



**HAL**  
open science

## Prehistoric Bantu-Khoisan language contact : A cross-disciplinary approach

Brigitte Pakendorf, Hilde Gunnink, Bonny Sands, Koen Bostoen

### ► To cite this version:

Brigitte Pakendorf, Hilde Gunnink, Bonny Sands, Koen Bostoen. Prehistoric Bantu-Khoisan language contact : A cross-disciplinary approach. *Language Dynamics and Change*, 2017, 7 (1), pp.1-46. hal-01960282

**HAL Id: hal-01960282**

**<https://hal.science/hal-01960282>**

Submitted on 29 Jul 2022

**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.

Pakendorf, Brigitte, Hilde Gunnink, Bonny Sands & Koen Bostoen (2017): Prehistoric Bantu-Khoisan language contact: A cross-disciplinary approach. *Language Dynamics and Change* 7/1: 1-46.

## **Prehistoric Bantu-Khoisan language contact: A cross-disciplinary approach**

### **Abstract**

Click consonants are one of the hall-marks of “Khoisan” languages of southern Africa. They are also found in some Bantu languages, where they are usually assumed to have been copied from Khoisan languages. We review the southern African Bantu languages with clicks and discuss in what way they may have obtained these unusual consonants. We draw on both linguistic data and genetic results to gain insights into the sociocultural processes that may have played a role in the prehistoric contact. Our results show that the copying of clicks accompanied large-scale in-marriage of Khoisan women into Bantu-speaking communities and took place in situations where the Khoisan communities may have had relatively high prestige. In the Kavango-Zambezi transfrontier region this must have taken place at an early stage of the Bantu immigration, possibly because small groups of food producers entering a new territory were dependent on the autochthonous communities for local knowledge.

### **Keywords:**

southern Africa; forager-farmer interactions; click consonants; molecular anthropology

## 1 Introduction

Southern Africa is characterized by the presence of two typologically distinct groups of indigenous languages: the Bantu languages and the so-called “Khoisan” languages. The Bantu languages form a close-knit genealogical unit that is widespread over sub-Saharan Africa. It is commonly assumed that this immense spread was accomplished by the gradual dispersal of expanding sedentary communities from the Grassfields region of Cameroon starting ~5,000 years ago and reaching southern Africa some 3,000 years later (Vansina, 1995: 189; Phillipson, 2005: 245ff; Blench, 2006: 126; Bostoen, 2007; Bostoen *et al.*, 2015; Grollemund *et al.*, 2015). In contrast to the closely related Bantu languages, the Khoisan languages of southern Africa belong to three separate language families, the genealogical relationship of which has yet to be demonstrated (Güldemann, 2014). These are Kx'a (Heine & Honken, 2010), Tuu (Güldemann, 2005), and Khoe-Kwadi (Güldemann, 2004; Güldemann & Elderkin, 2010) (roughly corresponding to what Greenberg (1963) called Northern, Southern, and Central Khoisan). Speakers of Kx'a and Tuu languages might be the descendants of the Late Stone Age peoples of southern Africa (Pickrell *et al.*, 2012; Schlebusch *et al.*, 2012), while the Khoe-Kwadi languages are hypothesized to have been later arrivals from eastern Africa (Güldemann, 2008). Whereas the majority of groups speaking known Khoisan languages are or were until recently foragers, some – most notably the Nama of Namibia – are pastoralists. There is archaeological evidence that pastoralists preceded agriculturalists in southern Africa by a few centuries (Mitchell, 2002; Pleurdeau *et al.*, 2012). The immigrating agriculturalists, who are generally considered to have been Bantu-speakers, can therefore be assumed to have come into contact with both hunter-gatherers and pastoralists who most probably spoke Khoisan languages (Güldemann, 2008).

One of the salient characteristics of all the languages belonging to the three Khoisan families is the presence of click consonants which generally have a high functional load; that is, they tend to occur in a high proportion of lexical items (Güldemann & Stoneking, 2008; Sands & Güldemann, 2009). Although clicks are cross-linguistically extremely rare, they are found in some Bantu languages of southern Africa in addition to the Khoisan languages. These consonants are not found in Bantu languages outside the area in which Khoisan languages are distributed. They therefore have not been reconstructed to Proto-Bantu (Meeussen, 1967: 83), and it is widely accepted that they were transferred to these Bantu languages from some Khoisan language(s) via contact (Voßen, 1997; Herbert, 2002; Maddieson, 2003; Güldemann & Stoneking, 2008; Sands & Güldemann, 2009). In this paper, we attempt to elucidate the possible prehistoric sociocultural contact situations that led to the copying of these highly rare and salient phonemes into Bantu languages by looking at both linguistic and molecular genetic data.

It should be noted from the outset that this article synthesizes data published previously: for the original results of genetic analyses of populations speaking Bantu languages with clicks readers are referred especially to Barbieri *et al.* (2013a) and Marks *et al.* (2015), while the linguistic details concerning the copying of clicks into some of the Bantu languages are discussed in depth in Gunnink *et al.* (2015). After a brief discussion of our cross-disciplinary approach in Section 1.1, we review the existence of click consonants in Bantu languages and discuss the considerable differences in size and origin of their click inventories in Section 2. These differences raise the question that we try to answer in this article, namely why some Bantu languages copied large numbers of click consonants and firmly integrated them into their phonology and lexicon, while others copied far fewer such consonants or none at all. In Section 3, we review the molecular anthropological evidence for prehistoric contact between Khoisan and Bantu-speaking populations. Interestingly, not only do Bantu-speaking populations with click languages show evidence for intimate contact with autochthonous populations, but so do populations whose languages have not incorporated clicks. In Section 4, we discuss the discrepancy between the linguistic and the genetic evidence for contact with Khoisan. We suggest two possible explanations: 1) the contact may not always have induced copying of salient click consonants, but could in some cases have led to structural changes which are more difficult to detect, and 2) the initial contact between Khoisan and Bantu languages might have led to the copying of clicks into more Bantu languages than those currently attested, with subsequent click loss taking place in some of them. In Section 5, we discuss the genetic evidence that the clicks in southeast Bantu click languages were copied in close interactions not with Khoisan foragers, but with pastoralists as hypothesized previously, and in Section 6 we discuss the genetic insights into the prehistoric contact between Khoisan peoples and the Fwe and Mbukushu, who speak southwest Bantu click languages. Section 7 concludes the paper by drawing together all the threads of the argument.

### ***1.1 Assumptions underlying our approach***

In our cross-disciplinary approach we assume that historical processes affecting peoples, such as migrations and especially contact with groups speaking different languages, shape both their genetic material (which we simplistically call “genes”) and the ways in which they speak (which we simplistically call “languages”). Whereas in humans genes are transferred strictly vertically, from parents to their offspring, languages can be transferred horizontally as well, with children learning languages not only from their parents, but also from their peers (Kerswill, 2006), resulting in multilingualism and potentially a complete shift in language. These fundamental differences between genes and languages mean that historical processes will leave different traces, so that studying each domain separately allows inferences about and insights into different aspects of the historical situation.

Combining the insights gained from genetic studies with those obtained from linguistic investigations thus has the potential of achieving a fuller understanding of the prehistoric events that shaped them (cf. Pakendorf, 2014a for a more elaborate discussion of this approach).

A fundamental problem is to know whether one is indeed comparing languages and genes that have undergone the same historical process. Since both are carried by humans who form groupings of varying size with a shared culture, language and self-identification, a common means of delimiting units of comparison is to use the self-identity of donors of genetic samples as a proxy for discrete “populations”, and to align these with the language identified as the donors’ native language. This approach is not unquestionable, since social anthropologists and historians have argued for a long time that ethnic identity is a fluid concept, especially in modern Africa, where individuals, who are generally multilingual, may affiliate themselves with different ethnolinguistic groups at different times depending on the situation or their interlocutor (Vail, 1989; Ranger, 1993; Lentz, 1995; Amselle & M'Bokolo, 1999; Tarimo, 2011). Accordingly, it cannot be taken for granted that ethnolinguistic groups are stable immutable units with one gene pool and one language to be shaped by historical processes. Nevertheless, analyses of data covering the entire human genome have shown that self-identified ethnolinguistic groups may form genetically coherent units, even in Africa (Pickrell *et al.*, 2012; Schlebusch *et al.*, 2012). If ethnolinguistic identity were merely based on the perceived interest of individuals when sampled, one would not expect that persons with the same self-identification form discrete clusters in principal components analyses. Despite the fluidity of ethnic and linguistic affiliation and the widespread multilingualism found in modern-day Africa, ancestry still plays an important role in shaping the ethnolinguistic identity and the intergenerational transmission of language, especially in the more rural parts of the continent, where most of the genetic samples considered here were collected. So even if the cross-disciplinary approach taken here is of needs overly simplistic, we consider it to be the best possible attempt for the time being to glean insights about prehistoric processes using complementary data from two different disciplines. We thereby hope to provide a solid basis for subsequent work that might take into account more complex scenarios.

## **2 Clicks in Bantu languages**

### ***2.1 Bantu languages with clicks***

Although it is fairly well-known that clicks may be used paralinguistically in many languages (including English) (Gil 2005), it is less well-known that there are a number of Bantu languages in which clicks occur as marginal phonemes, occurring in at most a handful of lexical items, often ideophones<sup>1</sup>. Clicks are marginal sounds in Kgalagadi (Dickens 1986, 29-30; Lukusa and Monaka 2008, 10; van der Merwe and Schapera 1943) and in the Ngwato variety of Tswana (Tlale 2005, 209-210). It is difficult to identify possible Khoisan sources, but some of these words may be the result of contact, such as Kgalagadi *mo-n!ú* 'big intestine' (Dickens 1986: 29), possibly copied from Naro [!ùu] 'large intestine' (Visser 2001, 166). However, marginal clicks also occur in Bantu languages outside of southern Africa, e.g. in Digo spoken on both sides of the Tanzanian-Kenyan border (Walsh 2006) and Ntandu spoken in the western DRC (Daeleman 1966, 44), as well as in non-African languages (e.g. Nathan 2001). Hence, clicks by themselves cannot be taken as evidence of contact with (southern African) Khoisan languages, contrary to widespread assumptions. Rather, they may have originated through independent innovations.

In contrast to these languages where clicks are marginal sounds, there are a number of other Bantu languages in which clicks constitute regular phonemes: they are found not only in a handful of lexemes, but in at least 60 and up to 2400 words, and clicks occur in many types of words, not just ideophones. These Bantu languages are spoken in two distinct areas of sub-Saharan Africa, namely the southeast and the southwest (Fig. 1). In the southeast, regular clicks are found in languages of the Nguni group (Xhosa, Zulu, Southern and Zimbabwean Ndebele, Swati<sup>2</sup>, and Phuthi) as well as in Southern Sotho. We here call this group of languages the “southeast Bantu” (SEB) languages. In the southwest, clicks are found in languages spoken in the Kavango-Zambezi transfrontier region of Angola, Zambia, Botswana, and Namibia, namely in Yeyi, Fwe, Mbukushu, Manyo, and Kwangali (Gunnink *et al.*, 2015). We use the term “southwest Bantu” (SWB) languages to designate them.

[Insert Fig.1 here]

Within the SWB and SEB languages, there are differences between regiolects or sociolects as to whether clicks occur as regular consonants, as marginal sounds or not at all. For example, Zimbabwean Ndebele and Ndebele spoken in what used to be the southern Transvaal (now the Gauteng and Mpumalanga provinces of South Africa) maintain clicks, while Northern Ndebele (spoken in what is now the Limpopo province) has replaced them with velar consonants (Ziervogel, 1959: 32-33). At the time of Ziervogel's research, some older speakers of the Northern variety had still retained clicks in terms for a few plant species, but nowadays clicks no longer occur in even these few items (Skhosana, 2009).

---

<sup>1</sup> Ideophones are “a vivid representation of an idea in sound” (Doke 1935: 118) that are “noted for their special sound patterns, distinct grammatical properties, and sensory meanings...” (Dingemans 2012: 654).

<sup>2</sup> Swati here includes Bhaca (Baca), which some might consider a separate language (Hallowes, 1942).

Similarly, the southernmost variety of Fwe uses clicks more frequently than the central variety, and the northernmost variety has no clicks at all (Gunnink, forthcoming).

## ***2.2 Origins of clicks in Bantu languages***

In contrast to frequent assumptions, clicks in Bantu languages cannot automatically be assumed to be the result of contact with Khoisan languages. There are three main sources of click words in Bantu languages that we have identified: 1) clicks could have been copied from other Bantu languages, 2) they could have been copied from Khoisan languages, and 3) they could have been innovated in particular lexical items in languages that have copied clicks as regular phonemes. This is exemplified by the occurrence of clicks in words of Bantu origin in the SEB and SWB languages (Herbert, 2002: 299; Gunnink *et al.*, 2015: 205) and even in copies from European languages, such as the Manyo word |umáte 'tomato', where a /t/ has been replaced with a dental click (Bostoen & Sands, 2012: 133; Gunnink *et al.*, 2015: 206).

With respect to the copying of clicks among Bantu languages, up to 50% of click words are shared between the SWB languages Manyo, Kwangali, and Mbukushu, with Manyo playing a central role in their diffusion (Gunnink *et al.*, 2015). Similarly, 376 Xhosa words with clicks (~16% of the total) also occur in Zulu and Southern Sotho (Bourquin, 1951). Clicks have also been copied from Nguni languages into other Bantu languages, though in these ultimate recipients clicks appear to be marginal, given the very limited numbers of lexical items that have been attested. Southern Sotho is an exception, since here clicks constitute regular phonemes (see Table 1). Some languages of Mozambique appear to have copied lexical items with clicks from Zulu: Tsonga (Doke, 1954: 180-182; Baumbach, 1974: 6-7, 25 33, 41, 45, Passy 1914); Ronga (Afido *et al.*, 1989: 111-118); Tswa (Persson 1932)); Ndaui (Mkanganwi, 1972; Afido *et al.*, 1989: 80-86) and Chopi<sup>3</sup> (Bailey, 1995). In Zimbabwe, clicks are found in some Midlands varieties of Karanga (Pongweni, 1990). The Mzimba variety of Tumbuka, spoken in Malawi, has clicks which appear to have been copied from either Zulu or Ngoni (Moyo, 1995). In addition, it appears that in one case, at least, a Khoisan language copied clicks from a Bantu language: the extinct !Ui language, ||Xegwi copied words with alveolar clicks from Swati, e.g. !ala<sup>4</sup> 'begin' (Lanham & Hallows, 1956: 46), after having lost inherited alveolar clicks (Sands, 2007).

Where there is evidence that click words were indeed copied from Khoisan languages, the SEB and SWB languages differ in whether these source languages were spoken by pastoralists or foragers. The SEB languages have copied significant numbers of lexical items

---

<sup>3</sup> According to Bailey (1995, 147), female and younger male speakers of Chopi tend to use velars in place of the clicks.

<sup>4</sup> In Bantu languages with an official orthography, clicks are represented with <c>, <q> and <x> respectively. In this paper, we represent all clicks, both in Khoisan and in Bantu words, with IPA symbols.



with clicks from now-extinct varieties of the Khoekhoe branch of the Khoe family which were spoken by pastoralists (e.g. Anders, 1937; Bourquin, 1951; Louw, 1977a, b). Evidence for loans into SEB from Tuu languages spoken by foragers is far more limited – possibly due to the lack of documentation of these forager languages. Languages belonging to the !Ui branch of Tuu are historically known to have been spoken in the Eastern Cape, and possible !Ui sources for certain Nguni words are attested (du Plessis, forthcoming). There are often unexplained phonological mismatches between the !Ui and SEB items, however, so that it is unclear if these are really copies. In contrast, the SWB groups are or were until recently in known contact only with foragers; these speak !Xuun, a Ju lect belonging to the Kx'a family, as well as Khwe, a language belonging to the West Kalahari branch of the Khoe-Kwadi family (Güldemann, 2014: 26). The SWB languages have adopted lexemes from various Ju lects, and also, but in smaller numbers, from Khwe, with a small number possibly stemming from Khoe languages spoken in Eastern and Central Botswana. In addition, a substantial proportion of the click words in the SWB languages cannot be traced to any known Khoisan language, suggesting that the contact may have taken place with languages that are now extinct (Gunnink *et al.*, 2015).

### ***2.3. Differences in click inventories***

The Bantu click languages differ considerably with respect to the inventory and functional load of their click phonemes (Table 1). In several Nguni languages and Yeyi, click inventory sizes are relatively large (comprising 10-30 phonemes), while Manyo, Kwangali, Mbukushu, and Fwe have only four or five distinct click phonemes. The number of words with clicks also differs considerably between the Southeast Bantu languages and Yeyi on the one hand (with up to 27% of the lexicon carrying a click consonant) and Manyo, Kwangali, Mbukushu and Fwe on the other, where only ~1% of the lexicon comprises a click (Gunnink *et al.*, 2015).

The estimates of the proportion of the lexicon carrying a click consonant are based on manual counts in dictionaries or word lists for the following languages: Zulu (Doke *et al.*, 1958), Xhosa (McLaren, 1968), Southern Ndebele (Shabangu & Swanepoel, 1989), Zimbabwean Ndebele (Pelling, 1966), Swati (Rycroft, 1981), Phuthi (Donnelly, 2007), Southern Sotho (Mabille *et al.*, 1961), Manyo (Möhlig and Shiyaka-Mberema, 2005), Kwangali (Kloppers *et al.*, 1994), and Mbukushu (Wynne, 1980). The estimate for Fwe is based on KB's and HG's fieldwork, while for Yeyi we cite the estimates provided by Sommer & Voßen (1992) (who base their estimate for Namibian Yeyi on the word list in Donnelly, 1990).

**Table 1: Click inventories and functional load of clicks in Bantu languages of southern Africa.**

Group	Language	# of click types <sup>5</sup>	# click phonemes <sup>6</sup>	Proportion of lexicon (%)
SEB (Nguni)	Zulu	3	15	22 <sup>7</sup>
	Xhosa	3	18	26.8
	Ndebele	0 (Northern)	0 (Northern)	0 (Northern)
		2 (Southern)	8 (Southern) <sup>8</sup>	6.6 (Southern)
		3 (Zimbabwean)	15 (Zimbabwean)	8.1 (Zimbabwean)
	Swati	1	4	12
Phuthi	3	12	7.7	
SEB	S. Sotho	1 <sup>9</sup>	3	3-5
SWB	Yeyi	2 (Namibia)	12 (Namibia)	10 (Namibia)
		4 (Botswana)	22 (Botswana)	15 (Botswana)
	Manyo	1	5	1.3
	Kwangali	1	5	2 <sup>10</sup>
	Mbukushu	1	4	0.8
	Fwe	1	4	~1

Sources: Zulu (Doke, 1947, Doke *et al.* 1958), Xhosa (McLaren, 1948, 1968), Ndebele (Pelling & Pelling, 1974; Bowerman & Lottridge, 2002; Skhosana, 2009), Swati (Ziervogel, 1952; Nussbaum *et al.*, 1969; Rycroft, 1981; Taljaard *et al.*, 1991), Phuthi (Donnelly, 2007), Southern Sotho (Mabilile *et al.*, 1961; Moeketsi, 1991), Namibian Yeyi (Seidel, 2008), Botswana Yeyi (Sommer & Voßen, 1992; Lukusa, 2009), Manyo (Möhlig, 1967; Möhlig & Shiyaka-Mberema, 2005), Kwangali (Dammann, 1957; Kloppers *et al.*, 1994), Mbukushu

<sup>5</sup> Click type refers to the place of articulation of the click, as well as the direction of airflow (central vs. lateral). Among five known click types, Bantu languages make use of four: dental <|>, lateral <||>, alveolar <!> and palatal <‡>. In the table, we count the number of phonemic click types found in each language, disregarding possible phonetic variation.

<sup>6</sup> Click phoneme refers to the click type together with other aspects of articulation, such as voicing, aspiration, or nasalization. Thus, /n!/ differs from /g!/ in the same way that /n/ differs from /d/, and each count as a separate consonant.

<sup>7</sup> Note that our estimates for click words in Zulu and Xhosa are much larger than previous estimates, e.g. Herbert 1990: 122 writes “It is estimated that approximately one-sixth of Xhosa words and one-seventh of Zulu words contain clicks [...]” – this would amount to 16.7% and 14.3% of click words in Xhosa and Zulu, respectively.

<sup>8</sup> This figure does not include the nasal lateral click which occurs in a few ideophones and de-ideophonic verbs (Skhosana, 2009: 54).

<sup>9</sup> Moeketsi (1991) lists a few ideophones with nasal lateral and dental clicks, but these are not included in the totals of Southern Sotho click types and phonemes.

<sup>10</sup> We find 67 words with clicks in Kwangali and Fwe and 64 in Mbukushu (Gunnink *et al.*, 2015: 198).

(Fisch, 1977; Wynne, 1980), Fwe (K. Bostoen field notes, Bostoen & Sands, 2012, H. Gunnink field notes).

It is clear from Table 1 that there are substantial differences among the SEB and SWB languages with respect to their clicks: Zulu, Xhosa (both SEB) and Yeyi (SWB) have large click inventories and these consonants occur in a large proportion of lexemes, while in Southern Sotho (SEB) and Manyo, Kwangali, Mbukushu, and Fwe (all SWB) click inventories are small and only a small proportion of the lexicon carries a click. Furthermore, within the SWB languages Manyo, Kwangali, and Mbukushu form a closely connected group: they share about a third of their click words, as mentioned above, with Manyo probably playing a central role in this intra-Bantu diffusion process (Gunnink *et al.*, 2015). Fwe stands apart from this little group: it shares only two click words with Manyo, one of which occurs in all five SWB languages.

The differences among the southern African Bantu languages are quite striking and raise the question which is at the heart of this article: why did some languages copy large numbers of click consonants and firmly integrate them into their phonology and lexicon, while others copied far fewer such consonants, and yet other languages didn't copy any clicks at all, or at most one or two lexemes? Since copied click consonants are a very particular outcome of language contact we next turn to a survey of molecular anthropological studies to elucidate the extent of contact between immigrating Bantu speech communities and autochthonous peoples speaking Khoisan languages.

### **3 Molecular anthropological evidence for population contact in southern Africa**

#### ***3.1 Khoisan genetic lineages***

Genetic contact between Bantu-speaking food producers and autochthonous foragers presumably speaking Khoisan languages can be relatively straightforwardly detected with analyses of mitochondrial DNA and the Y-chromosome. Mitochondrial DNA is passed on exclusively from mothers to their offspring, so that investigating the variation contained in mtDNA sequences permits one to trace back the maternal (pre)history of populations. The Y-chromosome, in contrast, is found only in males and is thus passed on only from fathers to their sons; this therefore offers a means of studying the paternal (pre)history of populations. Both the Y-chromosome and mtDNA are passed on largely intact from generation to generation; variation is introduced only through new mutations (Jobling & Tyler-Smith, 2003; Pakendorf & Stoneking, 2005). Such mutations can be used to define groups of molecules that descend from a common ancestor (called "haplogroups"), much as

shared innovations are used in historical linguistics to define genealogically related subgroups<sup>11</sup>.

In southern Africa, most Khoisan-speaking populations are characterized by high frequencies of specific mtDNA and Y-chromosome haplogroups, which are otherwise found only in populations that are very likely to have been in intimate contact with Khoisan-speaking groups (Wood *et al.*, 2005; Behar *et al.*, 2008; Soodyall *et al.*, 2008; Barbieri *et al.*, 2013b; Barbieri *et al.*, 2014a; Barbieri *et al.*, 2014b). The mtDNA haplogroups that are characteristic of Khoisan-speaking populations are labelled L0d and L0k, the Y-chromosomal ones are A2, A3b1, and B2b. The common ancestor of the different lineages belonging to each haplogroup can be dated with molecular methods: for L0d this is ~95,000 years, for L0k ~40,000 years (Barbieri *et al.*, 2013b), for A2 ~ 30,000 years, for A3b1 ~55,000 years, and for B2b ~60,000 years (Barbieri *et al.*, 2016). Given the old age of these mtDNA and Y-chromosome lineages and the fact that many of the Khoisan-speaking groups who until recently led a foraging way of life are genetically quite distinct from other peoples in Africa (Pickrell *et al.*, 2012; Schlebusch *et al.*, 2012), it is a reasonable assumption that these Khoisan-speaking foragers represent the descendants of the autochthonous peoples of southern Africa who were settled in the region before the immigration of food-producing peoples. Given the strong – though by no means exclusive – association of these autochthonous genetic lineages with Khoisan languages, we here assume that the autochthonous peoples the Bantu-speaking immigrants came into contact with 1,200-2,000 years ago would also have spoken Khoisan languages<sup>12</sup> with clicks. This assumption is further strengthened by the distribution of copied click consonants in the Bantu languages of southern Africa. We therefore use the label “Khoisan” interchangeably with “autochthonous” when discussing the genetic lineages.

The autochthonous haplogroups are not commonly found in Bantu-speaking populations, with the exception of those who live in the vicinity of Khoisan-speaking groups (Wood *et al.*, 2005; Batini *et al.*, 2011; de Filippo *et al.*, 2011; Barbieri *et al.*, 2013b), so that it is accepted that they were not part of the gene pool of the ancestors of the Bantu-speaking people who immigrated into southern Africa after ~2000 BP. The presence of mtDNA haplogroups L0d and/or L0k and Y-chromosome haplogroups A2, A3b1, and/or B2b in Bantu-speaking populations of southern Africa therefore provides clear evidence of physical contact with autochthonous peoples, i.e. of gene flow from Khoisan-speaking peoples into

---

<sup>11</sup> For a slightly more detailed explanation of these concepts, see Pakendorf (2014a, 2014b); for a thorough introduction to molecular anthropology, see Stoneking (forthcoming).

<sup>12</sup> It should be kept in mind that “Khoisan” is not a label for a linguistic genealogical unit; rather it subsumes the non-Bantu indigenous languages of southern Africa that are characterized by click consonants. Our assumption that the autochthonous foragers of southern Africa spoke Khoisan languages therefore does not imply that these languages would have necessarily been genealogically related to those spoken nowadays.

the Bantu-speaking populations<sup>13</sup>. In this way, the autochthonous haplogroups are the genetic equivalent of click consonants: both the haplogroups and the clicks identify instances of contact between immigrant speakers of Bantu languages and autochthonous speakers of Khoisan languages. An additional advantage of the genetic markers is that they can shed light on the extent to which the gene flow was sex-biased – that is, to what extent it was predominantly women or predominantly men of the autochthonous communities that contributed to the gene pool of Bantu speech communities. This is possible thanks to the sex-specific patterns of inheritance of mtDNA and the Y-chromosome.

### ***3.2 Molecular anthropological insights into prehistoric contact situations***

With the help of molecular anthropological studies, it is possible to distinguish between three types of prehistoric contact situation that may have characterized the interactions between Bantu-speaking food producers and autochthonous Khoisan-speaking peoples (cf. Pakendorf, 2014a: 631-634). First of all, the (near) absence of both mtDNA haplogroups L0d and L0k and Y-chromosomal haplogroups A2, A3b1, and B2b would be an indication that the click consonants entered the Bantu languages in the absence of intimate social contact with Khoisan-speaking peoples, a situation termed “culture contact” by Ross (2003: 192). This kind of contact situation involves what Van Coetsem (2000: 32, 53-54) calls “borrowing”, namely the adoption of lexemes from the source language by speakers who are dominant in the recipient language (and who might be only marginally familiar with the source language). However, the copying of phonemes is quite rare (Winford, 2010: 176) and requires fairly widespread bilingualism with prestige attached to the source language (Matras, 2009: 225; Lev-Ari *et al.*, 2014: 671).

Secondly, the presence of (substantial proportions of) both autochthonous mtDNA and autochthonous Y-chromosomal lineages in the Bantu-speaking populations would indicate that entire communities of Khoisan-speaking foragers gave up their native languages and shifted to the languages and ultimately the way of life of the immigrating food producers. In this case, the click consonants would be the result of “imposition” (Van Coetsem, 2000: 32, 53-54), that is, the process by which shifting speakers carry over phonemes from their dominant language into the recipient language (i.e. the language they are shifting to).

Lastly, the presence in the Bantu-speaking populations of either autochthonous mtDNA lineages or autochthonous Y-chromosomal lineages, but not both, would be an indication of sex-biased gene flow from Khoisan-speaking groups into the Bantu-speaking populations. In this case, it would have been either women or men who married into the farming

---

<sup>13</sup> The matter is more complicated than depicted here with respect to haplogroup B2b. This is not only common in Khoisan of southern Africa, but also in forager populations of the Central African rainforest. For simplicity's sake we here assume that all B2b chromosomes in southern African Bantu – of which there are only a few – come from Khoisan admixture.

communities. Such sex-biased intermarriage cannot be unambiguously associated with a specific linguistic contact situation. However, depending on the extent of gene flow detected, it might be an indication of long-term social contacts and concomitant long-term bilingualism between peoples, if stable marriage relations are established with a particular community. Long-term bilingualism can result in lexical and grammatical calques and might ultimately lead to large-scale structural changes, i.e. ‘metatypy’ according to Ross (2013: 37).

### ***3.3 Prehistoric contact between Bantu and Khoisan: a synthesis of the genetic data***

Unfortunately, genetic data to elucidate which of the three possible contact situations has taken place are available for the speakers of only some of the Bantu click languages, especially among the SWB group. As can be seen by the overview provided in the top half of Table 2, the amount of Khoisan lineages varies considerably across the groups speaking Bantu languages with clicks, but in all populations mtDNA lineages of Khoisan origin are much more frequent than Khoisan Y-chromosomal lineages. (As shown by the data in the bottom half of the table, which will be discussed in more detail further below, populations speaking Bantu languages without clicks also carry genetic lineages of Khoisan origins. Intimate physical contact alone can therefore not be sufficient for the copying of click consonants.)

**Table 2: Autochthonous genetic lineages in some Bantu populations of southern Africa and functional load of clicks in the languages they speak**

Population	Country	mtDNA		Y-chrom.		Load	Click	
		N	%	N	%		Group	Reference
Xhosa	S. Africa	54	44	57	1.8	high	SEB	1
Thembu	Lesotho	24	42	23	4.3	high	SEB	1
Zulu	S. Africa	54	30	54	1.9	high	SEB	1
Ndebele	Lesotho	33	33	52	3.8	intermed.	SEB	1
S. Sotho	Lesotho	287	30	267	11.2	low	SEB	1
Fwe	Zambia	33	24	26	0	low	SWB	2
Mbukushu	Zambia/ Namibia	20	15	12	0	low	SWB	2, 3, 4
Kgalagadi	Botswana	19	53	21	15	marginal		3, 4
Tswana	Botswana	17	30	18	0	zero		3, 4
Kalanga	Botswana	17	35	19	0	zero		3, 4
Herero	Namibia	30	16	29	17	zero		3, 4
Shanjo	Zambia	24	16	13	0	zero		2

N = number of individuals included in the genetic sample<sup>14</sup>

References: 1 = Marks *et al.* (2015); 2 = Barbieri *et al.* (2013a); 3 = Barbieri *et al.* (2014b) ; 4 = Bajić *et al.* (in preparation)

The data in Table 2 demonstrate that there was intimate physical contact between autochthonous peoples presumably speaking Khoisan languages and the ancestors of the Bantu-speaking groups who now speak languages with clicks. This indicates that click consonants were not copied in situations of casual contact – a conclusion which is well in accordance with the linguistic data. In casual linguistic encounters these phonemes, which are quite untypical of Bantu languages, are highly unlikely to have been copied, but would have been adapted to the phonology of the recipient language. A similar example of genetically detectable physical contact being associated with contact-induced linguistic changes can be found in northeastern Siberia: in this case, speakers of the Lamunkhin variety of Even, which has copied entire verbal paradigms from the neighbouring Turkic language Sakha (Yakut; Pakendorf, 2009, 2015), can be shown to have undergone fairly extensive Sakha gene flow in the paternal line (Duggan *et al.*, 2013). However, frequently linguistic contact takes place in the absence of genetic contact, as in the case of the Hungarians, who are genetically indistinguishable from their Indo-European neighbours but who speak a Uralic language. This indicates that the Central European population – mostly speakers of Slavic languages – that were settled in the territory of what is now Hungary shifted to a Uralic language in the absence of gene flow from the invading Magyars (Csányi *et al.*, 2008, Novembre *et al.* 2008, Semino *et al.* 2000, Tömöry *et al.* 2007).

As shown by the genetic survey, populations speaking Bantu languages with clicks are characterized by substantial proportions of Khoisan maternal ancestry (as shown by the high frequencies of mtDNA haplogroups L0d and/or L0k) and near absence of Khoisan paternal ancestry. This female-biased gene flow demonstrates that language shift and ultimate absorption of entire groups<sup>15</sup> of Khoisan into these Bantu-speaking populations is unlikely to have played a major role in the copying of click consonants, either. The genetic results thus show that it was the inmarriage of substantial numbers specifically of Khoisan-speaking

---

<sup>14</sup> It should be noted that sample sizes of 20 or more individuals allow for reliable inferences concerning the genetic structure of a population. Slightly lower sample sizes, as found for the Kgalagadi, Tswana, and Kalanga, are still acceptable. The Y-chromosomal samples of only 12 and 13 individuals for the Mbukushu and the Shanjo, respectively, are low, however, and some autochthonous gene flow in the paternal line cannot be excluded.

<sup>15</sup> With over 11% autochthonous Y-chromosome lineages the Southern Sotho are an exception to this rule. This might be the result of the historically attested incorporation of entire Khoisan groups into the Sotho community in the late 19th and early 20th century (Wright, 2007: 128). Whether this incorporation of Khoisan communities was accompanied by lexical copying is unknown. Interestingly, words with clicks in Southern Sotho are primarily from Nguni rather than from Khoisan (Doke & Mofokeng 1957).

women that led to the copying of click phonemes – and the larger the number of such women, the larger the impact of the copying appears to have been in terms of number of click phonemes and proportion of the lexicon affected (Table 2). In this context it is a great pity that genetic data for the Yeyi, whose language stands out among the SWB languages by the high functional load of its clicks, are currently lacking. Given the data presented here, one would expect the Yeyi to have incorporated larger numbers of autochthonous women than the other populations speaking SWB languages.

However, while the inmarriage of substantial numbers of Khoisan-speaking women might be a necessary condition for the copying of click consonants, it is certainly not sufficient, as an overview of Khoisan lineages in southern African Bantu-speaking populations shows (Fig. 2, Table 2 bottom). For example, the Kgalagadi have the highest frequency of haplogroup L0d recorded to date in any Bantu-speaking population (53%; *Barbieri et al.*, 2014a) as well as carrying 15% autochthonous paternal lineages (*Bajić et al.*, in preparation), but clicks in this language are extremely marginal, occurring in only one or two words depending on the variety<sup>16</sup> (van der Merwe & Schapera, 1943; Dickens, 1986). Similarly, the frequency of Khoisan mtDNA lineages in the Tswana and Kalanga (30% and 35%, respectively) is comparable to that found in the groups we here call SEB, and the frequency of haplogroups L0d and L0k in the Shanjo is on a par with that in the Mbukushu – and yet the Tswana, Kalanga and Shanjo do not speak languages with click consonants.

[Insert Fig.2A, 2B here]

To summarize: the genetic data show that many of the Bantu-speaking populations of southern Africa were engaged in close physical interactions with the autochthonous Khoisan-speaking peoples of the region, and that these interactions involved the incorporation specifically of Khoisan women into the Bantu speech communities. Furthermore, both the linguistic and the genetic data highlight the fact that the interactions between Khoisan-speaking foragers and immigrating Bantu-speaking food producers differed quite considerably across southern Africa. Not all the Bantu-speaking peoples entered into intimate contact with the indigenous peoples, and even where they did, such intermarriage did not always lead to the incorporation of click phonemes into the Bantu language. While clicks appear not only to have been copied but also maintained where relatively large numbers of autochthonous women were incorporated, the mere inmarriage of Khoisan women is obviously not a sufficient prerequisite for this process.

#### **4 Explaining the discrepancy between the linguistic and genetic data**

---

<sup>16</sup> Lukusa & Monaka (2008: 10) suggest that clicks may be more common in the Tjhauba lect spoken in the Okavango area.



How can we explain the discrepancy between the genetic and the linguistic traces that were retained of the contact between Bantu-speaking immigrants and autochthonous Khoisan-speaking groups, especially if we assume that the languages and the genes of the Bantu communities of southern Africa were shaped by the same event? One possible explanation is that the different contact events differed sociolinguistically and so led to different outcomes. In this case, the incorporation of Khoisan women would not always have led to the copying of highly salient click consonants, but may instead have sometimes had more subtle effects on the languages spoken by these communities. As demonstrated by Naumann & Bibiko (2015), southern Africa as a whole (comprising both Bantu and Khoisan languages) constitutes a well-supported linguistic area within sub-Saharan Africa based on phonological features. Such a degree of convergence of the Bantu and Khoisan languages cannot have come about merely through casual contact. Herero might constitute a specific case of contact-induced changes other than copied clicks. Meinhof (1910: 135) suggests that contact influence from Nama has led to a marked increase in frequency of noun-noun and noun-verb compounds, which are uncommon in Bantu languages, but quite common in Nama. This is reminiscent of the development of right-headed nominal compounds in some of the SWB languages under Khwe influence (Gunnink *et al.*, 2015) and the development of nominal suffixes in some of the Bantu languages of southern Africa (Güldemann, 1999). Interestingly Herero is spoken by the only Bantu community which has incorporated equal amounts of autochthonous maternal and paternal lineages<sup>17</sup> (Table 2); this indicates that here the admixture was not as sex-biased as in the other contact situations.

A second possibility, which we explore here, is that the contact events were indeed similar across southern Africa and had similar results in all the communities concerned, but that later events erased the traces of this outcome in some languages, i.e. that initially copied clicks were subsequently lost. Such “click loss” is known both from Bantu languages with clicks and from Khoisan languages. For instance, comparison with other Nguni languages shows that Swati appears to have lost its lateral click altogether and lost the contrast between dental and alveolar clicks (Ziervogel, 1952; Nussbaum *et al.*, 1969; Rycroft, 1981; Taljaard *et al.*, 1991). Northern Ndebele appears to have lost all its clicks (Herbert, 1990: 124), as seen in Table 1. Fwe also exhibits click loss in its northern variety, which borders on clickless Bantu languages such as Shanjo, but not its southern variety, which is spoken near other click languages such as Yeyi and Khwe. As discussed in detail by (Gunnink, forthcoming), this is shown by the fact that clicks in the southern variety correspond to velar consonants in the northern variety in a systematic way, even in those cases where comparative evidence shows that the clicks in southern Fwe replaced not a velar, but /ʒ/,

---

<sup>17</sup> It should be noted however, that Wood *et al.* (2005) find no autochthonous Y-chromosomal lineages among their sample of 24 Herero. This implies that the Herero are genetically substructured to some extent; it is currently unknown whether this genetic structure correlates with any linguistic differences.

/r/, /t/ and other consonants. The fact that these words, too, contain velar consonants in northern Fwe, and not the original consonant that occurred before click insertion, shows that northern Fwe used to have clicks, but lost them by replacing them with velar non-click consonants.

There are various reasons for the loss of click consonants: 1) natural phonological processes, 2) large numbers of speakers of non-click languages shifting to a click language, or 3) stigmatization of click languages and their speakers.

Several instances of click loss which might be due to natural phonological processes have been described for Khoisan languages. For instance, Ju lects spoken in the southeast have merged the Proto-Ju retroflex clicks with the central alveolar click type, while northern Ju lects have merged the retroflex clicks with the lateral click; only Grootfontein !Xuun still maintains a contrast between these clicks (Sands, 2010). Since contact with Bantu languages or Afrikaans cannot explain the direction of these mergers in the different Ju dialect areas, they are likely to have been internal changes. Similarly, Job (2011, 2014) describes click loss in the Sesfontein variety of Khoekhoe spoken in Northern Namibia. Although Sesfontein Khoekhoe is in contact with Afrikaans, so that one could suspect Afrikaans influence as the reason for this change, the same is true for varieties of the language that do not drop clicks, making internal changes more likely.

Click loss may also occur when significant numbers of speakers shift to a language with clicks not found in their L1, as in the cases of Old Ngoni, Lozi, Sowetan Zulu, and Fanagalo. Older people who identified as Ngoni in historic times spoke a language with clicks (as found in Zulu), while younger speakers reduced the system to a single click type (as found in Sotho) (Elmslie, 1891: 2; Spiss, 1904: 273) as a result of contact with Sotho serfs (Ngonyani, 2001). This variety of Old Ngoni was itself in contact with clickless Tanzanian languages spoken by peoples subjugated by the Ngoni who used "various combinations of consonants as substitutes for clicks" (Elmslie, 1891: 2). Here, the agentivity of the more numerous L2 learners appears to have been more decisive than the prestige of the Ngoni leaders, leading to click loss rather than maintenance. Lozi, the socioculturally dominant language of western Zambia, shows a similar case of click loss. This Sotho language lost its clicks when large numbers of speakers of the clickless language Luyana were incorporated into the speech community e.g. Southern Sotho [hò!álà] > Lozi [kùkálà] 'to begin' (Gowlett, 1989). The urban variety of Zulu spoken in Soweto has lost the contrast between dental and post-alveolar clicks found in Zulu as spoken in rural areas of KwaZulu-Natal due to the high number of L2 speakers who do not have this contrast in their L1 (Southern Sotho) (Gunnink, 2014: 164-165). Similarly, in Fanagalo, a Zulu-based pidgin spoken in South Africa not only by speakers of Bantu languages but also by Europeans, clicks are frequently replaced by /k/, especially by Europeans (Cole, 1953: 4). These examples show that it is possible that some populations speaking Bantu languages may have

copied clicks which were later lost because of subsequent incorporation of speakers of another language with fewer (or no) clicks.

Click loss in Khoisan languages has also been suggested to occur in situations of intense contact with sociopolitically dominant Bantu languages without clicks, since clicks are salient features of the “Bushman” languages that elicit ridicule and derision among the dominant Bantu speakers (Wilmsen & Voßen, 1990: 23-24). In this view, systemic click loss is “perceived as a means of legitimizing these languages and of weakening the identification of the speakers of these languages with an underclass” (Wilmsen & Voßen, 1990: 22). In Eastern Kalahari Khoe languages spoken in eastern Botswana, the alveolar and palatal clicks have been replaced by non-click consonants (Traill & Voßen, 1997: 25-26), with the most extensive replacement found in languages spoken in areas with archaeological evidence for close interactions between foragers and agropastoralists since ~1500 BP (Traill & Voßen, 1997: 35). However, this process may well have taken place as a Khoe-internal change as well as through contact with Bantu languages, since the language Ts'ixa of northern Botswana appears to have lost its clicks through contact with its close relative Shua (Fehn, forthcoming). Click loss was reported by Bleek (1927: 56) for Angolan !Xuun, where young men who were bilingual in the Bantu language Nyemba replaced the alveolar and the lateral click with non-click stop consonants, and it is also attested in Mupa (Angolan) !Xuun, even though not all speakers are bilingual in the Bantu language Kwanyama<sup>18</sup> (Fehn, forthcoming). It is thus possible that Bantu languages that had copied clicks from Khoisan languages at earlier stages of their history lost them later on due to social pressure from more dominant clickless languages.

It is important to note that click loss in situations of changing prestige need not have been the quasi-conscious process to rid the language of stigmatized segments suggested by Wilmsen & Vossen (1990). Adaptation of foreign segments in copies is influenced by the prestige of the donor language at the time of copying, with foreign sounds more likely to be retained in those semantic domains where the donor language has high prestige. Should the sociolinguistic situation change, foreign elements might simply undergo the usual adaptation to the phonology of the recipient language (Lev-Ari *et al.*, 2014: 675-676). For instance, in Tswana some plant names possibly copied from Khoisan have velar stops in place of the clicks in the potential Khoisan source lexemes, e.g. *mo-kala* 'camel thorn, *Acacia erioloba* (Cole, 1995: 214), cf. †Hoan //áǎ (Sands, field notes); G!ui //árà (Nakagawa, 2014). The same is true for Herero, where clicks in Khoekhoe borrowings are replaced by /k/ or /h/ (Meinhof, 1910: 131). It is currently not possible to distinguish whether these words were first copied with a click which was later changed to a non-click consonant in adaptation to the

---

<sup>18</sup> This is a variant of Ambo (aka Ovambo), speakers of which have been included in Y-chromosome studies, as seen in Fig.2B.

phonology of the Bantu, or whether the phonological adaptation took place at the time of copying.

Given the lack of historical data, we cannot conclusively identify the reasons that precluded the copying of clicks even where large numbers of Khoisan individuals were incorporated into the Bantu-speaking communities. However, by comparing the linguistic and genetic evidence for contact in those communities where they exist, namely the SEB and SWB groups, we can obtain further insights into the prehistoric contact between autochthonous Khoisan and immigrating Bantu.

## 5 Contact in the history of the southeast Bantu speech communities

For the Nguni languages it has been suggested that the copying of clicks took place through several centuries of very intimate contact between the Bantu agropastoralists and Khoekhoe herders, who would have been on an equal social footing and whose cattle-centred cultures would have been compatible with each other (Herbert, 2002). In addition, the proliferation of clicks in these Bantu languages is assumed to have been enhanced by a system of in-law taboo called *hlonipha*<sup>19</sup>, in which inmarrying women are expected to avoid pronouncing not only the names of several in-laws, but also composite syllables (Irvine & Gal, 2000; Herbert, 2002: 303-306). The insertion of clicks to modify taboo syllables also led to the occurrence of clicks in native Bantu vocabulary.

With the currently available genetic data it is difficult to assess the scenario of contact specifically with Khoekhoe pastoralists because the pastoralist Khoekhoe and foraging !Ui-speaking populations of South Africa may have been quite closely related genetically (Nurse *et al.*, 1985: 222; Schlebusch *et al.*, 2012: their Fig. 2C; Uren *et al.* 2016), thus making it hard to distinguish their relative contributions to the Bantu-speaking agropastoralists. This problem is exacerbated by the fact that both the Khoekhoe pastoralists as well as the !Ui foragers known from historical sources have been assimilated into other populations of South Africa. The so-called Coloured populations of South Africa might have incorporated both the descendants of Khoekhoe pastoralists as well as of former foragers (Nurse, Weiner, and Jenkins 1985, 83-89, 235), whereas the “Karretjie people” are probably partly descended from the |Xam-speaking foragers of the northern Cape (Traill 2007, 137). The †Khomani are in part descendants of !Ui-speaking foragers of the southern Kalahari as well as having incorporated Khoekhoe-speaking and Coloured groups (Schenck 2008, 90-94, 100-105). Thus, the small amount of autosomal gene flow<sup>20</sup> from Khoisan-speaking groups

---

<sup>19</sup> It is possible, however, that the system of language avoidance itself was an import from Khoisan, as was first suggested by Tony Traill, who noted that lexical avoidance is under-documented in Khoisan but does occur.

detected in South African Bantu-speaking populations is equally attributable to the Nama, the “Karretjie people”, or the †Khomani (Schlebusch *et al.* 2012: Fig. 2C).

Nevertheless, the mtDNA data show some evidence of inmarriage in the maternal line not only of Khoekhoe pastoralists, but also of forager groups, as summarized in Table 3. The table is based on Fig. 2 of Schlebusch *et al.* (2013), which presents the relationships among mtDNA sequences in a network. The table is split into three parts, presenting data on sequences that are directly shared with Khoisan populations or their probable descendants (A), on sequences that differ by one mutation from those found in Khoisan populations or their descendants (B), and on sequences that differ by two mutations (C), and it includes information on the probable or known subsistence of the groups, as far as possible. In the network produced by Schlebusch *et al.* (2013), some populations were grouped; these are the “Karretjie people” and the Karoo Coloured, the †Khomani and the Northern Cape Coloured, as well as the southeastern Bantu-speakers themselves. The latter comprise a sample of 22 Sotho and Tswana, five Swazi, and 36 Zulu and Xhosa. As can be seen from the table, the “southeastern Bantu-speakers” share more sequences directly with forager groups than with the Nama or Coloured populations. Sequences that are not directly shared, however, are separated by one or two mutations from sequences largely found in the descendants of South African forager groups and Coloureds, whose ancestry may include both foragers and Khoekhoe pastoralists, and Khoekhoe-speaking Nama pastoralists. This – admittedly very coarse-grained – picture thus points towards a likely maternal contribution not only of Khoekhoe pastoralists, but also of South African forager populations in the history of the SEB populations, although this conclusion is very tentative until it can be confirmed with genetic data covering more markers and more Khoisan populations.

**Table 3: Autochthonous mtDNA lineages in “southeastern Bantu-speakers” (Schlebusch *et al.* 2013: their Fig. 2)**

A: Sequences directly shared with Khoisan populations or probable descendants thereof

Seq #	Hg	forager descendants/mixed			pastoralist	forager			
		Karretjie/ K.Col (107)	C.Col (20)	†Khomani / N.C.Col (97)	Nama (28)	Ju ’hoan (42)	!Xuun (49)	Naro (2)	Xegwi/ Duma (4)

<sup>20</sup> The autosomes refer to all of the human chromosomes with the exception of the Y-chromosome found only in men and the X-chromosome present in two copies in women and one copy in men. MtDNA is an entirely different molecule found not in the cell nucleus like the chromosomes, but in separate cell structures called mitochondria.

1	L0k				+	+	+		
2	L0d3							+	
3	L0d1c							+	
4	L0d1b	+	+	+				+	
5	L0d1b					+			
6	L0dx				+				
7	L0d2a	+	+	+	+			+	+
8	L0d2a								+

Seq # = arbitrary number assigned by us to autochthonous sequences found in the “southeastern Bantu-speakers”

HG = Haplogroup to which a given sequence belongs

K.Col = Karoo Coloured

C. Col = Cape Coloured

N.C. Col = Northern Cape Coloured

In brackets = number of samples included in the group

B: Sequences one mutation distant from Khoisan populations or their probable descendants

Seq #	Hg	forager descendants/mixed			pastoralist Nama (28)	foragers			
		Karretjie/ K.Col (107)	C.Col (20)	‡Khomani/ N.C.Col (97)		Ju ’hoan (42)	!Xuun (49)	Naro (2)	Xegwi/ Duma (4)
9	L0d1b					+			
10	L0d1b			+					
11	L0d1b	+	+	+			+		
12	L0d2a	+		+					
13	L0d2a	+	+	+	+		+	+	
14	L0d2a	+	+	+	+		+	+	
15	L0d2a	+	+						

Legend as for Table A.

C: Sequences two mutations distant from Khoisan populations or their probable descendants

Seq #	Hg	forager descendants/mixed			pastoralist Nama (28)	forager			
		Karretjie/ K.Col (107)	C.Col (20)	‡Khomani/ N.C.Col (97)		Ju ’hoan (42)	!Xuun (49)	Naro (2)	Xegwi/ Duma (4)
16	L0d3	+		+					
17	L0d1a	+		+					

18	L0d2a	+	+	+	+	+	+
----	-------	---	---	---	---	---	---

Legend as for Table A.

It is thus possible that the Nguni peoples were in close contact with both socially equal pastoralists as well as with foragers. This is in good accordance with Hammond-Tooke’s suggestion (1998, 1999) that certain elements of ritual among the Nguni were borrowed from San foragers rather than from Khoekhoe-speaking pastoralists. It is notable that the linguistic evidence for contact points much more towards Khoekhoe than !Ui languages spoken by foragers as the source of clicks in the Nguni languages, which indicates that socioculturally the contact between the ancestors of the Nguni and the Khoisan-speaking foragers would have differed quite considerably from that in which Nguni and pastoralists were involved.

## 6 Contact in the history of the southwest Bantu speech communities

The small size of the click inventory and the low number of words carrying a click in Manyo, Kwangali, Mbukushu, and Fwe has been interpreted as an indication of fairly casual Khoisan-Bantu contact (Fisch, 1998: 11). However, there is evidence that the copying of these phonemes was in actual fact the result of intense contact. As we discuss in detail in Bostoen & Sands (2012) and in Gunnink *et al.* (2015), click phonemes have been integrated into words of Bantu origin – in contrast to what is claimed by Herbert (2002: 307) and in the absence of a system of in-law naming taboo as that known from the Nguni languages (cf. Seidel, 2009: 246). Furthermore, loanwords have been integrated not into the default noun class commonly used for copied items, but into semantically appropriate noun classes, a process called “paralexification” by Mous (2001). Both of these changes are indicative of conscious language manipulation not usually found in situations of casual contact (Mous, 2003: 223, 226-227). Furthermore, there is evidence of calquing and structural change, namely the development of head-final compounds untypical of Bantu languages, as a result of influence from Khwe. We therefore suggest that the evidence for paralexification and language manipulation through spread of clicks to Bantu words might be an indication that “speakers of SWB languages once valued Khoisan origins and positively identified with these” (Gunnink *et al.*, 2015). Unfortunately, we can assess this statement only for the Fwe of southwestern Zambia and the Mbukushu of southwestern Zambia and northeastern Namibia, for whom molecular anthropological data are available (Barbieri *et al.*, 2013a; Barbieri *et al.*, 2014b).

As shown in Table 2, the mtDNA analyses provide evidence for considerable levels of autochthonous geneflow in the maternal line into the Fwe (where haplogroups L0d and L0k are present at ~24% frequency) as well as into the Mbukushu (15% L0d and L0k). In contrast, the neighbouring southwest Zambian Bantu-speaking groups carry an average of only 3% of L0d/L0k lineages (Fig. 2), with the exception of the Shanjo, who do not speak a click language but have 16% L0d/L0k (Table 2) (Barbieri *et al.*, 2013a; Barbieri *et al.*, 2014b). Neither the Fwe nor the Mbukushu carry any Y-chromosomal lineages of autochthonous origin (Table 2). It should be pointed out, however, that both the Fwe and the Mbukushu are in closer geographical contact with Khwe than with other Khoisan-speaking groups (Gunnink *et al.*, 2015). Since the Khwe have high frequencies of Y-chromosomal haplogroups that are characteristic of Bantu-speaking populations (Knight *et al.*, 2003; Soodyall *et al.*, 2008; Bajic *et al.*, in preparation), paternal gene flow involving these lineages would not necessarily be detectable. However, it is unlikely that the Fwe and Mbukushu have experienced large amounts of paternal gene flow from the Khwe, since the Khwe carry a particular Y-chromosomal haplogroup called E1b1b in high frequency (Henn *et al.*, 2008; Bajic *et al.*, in preparation). This is practically absent among southern African Bantu (de Filippo *et al.*, 2011) and like their Bantu neighbours the Fwe and the Mbukushu lack this haplogroup. Nor is it likely that the Fwe and Mbukushu have been in contact with a hitherto unknown Khoisan-speaking population with very distinct paternal lineages, even if it did not carry the ‘signature’ autochthonous lineages, as they do not have higher levels of Y-chromosomal diversity than their neighbours.

The genetic data are particularly interesting with respect to the contact situation involving the Fwe for three reasons: 1) the Fwe carry divergent L0k lineages not found in any currently known Khoisan populations; 2) the Fwe carry several distinct autochthonous lineages that must have been incorporated via separate admixture events; and 3) the only mtDNA lineages shared between the Fwe and the Mbukushu and the Fwe and the Shanjo are autochthonous lineages.

Whereas the Fwe L0d lineages are closely related to lineages found in extant Khoisan-speaking populations, their L0k lineages belong to two very divergent branches of this haplogroup that are not found in present-day Khoisan, but only in Bantu, especially in Bantu-speaking populations of Zambia (Barbieri *et al.*, 2013b). The interpretation of these data is that this area, which is nowadays occupied only by Bantu-speaking food producers, used to be inhabited by foragers genetically distantly related to Khoisan-speaking foragers, but that these did not leave any descendants other than those that were incorporated into Bantu-speaking communities through in-marriage of women. It is notable that the Fwe carry the highest frequencies of these divergent L0k lineages found in any population studied to date. Thus, the ancestors of the Fwe must have admixed intensively with a population (or populations) that later on ceased to exist as a separate ethnolinguistic group – a scenario that



is in good accordance with the fact that many of the non-Bantu words with clicks in the SWB languages cannot be traced to a known Khoisan language, but might stem from a now extinct language (Gunnink *et al.*, 2015).

The autochthonous lineages in the Fwe are too diverse to have stemmed from only one or two founding mothers. Instead, of the five separate L0d/L0k lineages found in the Fwe sample included in the study, four are so divergent that they can only be the result of separate admixture events. Since mtDNA is transmitted only in the maternal line, any woman with autochthonous female ancestry who remained without offspring or who had only sons would not have passed on her mtDNA lineage to future generations beyond her sons. Such cases of intermarriage would therefore remain undetected in the current gene pool of the Fwe. Thus, the genetic data indicate that the contact between the ancestors of the Fwe and the autochthonous foragers must have been fairly intense, involving numerous autochthonous women; such intense contact may have taken place over a relatively long period of time.

The Fwe share L0d and/or L0k lineages not with Khoisan populations, but only with the Mbukushu and the Shanjo, who speak a closely related Bantu language without clicks. It is noteworthy that only these L0d and L0k lineages are shared between Fwe and Shanjo and Fwe and Mbukushu, not any lineages belonging to haplogroups found more widely in these and other Bantu-speaking populations. Interestingly, the Mbukushu do not share any L0d or L0k sequences directly with extant Khoisan-speaking populations, not even with the Khwe, with whom they are reported to interact very closely (Fisch, 2005).

In summary, the genetic data provide indications that the contact between the Fwe ancestors and the Khoisan communities may have taken place at a relatively early period of the Bantu immigration into southern Africa. This can be deduced from the fact that the majority of the autochthonous mtDNA lineages found in the Fwe belong to divergent branches of haplogroup L0k which are found only in the northern range of the L0k distribution. It is thus likely that the ancestors of the Fwe incorporated these lineages fairly early in their prehistory, upon their arrival in what is now Zambia. Historical linguistic research is in support of such a scenario of early contact between the Bantu-speaking ancestors of current-day Fwe and Khoisan speakers. The Western Botatwe subgroup, to which Fwe and other closely related languages spoken in Zambia's Western Province belong, is an early split from Proto-Botatwe (de Luna, 2010: 79).

Furthermore, both the Shanjo and the Mbukushu appear to have preferentially married Fwe women with Khoisan maternal ancestry: they share their L0d and L0k lineages with the Fwe rather than with Khoisan-speaking populations, making it unlikely that they incorporated these lineages through direct inmarriage of Khoisan women. Since the L0d/L0k sequences found in the Shanjo and Mbukushu comprise a subset of the sequences found in the Fwe, the direction of gene flow is more likely to have been from Fwe into Shanjo and

Mbukushu than vice versa. Intriguingly, however, these L0d and L0k sequences are the only sequences that the Shango and Mbukushu share with the Fwe; they do not share any sequences that belong to haplogroups commonly found in Bantu-speaking populations, even though such Bantu haplogroups are widespread in all three populations. This restricted amount of sharing is not due to genetic isolation of any of these populations, since all three share several non-autochthonous mtDNA sequences with other neighbouring Bantu-speaking populations (Barbieri *et al.*, 2013a). Rather, the ancestors of the Shango and the Mbukushu must have been preferentially marrying Fwe women with Khoisan maternal ancestry. There are various possible reasons for such preferential marriage patterns: women with Khoisan ancestry may have required less bridewealth payment, or they may have been more willing to leave the Fwe community for another one. However, given the fact that Fwe is a language which has maintained copied click consonants and even incorporated them into words of Bantu origin (Gunnink *et al.* 2015), it is also possible that Khoisan maternal ancestry was viewed positively.

## 7 Discussion and Conclusions

Clicks in Bantu languages clearly have a variegated history: contrary to what is widely assumed, they are not necessarily hard to copy (cf. Güldemann & Stoneking, 2008), nor are they always due to contact with Khoisan peoples, but they might be the result of independent innovations or of contact with a Bantu click language (cf. Section 2). Nevertheless, in several present-day Bantu languages spoken in southern Africa, clicks are indeed the most salient residue of past interactions with Khoisan languages, and their presence cannot be explained solely by sociocultural factors such as *hlonipha*, in contrast to Herbert's (2002) suggestion.

Various factors might have played a role in shaping the diverse linguistic effects of the contact between Bantu-speaking immigrants and autochthonous peoples speaking Khoisan languages that is detected in the genetic data. The genetically detectable physical contact may have taken place in different sociocultural situations, so that the linguistic outcome of this contact differed, with clicks being copied in some cases and structural changes, such as the compounds observed in Herero, taking place in others. It is therefore possible that more detailed linguistic investigations will uncover less noticeable contact-induced changes in the Bantu languages spoken by communities with genetic evidence of close interactions with Khoisan populations. Both copying of clicks and copying of syntactic structures (involving typologically dissimilar languages) implies close contact with extensive bilingualism. However, structural changes are indicative of a somewhat different kind of contact situation, since morphological or phonological forms are often emblematic of a particular lect and as such are more susceptible to speakers' control. Structural features are less emblematic and

thus more frequently copied even in situations where speech communities eschew language mixing (e.g. Aikhenvald, 2002, 213-216; Ross 2003, 189). Copying highly salient click consonants, in contrast, is unlikely to escape anyone's notice, so that copied click consonants are more likely to serve as flags of a new identity (Gunnink *et al.*, 2015).

Clicks may also have been copied early on but lost when the speech community incorporated large numbers of speakers of clickless languages, or when social relations changed and the clicks that were once associated with a domain of prestige were gradually adapted to the phonology of the Bantu language. It is also possible that some Bantu communities came into contact with Khoisan-speaking groups at a time when the social status of the Bantu and Khoisan communities had become imbalanced, so that click consonants were avoided as markers of the low-status language. This might explain the finding that in some possible loans from Khoisan languages in Tswana clicks were replaced with velars, as discussed in Section 4.

A further possibility is that the presence or absence of clicks in contemporary Bantu languages is not due to any sociocultural differences in the contact situation, but simply a matter of time, and that copied clicks are gradually eroded in all languages until they are fully lost. In this case, the contact undergone by the ancestors of the Manyo, Kwangali, Mbukushu, and Fwe would have taken place a substantially longer time ago than the contact experienced by the ancestors of the Yeyi, Xhosa, and Zulu, since these latter languages still have a large number of different click phonemes with a high functional load.

The ancestors of the Nguni peoples are indeed hypothesized to have immigrated to the area only within the last 1000 years (Mitchell, 2002: 287-288), so that the contact with Khoisan populations would have taken place relatively recently. However, the ancestors of Sotho-Tswana peoples are also hypothesized to have been among the late wave of migration that brought the Nguni-speaking groups to South Africa (Mitchell, 2002: 287-288), and the direct ancestors of the Tswana and Kgalagadi probably migrated to what is now Botswana as recently as 350 years ago (Kiyaga-Mulindwa, 1993: 386; Segobye, 1998). Given the high frequencies of autochthonous mtDNA lineages in Tswana and Kgalagadi (Table 2), they must have been in contact with groups presumably speaking Khoisan languages. Although this contact would have taken place relatively recently, these languages currently do not have clicks as regular consonants. There thus does not seem to be a simple correlation between time since contact and amount of clicks – although it will be possible to address this question directly in the future with more refined genetic methods and using data from the entire genome.

Nevertheless, the current genetic data already indicate that the contact between the ancestors of the Fwe and the autochthonous populations must have been among the earliest such contact events, as shown by the very divergent L0k lineages the Fwe carry in high frequency (cf. Section 6). Yet speakers of Fwe maintained clicks, which indicates that time

since contact is not the only factor determining the current distribution of clicks, but that social factors also play a role not only in the adoption of click consonants, but also in their subsequent maintenance.

Status-related social issues such as prestige of some kind (be that economic, ritual, or medicinal) may also have played a role in the outcome of Khoisan-Bantu contact. Nowadays, the Khoisan-speaking peoples of southern Africa who traditionally practised a foraging life-style occupy the lowest rungs of the social ladder (Wilmsen & Voßen, 1990; Rousset, 2003: 5; Bolaane, 2013: 217; Thiem & Jones, 2014: 350-351). This social inequity makes it unlikely that any Bantu-speaking community would flag intimate contact with Khoisan-speaking foragers by adopting click consonants, since in sociopolitically unequal relationships it is commonly the forager communities who adopt cultural elements from their dominant agropastoralist neighbours, rather than the other way round (Jolly, 1996: 279). For instance, Kgalagadi relations with Lala (former !Xóǀ speakers, cf. Monaka & Lepekoane, 2008) are far from equitable (Silberbauer & Kuper, 1966), and Kgalagadi speakers have not adopted clicks in spite of large-scale intermarriage with autochthonous foragers (see Table 2). However, the fact that clicks as highly salient loan phonemes did survive in certain Bantu languages and that their use was even extended to neologisms and inherited Bantu vocabulary, as discussed in detail by Gunnink *et al.* (2015), suggests that they may have indexed a positive identity in times when social relations between Khoisan- and Bantu-speakers were still more egalitarian.

In this respect the intriguing fact that the ancestors of the Mbukushu and the Shanjo appear to have preferentially married Fwe women with Khoisan maternal ancestry takes on a new light. It is of course possible that these women were preferred marriage partners because they required less bride price, as is currently observed in interactions between Bantu agriculturalists and hunter-gatherers in the rainforest (Verdu *et al.* 2013), or that they were more willing to marry into another community since their social networks in the community of origin were weaker. However, it is also possible that they were preferred as marriage partners because Khoisan ancestry was a valued asset. This may certainly have been so in the period of the first Bantu expansions in southern Africa, when Bantu-speaking newcomers were heavily dependent on the knowledge of autochthonous communities, especially Khoisan foragers, to survive in an unfamiliar habitat (cf. Vansina, 2004: 46-47). Such relatively egalitarian interactions have also been suggested for the early contact between Bantu-speaking food producers and Pygmy foragers in the Central African forest (Klieman, 2003). In this context, intermarriage with indigenous women may have been an efficient survival strategy. In early Bantu-speaking groups that were still small, the demographic and linguistic impact of these Khoisan-speaking mothers would have been very strong.

Furthermore, the indigenous populations may have had relatively high prestige in certain domains due to their close ties to the land and its spirits, or due to specific skills as hunters,

healers, rainmakers or metal workers. This has been suggested, for example, for the interactions of the Batwa foragers and Ekonda agriculturalists in the Congo Basin, where the Batwa were “...recognized as masters of the forest and owners of fire, that is, as aboriginal inhabitants whose intercession was necessary to the success of the smelt” (Herbert, 2012: 28). Similar considerations might have characterized the involvement of the !Xuun in the Kwanyama mining process (Herbert, 2012). It is notable in this respect that the words containing click consonants in the SWB languages belong particularly to the domain of hunting, fishing, and foraging or denote local fauna and flora (Gunnink *et al.*, 2015: 204-205), which might be an indication that the Khoisan peoples were accorded some prestige for their skills as hunters and foragers, even though they might have been less respected in other domains. Such a polyvalent relationship of dominant societies towards “peripatetic” communities who provide them with certain services is discussed in detail by Bollig (2005).

Our results thus indicate that the social interactions between the ancestors of the Bantu peoples speaking languages with clicks and the autochthonous Khoisan may have been more egalitarian at early stages of the Bantu expansion into southern Africa, in contrast to the frontier model proposed by Megaw (1977) and Alexander (1984). This model suggests that interactions between expanding food producers and autochthonous foragers would have been characterized by aggression and hostility during the early stages of the expansion of the food producers into new territory (called a “moving frontier” in the model) until the food producers had occupied all usable land. However, the ancestors of the Fwe appear to have entered into contact with autochthonous foragers at a very early stage of the Bantu expansion into southern Africa, as evidenced by the divergent L0k lineages they have incorporated, and should therefore still have been part of a moving frontier. Yet the genetic and linguistic results discussed in this article provide some indication that this contact involved close and possibly egalitarian interactions with the autochthonous forager groups rather than being characterized by hostility.

Only when Bantu speech communities were well established, grew bigger and gained political power and wealth may Khoisan identity and ancestry have lost social significance and were Khoisan speakers pushed towards the societal marginalization of which most remaining communities are still victim today. Such a warped social order may not only have discouraged the use of Khoisan languages themselves, but even the continuation of language features reminiscent of Khoisan when speaking a Bantu language. In such a Bantu speech community indexing Khoisan identity may not have been the best gateway to social success, even if many of its members had Khoisan ancestry. Clicks, the most salient linguistic traces of Khoisan ancestry, became less positively valued to the point that their functional load diminished in the Bantu languages or was even reduced to zero. This can be seen in Fwe: although the genetic data provide tentative indications that Khoisan maternal ancestry may once have been a valuable asset, these ties are losing their prestige. Clicks are being

progressively lost, and the northern variety spoken in the vicinity of clickless Bantu languages has lost its clicks entirely (cf. Section 4). This finding suggests that it should be possible to find words of Khoisan origin in which clicks have been replaced with non-click consonants in the languages spoken by communities with a large amount of Khoisan maternal ancestry. This might be the case for Tswana and Herero, as we discuss in Section 4, but more linguistic analysis is required to verify our hypothesis.

## Figure Legends

Figure 1: **Approximate distribution of Bantu languages in southern Africa** mentioned in the text with clicks as regular consonants (red), marginal clicks (green), and no click consonants (black).

Figure 2: **Proportion of autochthonous genetic lineages (black) in Bantu-speaking populations of southern Africa.** The populations speaking click languages are labelled in red. Abbreviations: AMB = Ambo (aka Ovambo); BIS = Bisa; CHO = Chopi; CHW = Chwabo; GAN = Ganguela; HER = Herero; HIM = Himba; KAL = Kalanga; KGA = Kgalagadi; KUN = Kunda; KUV = Kuvale; KWA = Kwamashi; MAK = Makhuwa; MKND = Makonde; MBU = Mbukushu; NDA = Ndau; NDE = Ndebele; NGU = Nguni; NKO = Nkoya; NYAK = Nyaneka; NYAJ = Nyanja; NYU = Nyungwe; OVI = Ovimbundu; RON = Ronga; SEN = Sena; SHNG = Shangaan; SHA = Shanjo; SHO = Shona; SOT = Sotho; SUB = Subiya; THE = Thembu; TNG = Tonga (Mozambique); TON = Tonga (Zambia); TOT = Totela; TSW = Tswana; XHO = Xhosa; ZUL = Zulu. A. **Mitochondrial DNA haplogroups L0d and L0k.** Data from: Salas *et al.* (2002); Castri *et al.* (2009); Coelho *et al.* (2009); de Filippo *et al.* (2010); Barbieri *et al.* (2013b); Barbieri *et al.* (2014a); Barbieri *et al.* (2014b); Marks *et al.* (2015). B. **Y-chromosomal haplogroups A3b1 and B2b.** Data from: Wood *et al.* (2005); Coelho *et al.* (2009); de Filippo *et al.* (2010); Marks *et al.* (2015); Bajić *et al.* (in preparation).

## References

- Afido, Pedro J., Gregório Firmino, John H. Heins, Samba Mbuub and Manuel Trinta (eds.). 1989. *I seminário sobre a padronização da ortografia de línguas Moçambicanas*. Maputo: Núcleo de Estudo de Línguas Moçambicanas (NELIMO), Faculdade de Letras, Universidade Eduardo Mondlane.
- Aikhenvald, Alexandra. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.

- Alexander, J.A. 1984. Early frontiers in southern Africa. In M. Hall, G. Avery, D.M. Avery, Wilson M.L. & Humphreys A.J.B. (eds.), *Frontiers: Southern African archaeology today*, 12-23. Oxford: B.A.R. International Series.
- Amselle, Jean-Loup and Elikia M'Bokolo (eds.). 1999. *Au coeur de l'ethnie: ethnie, tribalisme et état en Afrique*. Paris: La Découverte.
- Anders, H. D. 1937. Observations on certain sound changes in Xhosa derivatives from Khoisan. *South African Journal of Science* 33: 921-925.
- Bailey, Richard. 1995. Issues in the phonology and orthography of Chopi (ciCopi S 61). In Anthony Traill, Rainer Vossen & Megan Biesele (eds.), *The Complete Linguist: Papers in Memory of Patrick J. Dickens*, 135-170. Köln: Rüdiger Köppe.
- Bajić, Vladimir, Chiara Barbieri, Tom Güldemann, Christfried Naumann, Linda Gerlach, Falko Berthold, Hiroshi Nakagawa, Sununguko W. Mpoloka, Mark Stoneking and Brigitte Pakendorf. in preparation. Investigating the prehistory of "Khoisan"-speaking populations from southern Africa with large-scale Y-chromosome sequences.
- Barbieri, Chiara, Anne Butthoff, Koen Bostoen and Brigitte Pakendorf. 2013a. Genetic perspectives on the origin of clicks in southwestern Bantu languages. *European Journal of Human Genetics* 21: 430-436.
- Barbieri, Chiara, Tom Güldemann, Christfried Naumann, Linda Gerlach, Falko Berthold, Hiroshi Nakagawa, Sununguko W. Mpoloka, Mark Stoneking and Brigitte Pakendorf. 2014a. Unraveling the complex maternal history of Southern African Khoisan populations. *American Journal of Physical Anthropology* 153: 435-448.
- Barbieri, Chiara, Alexander Hübner, Enrico Macholdt, Shengyu Ni, Sebastian Lippold, Roland Schröder, Sununguko Wata Mpoloka, Josephine Purps, Lutz Roewer, Mark Stoneking and Brigitte Pakendorf. 2016. Refining the Y chromosome phylogeny with southern African sequences. *Human Genetics* 135: 541-553.
- Barbieri, Chiara, Mário Vicente, Sandra Oliveira, Koen Bostoen, Jorge Rocha, Mark Stoneking and Brigitte Pakendorf. 2014b. Migration and interaction in a contact zone: mtDNA variation among Bantu-speakers in southern Africa. *PLoS ONE* 9: e99117.
- Barbieri, Chiara, Mário Vicente, Jorge Rocha, Sununguko Wata Mpoloka, Mark Stoneking and Brigitte Pakendorf. 2013b. Ancient substructure in early mtDNA lineages of southern Africa. *American Journal of Human Genetics* 92: 285-292.
- Batini, Chiara, Gianmarco Ferri, Giovanni Destro-Bisol, Francesca Brisighelli, Donata Luiselli, Paula Sánchez-Diz, Jorge Rocha, Tatum Simonson, Antonio Brehm, Valeria Montano, Nasr Eldin Elwali, Gabriella Spedini, María Eugenia D'Amato, Natalie Myres, Peter Ebbesen, David Comas and Cristian Capelli. 2011. Signatures of the pre-agricultural peopling processes in sub-Saharan Africa as revealed by the phylogeography of early Y chromosome lineages. *Molecular Biology and Evolution* 28: 2603-2613.

- Baumbach, Erdmann J.M. 1974. *Introduction to the Speech Sounds and Speech Sound Changes of Tsonga*. Pretoria.
- Behar, D. M., R. Villems, H. Soodyall, J. Blue-Smith, L. Pereira, E. Metspalu, R. Scozzari, H. Makkan, S. Tzur, D. Comas, J. Bertranpetit, L. Quintana-Murci, C. Tyler-Smith, R. S. Wells and S. Rosset. 2008. The Dawn of Human Matrilineal Diversity. *American Journal of Human Genetics* 82: 1130-1140.
- Bleek, Dorothea Frances. 1927. The distribution of Bushman languages in South Africa. In Franz Boas, Otto Heinrich Dempwolff & Giulio Panconcelli-Calzia (eds.), *Festschrift Meinhof*, 55-64. Glückstadt and Hamburg: J J Augustin.
- Blench, Roger. 2006. *Archaeology, Language and the African Past*. Lanham: Altamira Press.
- Bolaane, Maitseo M.M. (ed.) 2013. *Chiefs, Hunters and San in the Creation of the Moremi Game Reserve, Okavango Delta; Multiracial Interactions and Initiatives, 1956-1979*. Osaka: National Museum of Ethnology.
- Bollig, Michael. 2005. Singing Smiths and Hunting Ritual Entrepreneurs: Transitions between Forager and Peripatetic Communities in Africa. In Joseph C. Berland & Aparna Rao (eds.), *Customary Strangers. New perspectives on Peripatetic Peoples in the Middle East, Africa and Asia*. New York: Praeger.
- Bostoen, Koen. 2007. Pots, words and the Bantu problem: On lexical reconstruction and early African history. *Journal of African History* 48: 173-199.
- Bostoen, Koen, Bernard Clist, Charles Doumenge, Rebecca Grollemund, Jean-Marie Hombert, Joseph Koni Muluwa and Jean Maley. 2015. Middle to Late Holocene Paleoclimatic Change and the Early Bantu Expansion in the Rain Forests of West Central-Africa. *Current Anthropology* 56: 354-384.
- Bostoen, Koen and Bonny Sands. 2012. Clicks in south-western Bantu languages: contact-induced vs. language-internal lexical change. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics Cologne 2009*, 129-140. Köln: Rüdiger Köppe.
- Bourquin, Walther. 1951. Click-Words which Xhosa, Zulu and Sotho have in common. *African Studies* 10: 59-81.
- Bowern, Claire and Victoria Lotridge (eds.). 2002. *Ndebele*. München: Lincom Europa.
- Castri, L., S. Tofanelli, P. Garagnani, C. Bini, X. Fosella, S. Pelotti, G. Paoli, D. Pettener and D. Luiselli. 2009. mtDNA variability in two Bantu-speaking populations (Shona and Hutu) from Eastern Africa: Implications for peopling and migration patterns in sub-Saharan Africa [May 7]. *Am J Phys Anthropol* 140: 302-311.
- Coelho, M., F. Sequeira, D. Luiselli, S. Beleza and J. Rocha. 2009. On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. *BMC Evol Biol* 9: 80.



- Cole, Desmond Thorne. 1953. Fanagalo and the Bantu languages in South Africa. *African Studies* 12: 1-9.
- Cole, Desmond Thorne. 1995. *Setswana - animals and plants (Setswana - ditshedi le dithlare)*. Gaborone: The Botswana Society.
- Csányi, B., E. Bogácsi-Szabó, Gy. Tömöry, Á. Czibula, K. Priskin, A. Csősz, B. Mende, P. Langó, K. Csete, A. Zsolnai, E. K. Conant, C. S. Downes and I. Raskó. 2008. Y-chromosome analysis of ancient Hungarian and two modern Hungarian-speaking populations from the Carpathian Basin. *Annals of Human Genetics* 72: 519-534.
- Daeleman, Jan. 1966. *Morfologie van naamwoord en werkwoord in het Kongo (Ntandu) met ontleding van het foneemsysteem*. PhD, Katholieke Universiteit Leuven.
- Dammann, Ernst. 1957. *Studien zum Kwangali: Grammatik, Texte, Glossar*. Hamburg,: Cram, De Gruyter.
- de Filippo, Cesare, Chiara Barbieri, Mark Whitten, Sununguko Wata Mpoloka, Ellen D. Gunnarsdóttir, Koen Bostoen, Terry Nyambe, Klaus Beyer, Henning Schreiber, Peter de Knijff, Donata Luiselli, Mark Stoneking and Brigitte Pakendorf. 2011. Y-chromosomal variation in Sub-Saharan Africa: insights into the history of Niger-Congo groups. *Molecular Biology and Evolution* 28: 1255 - 1269.
- de Filippo, Cesare, Patricia Heyn, Larry Barham, Mark Stoneking and Brigitte Pakendorf. 2010. Genetic perspectives on forager-farmer interaction in the Luangwa Valley of Zambia. *American Journal of Physical Anthropology* 141: 382-394.
- de Luna, Kathryn M. 2010. Classifying Botatwe: M.60 and K.40 Languages and the Settlement Chronology of South Central Africa. *Africana Linguistica* 16: 65-96.
- Dickens, Patrick John. 1986. *Qhalaxarzi phonology*. MA thesis, University of the Witwatersrand.
- Dingemans, Mark. 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6. 654–672.
- Doke, Clement Martyn. 1935. *Bantu linguistic terminology*. London: Longmans, Green & Co.
- Doke, Clement Martyn. 1947. *Text-book of Zulu grammar*. Johannesburg: Longmans, Green & Co.
- Doke, Clement Martyn. 1954. *The Southern Bantu languages*. London: Oxford University Press for the International African Institute (IAI).
- Doke, Clement Martyn, D. McK. Malcolm & Jonathan M.A. Sikakana. 1958. *English-Zulu Dictionary*. Johannesburg: Witwatersrand University Press.
- Doke, Clement Martyn & S. Machabe Mofokeng. 1957. *Textbook of Southern Sotho Grammar*. Johannesburg: Longmans.
- Donnelly, Simon Scurr. 1990. *Phonology and Morphology of the Noun in Yeeyi*. B.A. Honours thesis, University of Cape Town.

- Donnelly, Simon. 2007. *Aspects of Tone and Voice in Phuthi*. PhD thesis, University of Illinois.
- du Plessis, Menán. forthcoming. Notes on Qing's own language. In Jeremy Hollmann, Menán du Plessis, José M. de Prada-Samper, Jill Weintraub, Wintjes Justine & John Wright (eds.), *On the Trail of Qing and Orpen*. Johannesburg: Witwatersrand University Press.
- Duggan, Ana T., Mark Whitten, Victor Wiebe, Michael Crawford, Anne Butthof, Victor Spitsyn, Sergey Makarov, Innokentiy Novgorodov, Vladimir Osakovsky and Brigitte Pakendorf. 2013. Investigating the Prehistory of Tungusic Peoples of Siberia and the Amur-Ussuri Region with Complete mtDNA Genome Sequences and Y-chromosomal Markers. *PLoS One* 8: e83570.
- Elmslie, Walter Angus. 1891. *Introductory Grammar of the Ngoni (Zulu) Language, as Spoken in Mombera's Country*. Aberdeen: G. & W. Fraser, Belmont Works.
- Fehn, Anne-Maria. forthcoming. Click replacement and loss: an overview with new data from Kalahari Khoe and Ju. In Bonny Sands (ed.), *Handbook of Click Languages*. Leiden: Brill.
- Fisch, Maria. 1977. *Einführung in die Sprache der Mbukushu*. Windhoek: S.W.A. Wissenschaftliche Gesellschaft.
- Fisch, Maria. 1998. *Thimbukushu grammar*. Windhoek: Out of Africa Publishers.
- Fisch, Maria. 2005. *The Mbukushu in Angola (1730-2002). A History of Migration, Flight and Royal Rainmaking*. (History, cultural traditions and innovations in Southern Africa 11). Köln: Rüdiger Köppe.
- Gil, David. 2005. Paralinguistic Usage of Clicks. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 572–575. Oxford New York: Oxford University Press.
- Gowlett, Derek F. 1989. The parentage and development of Lozi. *Journal of African Languages and Linguistics* 11: 127-149.
- Greenberg, Joseph H. 1963. *The languages of Africa*. Bloomington: Indiana University Center in Anthropology, Folklore and Linguistics, and The Hague: Mouton.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti and 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: 13296-13301. 10.1073/pnas.1503793112.
- Güldemann, Tom. 1999. Head-initial meets head-final: nominal suffixes in eastern and southern Bantu from a historical perspective. *Studies in African linguistics* 28: 49-91.
- Güldemann, Tom. 2004. Reconstruction through 'de-construction': the marking of person, gender, and number in the Khoe family and Kwadi. *Diachronica* 21: 251-306.

- Güldemann, Tom. 2005. "Tuu" - a new name for the Southern Khoisan family. In Tom Güldemann (ed.), *Studies in Tuu (Southern Khoisan)*, 3-10. Leipzig: Institut für Afrikanistik, Universität Leipzig.
- Güldemann, Tom. 2008. A linguist's view: Khoe-Kwadi speakers as the earliest food-producers of southern Africa. *Southern African Humanities* 20: 93-132.
- Güldemann, Tom. 2014. 'Khoisan' linguistic classification today. In Tom Güldemann & Anne-Maria Fehn (eds.), *Beyond Khoisan : Historical relations in the Kalahari basin*, 1-40. Amsterdam; Philadelphia: John Benjamins.
- Güldemann, Tom and Edward D. Elderkin. 2010. On external genealogical relationships of the Khoe family. In Matthias Brenzinger & Christa König (eds.), *Khoisan languages and linguistics: Proceedings of the 1st International Symposium January 4-8, 2003, Riezlern/Kleinwalsertal*, 15-52. Köln: Rüdiger Köppe.
- Güldemann, Tom and Mark Stoneking. 2008. A Historical Appraisal of Clicks: A Linguistic and Genetic Population Perspective. *Annual Review of Anthropology* 37: 93-109.
- Gunnink, Hilde. 2014. The Grammatical Structure of Sowetan Tsotsitaal. *Southern African Linguistics and Applied Language Studies* 32: 161-171.
- Gunnink, Hilde. forthcoming. Click loss and click insertion in Fwe. In Bonny Sands (ed.), *Handbook of Click Languages*. Leiden: Brill.
- Gunnink, Hilde, Bonny Sands, Brigitte Pakendorf and Koen Bostoen. 2015. Prehistoric language contact in the Kavango-Zambezi transfrontier area: Khoisan influence on southwestern Bantu languages. *Journal of African Languages and Linguistics* 36: 193-232.
- Hallowes, D.P. 1942. *A grammar of Baca*. MA thesis, University of the Witwatersrand.
- Hammond-Tooke, W. D. 1998. Selective Borrowing? The Possibility of San Shamanistic Influence on Southern Bantu Divination and Healing Practices. *The South African Archaeological Bulletin* 53: 9-15.
- Hammond-Tooke, W. D. 1999. Divinatory Animals: Further Evidence of San/Nguni Borrowing? *The South African Archaeological Bulletin* 54: 128-132.
- Heine, Bernd and Henry Honken. 2010. The Kx'a Family. *Journal of Asian and African Studies* 79: 5-36.
- Henn, Brenna M., Christopher Gignoux, Alice A. Lin, Peter J. Oefner, Peidong Shen, Rosaria Scozzari, Fulvio Cruciani, Sarah A. Tishkoff, Joanna L. Mountain and Peter A. Underhill. 2008. Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa. *Proceedings of the National Academy of Sciences of the United States of America* 105: 10693-10698.
- Herbert, Eugenia W. 2012. Elusive Frontiers: Precolonial Mining in Sub-Saharan Africa. In Katja Werthmann & Tilo Grätz (eds.), *Mining frontiers in Africa : anthropological and historical perspectives*, 23-32. Köln: Rüdiger Köppe.

- Herbert, Robert K. 1990. The relative markedness of click sounds: evidence from language change, acquisition and avoidance. *Anthropological linguistics* 32: 120-138.
- Herbert, Robert K. 2002. The sociohistory of clicks in southern Bantu. In R. Mesthrie (ed.), *Language in South Africa*, 297-315. Cambridge: Cambridge University Press.
- Irvine, Judith T. and Susan Gal. 2000. Language Ideology and Linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of Language: Ideologies, Politics, and Identities*, 35-83. Santa Fe: School of American Research Press.
- Job, Sylvanus. 2011. *Features of a Sesfontein dialect of Khoekhoegowab as manifested in a corpus analysed by means of the Field Linguist's Toolbox programme*. BA thesis, University of Namibia, Windhoek.
- Job, Sylvanus. 2014. Features of a Sesfontein dialect of Khoekhoegowab. In *Linguistisches Kolloquium des Seminars für Afrikawissenschaften*. Humboldt-Universität zu Berlin.
- Jobling, M.A. and C. Tyler-Smith. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics* 4: 598-612.
- Jolly, Pieter. 1996. Symbiotic Interaction Between Black Farmers and South-Eastern San: Implications for Southern African Rock Art Studies, Ethnographic Analogy, and Hunter-Gatherer Cultural Identity. *Current Anthropology* 37: 277-305.
- Kerswill, Paul. 2006. Migration and language. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Sociolinguistics. An International Handbook of the Science of Language and Society*, 2271-2285. Berlin: Mouton-De Gruyter.
- Kiyaga-Mulindwa, David. 1993. The Iron Age Peoples of East-Central Botswana. In Thurstan Shaw, Peter Sinclair, Bassey W. Andah & Alex Ikechukwu Okpoko (eds.), *The Archaeology of Africa : food, metals, and towns*, 386-390. London ; New York: Routledge.
- Klieman, Kairn A. 2003. *"The Pygmies were our compass": Bantu and Batwa in the history of west central Africa, early times to c. 1900 C.E* (Social history of Africa). Portsmouth, NH: Heinemann.
- Kloppers, J. K., D. Nakare, L. M. Isala and A. W. Bredell. 1994. *Bukenkango Rukwangali-English = English-Rukwangali dictionary*. Windhoek, Namibia: Gamsberg Macmillan.
- Knight, Alec, Peter A. Underhill, Holly M. Mortensen, Lev A. Zhivotovsky, Alice A. Lin, Brenna M. Henn, Dorothy Louis, Merritt Ruhlen and Joanna L. Mountain. 2003. African Y chromosome and mtDNA divergence provides insight into the history of click languages. *Current Biology* 13: 464-473.
- Lanham, Leonard Walter and D. P. Hallows. 1956. Linguistic relationships and contacts expressed in the vocabulary of eastern Bushman. *African studies* 15: 45-48.
- Lentz, Carola. 1995. „Tribalismus“ und Ethnizität in Afrika — ein Forschungsüberblick. *Leviathan* 23: 115-145.

- Lev-Ari, Shiri, Marcel San Giacomo and Sharon Peperkamp. 2014. The effect of domain prestige and interlocutors' bilingualism on loanword adaptations. *Journal of Sociolinguistics* 18: 658-684.
- Louw, Jacobus Abraham. 1977a. The adaption of non-click Khoi consonants in Xhosa. In Anthony Traill (ed.), *Khoisan linguistic studies 3*, 74-92. Johannesburg: University of the Witwatersrand.
- Louw, Jacobus Abraham. 1977b. Clicks as loans in Xhosa. In Jan W. Snyman (ed.), *Bushman and Hottentot linguistic studies 1975.*, 82-100. Pretoria: University of South Africa (UNISA).
- Lukusa, Stephen T. M. 2009. *Shiyeyi-English dictionary* (Languages of the world. Dictionaries 42). München: Lincom Europa.
- Lukusa, Stephen T. M. and Kemmony C. Monaka. 2008. *Shekgalagari grammar: A descriptive analysis of the language and its vocabulary* (CASAS book series). Cape Town: Centre for Advanced Studies of African Society.
- Mabille, Adolphe, H. Dieterlen and R.A. Paroz. 1961. *Southern Sotho-English Dictionary*. Morija: Morija Sesuto Book Depot.
- Maddieson, Ian. 2003. The sounds of the Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 15-41. London; New York: Routledge.
- Marks, Sarah J., Francesco Montinaro, Hila Levy, Francesca Brisighelli, Gianmarco Ferri, Stefania Bertocini, Chiara Batini, George B.J. Busby, Charles Arthur, Peter Mitchell, Brian A. Stewart, Ockie Oosthuizen, Erica Oosthuizen, Maria Eugenia D'Amato, Sean Davison, Vincenzo Pascali and Cristian Capelli. 2015. Static and moving frontiers: the genetic landscape of Southern African Bantu-speaking populations. *Molecular Biology and Evolution* 32: 29-43.
- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.
- McLaren, James. 1948. *A Xhosa grammar*. London: Longmans, Green & Co.
- McLaren, James. 1968. *A New Concise Xhosa-English Dictionary*. Cape Town: Longmans.
- Meeussen, Achille E. 1967. Bantu Grammatical Reconstructions. *Africana Linguistica* 3: 79-121.
- Megaw, J. V. S. 1977. *Hunters, Gatherers and First Farmers beyond Europe: An Archaeological Survey*. Leicester: Leicester University Press.
- Meinhof, Carl. 1910. *Grundriß einer Lautlehre der Bantusprachen nebst Anleitung zur Aufnahme von Bantusprachen - Anhang: Verzeichnis von Bantuwortstämmen*. Berlin: Dietrich Reimer (Ernst Vohsen).
- Mitchell, Peter. 2002. *The Archaeology of Southern Africa*. Cambridge: Cambridge University Press.
- Mkanganwi, K.G. 1972. The relationships of coastal Ndaui to the Shona dialects of the interior. *African Studies* 31: 111-137.

- Moeketsi, Rosemary H. 1991. *English-Southern Sotho-Afrikaans: the concise trilingual pocket dictionary*. Parklands: Ad Donker.
- Möhlig, Wilhelm Johann Georg. 1967. *Die Sprache der Dciriku. Phonologie, Prosodologie und Morphologie*. PhD thesis, Universität zu Köln.
- Möhlig, Wilhelm Johann Georg and Karl Peter Shiyaka-Mberema. 2005. *A dictionary of the Rumanyo language : Rumanyo-English / English-Rumanyo : including a grammatical sketch* (Southern African languages and cultures v. 2). Köln: Rüdiger Köppe.
- Monaka, Kemmony C. and Patricia Lepekoane. 2008. How the !Xóõ became the Balala: Socio-economic and linguistic factors considered. In Sonja Ermisch (ed.), *Khoisan Languages and Linguistics. Proceedings of the 2nd International Symposium January 8-12, 2006, Riezlern/Kleinwalsertal*, 265-278. Köln: Rüdiger Köppe Verlag.
- Mous, Maarten. 2001. Paralexification in language intertwining. In Norval Smith & Tonjes Veenstra (eds.), *Creolization and contact*, 113-123. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Mous, Maarten. 2003. *The making of a mixed language : the case of Ma'a/Mbugu* (Creole language library 26). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Moyo, C. T. 1995. Language contact and language change: the case for chiTumbuka in Northern Malawi. *South African Journal of African Languages* 15: 186-191.
- Nakagawa, Hiroshi. 2014. G|ui Dictionary. Ms. Version April 2014.
- Nathan, Geoffrey S. 2001. Clicks in a Chinese nursery rhyme. *Journal of the International Phonetic Association* 3: 223-228.
- Naumann, Christfried and Hans-Jörg Bibiko. 2015. Southern Africa as a phonological area. Paper presented at "Speaking (of) Khoisan", a Symposium Reviewing Southern African Prehistory, MPI-EVA, Leipzig.
- Ngonyani, Deogratias S. 2001. The evolution of Tanzanian Ngoni. *Sprache und Geschichte in Afrika* 16/17: 321-354.
- Novembre, J., T. Johnson, K. Bryc, Z. Kutalik, A.R. Boyko, A. Auton, A. Indap, et al. 2008. Genes mirror geography within Europe. *Nature* 456(7218). 98-101.
- Nurse, G.T., J. S. Weiner and Trefor Jenkins. 1985. *The Peoples of Southern Africa and their Affinities*. Oxford: Clarendon Press.
- Nussbaum, Loren V., Sandlane D. Zwane and T. T. Ginindza. 1969. *Understanding and Speaking SiSwati*. Washington DC: Center for Applied Linguistics.
- Pakendorf, Brigitte. 2009. Intensive Contact and the Copying of Paradigms: An Even Dialect in Contact with Sakha (Yakut). *Journal of Language Contact Varia* 2: 85-110.
- Pakendorf, Brigitte. 2014a. Historical linguistics and molecular anthropology. In Claire Bowerman & Bethwyn Evans (eds.), *The Routledge Handbook of Historical Linguistics*, 627-641. London: Routledge.

- Pakendorf, Brigitte. 2014b. Molecular anthropological perspectives on the Kalahari Basin Area. In Tom Güldemann & Anne-Maria Fehn (eds.), *Beyond 'Khoisan'. Historical relations in the Kalahari Basin*, 45-68. Amsterdam: John Benjamins.
- Pakendorf, Brigitte. 2015. A comparison of copied morphemes in Sakha (Yakut) and Èven. In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed Morphology*, vol. 157-187. Berlin, Boston, Munich: De Gruyter Mouton.
- Pakendorf, Brigitte and Mark Stoneking. 2005. Mitochondrial DNA and human evolution. *Annual Review of Genomics and Human Genetics* 6: 165-183.
- Passy, Paul. 1914. La langue Thonga. *Miscellanea Phonetica* 1. 27–32.
- Pelling, J.N. 1966. *A Practical Ndebele Dictionary*. Bulawayo: Daystar.
- Pelling, James N. & Pamela Pelling. 1974. *Lessons in Ndebele*. Salisbury: Rhodesia Literature Bureau [by] Longman Rhodesia.
- Persson, J.A. 1932. *Outlines of a Tswa Grammar, with Practical Exercises*. Cleveland: Central Mission Press.
- Phillipson, David W. 2005. *African archaeology*. Cambridge; New York: Cambridge University Press.
- Pickrell, Joseph K., Nick Patterson, Chiara Barbieri, Falko Berthold, Linda Gerlach, Tom Güldemann, Blesswell Kure, Sununguko Wata Mpoloka, Hiroshi Nakagawa, Christfried Naumann, Mark Lipson, Po-Ru Loh, Joseph Lachance, Joanna Mountain, Carlos D. Bustamante, Bonnie Berger, Sarah A. Tishkoff, Brenna M. Henn, Mark Stoneking, David Reich and Brigitte Pakendorf. 2012. The genetic prehistory of southern Africa. *Nature Communications* 3:
- Pleurdeau, David, Emma Imalwa, Florent Détoit, Joséphine Lesur, Anzel Veldman, Jean-Jacques Bahain and Eugène Marais. 2012. “Of sheep and men”: earliest direct evidence of caprine domestication in southern Africa at Leopard Cave (Erongo, Namibia). *PLoS One* 7: e40340.
- Pongweni, Alec J. C. 1990. *Studies in Shona phonetics: an analytical review*. Harare: University of Zimbabwe.
- Ranger, Terence. 1993. The invention of tradition revisited: the case of colonial Africa. In Terence Ranger & Olufemi Vaughan (eds.), *Legitimacy and the State in Twentieth-Century Africa*, 62-111. Oxford: Palgrave Macmillan.
- Ross, Malcolm. 2003. Diagnosing prehistoric language contact. In Raymond Hickey (ed.), *Motives for language change*, 174-198. Cambridge: Cambridge University Press.
- Ross, Malcolm. 2013. Diagnosing Contact Processes from their Outcomes: The Importance of Life Stages. *Journal of Language Contact* 6: 5-47.
- Rousset, Karine. 2003. *To be Khwe means to suffer: Local dynamics, imbalances and contestations in the Caprivi Game Park*. MA thesis, University of Cape Town.

- Rycroft, David K. 1981. *Concise SiSwati Dictionary: SiSwati-English/English-SiSwati*. Pretoria: J.L. van Schaik.
- Salas, Antonio, Martin Richards, Tomás De la Fe, María-Victoria Lareu, Beatriz Sobrino, Paula Sánchez-Diz, Vincent Macaulay and Ángel Carracedo. 2002. The making of the African mtDNA landscape. *American Journal of Human Genetics* 71: 1082-1111.
- Sands, Bonny. 2007. The contribution of language documentation to historical phonology. In Peter K. Austin, Oliver Bond & David Nathan (eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory*, 209-219. London: SOAS.
- Sands, Bonny. 2010. Juu subgroups based on phonological patterns. In Matthias Brenzinger & Christa König (eds.), *Khoisan Languages and Linguistics. Proceedings of the 1st International Symposium, January 4-8, 2003, Riezlern/Kleinwalsertal*, 85-114. Köln: Rüdiger Köppe.
- Sands, Bonny and Tom Güldemann. 2009. What click languages can and can't tell us about language origins. In R.P. Botha & C. Knight (eds.), *The Cradle of Language*, 204-218. Oxford ; New York: Oxford University Press.
- Schenck, Marcia. 2008. Land, Water, Truth, and Love - Visions of Identity and Land Access: From Bain's Bushmen to Khomani San. B.A. Honours Thesis. <https://ida.mtholyoke.edu/xmlui/handle/10166/735> (18 July, 2016).
- Schlebusch, Carina M., Marlize Lombard and Himla Soodyall. 2013. MtDNA control region variation affirms diversity and deep sub-structure in populations from southern Africa. *BMC Evolutionary Biology* 13:
- Schlebusch, Carina M., Pontus Skoglund, Per Sjödín, Lucie M. Gattepaille, Dena Hernandez, Flora Jay, Sen Li, Michael De Jongh, Andrew Singleton, Michael G. B. Blum, Himla Soodyall and Mattias Jakobsson. 2012. Genomic Variation in Seven Khoisan Groups Reveals Adaptation and Complex African History. *Science* 338: 374-379.
- Segobye, Alinah. 1998. Early Farming Communities. In Ditswa Mmung (ed.), *The Archaeology of Botswana*, 101-114. Gaborone: Pula Press and The Botswana Society.
- Seidel, Frank. 2008. *A Grammar of Yeyi. A Bantu Language of Southern Africa*. (Grammatical Analyses of African Languages. 33). Köln: Rüdiger Köppe.
- Seidel, Frank. 2009. Layered Language Genesis in the 'Catch Basin' of the Linyanti and Okavango Swamps: The Case of Yeyi. *Sprache und Geschichte im Afrika* 20: 231-264.
- Semino, O., G. Passarino, L. Quintana-Murci, A. Liu, J. Beres, A. Czeizel & A.S. Santachiara-Benerecetti. 2000. MtDNA and Y chromosome polymorphisms in Hungary: inferences from the palaeolithic, neolithic and Uralic influences on the modern Hungarian gene pool. *Eur J Hum Genet* 8(5). 339-46.
- Shabangu, Thos M & J. J. Swanepoel. 1989. *Isihlathululimezwi: an English-South Ndebele Dictionary*. Cape Town: Maskew Miller Longman.



- Silberbauer, George B. and Adam J. Kuper. 1966. Kgalagari masters and Bushmen serfs: some observations. *African Studies* 25: 171-179.
- Skhosana, Philemon Buti. 2009. *The linguistic relationship between southern and northern Ndebele*. PhD thesis, University of Pretoria.
- Sommer, Gabriele and Rainer Voßen. 1992. Schnalzwörter im Yei (R.41). *Afrika und Übersee* 75: 1-42.
- Soodyall, Himla, Heeran Makkan, Philip Haycock and Thijessen Naidoo. 2008. The genetic prehistory of the Khoe and San. *Southern African Humanities* 20: 37-48.
- Spiss, Cassian. 1904. Kingoni und Kisutu. *Mittheilungen des Seminars für Orientalische Sprachen* 7: 270-414.
- Stoneking, Mark. forthcoming. *An Introduction to Molecular Anthropology*. Malden MA: Wiley-Blackwell.
- Taljaard, P. C., J. N. Khumalo and Sonja E. Bosch. 1991. *Handbook of Siswati*. Pretoria: J. L. van Schaik.
- Tarimo, Aquiline. 2011. *Ethnicity, Citizenship and State in Eastern Africa*. Oxford: African Books Collective.
- Thiem, Maarit and Brian T. Jones. 2014. Kavango Region. 'Scraping the Pot': San in Namibia Two Decades After Independence. In Ute Dieckmann, Maarit Thiem, Erik Dirks & Jennifer Hays (eds.), *Land, Environment and Development Project of the Legal Assistance Centre and Desert Research Foundation of Namibia*, 329-364. Windhoek: Land, Environment and Development Project of the Legal Assistance Centre and Desert Research Foundation of Namibia.
- Tlale, One. 2005. *The phonetics and phonology of Sengwato, a dialect of Setswana (Botswana)*. PhD thesis, UMI, University of Georgetown.
- Tömöry, Gyöngyvér, Bernadett Csányi, Erika Bogácsi-Szabó, Tibor Kalmár, Agnes Czibula, Aranka Csoz, Katalin Priskin, et al. 2007. Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations. *American Journal of Physical Anthropology* 134(3). 354–368. doi:10.1002/ajpa.20677.
- Trail, Anthony. 2007. !Kha-ka Hhouiten Hhouiten 'The rush of the storm': The linguistic death of |Xam. In *Claim to the Country: the Archive of Lucy Lloyd and Wilhelm Bleek*, edited by P. Skotnes. Athens: Ohio University Press.
- Trail, Anthony and Rainer Voßen. 1997. Sound change in the Khoisan languages: new data on click loss and click replacement. *Journal of African languages and linguistics* 18: 21-56.
- Uren, Caitlin, Minju Kim, Alicia R. Martin, Dean Bobo, Christopher R Gignoux, Paul D. van Helden, Marlo Moller, Eileen G. Hoal & Brenna M Henn. 2016. Fine-scale human population structure in southern Africa reflects ecological boundaries. <http://biorxiv.org/content/early/2016/02/03/038729>.

- Vail, Leroy (ed.) 1989. *The Creation of Tribalism in Southern Africa*. Berkeley and Los Angeles: University of California Press.
- Van Coetsem, Frans. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Heidelberg: C. Winter Universitätsverlag.
- van der Merwe, D. F. and Isaac Schapera. 1943. *A comparative study of Kgalagadi, Kwena and other Sotho dialects*. Cape Town: University of Cape Town.
- Vansina, Jan. 1995. New Linguistic Evidence and the Bantu Expansion. *Journal of African History* 36: 173-195.
- Vansina, Jan. 2004. *How societies are born: governance in West Central Africa before 1600*. Charlottesville: University of Virginia Press.
- Verdu, Paul, Noémie S. A. Becker, Alain Froment, Myriam Georges, Viola Grugni, Lluís Quintana-Murci, Jean-Marie Hombert, et al. 2013. Sociocultural behavior, sex-biased admixture, and effective population sizes in Central African Pygmies and non-Pygmies. *Molecular Biology and Evolution* 30(4). 918–937. doi:10.1093/molbev/mss328
- Visser, Hessel. 2001. *Naro dictionary : Naro-English, English-Naro*. Gantsi, Botswana: Naro Language Project.
- Voßen, Rainer. 1997. *Die Khoe-Sprachen: ein Beitrag zur Erforschung der Sprachgeschichte Afrikas* (Quellen zur Khoisan-Forschung.). Köln: Rüdiger Köppe.
- Walsh, Martin T. 2006. A click in Digo and its historical interpretation. *Azania* 41: 159-166.
- Wilmsen, Edwin N. and Rainer Voßen. 1990. Labour, Language and Power in the Construction of Ethnicity in Botswana. *Critique of Anthropology* 10: 7-37.
- Winford, Donald. 2010. Contact and borrowing. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 170-187. Oxford: Wiley-Blackwell.
- Wood, Elizabeth T., Daryn A. Stover, Christopher Ehret, Giovanni Destro-Bisol, Gabriella Spedini, Howard McLeod, Leslie Louie, Mike Bamshad, Beverly I. Strassmann, Himla Soodyall and Michael F. Hammer. 2005. Contrasting patterns of Y chromosome and mtDNA variation in Africa: Evidence for sex-biased demographic processes. *European Journal of Human Genetics* 13: 867-876.
- Wright, John. 2007. Bushman Raiders Revisited. In Pippa Skotnes (ed.), *Claim to the Country. The archive of Lucy Lloyd and Wilhelm Bleek*, 119-129. Athens: Ohio University Press.
- Wynne, R. C. 1980. *English-Mbukushu dictionary*. Amersham: Avebury Publishing Co.
- Ziervogel, Dirk. 1952. *A grammar of Swazi (siSwati)*. Johannesburg: Witwatersrand University Press.
- Ziervogel, Dirk. 1959. *A Grammar of Northern Transvaal Ndebele*. Pretoria: van Schaik.

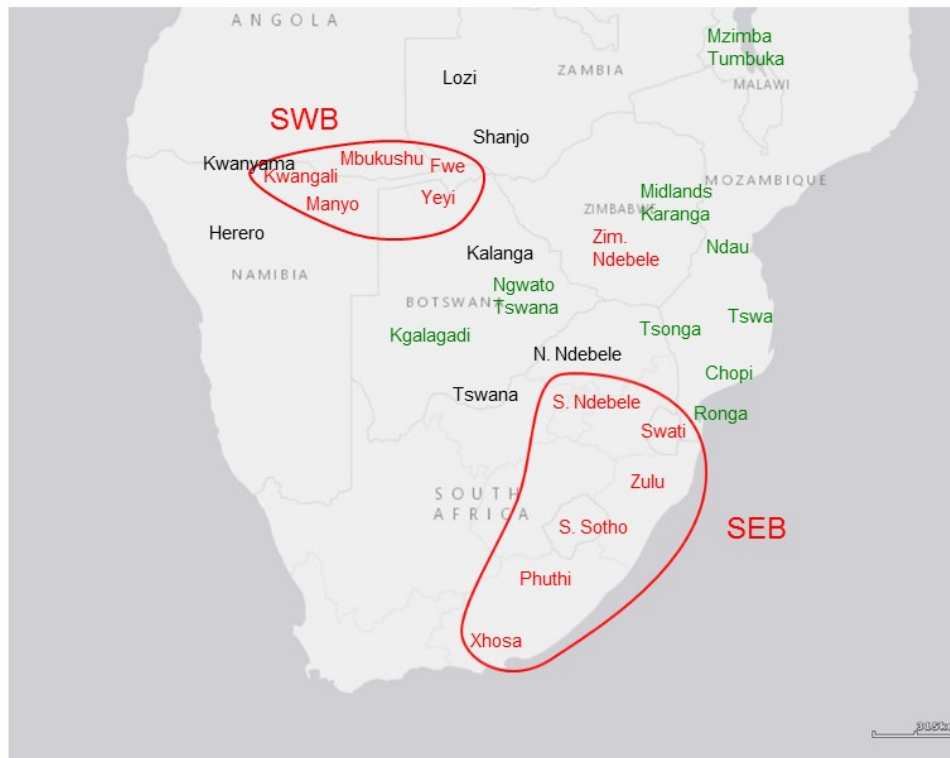
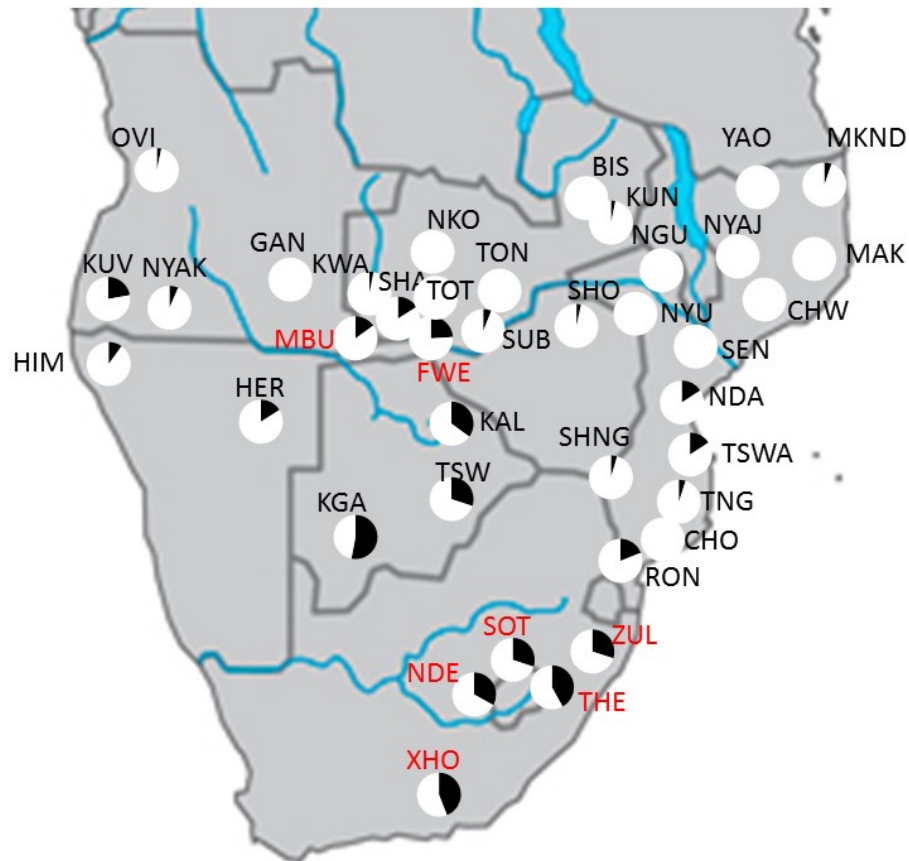


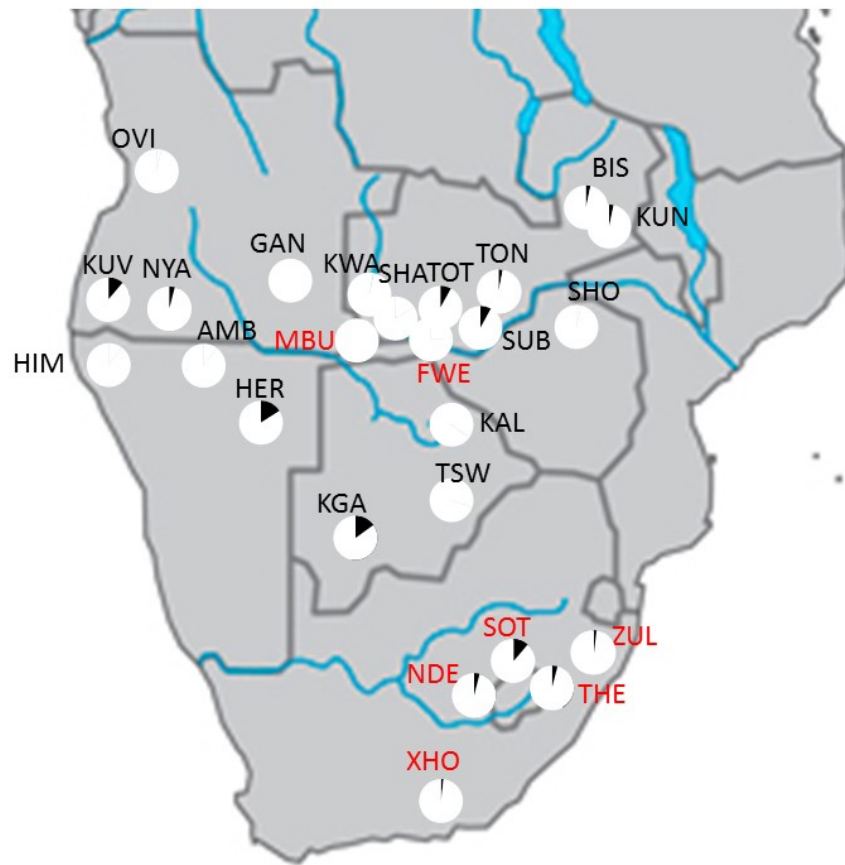
Figure 1 *Approximate distribution of the Bantu languages in southern Africa mentioned in the text: with clicks as regular consonants (red), with marginal clicks (green), and without click consonants (black)*



**Figure 2a** Proportion of autochthonous genetic lineages (black) in Bantu-speaking populations of southern Africa. The populations speaking click languages are labelled in red. Abbreviations: amb = Ambo (aka Ovambo); bis = Bisa; cho = Chopi; chw = Chwabo; gan = Ganguela; her = Herero; him = Himba; kal = Kalanga; kga = Kgalagadi; kun = Kunda; kuv = Kuvale; kwa = Kwamashi; mak = Makhuwa; mknd = Makonde; mbu = Mbukushu; nda = Ndau; nde = Ndebele; ngu = Nguni; nko = Nkoya; nyak = Nyaneka; nyaj = Nyanja; nyu = Nyungwe; ovi = Ovimbundu; ron = Ronga; sen = Sena; shng = Shangaan; sha = Shanjo; sho = Shona; sot = Sotho; sub = Subiya; the = Thembu; tng = Tonga (Mozambique); ton = Tonga (Zambia); tot = Totela; tsw = Tswana; xho = Xhosa; zul = Zulu. a. Mitochondrial dna haplogroups L0d and L0k.

data from: salas et al. (2002); castri et al. (2009); coelho et al.

(2009); de filippo et al. (2010); barbieri et al. (2013b); barbieri et al. (2014a); barbieri et al. (2014b); marks et al. (2015)



**Figure 2b** Y-chromosomal haplogroups A3b1 and B2b

data from: wood et al. (2005); coelho et al. (2009); de filippo et al. (2010); marks et al. (2015); bajić et al. (in preparation)