugara [Tivoid; = Ugarə], 243, 245, 272
 Ugarə, *see* Ugara [Tivoid]
 Ukaan [Benue-Congo ?], 241
 Umbundu (R11), 81, 117, 432, 437, 438, 440, 452, 455, 616
 Unguja, *see* Swahili (G42d)
 Ur-Bantu, xiv, xvi
 Uto-Aztecan, *see* Yaqui
 Ūt-Ma'in [Kainji], 135
 Venda (S21; Tshivenda), 82, 86, 500, 621
 Vengo, *see* Babungo [Ring Grassfields]
 Vidunda (G38), 621
 Vili (H12L; Civili), 76, 83, 409, 553, 554, 573
 Viti [North-Eastern Grassfields], 261, 272
 Volta-Congo [Niger-Congo subgroup], 10, 394
 Vove (B305), 110

Language index

- Vunjo-Chaga (E622C), 440, 619, 623, 624, 629
- Vute [Mambiloid], 113, 121, 126–128, 130, 131, 134, 136, 139, 141, 142, 146–150, 156, 158, 161, 164, 165, 238, 240, 243, 248, 249, 252, 254, 255, 269, 272, 445, 487, 510
- Wambo (R20; Oshiwambo), 73, 74
- Wanga (JE32a), 284
- Wawa [Mambiloid], 142, 252, 272, 487, 510
- WCB, *see* West-Coastal Bantu
- Weh [Ring Grassfields], 261, 272, 702, 703
- West African languages, 239, 299, 389, 546, 547
- West Kele (B22a), 33, 90
- West Nyala (JE18), 282, 284, 286, 287
- West-Coastal Bantu (WCB), xxv, 4, 40, 41, 65, 73, 179, 186, 324, 354, 358–360, 363, 366, 370, 372, 374, 411, 431, 511, 518, 555
- West-Western Bantu (WWB), xviii, xix, xxv, xxviii, 107, 109, 117, 154, 325, 328, 374, 431, 432, 449, 452, 454, 457, 507, 614, 638, 646, 647, 653
- Western Serengeti (JE45 languages), 594, 606, 609, 617, 634, 638, 645–647
- Western Sudanic [Nilo-Saharan], 240
- Wider Grassfields, 126, 128, 131, 135, 142, 149, 150, 156, 158, 161, 164, 260
- Wolof [Atlantic], 320, 328
- Wovia (A222), 364
- Woyo (H16dK; Kiwoyo), 83, 353, 354, 552
- Wumbvu (B24), 86, 90
- Wushi [Ring Grassfields], 261, 272
- Wuumu (B78), 91
- WWB, *see* West-Western Bantu
- Xhosa (S41; isiXhosa), 90, 91, 513, 589, 595, 607, 609, 613, 624
- Xitsonga, *see* Tsonga (S53)
- Yaka (H31; Kiyaka), 83, 431, 454, 552, 553, 618, 622
- Yamba [North-Eastern Grassfields], 243, 261, 272
- Yambasa (A62; Yambassa), 5, 9, 13, 18, 27, 72, 112, 150
- Yambassa, *see* Yambasa (A62)
- Yambeta (A462), 5, 14, 27, 30, 43, 45
- Yangben (A62A), 5, 13, 14, 43, 45, 110, 117, 711, 729
- Yanzi (B85), 110
- Yao (P21; Ciyao), 60, 74–76, 80, 84, 91, 95, 286, 359, 361, 362, 475, 512, 574, 606, 607, 619, 621
- Yaqui [Uto-Aztecan], 319
- Yasa (A33a), 5, 30, 42, 72, 86, 181, 186, 191–194, 198, 205, 220, 223, 224
- Yasanyama (a language from the Upper Tshuapa, presumably zone D or C), 723
- Yela (C74), 517
- Yemba [Bamileke; = Dschang], 128, 133, 135, 137, 140, 141, 150, 156, 158, 161, 164, 261, 262, 269, 272, 643
- Yemne-Kimbi [Bantoid subgroup], 105, 122–126, 128, 135, 137,

- 142, 145, 147–150, 156, 158,
161, 164, 165, 242–244, 249,
256, 257, 261, 272, 510
- Yeni [Mambiloid], 252, 272
- Yeyi (R41), 625, 626
- Yilumbu, *see* Lumbu (B44)
- Yipunu, *see* Punu (B43)
- Yombe (H16c; Kiyombe), 83, 117, 616
- Yoruba [Benue-Congo], 111
- Za', *see* Biya [Yemne-Kimbi]
- Zakaan, *see* Fa' (A51)
- Zali (H16cZ; Cizali), 553, 554
- Zande [Adamawa-Ubangi], 111, 118,
510
- ZeZuru (S12), 90, 91
- Zhoa [Ring Grassfields], 261, 272
- Ziba (JE22D), 436, 691
- Zombo (H16hK; Kizombo), 601, 617,
622, 634, 635, 637, 638, 646
- zone A languages, xxviii, xxix,
xxxiii, 4, 6, 8, 12, 22, 27,
62–64, 67, 69–72, 77–79, 81,
85–89, 91, 92, 107, 111–114,
116, 118, 120, 122, 126, 127,
129, 138, 142, 146, 147, 149,
151, 175, 177, 185, 186, 191,
211, 223, 225, 226, 236, 237,
239, 242, 243, 249, 262, 270,
271, 298, 324, 334, 363–365,
370, 397, 400, 404, 409, 412,
425, 426, 428, 431, 438, 446,
448–450, 473, 488, 499, 500,
507, 508, 510, 513, 516, 518,
522–525, 618, 627, 629, 631,
641, 644, 688, 691, 694, 695,
699, 707, 729
- zone B languages, xxviii, 4, 7, 8, 62–
64, 67, 69–73, 78, 79, 81, 85–
89, 91, 92, 107, 109, 111–114,
116, 118, 120, 142, 146, 147,
175, 177, 181, 185, 186, 191,
211, 222, 225, 321, 324, 329,
365, 397, 400, 404, 411, 412,
425, 426, 428, 438, 449, 473,
482, 488, 507, 508, 518, 538,
554, 557, 558, 570, 627, 629,
688, 691, 694, 712, 729
- zone C languages, 38, 63, 67, 69, 70,
72, 73, 77–79, 81, 86–88, 107,
109, 111–114, 116, 120, 147,
175, 324, 329, 365, 366, 370,
400, 404, 411, 425, 426, 428,
438, 443, 449, 451, 452, 482,
507, 518, 554, 570, 627, 629,
644, 648, 669, 682, 684, 688–
691, 694, 701, 712, 723
- zone D languages, 67, 72, 74, 77, 79,
81, 86, 88, 89, 107, 109, 111,
113, 114, 116, 120, 321, 324,
329, 356, 361, 365, 398, 400,
404, 409, 425, 426, 428, 438,
443, 449, 507, 518, 615, 627,
644, 682, 687, 691, 712, 720,
723
- zone E languages, 67, 69, 70, 74, 77–
79, 86, 88, 89, 116, 117, 321,
324, 329, 361, 398, 399, 411,
425, 438, 442, 484, 507, 516,
557, 687, 712, 720
- zone F languages, 67, 70, 74, 77, 78,
86, 88, 89, 324, 329, 351, 399,
404, 425, 438, 482, 507, 558,
712
- zone G languages, 63, 67, 69, 74, 77,
78, 86, 88, 89, 92, 116, 321,
324, 329, 365, 399, 404, 411,
438, 508, 516, 687, 720

Language index

- zone H languages, 4, 67, 73, 77, 86, 88, 89, 107, 109, 324, 329, 365, 400, 404, 409, 425, 432, 438, 449, 452, 484, 507, 508, 538, 554, 557, 558, 567, 570, 600, 601, 618, 712
- zone J languages (= Great Lakes Bantu, Interlacustrine Bantu), xxix, 67, 73, 74, 77, 78, 89, 117, 283, 284, 294, 298, 303, 315, 321, 324, 329, 351, 359, 361, 365, 369, 398, 399, 404, 432–434, 436, 438, 441, 452, 455, 484, 498, 507, 549, 551, 569, 615, 617, 630–632, 682, 689, 691, 692, 712, 721
- zone K languages, 67, 69, 73, 78, 89, 114, 117, 324, 329, 351, 365, 399, 409, 438, 507, 554, 559, 560, 618
- zone L languages, 67, 73, 77, 88, 89, 116, 117, 181, 315, 324, 329, 351, 359, 361, 365, 399, 409, 425, 430–432, 438, 507, 508, 689
- zone M languages, 67, 74, 77, 88, 89, 117, 315, 321, 324, 329, 351, 365, 399, 404, 425, 438, 507, 508
- zone N languages, 67, 71, 74, 78, 81, 88, 321, 324, 329, 365, 399, 404, 438, 484, 508, 516, 687, 720
- zone P languages, 63, 67, 71, 74, 78, 81, 88, 89, 315, 321, 324, 329, 365, 399, 404, 507, 516, 570, 687, 720
- zone R languages, 67, 73, 78, 81, 88, 89, 324, 329, 365, 399, 409, 482, 507, 508, 691
- zone S languages, 67, 69, 73, 74, 77, 78, 81, 88, 92, 116, 315, 321, 324, 329, 363, 365, 399, 409, 482, 484, 507, 508, 516, 615, 622, 628, 629
- Zulu (S42; isiZulu), xvi, 82, 83, 363, 392, 440–443, 501, 502, 512, 562, 563, 569, 571, 574, 613, 614, 624, 723

Subject index

- ablative, 256
- accretion, xxiv, xxxiv, 320, 488, 675–679, 681, 682, 686, 688, 691, 692, 694, 696, 700, 703, 706, 707, 715, 720
- accusative, 429, 692
- addressee, 321, 410, 415, 583, 683, 719
- adjective, 572, 591, 706, 711
- adnominal, 466, 467, 470, 471, 474, 475, 486, 706, 724, 725
- adposition, 237, 319
- advance verb construction, xx, xxxiii, 537, 539, 545, 560, 569
- adverb, 152, 269, 318
- affirmative, 187, 213, 216
- affix, 175, 317, 318, 323, 346, 364, 365, 399, 410, 505, 516
- affricate, 67, 70, 72, 73, 75, 78, 88, 368, 370, 727
- affrication, 19, 34
- agent, 155, 312, 353, 526
- agent noun, 549
- agent phrase, 297
- agentive, 298, 302, 303, 549
- agglutinative, xxv, xxx, xxxiii, 6, 107, 223, 237, 390, 391, 402, 413, 453, 499, 518, 519
- agreeing inversion, 603, 605, 607–609, 613, 632–634, 639, 642, 646, 651
- agreement, xxiii, xxiv, xxxii, xxxiv, 66, 226, 364, 395, 402, 452, 465–470, 474–486, 488–490, 502, 516, 551, 584, 587–589, 591, 593, 602–608, 612, 614, 627–633, 642, 650–652, 687, 694, 697, 704, 711, 717, 722, 724
- allative, xxiii, 253, 256, 319, 323
- allomorph, 79, 90, 91, 114, 139, 195, 208, 239, 260, 344, 349, 350, 360, 362, 363, 365, 366, 621, 687, 693, 706, 711, 720, 722, 729
- allomorphy, 66, 208, 251, 254, 344, 360, 363
- allophone, 29, 35, 37, 61, 76, 92, 94
- allophony, 66
- alveolar, 36, 256, 406, 727
- alveolar fricative, 73
- alveolar nasal, 404, 407
- amplexive morpheme, 715, 722
- analogical levelling, xxiii, 482–485, 707
- analogy, xxii, 34, 73, 80, 82, 83, 91, 130, 148, 177, 183, 331, 366, 552, 630
- analytic, xxxiii, 113, 117, 118, 143, 145, 400, 488, 500, 514, 515, 517–519, 523–526
- analyticity, xxv, 118, 237, 515

Subject index

- anaphoric, 444, 471, 472, 523–525, 556, 684, 688, 708, 719
- anastasis, xx, xxxiv, 637, 649
see also subject inversion
- ancestor, vi, xiv, xxi, xxii, xxvii, xxviii, xxx, 3, 10, 37, 63, 64, 113, 120, 121, 129, 132, 136, 145, 146, 321, 354, 358, 364, 366, 372, 373, 388, 393, 394, 490, 499, 524, 539, 714
- animacy, 424, 426–428, 473, 724
- animate, 89, 317, 423, 428, 437, 455, 456, 515, 715, 717
- animate goal, 319–322
- anterior, xx, 34, 116, 155, 197, 227
- anteriority, 136
- anti-causative, 247, 268
- antipassive, 353
- apical nasal, 448
- applicative, vii, xvi, xxiii, xxix–xxxix, 121, 204, 227, 237, 238, 251, 253, 266–268, 295, 296, 309–329, 331–333, 335, 344, 346, 348, 351, 362, 366, 369–371, 373, 457, 491
- applicative-like, 323
- archaic, xxiii, 179, 225, 282, 414, 448, 456, 513, 514, 637, 639, 640, 642, 647, 649
- archaic heterogeneity, xxiii, xxxiv, 66, 71, 414
- archaism, xx, xxxiv, 218, 223, 225, 637
- areal, xxiv, xxv, 30, 31, 66, 92, 114, 125, 127, 181, 191, 222, 223, 242, 270, 352, 389, 390, 393, 395, 397, 398, 401, 546, 560, 613, 637, 644–647, 649
- argument focus, 126, 134–136, 139
- argument indexing, xxiv, 389, 393, 395, 412
- aspect, xxiii, xxv, xxvii, xxx, xxxiii, 106–109, 118, 120, 127, 129, 134, 138, 140–143, 145, 146, 151, 153–155, 175, 177–180, 182, 188–190, 237, 242, 251, 326, 335, 388, 389, 457, 676, 715, 725, 727
- aspect-prominent, xxvii, 109, 121, 124, 125, 127, 129–131, 135, 141, 142, 145, 146, 148, 154
- aspectual meaning, 109, 258, 326, 331, 333
- aspectual-like, 326, 333
- aspirated, 11
- assimilation, 282, 312, 479, 728
- associative, 238, 359, 722
- asymmetric, 430, 439–441, 443, 453, 454
- atemporal, 207, 208
- attenuative, 258
- augment, xxxiv, xxxv, 79–82, 302, 471, 472, 485, 491, 526, 572, 596, 597, 620, 652, 668, 677, 686, 687, 697, 700, 701, 705–707, 715, 717, 720, 729
- autocausative, 204
- Autosegmental Phonology, vii
- auxiliary, 133, 134, 141, 145, 152, 226, 269, 392, 414, 415, 457, 521, 526, 555, 556, 558, 559, 563, 572, 698, 717, 723
- auxiliary focus, 539
- Bantu Expansion, 10, 77, 87
- Bantu Frication, 7, 19, 34, 41
- Bantu Grammatical Reconstructions (BGR), vii–ix, xii, xv, xxii,

- xxvi, 59, 61, 64, 238, 281,
465, 467, 470, 481, 482, 486,
488, 489, 491, 668, 672, 685,
688, 694–696, 703, 709, 710,
712, 714, 715, 720
- Bantu Lexical Reconstructions
(BLR), xxvi, xxvii, 8, 14–30,
35–38, 41, 44, 45, 60–64,
68–71, 73, 76–79, 84, 86–90,
93, 94, 136, 180, 205, 211,
217, 254, 284, 286, 292, 293,
298, 321, 365, 450, 668, 688,
690, 710–712, 725, 726, 728,
729
- Bantu relative agreement (BRA) cy-
cle, xxiv, xxxii, xxxv, 467,
476–482, 485–487, 490, 491
- Bantu Spirantisation, 7, 84, 94
- basic vocabulary, xviii, 106, 244, 430,
675
- benefactive, xxiii, 254, 255, 317–319,
321, 323
- beneficiary, 310, 317, 318, 320–324,
328, 348, 429, 430
- BGR, *see* Bantu Grammatical Recon-
structions
- bibliometrics, xiii
- biclausal, 695, 696, 699
- bifurcative, 258
- bilabial, 17, 89, 406, 408, 410, 447
- bilabial nasal, 396, 448
- bilabial stop, xv, 14
- bimoraic, 300, 301
- bimorphemic, 113, 117, 299
- bipartition, 402
- BLR, *see* Bantu Lexical Reconstruc-
tions
- borrowing, xxi, xxvi, 28, 37, 39, 66,
148, 246, 266, 352, 356, 675,
726, 728
- bottom-up, xiv, xxiv, xxv, 40
- bounded existence, 582–585, 596,
651
- calquing, 148
- Cameroon Volcanic Line, xxvii, 120,
124, 125, 146, 149–151, 153
- canonical, xxviii, 175, 176, 181, 188,
189, 191, 193, 198, 200, 201,
203, 205, 207, 210–212, 214,
215, 217–220, 222–227, 394,
496, 502, 512, 513, 515, 523,
551, 552, 555, 556, 558, 581,
586, 591, 594, 634, 635, 638,
645, 646, 725
- CARCP template (causative-
applicative-reciprocal-
causative-passive tem-
plate), 346, 348, 351, 354,
360, 373
- case marker, 319
- CAT sequence (causative-
applicative-transitive
sequence), 346, 373
- cataphoric, 676, 718
- cataphorically, 718
- causative, xxviii–xxx, 152, 183, 190–
192, 194, 195, 199, 204, 227,
238, 247, 249, 251, 255,
256, 258, 262, 264, 266–269,
271, 281–284, 287, 288, 290,
291, 293–297, 299–303, 319,
326, 333, 344–346, 350, 352,
354–358, 360, 362, 367–373,
429, 457, 491
- cause, 322
- cessive, 263
- chaining, 353

Subject index

- circumstance, 323
- class, 41, 90, 95, 216, 236, 374, 450, 457, 489, 491, 562, 629, 641, 650, 652, 681–683, 697, 705, 708, 718, 719, 724, 725
- class 1, 16, 69, 84, 89, 93, 187, 213, 214, 302, 423, 439, 443, 445, 446, 470, 473, 483–485, 552, 587, 621, 631, 640, 641, 673, 683, 685–687, 690, 691, 693, 697, 702–704, 706, 708, 715, 716, 718–720, 724, 725, 729
- class 1a, xxxiv, 78, 86, 668, 672, 697
- class 1/2, 69, 76, 86, 423, 428, 443, 449
- class 2, 29, 69, 408, 423, 446, 447, 473, 475, 479, 685, 686, 693, 719
- class 3, 16, 18, 69, 472, 718, 719
- class 4, 69, 472
- class 5, 38, 68, 69, 73, 78–86, 483, 685, 691, 705, 707, 708, 715, 718–720, 724
- class 5/6, 16, 69, 77, 79, 80, 83–85, 90, 711
- class 6, 69, 78, 79, 81–84, 86, 446, 708
- class 7, xxxiv, 69, 91, 668, 670, 681, 683, 687, 689, 690, 694, 697, 710–713, 718–720
- class 8, 69, 551, 642
- class 9, 27, 38, 69, 72, 79, 84, 87–91, 93, 484, 687, 713, 722, 723, 729
- class 9b, 722
- class 9/10, 38, 88, 91, 92
- class 10, 69, 79, 84, 90, 91, 93, 687, 729
- class 10b, 72
- class 11, 69, 79, 90, 91, 215
- class 11/10, 38, 75, 77, 90–92
- class 12, 472, 705, 715, 722
- class 13, 79, 472
- class 14, 69, 91, 92, 473, 549, 550, 562
- class 15, 72, 549–551, 556, 562
- class 16, xv, xvi, xxxiv, 593, 604, 606, 607, 619, 627–631, 640, 641, 668, 685, 688–690, 693, 699, 705, 710–713, 718, 720, 721, 723
- class 17, xxxiv, 593, 624, 627–629, 631, 640, 668, 690, 694
- class 18, xxxiv, 214, 556, 562, 593, 604, 625, 629, 631, 668, 690, 699, 705, 721
- class 19, 89, 683
- class 23/25, 593, 631
- class 24, 687, 722
- class agreement, 470, 719
- class marker, 89, 468, 473, 720, 724
- class prefix, 69, 90, 270, 471, 670, 686, 706, 707, 710, 718, 719, 723, 724, 727
- classification, xviii, 4, 5, 7, 106, 123, 239, 240, 242, 244, 246, 318, 366, 414, 467, 499, 508, 510, 511, 513, 524, 680, 724
- clause, xx, xxiii, xxix, xxx, xxxiv, 152, 174, 176, 181, 309, 310, 312, 313, 316, 320, 322, 327, 333, 334, 388, 400, 427–429, 441, 476, 479, 501, 521, 523, 538, 539, 552–554, 557, 676, 683, 699, 721
- clause-final, 297
- clause-initial, 500, 570, 682, 683
- cleft, xxxiv, xxxv, 541, 547, 553, 554, 668, 670, 676, 682, 683, 693, 695–703, 716, 718, 719
- cleft sentence, 503
- cleft-like, 548, 552, 557, 567, 570
- clefted *wh*-question, 504
- clitic, 299, 650, 721

- CM, *see* comparative method
 coalescence, 81, 198, 355, 356, 728
 cognacy, 28, 236, 317, 395
 cognate, xiv, xv, xvii, xxviii, 5, 22,
 23, 72, 121, 125, 135, 136, 139,
 143, 254, 258, 266, 269, 295,
 316, 326, 359, 367, 396, 398,
 408, 410, 411, 456, 471, 472,
 540, 556, 718
 cognate object construction, 540
 cohortative, 255
 colexification, 693, 694, 699, 701, 714
 collective, 353, 708
 comitative, 267, 353, 562, 569, 587,
 588, 592, 599–601, 607, 609,
 614, 617–627, 633, 652
 comitative copula, xxxiv, 600–602,
 606–609, 613–615, 617–619,
 622–624, 630, 632, 633, 636,
 647, 650, 651
 compact predicate hypothesis, 390,
 391
 comparative data, 62, 64, 143, 237,
 345, 369
 comparative evidence, x, xxiii, xxvi,
 xxx, 134, 186, 411, 693, 695
 comparative method (CM), xiv, xv,
 xvii, xviii, xxii, xxiii, xxv,
 xxix, xxxv, 3, 4, 40, 67, 106,
 152, 271, 394, 673
 comparative series (C.S.), xv, xvii, 8,
 14–28, 30, 35–38, 40, 41, 44,
 45, 63, 64, 69–71, 77, 79, 80,
 82, 84, 86–92, 315, 671
 comparative study, xxv, xxxiii, 174,
 177, 282, 354, 358, 367, 470
 compensatory lengthening, 290
 complementiser, 319, 502
 completely, 248, 254
 completeness, xxx, 310, 325, 326, 333
 completive, 253, 254, 264
 compositional, xxx, 114, 344, 345,
 347–354, 356–358, 364, 370,
 371
 concord, xxxiv, 236, 240, 514, 523,
 525, 551, 567, 593, 624, 641,
 642, 647–649, 652
 concurrent object, 430, 438, 439, 441
 conditional, 116, 117, 139, 152, 373
 conjoint, 126, 127, 134, 491, 571, 572,
 596, 597, 652
 conjunctive, 143, 152, 154, 389, 526
 connective, xxxii, 373, 471, 473, 476,
 478, 652, 715, 722–724, 729
 consecutive, 116, 117
 consonant-final, 198
 consonant-initial, 69, 198, 412, 725,
 727
 constituent order, xxxv, 525, 668,
 698, 715
 construction, xvii, xviii, xxxiv, 120,
 205, 214, 310, 311, 313–316,
 319, 320, 324, 325, 331, 335,
 397, 427, 466, 476, 477, 479,
 480, 482, 485, 486, 505, 507,
 538–540, 546, 552, 554, 557–
 559, 567, 570, 571, 583, 584,
 586, 591, 599, 605, 632, 640,
 648, 668, 670, 676, 678, 682,
 683, 685, 689–691, 695–703,
 705–709, 711, 714–716, 719,
 721–723, 727
 constructional origin, xxxv, 668, 702,
 703, 715
 constructionalisation, 587, 591, 594,
 613
 contact-induced, xxix, 270, 413, 649,
 650, 652

Subject index

- contactive, 238, 268, 344
continuant, 34, 35
continuation, 348
continuity, 351
continuous, 127, 138–140, 152, 264
converb, 572
conversion, xxxiv, 697, 701
copula, xxxv, 457, 525, 526, 572, 586, 587, 592, 593, 598, 599, 609, 614, 617, 620, 623, 627, 636–638, 640, 642, 649–652, 668, 677, 682–687, 696–703, 706, 714, 715, 717, 718, 720, 721, 723, 724
copulative, 415
coronal, 28, 29, 34, 84, 360
coronal affricate, 34
coronal fricative, 95
coronal stop, 9, 12, 34, 37
CPH (causative-passive high tone), 282–294, 297–303
C.S., *see* comparative series
cyclicity, xxiv

dative, 318, 319, 430
de-intensifier, 247
defocused, 524
degrammaticalisation, xxv
deictic, xxxv, 137, 253, 323, 332, 423, 447, 450, 668, 671, 673, 676, 677, 680, 682–685, 688, 689, 695, 696, 699, 700, 703, 705, 715, 718–722
demonstrative, 69, 373, 471, 472, 478, 479, 487, 491, 526, 572, 642, 652, 671, 676–678, 682–685, 688, 699, 706, 708, 715, 717–719, 721, 722, 724
demonstrative pronominal, 680, 683
demonstrative pronoun, 500
denasalisation, 408, 409, 446
derivation, 39, 130, 254, 288, 292, 299, 301, 312, 313, 315, 324, 344, 388, 708
derivational suffixes, xxviii, xxix, 154, 180, 182, 196, 310, 316, 324, 332, 333, 344–346, 350, 351, 365, 366, 373
desyntactisation, xxv
determiner, 652, 684, 687, 705, 707–709, 715, 720, 724, 725
detopicalised constituent, 594, 609
detransitiviser, 262
devoicing, 7
diachronic, vii, xv, xxii–xxiv, xxix, xxxiii, 6, 7, 10–12, 29, 32, 33, 39, 74, 140, 184, 188, 288, 290, 309–311, 316, 319, 326, 332, 333, 362, 373, 393, 411, 471, 473, 571, 668, 669, 673–676, 682, 701, 713
diachrony, 33, 677, 681
dialect, xviii, 27, 67, 74, 106, 133, 136, 137, 188, 243, 266, 400, 401, 669, 686, 693, 696, 698, 700, 705, 713, 722
diffusion, xxvi, 40, 129, 130, 147, 637, 644–647, 649
diminutive, 89, 227, 251, 264, 715, 722, 723
diminutiviser, 262
directionality, xxii, 66, 92, 236, 253, 255, 270, 318, 319, 723
discourse, xxix, xxxii, xxxv, 313, 318, 320, 324, 325, 327, 332, 333, 335, 402, 427, 428, 430, 434, 451, 452, 454–456, 539, 560, 588, 589, 596, 668, 676

- discourse-driven, 496, 514, 524
discourse-referential, 683, 684, 715, 718
discourse-related, 313, 316
disjoint, 126, 127, 134, 440, 457, 559, 571, 572, 596, 652
disjunctive, 118, 119, 143, 153, 154
distal, 137, 373, 652, 678, 682–684, 715, 718, 719, 721
distant future, 112, 130, 153
distant past, 112, 116, 130, 137, 153, 218
distantive, 253
distributive, 258, 262, 264
disyllabic, 136, 203, 205, 347, 502, 504, 710, 712
ditransitive, 237
divergence, xviii, xix, 64, 65, 181, 197
double reflexes, xxvi, 4, 7–13, 16, 18, 21, 26, 28, 29, 33, 35, 39, 40, 74
doublet, 69, 82, 88, 540, 541, 551, 565
doubling, xx, 263, 288, 289, 441, 542–544, 546, 559
Duke of York, 33
duration, 348, 351
durative, 138, 153, 258, 652
Dynamic Syntax, vii
echo-question, 671, 692
EL, *see* existential locational
enclitic, 298–300, 438, 439, 443, 446, 447, 457, 516, 518–520, 556, 584, 606, 607, 631
endophoric, 683, 715
epenthesis, 194, 727–729
 epenthetic consonant, 203, 204, 729
 epenthetic vowel, 205–207, 220, 621
ethical dative, 317, 319
etymology, 85, 205, 208, 217, 254, 329, 334, 617, 689, 722
ex-situ, 542, 545–547
excess, 315, 325, 326
excessive, 264, 372
exclusive, 415, 717
exclusive focus, 594, 596
existential construction, xxxiii, 586, 592, 595, 613, 630, 636, 637, 641, 647, 649, 651
existential locational (EL), xx, xxiii, xxiv, xxxiii, xxxiv, 587–589, 591–593, 596, 597, 601, 602, 604–608, 613, 614, 622, 623, 625, 638, 639, 646–652
exophoric, 584, 589
expletive, 524, 584, 587, 602–604, 614, 627–632, 637, 640, 641, 648, 651, 653
expletive inversion, 604, 605, 607, 631, 632, 639, 648, 651, 652
exponence, 106
extensive, 238, 264, 353, 359
extra-clausal, 401, 542, 569
factative, 178, 187, 188, 191
family tree, xix, xx, xxiv, 62, 64, 67, 88, 298, 345, 352, 354, 510, 511, 570
fauna, 691, 695
figure inversion, 594, 595, 602–604, 606, 613, 634, 636, 637, 640, 649, 651
final consonant, 182, 209, 210, 256, 312
final vowel, xvii, xxviii, xxix, 28, 107, 114, 153, 154, 173, 174, 176, 178, 180, 190, 191, 193, 204,

Subject index

- 205, 207, 210, 214, 219, 220,
222, 223, 225–227, 266, 282,
283, 286, 290, 314, 344, 345,
355, 356, 361, 364, 368, 369,
373, 415, 457, 491, 526, 549,
572, 653, 717, 720, 728
- finite verb, xxxiii, 402, 466, 486, 538,
539, 542, 543, 549, 550, 552,
553, 555, 559, 562, 563, 565,
566
- flora, 77, 691, 695
- focalising, 327, 328
- focus, xxvii, xxxiii, 10, 64, 107, 111,
119, 121, 125, 143, 145, 149,
151, 153, 154, 174, 177, 180,
205, 244, 314, 325, 327–329,
331, 334, 454, 457, 496, 500,
522, 523, 525, 539–547, 549–
552, 554–558, 560, 562, 563,
565–567, 571, 572, 589, 592,
595–597, 608, 633, 653, 676,
677, 681, 682, 684, 696, 697,
703, 708, 718, 721, 724
- focus-sensitive, xxxiii, 328, 552, 562,
569
- formative, 468, 491
- fortis, 10–13, 28, 32, 39
- fortition, 67, 362, 408, 409, 411
- frequentative, 258, 291, 326, 350, 352
- fricative, 9, 67, 70, 75, 81, 87, 88, 95,
299, 355, 356, 368–370, 710,
727
- fricativisation, 95
- fronting, 19, 506, 545, 555, 556, 566,
567, 569, 698, 699, 715, 728
- fusion, xxv, 74, 189, 191, 372, 373, 390,
393, 400, 402, 407, 411, 414,
650, 709, 714
- future tense, xxxiii, 111, 116, 119, 131–
134, 141–143, 147, 153, 187,
188, 212, 213, 216–218, 227,
373, 572, 676, 698
- geminatation, 82
- genealogical classification, xiv, xxi,
123, 413, 498
- general location, 310, 312–314, 320,
322, 323, 327, 334
- generic existence, 582–588, 613, 629
- generic existential, 584, 587, 591, 613
- genetic, xiv, 7, 24, 123, 127, 149, 236,
239–242, 251, 254, 256, 265,
318, 490, 675
- genitive, 473, 572, 601, 625, 692
- glide, 9, 62, 66, 67, 70, 71, 75–78, 82,
84, 87, 91, 94, 95, 183, 194,
290, 299, 412, 727
- glide formation, 69, 74, 76, 77
- glottal, 9, 266
- glottal fricative, 73
- glottal stop, 187–189
- goal, 327
- Government and Binding, vii
- grammar, v, vii–ix, xiv, xv, xxi–xxiii,
xxvi, 62, 254, 265, 325–327,
414, 427, 470, 490
- grammaticalisation, xiv, xxiii, xxv,
117, 120, 133, 151, 319, 400,
554, 555, 568
- habitual, 109, 119, 127, 138, 146, 153–
155, 256, 258, 351, 373, 415,
572, 582, 596, 653
- harmony, 11, 12, 201, 218, 349, 366
- have-verbs, 600–602, 607, 615, 623,
651
- head-marking, 514–516

- historical linguistics, x, xvii, xxvi, xxviii, xxx, 3, 241, 332, 393, 668, 671, 674
- homeland, v, vi, viii, xviii, xxv, xxviii, 333, 395, 397, 400, 404, 499, 507, 524, 546, 569, 571, 636
- homorganic nasal, 110, 139, 153, 406, 426
- hortative, 139, 153
- human goal, 310, 320, 323, 324, 333
- identificational copula, 617, 685, 723
- ideophone, 75
- illative, 253
- imperative, 72, 76, 185, 195, 196, 207, 208, 212, 213, 216, 255, 285, 406, 446, 447, 727
- imperfect, 212
- imperfective, xxvii, 108–110, 114, 119, 126, 127, 130, 131, 133, 134, 136, 138–142, 145–147, 153–155, 178, 179, 187, 188, 213, 214, 216, 218, 227, 374, 415, 491, 558, 568, 572, 615, 653, 717
- implosive, 11, 18, 33
- impositive, 238, 268, 298, 303, 344, 366, 368, 371
- in-situ, xx, xxxiii, 541, 542, 544–547, 549, 562, 565, 569, 699, 715
- inanimate, 319, 427, 428, 434, 443, 449, 452, 455, 457, 572, 653
- inceptive, 227, 254, 264
- inchoative, 264
- incompletive, 138, 153, 227
- indexing, 391, 423, 428, 429, 433, 440, 443, 454, 469, 485, 650
- inessive, 556, 572
- infinitive, xxxiii, 72, 74, 179, 180, 207–209, 211, 212, 214–216, 219, 311, 319, 373, 415, 457, 526, 538, 539, 545, 549–551, 553, 554, 556–560, 562, 569, 572, 715, 717, 722
- infinitive prefix, 73, 76, 85, 95
- infix, 200, 396, 423, 424, 468, 491, 516, 709
- inflection, xxvii, 155, 291, 562, 597, 598, 615
- information structure (IS), xxiii, 506, 512, 522–524, 538, 540–546, 555, 557, 563, 565, 566, 568, 570, 571, 573, 594, 597, 612
- initial, xx, xxvi, xxvii, 16, 17, 30, 32, 36, 37, 60, 61, 67, 70, 73, 74, 76, 79, 80, 82, 84–86, 88, 89, 93, 94, 106, 129, 133, 138, 146, 148, 151, 176, 181, 182, 185, 188, 192, 203, 206, 210, 224, 290, 345, 354, 372, 388, 401, 402, 408–411, 432, 448, 456, 468–471, 475, 479, 482–485, 491, 539–542, 550, 551, 553, 558, 560, 561, 564, 568, 570, 572, 589, 597, 676, 679, 685–687, 689, 691, 697, 701–703, 714, 718, 720, 722, 725, 729
- initial consonant, xxvii, 38, 41, 61, 62, 69, 70, 85, 110, 208, 212, 256, 406, 408, 411, 426, 449, 710, 711, 726, 727, 729
- initial glide, 67, 76
- initial nasal, 89, 406, 685
- initial position, xxxiii, 94, 224, 638, 645, 686
- initial stop, 69, 70

Subject index

- initial vowel, xxvi, 61, 62, 72, 79, 80,
84, 93, 176, 193, 207, 362,
686, 712
- innovation, xx–xxii, xxiv, xxv,
xxvii–xxxiv, 7, 21, 64–66,
71, 72, 74, 84, 89, 108, 111,
117, 120, 121, 123, 129, 130,
134, 141, 145, 149, 151, 177,
202, 206, 214, 216, 223, 226,
254, 258, 286–288, 291, 311,
316, 320, 321, 323, 325, 327,
344, 352, 355, 360, 366,
367, 369–373, 390, 395, 397,
398, 400, 403, 411, 413, 414,
449, 451, 454, 471, 476, 478,
486–488, 490, 499, 513, 524,
525, 569, 571, 630, 634, 637,
640, 642, 647–649, 652, 674,
704, 729
- insistive, 584
- instrument, 268, 320, 322, 323, 453,
457
- instrumental, 18, 267, 269, 315, 318,
321–323, 353
- intensification, 263, 264, 326
- intensifier, 247, 260, 266
- intensity, 264, 268, 315, 325, 348, 357,
372
- intensive, xxx, 247, 248, 263, 295,
303, 350, 353, 359, 360, 371,
372, 491
- intentionality, 325
- internal classification, xviii, xxi,
xxvi, 239, 311, 372, 488, 490
- interrogative construction, 677, 695,
696, 699, 701, 703–705, 711,
713, 714
- interrogative modifier, 680, 682, 686,
689–693, 695, 697, 699, 700,
702, 704–707, 709, 713, 714
- intersubjective, 683, 684, 715, 718
- intervocalic, 14, 25, 60, 361, 362, 621,
622, 711, 727
- intralingual, 612, 639, 647
- intransitive, 41, 205, 256, 260, 268,
296–298, 303, 315, 350, 351,
354, 373, 428, 521, 601
- intransitive-medial, 296
- intransitivity, 297, 351
- inverse location, 582–588, 591, 597,
645
- inversion, 284, 287, 303, 469, 496,
505, 507, 594, 595, 602, 608,
630, 637, 640, 641, 649
- inversive, 268, 296, 351
- inverted construction, xxxiv, 639,
640, 647, 648
- irrealis, 153, 415
- irregularity, xv, 40, 217, 289, 290, 705
- IS, *see* information structure
- isogloss, 64, 224, 225, 241
- iterative, 109, 119, 146, 153–155, 227,
256, 258, 260, 262–265, 269,
270, 296, 350, 351, 353
- iterativity, xxx, 310, 326, 333, 348
- itive, 108, 116–119, 137, 152, 155, 323,
332, 333
- labial, 9, 28, 34, 208, 362, 706
- labial plosive, 408
- labial-velar, 210
- labial-velar glide, 729
- labial-velar stop, 637
- language contact, 363, 514, 609, 612,
638
- lenis, 10–13, 20, 28, 32, 33, 39
- lenition, xxvi, 67, 73, 362, 729
- lexeme, 11, 61, 66, 67, 71, 92, 389

- Lexical Functional Grammar, vii
- lexical verb, 389, 393, 397–399, 402, 429, 521, 543, 545, 606, 607
- lexicalisation, 85, 335, 350, 370, 598, 606, 607
- lexicon, xiv, 3, 7, 24, 39, 62, 67, 129, 247, 248, 325, 353, 364, 672
- lexicostatistics, 6, 123, 129, 244, 669
- light verb, 541, 543, 544, 555, 563, 564
- location, xxiii, xxx, 310, 313, 317, 318, 320–323, 325, 328, 333
- locative, xv, xxxiv, xxxv, 120, 311, 313, 318, 321–323, 325, 327–329, 332, 374, 435, 491, 526, 555, 556, 572, 582, 586, 588, 591, 593, 597, 599, 601–607, 609, 612, 614, 615, 621, 624, 625, 627–634, 636–642, 646–648, 650–653, 668, 685, 687, 689–691, 693, 708, 717, 718, 722, 723
- locative agreement, xxxiv, 593, 629, 631, 641, 648, 649, 651
- locative class, 602, 603, 606, 629, 630, 648, 650, 681, 689, 693, 694, 705, 718
- locative class agreement, 604, 631
- locative copula, xxxiii, 597–602, 607, 608, 612, 614, 615, 617–619, 623, 624, 632, 633, 647, 649, 687
- locative enclitic, 584, 589, 591, 592, 598, 604, 606, 607, 612, 630–633, 639, 645, 646
- locative expletive, 605, 614, 629, 631, 641
- locative interrogative, 670, 672, 682, 685, 686, 688, 689, 691, 693, 694, 705, 715, 721, 724
- locative inversion, xxxiii, 602, 604, 605, 607–609, 613, 632, 634, 640, 646, 647, 649–651
- locative marking, 593, 625, 652
- locative noun, 593, 627, 628
- locative phrase, 312–314, 327, 335
- locative proform, 586, 587, 589, 591, 604–608, 630, 632, 635, 638
- locative system, 627, 640, 641
- locative/possessive copula, 601, 602, 607, 615
- logophoric, 709, 717, 725
- Lolemi*, ix, 470
- macrostem, 391, 392, 402, 412
- maleficiary, 317
- manner, 318, 323, 718
- marker, 110, 111, 116, 117, 133, 139, 140, 146, 152, 192, 211, 216, 258, 271, 297, 311, 317, 319, 323, 327, 347, 351, 353–356, 359, 364, 406–410, 423, 445, 457, 467, 469, 471, 472, 474, 479, 526, 551, 557, 562, 569, 584, 587, 588, 592, 599–607, 617–626, 632, 633, 638, 639, 646, 651, 670, 676, 677, 681, 683, 685–688, 694, 697, 700, 703, 711, 713, 715, 718, 719, 721, 724, 729
- mass, 707, 708
- maximality constraint, 269, 692
- Meeussen's Rule, 287–289
- Meinhof's Rule, 88, 92
- metatony, 197
- metatypy, 270
- middle suffix, xxx, 192, 345, 353, 355, 357, 358, 362, 364–366, 372
- Minimalism, vii

Subject index

- minimality constraint, 189, 692
Mirror Principle, 346, 348, 354
modality, *see* mood/modality
modifier, 313, 349, 470, 473, 502, 678,
683, 694, 699, 706–708, 718,
725
MOM, *see* multiple object marking
monoclausal, 512, 570, 695, 696, 699
monophonic, 425, 426, 432, 438–440,
449, 451
monosyllabic, 83, 182, 189, 196, 205,
447, 453, 456, 504, 621, 699
mood/modality, xxiii, xxv, xxvii,
xxx, 175, 177, 180, 182, 255,
388, 389
mora, 284–286, 288, 290, 293, 296,
301, 303
morphologisation, xxv, 208, 226
morphology, vii, xxii, xxv, xxvii,
xxviii, xxx, xxxiv, 7, 106,
107, 116, 148, 149, 151, 173,
175, 177, 179, 180, 184, 185,
190, 194, 202, 215, 219, 221,
223, 236–238, 242, 260, 270,
311, 314, 316, 320, 322, 325–
328, 331, 333, 334, 347, 348,
365–367, 371, 372, 390, 401,
470, 500, 514, 518, 519, 523–
526, 586, 588, 591, 650, 693,
697, 698, 715, 724
morphophonological, 183, 208, 372,
515
morphosyntax, xxiii, xxx, xxxv, 397,
561, 591, 668, 697, 699, 715
morphotactic, 344, 401, 402
multilingualism, 148, 246, 514
multiple object marking (MOM), xx,
xxviii, xxxi, xxxii, xxxv,
424–441, 443, 446–448,
450–457
narrative, 116, 135
narrow focus, xxx, 310, 313, 323–325,
327, 328, 333, 596, 597
nasal, xxxi, 11, 13, 20, 38, 41, 68, 69, 72,
73, 76, 87–92, 94, 138, 139,
153, 374, 685, 696, 722, 723,
729
nasal assimilation, 92
nasalisation, 208
near future, 112, 130, 153, 216, 285
near past, 116, 130, 155, 218
near-addressee, 683, 699, 715, 717
negation, 153, 174, 176, 489, 526, 591,
653
negative, 143, 174, 181, 187, 188, 190,
191, 213, 215, 216, 218, 227,
415, 446, 457, 488, 547, 572,
591, 598, 615
Neogrammarian, xv, xxvi
neuter, xxx, 238, 268, 344, 349, 350,
364, 366, 368, 371, 373, 678,
680, 684, 717
neutro-passive, 271, 296, 304, 361
no object marking (NOM), xxxi,
xxxv, 424, 426, 431, 443–
451, 453, 454, 456, 457
node, xix, xx, xxiii, xxviii, 65, 69, 76,
84, 87–89, 91, 93, 94, 108,
109, 120, 121, 177, 186, 222,
226, 310, 311, 316, 317, 320,
328, 332, 352, 356, 358, 360,
361, 366, 370, 372, 430–451,
453–456, 488, 490, 499, 513,
524, 570, 647, 649–651
NOM, *see* no object marking

- nominal, 69, 91, 206, 236, 237, 242, 253, 302, 405, 423, 444, 472, 539, 555, 557, 628, 640, 670, 672, 677, 680, 681, 685, 689, 690, 693, 694, 696, 699, 700, 706, 707, 710, 713, 714, 719, 724, 727, 729
- nominal ground, 582–584, 587–589, 591, 593, 630, 631
- nominal predicative marker, 686, 723
- nominal prefix, 468, 491, 593, 670, 672, 707, 710, 711, 723
- nominalisation, xxxiv, xxxv, 302, 303, 413, 549, 557, 570, 572, 668, 697, 701, 702, 704, 706, 707, 713, 724
- nominaliser, xxxiv, xxxv, 298, 303, 677, 686, 687, 695, 697, 700, 701, 705–707, 713, 715, 717, 719, 720, 722, 724
- nominalising morphology, 695, 702, 705, 713
- nominative, 429
- non-actor, xxx
- non-distal, 684, 700, 718
- non-finite verb, xxxiii, 540, 542–544, 546, 549, 551, 553, 561–566, 569
- non-inverted construction, xxxiv, 594, 608, 609, 613, 636, 637, 647
- non-selective interrogative pronominal (NSIP), xxiii, xxxiv, 668, 669, 678, 680, 682, 683, 685, 690, 691, 694, 699, 701, 705–707
- noun, xvi, xxxii, 41, 62, 68–70, 77, 86, 206, 236, 302, 401, 402, 470–472, 474, 477, 500, 502, 503, 505, 515, 549, 620, 670, 676, 680, 682, 694, 695, 697, 705–708, 717, 721, 723–727
- noun class, xxxv, 38, 206, 227, 228, 236, 239, 242, 246, 265, 311, 396, 398, 415, 428, 473, 487, 491, 500, 502, 527, 573, 593, 649, 650, 652, 668, 677, 683, 690, 697, 705, 707–709, 712, 713, 717, 719, 724, 725, 729
- noun phrase, xxx, xxxii, xxxiii, xxxv, 311, 335, 401, 457, 466, 467, 491, 538, 556, 561, 593
- NSIP, *see* non-selective interrogative pronominal
- numeral, 635, 637, 653, 712
- numeral prefix, 39, 468, 491
- object indexation, xxi, xxv, xxxi, 403, 406, 448, 456
- object marker, 153, 216, 255, 287, 402, 457, 500, 502, 514, 516–520, 523, 526, 556, 606, 619, 653
- object marking system, 237, 423–451, 453–457
- object prefix, xx, xxxii, 216, 227, 374, 411, 468, 491, 515, 516, 518, 607
- object pronoun, xxxi, xxxv, 153, 389, 391, 401, 516, 518
- object relative clause, 500–505
- object role, 429, 441
- object-type languages, vii
- occlusive, 210
- OM, *see* object marking system
- OM indexing, 427, 428, 442, 443, 449, 452, 454, 455

Subject index

- onset, 16, 67, 72, 73, 76, 80–84, 89, 91–93, 412, 426
- operator focus, 540, 541, 551, 552, 557, 563, 566–568
- Optimality Theory, vii
- osculant, 17, 26, 38
- palatal, 20, 208, 210, 406, 727, 729
- palatal glide, 67, 94, 710, 727
- palatal nasal, 88–92, 404, 407
- palatal stop, 75, 78, 88
- palatalisation, 8, 18, 19, 79, 363
- palato-alveolar, 20
- paradigmatic reduction, 473
- partial series (ps.), 16, 22, 25, 35, 41, 44, 77
- participant, xxv, 255, 310, 313, 317, 388, 396, 469, 684
- participial, 116, 486
- passive, xx, xxix, xxx, 35, 121, 153, 177, 183, 191–193, 204, 211, 215, 217, 221, 225, 238, 268, 281–284, 287, 288, 290–298, 301–303, 344, 345, 350, 352, 360–367, 371, 372, 374, 441–443, 453, 457, 572, 653
- passivisation, 321, 327, 430, 440–442
- past tense, 109, 111, 112, 116, 119, 127, 129–135, 137, 138, 141, 145, 147–149, 151, 153, 174, 181, 182, 184, 187, 188, 190, 195, 205, 207, 208, 210, 212–215, 227, 228, 374, 526
- path, 320, 322
- patient, 312, 649
- paucal, 264
- PCF, *see* predicate-centred focus
- perfect, xxvii, 127, 155, 205, 218, 290, 293, 294, 298, 303, 358, 466, 469, 480, 572, 615, 617, 642, 653
- perfective, 109, 114, 119, 126, 127, 130, 134–136, 138, 139, 141, 142, 145–147, 153–155, 174, 178, 179, 187–189, 191, 197, 198, 213, 217, 227, 266, 291, 299, 300, 303, 457, 491, 526, 572, 598, 653, 687, 717
- perfectivity, 136, 326
- periphrasis, 237, 555, 556
- persistence, 315, 325, 348
- persistive, 116, 118, 119, 140, 315, 572
- petrified, 266, 328
- phasal, 254, 489
- phoneme, xxvi, xxvii, 4, 8, 13, 29, 32, 33, 35–37, 39, 60, 87, 94, 95
- phonemic, xxvi, 12, 13, 88, 92, 94
- phonetic content, 33, 36
- phonological change, xv, xxii, 82, 225
- phonological mergers, 180, 310, 316, 332, 333
- phonological processes, 84, 180, 197–199, 226, 355, 402
- phonology, vii, xiv, xv, xxii, xxvii, 6, 7, 10, 62, 93, 176, 209, 237, 246, 325, 362
- phonotactic pattern, 716, 725
- phonotactics, 179, 315, 348
- phrase-final, 288–290, 295, 296
- phrase-internal, 287
- phraseme, 114, 345, 348, 349, 354, 355, 358, 360, 362, 368–372
- phraseologisation, xxx, 357, 358, 360, 368, 370, 372
- phylogenetics, xviii, 4, 64–66, 68, 71, 107, 121, 177, 222, 358, 360,

- 370, 499, 524, 528, 552, 647,
651, 714
- phylogeny, xviii, xix, 64, 106, 107,
109, 120, 124, 129, 311, 324,
352, 356, 366, 372, 397, 430–
432, 438, 443, 448, 513, 646
- pivot, 349, 350, 452, 592
- plosive, 11, 408, 409
- pluractional, xxx, 247, 251, 266, 347,
349–351, 371, 374, 653
- plurality, 253, 260, 270, 347
- plurative, 263
- polarity, xxv, xxvii, xxx, xxxiii, 175,
177, 180, 388, 389, 489, 550
- polysemy, 323, 353, 355, 600–602,
623, 625, 626, 651
- positional, 192, 204, 238, 258, 271
- positive, 143
- possessee, 599, 625, 632
- possession, 582, 600, 601
- possessive, 396, 407, 471–473, 526,
572, 597, 599–601, 619, 622,
625, 626, 651, 653, 671, 685,
715, 717, 722, 724, 725
- possessive construction, 599–601,
623–626
- possessive relative clause, 501
- possessor, 317, 413, 472, 473, 599, 601,
625, 632, 650
- post-clitic, 136
- post-copula, 702, 704, 705
- post-final, 176, 214, 388, 410, 592, 593,
607, 634
- post-infinitive, 570
- post-initial, 176, 180, 182, 388, 488,
491
- post-nasal, 91, 93
- post-radical, 239, 282, 283
- post-stem, 517
- postposing, 502, 503, 542, 564, 565,
697, 718
- postverbal, xxviii, xxxi, xxxii, 136,
154, 206, 224, 225, 258, 327,
401, 405, 424, 425, 440, 441,
443–447, 450, 454–456, 476,
478–480, 496, 497, 500, 502,
503, 505–507, 510, 515, 516,
518, 520–523, 525, 526, 559,
586, 588, 589, 594, 603, 604,
614, 699
- potential, 121, 146, 353, 356, 407, 415
- pragmatic function, 313, 315, 328,
331, 335, 523
- pre-clitic, 134
- pre-copula, 702
- pre-final, xxvii, 108, 155, 176, 182, 225,
291, 345, 351, 388, 457, 526
- pre-glottalised, 18
- pre-initial, 176, 388, 411, 468, 470, 475,
479, 486, 488, 490, 604, 606,
632, 638, 646
- pre-nasalisation, 35, 38
- pre-nasalisated stop, 7, 8, 13, 28
- pre-prefix, *see* augment
- pre-radical, 176, 388, 412
- pre-root, 288
- pre-stem, xxiii, xxvii, 107–111, 116–
119, 135, 137, 142, 143, 145–
147, 149, 151, 388, 400–402,
410, 412, 414, 467, 488, 489,
500, 516–520, 523
- precessive, 215
- predicate, xxx, xxxi, xxxiii, 313, 388,
390–392, 396, 397, 401–403,
405, 412–414, 423, 538–540,
543–545, 552, 555–557, 560,
570, 587, 589, 590, 597, 604,

Subject index

- 650, 681, 696, 700, 715, 717,
722, 723
- predicate arguments, xxxi, 412
- predicate cleft, 540, 541, 546, 557, 558
- predicate partition, 540, 542, 571
- predicate structure, xxiv, 389, 390,
392, 538
- predicate-centred focus (PCF),
xxxiii, 538, 539, 541, 543,
545, 551–553, 558–560, 562,
563, 565–568, 570–572
- predication, 391, 557, 582, 583, 597,
670, 687, 697, 699, 717
- prefix, xv, xxi, xxxi, xxxii, 11, 18, 38,
41, 69, 72, 73, 76, 78–83, 85,
87, 89–91, 117, 120, 125, 133,
187, 205, 206, 214, 355, 356,
364, 374, 401, 402, 405, 406,
411, 412, 414, 466–468, 471,
474–477, 479, 481–485, 516,
518, 551, 556, 593, 606, 628,
640, 642, 668, 673, 685, 687,
691, 693, 694, 697, 702, 706–
709, 711–713, 718, 721, 722,
727, 729
- prehistory, xvii
- preposition, 205, 237, 303, 315, 316,
320, 593, 627, 628, 640, 653,
685, 715, 722–724
- prepositional phrase, 311, 322, 325,
335
- present tense, 109, 134, 187, 212–218,
228, 526, 584, 598, 599, 616,
619, 649, 697, 717, 723
- presentational construction, 584,
628, 634
- presentational presentative, 583, 584
- preterite, 108, 110, 153, 155
- preverbal, 516
- preverbal object, 516, 518, 523
- preverbal position, 490, 586, 594,
609, 634, 638, 646, 648
- preverbal pronoun, 518
- preverbal subject, 479, 506, 508
- prevocalic, 80, 95
- proclitic, 406, 413, 414, 471, 558, 606,
620
- progressive, xxxiii, 64, 119, 120, 127,
138, 140, 141, 146, 153–155,
205, 228, 374, 415, 457, 526,
552, 555, 558–560, 568, 570,
572
- pronominal, 392, 393, 396, 397, 400,
401, 405, 407, 413, 473, 480,
524, 551, 672, 678, 683, 685–
687, 695, 699, 700, 709, 720,
723, 729
- pronominal form, 398, 407, 432
- pronominal object, xxxi, xxxv, 206,
449, 450
- pronominal prefix, xxxii, xxxv, 93,
374, 465, 466, 468, 470, 471,
474, 479, 483, 491, 672, 673,
687, 713, 716, 725, 729
- pronominal subject, 402, 481, 482
- pronominalisation, 321
- pronoun, xxv, xxx–xxxii, 85, 143,
205, 226, 228, 393, 394,
396, 397, 399–402, 405–410,
412–414, 423, 427, 457, 468,
471–473, 475, 476, 478, 480,
487–489, 491, 502, 514–516,
518–520, 522, 525, 526, 556,
572, 621, 653, 719
- proper name, 572, 691, 695
- prosodic constraint, 175, 202, 224,
225, 504
- prosodic restriction, 202, 373

- prosody, 544, 545
 proto-applicative, 329
 proto-form, 89, 355, 362, 471
 proto-language, xv, 62, 65, 111, 180,
 316, 318, 320, 327, 410, 412,
 466, 477, 488
 proto-phoneme, 4, 13, 14, 29, 33, 61,
 72, 92
 proximal, xxxiii, 415, 472, 552, 570,
 572, 678, 684, 699, 718, 719,
 721, 724
 ps., *see* partial series
 pseudo-applicative, 315, 335
 pseudo-cleft, 503, 564, 682, 701
 purpose, 319, 320, 322, 323
 purposive, 415

 radical, 74, 176, 184, 388, 424
 reanalysis, xxii–xxiv, xxviii, xxx, 12,
 24, 69, 81, 88, 89, 262, 269,
 345, 447, 539
 reason, 319, 322, 325
 recipient, 317, 318, 320–323, 328, 429,
 430
 reciprocal, xxx, 153, 192, 193, 204,
 238, 251, 258, 262–264, 266–
 269, 271, 291, 293–296, 344–
 346, 348, 350, 352–356, 358–
 360, 364, 370–372, 374
 reciprocity, 348, 353–358, 360
 reconstructed verb, 208, 323, 331
 recycling, 134, 148
 reduction, xxiv, xxviii, xxxiv, 138,
 177, 181, 184, 188, 189, 194,
 196, 202, 210, 219, 222–224,
 450–452, 479–481, 588, 591,
 642, 648, 649, 652, 672, 675–
 679, 681, 682, 685, 688, 690–
 692, 699, 702, 705
 reductive, 264
 reduplicate, 292, 294, 348, 439, 692
 reduplication, 256, 258, 260, 291–294,
 423, 682, 691, 692
 reference marker, 720
 referential classification, xix, 107,
 236, 284, 430, 601
 reflex, ix, xvi, 9, 16–20, 22, 25–30, 32–
 36, 39, 72–76, 80, 82, 85, 87,
 185, 196, 200, 211, 214, 217,
 218, 254, 311, 312, 315, 320,
 321, 327, 332–334, 353–358,
 360, 362, 367–370, 395, 408,
 410, 414, 453, 467, 479, 483,
 526, 549, 554, 615, 617, 619,
 620, 623, 691, 694, 695, 708,
 711–713, 718, 722, 727, 729
 reflexive, 24, 73, 85, 256, 262–264,
 266, 268, 269, 319, 355, 356,
 415, 425, 428, 457, 518, 709,
 715
 reflexivity, 236
 relational, 258
 Relational Grammar, vii
 relative, xx, xxi, xxxi–xxxiii, 18, 34,
 155, 181, 190, 270, 424–426,
 428, 429, 432, 437, 439, 440,
 448, 451, 467, 471, 477–479,
 481–486, 488, 489, 491, 496,
 497, 500, 502, 505, 526, 572,
 653, 697–700, 719, 722, 723
 relative agreement, xxxii
 relative clause, xxiv, xxxii, xxxv,
 465–467, 470, 473, 474, 476–
 478, 480, 483, 485, 487, 488,
 496–503, 505–507, 510–515,
 519–522, 524, 525, 668, 681,
 696, 698, 700

Subject index

- relative construction, xxiv, xxxii, 466, 477, 483
relative form, 698–700, 717, 724
relative marker, 457, 500–502
relative prefix, 474, 479, 483, 491, 697
relative verb, xxiv, xxxii, 466–470, 474–477, 480–490, 497, 501, 502, 506
relativiser, xxxii, 476–483, 485, 487, 488, 490, 491, 502–505, 514, 677, 682–685, 696, 699, 700, 717–719
relic, xxv, xxxi, 65, 85, 180, 196, 206, 323, 327–329, 359, 448, 451, 456, 699, 721
remote past, 112, 153, 491, 526
repetition, 315, 351, 552
repetitive, 238
repetitiveness, 325
restrictive focus, 549, 572
resultant state, 116
resultative, 247, 251, 258, 264, 698
retention, xxi, xxvi, xxix, xxxi, 6, 16, 38, 83, 108, 114, 118, 123, 316, 345, 356, 361, 397, 401, 404, 470, 569, 648, 650, 652, 668, 721, 723, 729
retrospective, xxvii, 111, 114, 116, 120, 121, 127, 138, 154, 155, 197, 228
reversible, 238, 260, 263, 295, 344, 351
reversible-intensive, 296
rhematic location, 582–584, 586, 587, 591, 613, 651
rigid word order, 496, 515, 608, 642, 643, 649, 652
root, xxxv, 12, 26, 60–62, 65, 69, 76, 78–82, 89, 91, 121, 130, 140, 154, 176, 179, 182, 188, 189, 194, 195, 197–199, 204–206, 209, 211, 212, 247, 269, 282, 286–288, 290–293, 296, 297, 299, 301, 303, 310–314, 320, 329, 331, 344, 346, 348–350, 356, 363, 402, 429, 468, 703, 716
root vowel, 114, 139, 199, 255, 357, 364, 366
root-final, 184, 203, 209
root-initial, 74, 75, 362, 727
Sanaga River Basin, 120, 121, 124, 146, 149–151
selective interrogative pronominal (SIP), 668, 671, 680, 683, 685, 686, 690, 691, 701, 704, 706, 707, 709, 717
semantics, xxix, 107, 111, 112, 141, 144, 175, 181, 187–189, 197, 212, 213, 236, 238, 251, 256, 262, 263, 269, 270, 313, 317, 335, 347, 351, 357, 359, 371, 568, 570, 602, 617, 629, 630, 676, 677, 680, 689, 690, 693, 704, 715
sentence-final, 682, 686, 715
sentence-initial, 594, 595, 635, 686, 697–699, 701, 713, 715
separative, 247, 253, 260, 268, 323, 332, 333, 351, 368, 371, 373
serial verb construction, 130, 133, 151, 238, 269
serial verb tone, 299
shared innovation, xvii, 362, 366, 430
shared retention, x, xvii, xxiv, xxxii–xxxiv
single object, 441, 450

- single object marking (SOM), xxxi, xxxv, 424, 426–428, 430, 431, 439–443, 447–449, 451–457
- single object role, 441, 442
- singulative, 248, 708, 715
- SIP, *see* selective interrogative pronominal
- situative, 118, 119, 140, 653
- SoA, *see* state-of-affairs focus
- sociative, 353
- SOM, *see* single object marking
- sonorant, 208, 210
- sound change, xiv, xv, xxvi, 40, 66, 67, 95, 185, 362, 406, 407, 411
- SOV (Subject Object Verb word order), 515, 554
- spatial goal, xxiii, xxx, 310, 317–324, 333, 334
- specialised EL verbs, 602, 607, 615, 619, 624, 651
- specific location, 320, 322
- speech-act participant, xxxi, 393, 394, 398, 403, 411
- spirant, 70
- spirant devoicing, 355, 356
- spirantisation, 356, 362, 363, 369, 549
- spirantise, 34, 84, 299, 370
- split predicate hypothesis, 390, 391, 650
- split predicate structure, xxv, xxx, 395, 402
- stabiliser, 299
- STAMP morpheme (subject/tense/aspect/modality/polarity morpheme), 389, 395, 399, 400, 402, 413–415
- state-of-affairs focus (SoA focus), xxxiii, 538, 540–543, 545, 547–550, 552, 554, 555, 557–559, 561, 562, 564, 565, 567, 569, 570, 573
- stative, 155, 180, 195, 238, 268, 295, 298, 303, 304, 344, 349, 573, 599, 653
- stativiser, 262
- stem, xxxiv, 15, 16, 23, 24, 28, 38, 40, 41, 69, 73, 76, 79, 81, 85, 86, 89–91, 114, 130, 133, 155, 174, 178, 182, 183, 188, 190–192, 196, 202, 203, 209, 212, 216, 225, 238, 242, 256, 282, 283, 288, 291–294, 299, 300, 315, 329, 359, 400, 402, 472, 473, 549, 621, 650, 668, 672–674, 678–680, 684, 685, 688–690, 693, 694, 699–703, 705, 706, 708, 710–714, 718–721, 724, 727
- stem-final, 176, 180, 181, 184, 201, 210
- stem-initial, xxvi, 27, 87, 296, 710, 711, 728
- strengthening, 35–38, 67, 73, 76, 82, 83, 86, 91, 92, 95, 727
- strong reflex, 19, 23, 25–29, 31, 34, 35, 38, 71, 79, 92
- sub-classification, xviii, 123, 240, 490
- subject, xxiv, xxxi, xxxii, xxxv, 8, 154, 176, 177, 201, 216, 219, 224, 264, 282, 291, 313, 314, 358, 388–391, 396, 401–403, 407, 410–413, 415, 423, 426, 428, 429, 441, 453, 466–470, 474–483, 485, 486, 488, 489, 491, 496, 497, 500–506, 510, 512, 514, 520, 521, 523, 525, 539, 551, 553, 567, 572, 586, 587, 589, 595, 599, 602–604, 614,

Subject index

- 619, 621, 623, 624, 627–629,
631, 634, 640–642, 650, 651,
699, 700, 702, 703, 706–709,
717, 720
- subject agreement, 402, 467, 480,
496, 501, 515, 586, 589, 630,
642
- subject doubling, 497, 510
- subject indexation, xxv, xxxv, 401,
403, 479, 668, 670, 703, 721
- subject inversion, xx, xxxiv, 496,
504, 506, 507, 586, 595, 606,
637
- subject marker, xxxiv, xxxv, 187, 414,
424, 457, 491, 514, 527, 584,
587, 591–593, 601–607, 612,
614, 625, 627–633, 640–642,
647, 648, 650, 651, 653, 687,
703, 716, 720
- subject prefix, xxxii, 154, 187, 189,
213, 214, 228, 374, 411, 468,
584, 686, 687, 716, 718, 720
- subject pronoun, 154, 212, 391, 395,
402, 405, 510, 650
- subject relative, 466, 470, 474–476,
478–483, 485, 489, 499, 500,
506, 507, 513
- subject relative clause, 467, 474, 475,
477, 479, 482–485
- subject role, 429
- subject-object inversion, 649
- subject-verb inversion, 502
- subject-verb word order, 519, 520
- subjunctive, 116, 117, 119, 174, 207,
208, 212, 215, 216, 228, 285,
457, 572, 586
- subordinate, 207
- substitutive, 423, 673, 688, 720
- suffix, xvi, xx, xxiii, xxvii–xxx, 114,
121, 136, 138–140, 145, 147,
155, 174, 178, 179, 183–185,
188, 190, 192–199, 201, 206–
208, 210, 212–215, 217, 218,
221, 224, 237, 253, 258, 269,
271, 287, 291, 298, 303, 309,
310, 313–318, 320–324, 328,
332, 333, 335, 343–349, 351,
353, 355, 356, 358–372, 516,
593, 602, 615, 653
- suffixal phraseme, xxx, 358, 359, 364,
365, 370
- suffixation, 209, 323, 423, 489
- suppletive, 584
- SVO (Subject Verb Object word or-
der), 495, 499, 510, 512, 514,
515, 523, 642
- syllable, 17, 41, 67, 77, 93, 179, 202,
219, 260, 282, 286–291, 293,
297, 300, 301, 303, 315, 355,
360, 425, 426, 450, 674, 677,
684, 703
- synchronic, xiv, xvii, xxii, xxiii,
xxxiii, 13, 19, 39, 74, 180–
184, 193, 194, 200, 203, 205,
211, 223, 236, 238, 239, 247,
266, 270, 288, 300, 309–311,
315, 316, 318, 321, 324, 325,
328, 329, 332, 333, 347, 348,
351, 358, 359, 425, 457, 490,
539, 569, 571, 601, 627, 651,
671, 682–684, 686, 694, 699,
704, 706, 711, 715, 718, 720
- synonym, 348
- synonymous, xvii
- synonymy, 329
- syntax, vii, xxii, xxv, xxix, 236, 317,
327, 332, 335, 347, 371, 401,

- 470, 496, 500, 504, 514, 515,
523, 525, 538, 560, 715
- synthetic, xxxiii, 113, 117, 118, 143, 145,
488, 514–516, 519, 520, 523–
525, 650
- syntheticity, 118, 237, 515
- S/A (subject/agent, as semantic role),
xxxiii, 401, 402, 415, 553,
554, 559, 561, 570
- temporal, xxxiii, 106, 112, 131, 138,
154, 187, 319, 552, 570, 722
- tense, xxiii, xxv, xxvii, xxx, 71, 74,
105–111, 119–121, 124–127,
129–134, 138, 141–143, 145–
149, 151, 152, 154, 155, 174,
175, 177, 178, 180, 182, 193,
194, 196, 388, 389, 489, 491,
527, 597, 676
- tense formulae, xxvii, 107
- tense marker, 117, 130, 140, 151, 292,
457
- tense-prominent, 109, 110, 146
- tense/aspect focus, 568
- tense/aspect marker, xxvii, 106, 107,
124, 155, 237, 247, 290, 527,
555, 556, 573, 598
- tense/aspect/mood marker, xxxiii,
154, 176, 285, 356, 447, 457,
540, 599, 650
- tense/aspect/mood/polarity marker,
xxviii, xxxi, xxxv, 180, 181,
186, 187, 194, 196, 197, 209–
211, 214, 220, 224–226, 228,
388, 393
- tentative, 238
- term focus, 539, 547, 548, 551, 555,
567
- thematic location, 582–584, 586, 591,
613
- thoroughness, 310, 325, 333
- three-syllable, 210
- time, 322
- tonal conditioning, xxvi, 12, 13, 15–
17, 32, 34
- tone, vii, xxix, 12, 13, 15–18, 20, 22, 25,
28, 32, 35, 40, 41, 92, 109, 111,
133, 136, 137, 140, 149, 153,
174, 181, 182, 189, 190, 196,
197, 205–207, 213, 214, 227,
260, 263, 281–285, 287–292,
294–304, 344, 390, 396, 450,
465, 483, 484, 486, 526, 589,
597, 599, 672, 685–687, 690,
695, 697, 699, 700, 702, 703,
705–707, 709, 713, 718, 719,
722
- tone extensions, 283, 294, 295, 297,
303
- tone plateauing, 300, 695
- tone shift, 285, 303
- tone spreading, vii, 285, 286, 301, 303
- tone-bearing unit, 286, 290, 296, 703
- top-down, xxiv, xxv
- topic, 106, 126, 177–179, 185, 191, 313,
401, 424, 427–429, 441, 454,
457, 467, 470, 496, 502, 506,
519, 523, 541, 542, 544, 546,
550, 551, 553, 560, 563, 565,
566, 572, 573, 596, 602
- topicalisation, 327, 427, 432, 700
- topicalised object, 427
- topicality, xxxii, 407, 424, 426, 427,
429, 441, 454, 457
- toponym, 691, 695
- totality, 253, 254

Subject index

- transitive, 205, 258, 268, 297, 314,
332, 346, 350, 351, 357, 371,
373, 423, 429, 521, 554, 599,
600, 623
- transitiviser, 255, 260, 270
- transitivity, 351, 426, 453
- tree, xviii–xx, xxviii, xxxiii, 20, 23,
44, 64–66, 106, 129, 358, 366,
487, 499, 514, 524, 525
- triplicate, 348
- truth focus, 541, 543, 544, 551, 552,
557, 560, 562, 563, 568
- two-syllable, 182
- typology, vii, viii, xiv, xix, xxiii–xxv,
xxxi, 239, 371, 389, 393, 450,
451, 514, 516, 519, 540, 583,
591–593, 595, 613, 668, 669,
673–675, 683, 701, 713, 715,
716
- ubiquitiser, 260, 262
- univerbation, xxxiv, 677, 695, 696,
701, 702, 706
- valence, 176, 237, 255, 260, 311, 326
- valence-changing, 255, 263–265, 269
- valence-decreasing, 255, 348, 351,
602, 653
- valence-increasing, xxix, 247, 255,
309, 310, 317, 324, 328, 333,
334, 348, 351
- valence-neutral, 348
- valence-related, 316, 326
- valency, *see* valence
- variation, viii, xiv, xvii, xxii–xxiv,
xxxii, xxxiii, 14, 35, 41, 73,
75, 76, 84, 87, 88, 95, 106, 108,
126, 127, 130, 148, 180, 181,
202, 290, 292, 319, 361, 364,
366, 368, 397, 399, 401, 411,
414, 424, 426, 428, 442, 450,
452, 467, 473, 476, 481, 485,
488, 496, 497, 499, 539, 542,
547, 549, 550, 553, 556, 560,
561, 563, 564, 566, 568–570,
582, 592, 595, 598, 601, 602,
607–609, 612, 614, 615, 619,
620, 627, 639, 641, 647, 651,
668, 669, 671, 672, 674, 679,
699, 713
- velar, 28, 208, 360, 729
- velar fricative, 411
- velar plosive, 139, 411, 412
- velar stop, 4
- ventive, 247, 323, 332, 333
- verb doubling, xxxiii, 540, 542, 549,
568, 570
- verb extensions, xxviii, 145, 177, 182,
183, 194, 201, 203, 236–239,
246–249, 251, 253–256, 258,
260, 262–271, 281–283, 294,
295, 298, 316, 317, 326, 331,
332
- verb focus, xx, 126, 134–136, 139, 542,
543
- verb form, xx, xxi, xxviii, xxx–xxxii,
119, 135, 174, 180, 184, 188,
197, 198, 208, 282, 287, 331,
361, 424, 466–468, 471, 479,
482, 485, 486, 488, 650, 697–
699
- verb morphology, xxvii, 121, 176, 180,
183, 213, 223, 253, 354, 414,
489, 514, 515, 518, 520, 524
- verb phrase, 141, 523, 542, 557
- verb position, 328, 525, 594
- verb postposing, 541, 565, 569

- verb prefix, 135, 138, 208, 226, 468–470, 474, 483, 489, 491, 503, 558, 685, 722, 723, 729
- verb preposing, 541, 546, 565, 566
- verb root, xxx, 139, 153, 175, 178, 179, 187, 200, 212, 213, 215, 254, 255, 295, 297, 310–315, 317–320, 322, 325, 326, 328, 332–334, 345, 390, 423, 426, 450, 456, 489
- verb serialisation, *see* serial verb construction
- verb stem, xvi, 72, 85, 87, 139, 144, 155, 176, 178, 179, 182, 183, 186, 188, 189, 191, 193, 194, 198, 215, 236, 256, 282, 291, 292, 294, 295, 300, 311, 315, 344, 390–393, 395, 406, 424, 432, 468, 562
- verb structure, 6, 117, 118, 178, 181, 189, 201, 202, 388, 389, 499, 543, 544, 546
- verb suffix, xxviii, 178, 180, 184, 299, 323, 333, 345, 400, 410
- verb template, xxv, 174, 344
- verb topic, 542, 543
- verb-final, xxvii, 113, 177, 179–181, 185, 190, 191, 195, 196, 205, 207, 208, 219–221, 501
- verb-initial, 60, 593, 602–604
- verbal derivation, xxviii–xxx, 114, 332, 343–347, 352, 359, 371, 372
- verbal element, 593, 597, 600, 621, 625, 627, 632, 635, 651
- verbal extension, *see* verb extensions
- verbal noun, 120, 541, 549, 550, 553, 555, 562, 573
- verbal system, 121, 135, 142, 146, 173, 174, 191, 226
- voiced, xxvi, 7, 9–14, 17, 18, 26–29, 34, 36, 38, 40, 73, 92, 355, 369, 370, 408, 411
- voiceless, xv, xxvi, 9–14, 18, 24, 26, 28, 34, 36, 40, 355, 368, 411, 727
- vowel, xxix, xxxi, xxxv, 7, 8, 10, 13, 15, 16, 18, 20, 23, 28, 36, 38, 40, 41, 67, 69, 74, 75, 77–81, 83, 89, 91, 92, 94, 95, 114, 117, 136, 138, 139, 144, 151, 153–155, 174, 176, 178, 179, 181, 183, 184, 187–198, 200–203, 205, 206, 208–211, 213–215, 217, 219, 220, 223–225, 253, 255, 263, 266, 282, 287, 290, 291, 297, 298, 312, 314, 344, 345, 349, 355–357, 361–369, 372, 374, 395, 396, 404, 406, 409–413, 447, 453, 468, 471, 479, 490, 491, 620–622, 674, 686, 688, 699, 706, 707, 710–712, 716, 717, 719, 728, 729
- vowel copy, 114, 154, 155
- vowel harmony, 180, 185, 193, 198, 199, 201, 210, 214, 215, 218, 363, 637
- vowel length, 8, 11, 12, 38, 711, 725
- vowel loss, 181, 406
- vowel quality, 409–411, 686
- vowel suffix, 195, 622
- vowel-final, 194, 195
- vowel-initial, 16, 17, 60–62, 69, 70, 76, 78–82, 84, 91, 93, 672, 694, 710, 716, 726–728
- VP focus (verb phrase focus), 544

Subject index

- weak reflex, 19–25, 35
- weakening, 24, 25, 34, 36, 37, 73, 362
- wh*-question, 325, 327, 328, 503
- word order, xxiii, xxxiii, 495, 496, 498, 499, 502, 504, 506, 510, 512–515, 519–525, 545, 556, 584, 586, 591, 593, 594, 605, 607, 608, 613, 614, 634, 636, 638, 642, 643, 645, 651, 697
- word-final, 204, 301, 470, 728
- word-initial, 727
- word-internal, 402
- words-and-things method, vii
- Y-absorption, 355

On reconstructing Proto-Bantu grammar

This book is about reconstructing the grammar of Proto-Bantu, the ancestral language at the origin of current-day Bantu languages. While Bantu is a low-level branch of Niger-Congo, the world's biggest phylum, it is still Africa's biggest language family. This edited volume attempts to retrieve the phonology, morphology and syntax used by the earliest Bantu speakers to communicate with each other, discusses methods to do so, and looks at issues raised by these academic endeavours. It is a collective effort involving a fine mix of junior and senior scholars representing several generations of expert historical-comparative Bantu research. It is the first systematic approach to Proto-Bantu grammar since Meeussen's *Bantu Grammatical Reconstructions* (1967). Based on new bodies of evidence from the last five decades, most notably from northwestern Bantu languages, this book considerably transforms our understanding of Proto-Bantu grammar and offers new methodological approaches to Bantu grammatical reconstruction.

ISBN 978-3-96110-406-2



9 783961 104062