



HAL
open science

Swahili verbs and the value of abstractive accounts for agglutinating inflection

Louise Esher

► **To cite this version:**

Louise Esher. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure*, 2022, 15 (2), pp.148-173. 10.3366/word.2022.0204 . hal-03879313

HAL Id: hal-03879313

<https://hal.science/hal-03879313>

Submitted on 19 Dec 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.

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

Swahili verbs and the value of abstractive accounts for agglutinating inflection

Abstract

Within ‘word-based’, ‘paradigm-based’ or ‘abstractive’ models of inflectional systems (Blevins 2006, 2016), only full inflected wordforms are considered primitives; subword strings are treated not as distinct entities, but as abstract generalisations inferred by speakers across multiple inflected forms. These models stand in contrast to ‘constructive’ approaches, which proceed from individual, distinct subword units to full words.

An argument consistently adduced in favour of abstractive approaches is that they afford a descriptive advantage regarding ‘fusional’ systems characterised by pervasive non-canonical exponence, such as cumulative exponence, extended exponence, and morphomic structure (Stump 2016: 17–30). Via an exploration of inflectional phenomena including non-canonical exponence and arbitrary distributional regularities in the verb inflection of standard Swahili, a language usually described as exemplifying ‘agglutinating’ inflection and amenable to constructive, morpheme-based, analyses, this paper will argue that abstractive systems are equally applicable to ‘agglutinating’ inflection, offering greater empirical plausibility and in some cases descriptive advantage.

Keywords

inflectional typology; agglutinating inflection; non-canonical exponence; morphemes; abstractive morphology; paradigm-based theories

1. Introduction

Paradigm-based theories of inflection are in principle applicable to any inflectional system, and empirically plausible since they reflect speakers' real-life exposure to full inflected forms (see e.g. Blevins 2006, 2016, O'Neill 2014). In practice, such theories are framed as necessary for inflectional systems of FUSIONAL types, with pervasive non-canonical exponence and irreducibly morphological (MORPHOMIC) structure; by contrast, for systems classed as AGGLUTINATING, morpheme-based analyses may be considered descriptively adequate (Deen & Hyams 2006, Stump 2016: 14–17, Meadows 2020; see also Katamba 1978 for expressions of this view and an early challenge to it). Since the concurrent theories are founded on incompatible claims about the cognitive organisation of inflection and the morphological primitives which speakers manipulate, the apparently open question of whether agglutinating systems should be analysed within paradigm-based or morpheme-based frameworks is not anodyne.

This paper argues that paradigm-based approaches are viable and of descriptive usefulness for agglutinating systems, taking as a case study verb inflection in standard Swahili (North-East Coast Bantu), a language consistently cited as a canonical exemplar of agglutinative morphology.

As is the case for other Bantu languages, Swahili verb inflection is conventionally analysed as involving concatenation of exponents according to a position-class template (Meussen 1967, Nurse 2008, Stump 2016, Goldsmith & Mpiranya forthcoming). The template comprises a series of slots, and an inventory of the exponents which can appear within each slot; the slots are sometimes given labels denoting morphological features or feature values.

An illustrative example is shown in Table 1 for the synthetic verb forms *wanasoma* [wana'soma] 'they are reading' and *hawasomi* [hawa'somi] 'they do not read'. As will be explored in more detail throughout the paper, while Swahili verb forms can readily be segmented into distinct substrings, and these substrings assigned to position classes, the relationship between individual substrings or position classes and meaning is not canonical: extended exponence, cumulative exponence, empty morphs and the lack of distinctive exponents for a given feature value are all common. Moreover, observing the distribution of exponents across the array of inflected forms for a given lexeme, or across different groups of lexemes, reveals arbitrary and conventionalised distributional patterns: in other words, morphomic structure (for which see e.g. Aronoff 1994, Round 2015, Stump 2016, Maiden 2018).

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
 Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

| pre-SM | subject marker | secondary negative | tense/aspect | object marker | lexical root | final vowel | post-FV |
|-----------|----------------|--------------------|--------------|---------------|--------------|-------------|---------|
| | <i>wa</i> | | <i>na</i> | | <i>som</i> | <i>a</i> | |
| | 3PL | | PRS.PROG | | ‘read’ | ‘FV’ | |
| <i>ha</i> | <i>wa</i> | | | | <i>som</i> | <i>i</i> | |
| NEG | 3PL | | | | ‘read’ | PRS.NEG | |

Table 1. Illustrative segmentation of *wanasoma* ‘they are reading’ and *hawasomi* ‘they do not read’ according to the conventional position-class template. For convenience of exposition, the empty “derivational extension” slot is omitted.

The existence of morphomic structure in Swahili verb inflection, including inflectional classes and paradigmatic distribution patterns, may come as a surprise. It is a commonplace of the linguistic literature that Swahili constitutes an established example of agglutinating morphology: for example, Evans (2019: 100) mentions “many languages close to the canonical one-form one-meaning agglutinative type (e.g. as exemplified by Swahili, Turkish or Warlpiri)”, citing Swahili as a classic and instantly recognisable example; while Crysmann & Bonami (2017: 144) specifically select Swahili for comparison with Estonian, as an example of an inflectional system which is “transparent” and “compatible with an atomistic view that associates function more directly with individual exponents that serve to express them”. As such, Swahili inflection is considered a proper object for morpheme-based analyses: familiar morphology textbooks such as Spencer (1991: 231) and Aronoff & Fudeman (2005: 29–30) provide Swahili datasets for students to practise morphemic analysis on, exhaustively segmenting word forms into atomic units which each consistently associate a unique form with a unique meaning (see e.g. Anderson 2015: 15–16). The introductory chapter of Stump (2016) selects the Swahili verb form *alikusukuma* [alikusu'kuma] ‘he pushed it’ to illustrate different morphological theories, and characterises this form as “obviously amenable to a morpheme-based analysis” (2016: 16); though note that Stump (2016: 41) will ultimately argue that “even highly agglutinating languages such as Turkish present inflectional paradigms whose structure is non-canonical”, thus requiring analysis within a paradigm-based theory of inflection (a view which this paper shares). The literature on morphemes, meanwhile, is for the most part concerned with Indo-European languages of the ‘fusional’ type with pervasive non-canonical exponence (Maiden et al. 2011, Cruschina et al. 2013, Maiden 2018, Enger 2021¹), though the typological survey of Bach (2018) confirms that inflectional classes are widespread across genetically and typologically distinct languages, and Round (2013, 2015) identifies one specific type of morpheme in the Tangkic language Kayardild. The findings of this paper demonstrate that morphomic structure is also present in Swahili, albeit to a limited extent, indicating that, while the

¹ Herce’s crosslinguistic sample of metamorphomic phenomena, comprising instances of “morphosyntactically unnatural sets of paradigm cells that systematically share (some of) their exponence” (2020: 12), looks well beyond Indo-European (2020: 187–189), while remaining dominated by fusional languages; Bantu languages do not feature. Herce (2020: 185) casts this skewing as a typological inevitability: “[i]t goes without saying that highly isolating or highly agglutinative languages will lack morphemes more frequently than the cross-linguistic average whereas highly fusional, morphologically complex languages will constitute the best breeding ground for morphemes. For this reason, languages and language families with these characteristics will be overrepresented here”. The conclusion of the present study with respect to ‘agglutinating’ inflectional systems (§4.1) will be somewhat different: such systems are likely to exhibit *simpler* metamorphemes than ‘fusional’ inflectional systems, but are no less likely to *exhibit* metamorphemes overall.

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

richness and complexity of morphomic structure may be constrained by inflectional typology, the presence of such structure is not.

2. Swahili viewed from inflectional typology and morphological theory

Within WORD-BASED, PARADIGM-BASED and ABSTRACTIVE models of inflectional systems (Blevins 2006, 2016, O’Neill 2014), only full inflected wordforms are considered primitives; subword strings are treated not as distinct entities, but as abstract generalisations inferred by speakers across multiple inflected forms. These models stand in contrast to CONSTRUCTIVE approaches, which proceed from individual, distinct subword units to full words. Abstractive, paradigm-based approaches are typically considered to afford a descriptive advantage in accounting for ‘fusional’ systems characterised by pervasive non-canonical exponence, such as cumulative exponence, extended exponence, and morphomic structure (Stump 2016: 17–30), since paradigm-based accounts are equipped to capture the existence of predictable, systematic and recurrent patterns of relationship between inflectional forms, independently of the relative canonicity of form-function relationships at a subword level.

For inflectional systems with differing characteristics, a significant open question remains: are such systems amenable to description within abstractive models, and is any descriptive advantage conferred by abstractive models as opposed to constructive models?

In response, this paper explores aspects of the inflectional system of standard Swahili. Originally based on the variety of Unguja on the island of Zanzibar, this is the variety usually codified in grammars for foreign learners, such as Ashton (1944) whose grammar is the source for modern studies in theoretical morphology such as Stump (1993, 2016) and Crysmann and Bonami (2017). As well as being a popular example of ‘agglutinative’ inflection in the morphological literature, standard Swahili offers a particularly straightforward example of the Bantu position-class system, since contrast between inflected verb forms is purely segmental, with no prosodic element: unlike the majority of Bantu languages, Swahili is not a tone language, and stress in Swahili verbs is fixed, consistently falling on the penult (Vitale 1975).

For convenience of exposition, this paper follows the version of the position-class system (1) outlined by Nurse (2008: 40). In this template, a generalised framework for describing verb inflection across Bantu languages, the slots are given descriptive labels referring to the type of material which typically fills them. For some, the label directly identifies morphosyntactic or morphosemantic feature values: SUBJECT MARKER (SM), TENSE/ASPECT (TA), OBJECT MARKER (OM). For others, the label refers in part or exclusively to position: the FINAL VOWEL (FV) and SECONDARY NEGATIVE (NEG₂) slots refer to the phonological shape and morphosyntactic function of the exponents appearing in the slot, but also the position of these exponents; the PRE-SM and POST-FV slots are catch-all slots containing any material preceding the subject marker or following the final vowel. The EXTENSION slot contains material associated with contrasts of valency and argument structure (see also Hyman 2003).

(1) Pre-SM + SM + NEG₂ + TA + OM + root + extension + FV + post-FV

It should be noted that, within the Bantu descriptive tradition, the slots have not always been labelled in terms of meaningful features and values, nor is such labelling a necessity. In the position-class template initially developed by Meeussen (1967) in the context of comparative-historical reconstruction, all slots of the template outside the lexical root and its derivational extensions are labelled exclusively in terms of position rather than in terms of meaningful values. Stump’s (1993,

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
 Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

2001) formal theoretical discussion of Swahili negative forms similarly labels the position-class slots with abstract numerical indices referring simply to distance from the root.

2.1 Swahili in morpheme-based analyses

A prominent concern of current morphological theorists is to develop holistic descriptions of entire inflectional systems, quantifying the overall complexity of the systems, exploring their organisational properties and identifying predictive generalisations about the behaviour of inflected forms (see e.g. Stump & Finkel 2013, Bonami 2014, Maiden 2018, Beniamine 2021, Beniamine et al. 2021, Feist & Palancar 2021, Palancar 2021). By contrast, to the extent that Swahili inflection figures in morphological discussion, it is in the context of illustrating morpheme-based approaches.

A salient example can be found in the opening chapter of Stump (2016), which introduces the reader to basic theoretical concepts including the distinction between morpheme-based and paradigm-based accounts of inflectional systems, and to Stump’s (2001: 1–3) taxonomy of theoretical accounts (compare also Anderson 2015: 26–32). One contrast within this taxonomy relates to how exponents and morphological feature content interact in inflection: INCREMENTAL theories assume that “exponents introduce morphosyntactic properties”, whereas REALISATIONAL theories assume that “morphosyntactic properties license the introduction of their exponents” (Stump 2016: 14). The other relates to how the relationships between exponents and properties are encoded in the grammar: LEXICAL theories list combinations of exponents and properties in the lexicon, whereas INFERENTIAL theories express relationships between wordforms in terms of rules or mappings (Stump 2001: 1); in Stump (2016: 13) this distinction is recast as an opposition between morpheme-based and paradigm-based approaches, referring to the principal organisational structures assumed to be necessary for morphology.

By way of illustration, Stump (2016: 14–16) contrasts incremental morpheme-based, realisational morpheme-based and realisational paradigm-based analyses of the Swahili verb form *alikusukuma* ‘he pushed it’. A position-class analysis of this form is shown in Table 2 (note that for the purposes of Stump’s exposition, the final vowel is integrated with the lexical root).

| pre-SM | subject marker | secondary negative | tense/aspect | object marker | lexical root | final vowel | post-FV |
|--------|----------------|--------------------|--------------|---------------|--------------|-------------|---------|
| | <i>a</i> | | <i>li</i> | <i>ki</i> | <i>sukum</i> | <i>a</i> | |
| | 3SG(1) | | PST | 3SG(7) | ‘push’ | ‘FV’ | |

Table 2. Position class analysis of *alikusukuma* ‘he pushed it’, with subject of class 1 (*m/wa* class, singular) and object of class 7 (*ki/vi* class, singular).

Both morpheme-based analyses share an assumption which Stump (2016: 13) terms the REPRESENTATIONAL DETERMINISM HYPOTHESIS, namely that “[t]he grammatical and semantic content of an inflected word form is fully determined by its representation as a combination of morphemes”. Within the incremental approach, each prefix is associated with a property set in the lexicon, and as one prefix at a time is added to the verb stem, the prefixes individually contribute morphosyntactic property information. Alternatively, within the realisational approach, the morphosyntactic property values are specified directly in syntactic structure, and each prefix is inserted into this structure at the node matching its properties. Under both views, the internal structure of the inflected word form – the presence of each constituent exponent in an inflected word form, and also the relative ordering of

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

exponents – is compositionally attributable to distinct morphological property values individually associated with exponents.

Stump contrasts both morpheme-based approaches with a paradigm-based analysis in which each inflected form of a lexeme is understood to exist in relation to the other inflected forms of that lexeme, and in which relationships between form and function hold at the level of the inflected form. Under such an approach, the form *alikusukuma* is directly associated with the full property set {SUBJ: {m/wa sg}, OBJ: {ki/vi sg}, pst} for the lexeme SUKUMA ‘push’ (Stump 2016: 15–16).

This demonstration lays the ground for a discussion of non-canonical phenomena in exponence (e.g. underdetermination, deponency, overlapping exponence, extended exponence, amorphousness), for which a paradigmatic approach is shown to be of descriptive advantage, and in which examples are drawn principally from fusional languages such as Latin, Bulgarian and Ancient Greek (though note that the monograph will later discuss isolated instances of non-canonical exponence in both Swahili, for which see §3.1 and §3.4, and another agglutinating language, Turkish). Regarding analysis of the form *alikusukuma*, Stump concludes:

Proponents of morpheme-based approaches to inflectional morphology argue that paradigms are epiphenomenal because they play no essential role in analyses such as those sketched in Figures 1.2 and 1.3. But these analyses are ultimately not at all probative, because *alikusukuma* is so obviously amenable to a morpheme-based analysis. Its form-content relations are very simple: its morphology is purely agglutinative, every part of its content receives overt expression, it exhibits neither extended nor overlapping exponence, and its meaning is compositionally computable from its form. In short, *alikusukuma* behaves as an instance of canonical inflection. (Stump 2016: 16–17).

2.2 Theoretical consequences

In exploring the nature of morphological units, objects and structures which should be identified for Swahili, this paper speaks to a broader theoretical question about the relationship between inflectional typology, inflectional theory and morphological cognition.

In an overview of morphological diversity and its articulation with linguistic universals, Matthews (1974: 17–18) raises the question of how desirable it is for languages with typologically contrasting morphology to be described within a single model, with the inference that variation in the choice of models is not only established practice but also empirically justified in at least some cases:

But is each model just as suitable, in fact, for any group of languages? In practice, different types will often be described in different ways, and not merely through inculcated theoretical prejudice (1974: 18).

Blevins (2016: 10) collates several earlier expressions of this view, including the following statement by Robins (1959: 144) to the effect that:

It may also be that while each of the models discussed in this paper is feasible with every language, one of them is more appropriate with certain languages (1959: 144, cited in Blevins 2016: 10).

For languages with inflectional systems classed as ‘fusional’, a strong case can be made for the descriptive inadequacy of morpheme-based approaches, and thus for preferring paradigm-based approaches (Blevins 2016: 4, 51–57, 69–76). Such approaches are not merely compatible with a wide range of non-canonical inflectional phenomena (Stump 2016: 29–30), but actively facilitate accounting

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

for pervasive distributional regularities such as systematic syncretism or morphomic distribution patterns as inherent properties of a given inflectional system (Blevins 2016). As such, paradigm-based approaches confer a tangible descriptive advantage.

Furthermore, paradigm-based approaches which treat inflectional structure as a network of implicative relationships between forms can be seen to approach greater empirical plausibility. In such theories, the fundamental cognitive units assumed for speakers are full inflected wordforms (as encountered in speech); segmentations are viewed as multiple concurrent generalisations across attested inflected forms, removing the need to impose a single, definitive segmentation (Blevins 2006, 2016: 28–29, 74–79).

For languages with systems classed as ‘agglutinating’, such as Swahili, the morphological literature has yet to address the question of whether morpheme-based approaches or paradigm-based approaches are empirically preferable.² Stump (2001: 139–144, 2016: 15–26) illustrates aspects of a paradigm-based analysis for Swahili, showing that such an analysis is possible, and also demonstrates how a paradigm-based analysis can straightforwardly capture an instance of directional syncretism in Turkish (2016: 175–179). Yet it does not incontrovertibly emerge from any of these examples that a paradigmatic approach is overall more successful in accounting for the inflection of the relevant language: as the author makes clear, multiple Swahili and Turkish word forms are entirely amenable to description within morpheme-based approaches, and these approaches do include mechanisms designed to handle isolated instances of cumulative exponence and allomorphy. Both types of approach are thus presented as available and potentially suitable for the description of agglutinating inflection.

The choice between them is not innocent of consequences, insofar as modern theoretical accounts go beyond developing formalisms that reproduce existing systems, to embody claims about the cognitive representation of morphology and the units which speakers manipulate. Paradigm-based approaches assume that the basic cognitive units of morphology are full inflected wordforms, which speakers produce via a process of relational inference, whereas morpheme-based approaches assume that the basic cognitive units of morphology are morphemes, and that speakers produce full wordforms by concatenating morphemes (see e.g. Blevins 2006:537). Thus, treating ‘fusional’ inflection as paradigm-based and ‘agglutinating’ inflection as morpheme-based is not necessarily confined to a difference of theoretical opinion, but also manifests an empirical claim that morphological cognition is organised differently in speakers of different types of language, with a sharp distinction between the two groups of systems.

Such a distinction sits uneasily with the typical view of typological and descriptive morphology that categories such as ‘agglutinating’ and ‘fusional’ inflection are not discrete categories, but rather “quantitative and relative” (Brown 2010: 487–488). In a specifically Bantu context, Nurse (2008: 28) sets out a view reminiscent of the Canonical Typology approach (Corbett 2009), casting the categories of ‘inflectional’ (i.e. ‘fusional’ in the terms of the present study) and ‘agglutinating’ as conceptual extremes which are useful yardsticks for characterising inflectional systems but rarely if ever encountered in real life:

² Katamba (1978) adduces a number of examples of what would today be classed as non-canonical phenomena in Bantu languages, particularly Luganda and Swahili, arguing that these phenomena cannot adequately be captured within Item-and-Arrangement approaches, though without drawing any firm conclusion on the suitability of the ‘agglutinating’ label. Nevertheless, morpheme-based analyses continue to be presented as mainstream for Swahili, often on the basis of reduced data sets.

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

Inflectional and agglutinating are often presented as a contrast [...] In agglutinating languages the strings of morphemes are longer and the morphemes are relatively transparent, having a single shape (no or few allomorphs, often phonologically conditioned) and one meaning, while in inflectional languages the morphemes are often opaque, with multiple allomorphs and fused meanings. In fact, there is a cline, with ideal types at either end and many languages along the cline. Since Bantu languages are nearer the agglutinating end, [...] they are regarded here as agglutinating, even though none is strictly so (Nurse 2008: 28).

Assuming that there is such a cline, but that different types of system, analysed within different theories, require different and largely incompatible cognitive structures, leads to a conceptual problem: it is empirically implausible that, at some point along the cline, morphological cognition should switch abruptly from manipulating morphemes to manipulating full inflected wordforms. The solution which will be proposed here is to switch attention from ‘morphemes’ to MORPHS, or to Blevins’ concept of RECURRENT PARTIALS (2016: 74–75): as an inflectional system more closely approaches the agglutinating pole, subword units do indeed become increasingly perceptible and segmentation of inflected forms less ambiguous; however, the segmentation is relevant only to units of form, and does not identify morphosyntactic-content-bearing morphemes (compare e.g. Blevins 2016: 225–227, Beniamine et al. 2021).

3. Segmenting verb forms: finding parts, ascribing meaning

Via the process of segmenting verb forms into exponents which can be assigned to slots in the overall template, and assigning meaning to these exponents, this section reviews a variety of non-canonical phenomena occurring in Swahili verb inflection, certain of which have received some measure of analysis in existing morphological literature; the typology of non-canonical exponence is based on Stump (2016: 31–42). Together, these phenomena indicate that segmentation via generalisation across word forms is readily achievable and uncontroversial, whereas relationships between form and meaning are more complex and evade the Representational Determinism Hypothesis (see §2.1 above). An illuminating counterpart to this theoretical observation is offered by an experiment conducted by Goldsmith & Mpiranya (forthcoming), in which linguistics students and computational algorithms consistently succeeded in segmenting Swahili verb forms without access to the meaning of the forms.

3.1 Allomorphy, extended and cumulative exponence

A first example of non-canonical exponence in Swahili is adduced by Stump (2016: 40), concerning inflection for polarity in the past tense. In contrast to the canonical expectation that the exponent of a given property should be constant across the different inflectional forms associated with that property (i.e., should not display allomorphy), Stump points out that in Swahili “the property ‘past’ exhibits two distinct exponents in the inflection of a Swahili verb, *li-* in affirmative forms, *ku-* in negative forms” (2016: 40). This phenomenon is illustrated in Table 3 for the verb *-fanya* [‘fajɒ] ‘do, make’.

In addition to allomorphy, the Swahili past forms also illustrate two further non-canonical properties: cumulative exponence and extended exponence. Precisely because of the formal contrast between *-li-* and *-ku-*, correlated with a contrast of polarity, neither of these exponents can be associated biuniquely with the feature content {TENSE: PST}. Instead, *-li-* is associated with the feature content {TENSE: PST, POLARITY: AFF}, while *-ku-* is associated with {TENSE: PST, POLARITY: NEG}: cumulative exponence, in contrast to the canonical situation in which distinct individual properties associated with

an inflectional form should find distinct exponents. Furthermore, *-ku-* co-occurs with a series of exponents *si-*, *h-*, *ha-* characteristic of negative forms in the majority of the inflectional system: thus, in any individual past tense form, negation has multiple exponents, rather than a single exponent as would be canonical. The distribution of *si-*, *h-* and *ha-* is specifically addressed in §3.4 below.

| | pre-SM | subject marker | secondary negative | tense/aspect | object marker | lexical root | final vowel | post-FV |
|-----|-----------|----------------|--------------------|--------------|---------------|--------------|-------------|---------|
| 1SG | | <i>ni</i> | | <i>li</i> | | <i>fany</i> | <i>a</i> | |
| 2SG | | <i>u</i> | | <i>li</i> | | <i>fany</i> | <i>a</i> | |
| 3SG | | <i>a</i> | | <i>li</i> | | <i>fany</i> | <i>a</i> | |
| 1PL | | <i>tu</i> | | <i>li</i> | | <i>fany</i> | <i>a</i> | |
| 2PL | | <i>m</i> | | <i>li</i> | | <i>fany</i> | <i>a</i> | |
| 3PL | | <i>wa</i> | | <i>li</i> | | <i>fany</i> | <i>a</i> | |
| 1SG | | <i>si</i> | | <i>ku</i> | | <i>fany</i> | <i>a</i> | |
| 2SG | <i>h</i> | <i>u</i> | | <i>ku</i> | | <i>fany</i> | <i>a</i> | |
| 3SG | <i>h</i> | <i>a</i> | | <i>ku</i> | | <i>fany</i> | <i>a</i> | |
| 1PL | <i>ha</i> | <i>tu</i> | | <i>ku</i> | | <i>fany</i> | <i>a</i> | |
| 2PL | <i>ha</i> | <i>m</i> | | <i>ku</i> | | <i>fany</i> | <i>a</i> | |
| 3PL | <i>ha</i> | <i>wa</i> | | <i>ku</i> | | <i>fany</i> | <i>a</i> | |

Table 3. Past affirmative and negative forms for *-fanya* ‘do, make’. For discussion of the pre-SM and subject markers in the negative, see §3.4.

A similar distributional pattern involving cumulative exponence, extended exponence and allomorphy is developing for the future indicative (Table 4). Ordinarily, both affirmative and negative future indicative forms are characterised by the presence of the element /*ta*/ in the ‘tense/aspect’ slot (e.g. *watafanya* ‘do.FUT.AFF.3PL’, *hawatafanya* ‘do.FUT.NEG.3PL’). However, for some speakers, /*ta*/ is confined to the affirmative future indicative, while the negative future indicative instead presents /*to*/ in the ‘tense/aspect’ slot (e.g. *watafanya* ‘do.FUT.AFF.3PL’, vs. *hawatofanya* ‘do.FUT.NEG.3PL’). In these speakers’ system, both *-ta-* and *-to-* exhibit cumulative exponence of tense and polarity, while *-to-* also participates in extended exponence of polarity together with *si-* or *h(a)-*.

A noteworthy aspect of this development is the probable source of the form *-to-* (Mpiranya 2015: 43, 58). Swahili infinitives inflect for polarity and the negative infinitive is characterised by the presence of *-to-* (e.g. *kufanya* ‘do.INF.AFF’, *kutofanya* ‘do.INF.NEG’), which is not found elsewhere in the inflectional system:³ as such, the negative infinitive offers a plausible model for the analogical replacement of *-ta-* by *-to-* in the negative future, based on similarity of feature values and of phonological form. Significantly, although in the negative infinitive a morphemic analysis would categorise *-to-* as an exponent of negation alone, or as a cumulative exponent of negation and non-finiteness, the introduction of *-to-* into the negative future does not result in an interpretation of non-finiteness or double negation, nor does it induce the loss of future temporal reference or the creation of forms lacking an initial exponent of negation (e.g. **watofanya* ‘do.FUT.NEG.3PL’). The development

³ The probable source of the negative infinitive is grammaticalisation of a periphrasis in which the verb *kutoa* ‘take out, put out’ preceded a lexical infinitive. In nineteenth-century grammars, the structure *kutoa kupenda* ‘take_out.INF.AFF like.INF.AFF’ is still attested alongside the forms *kutokupenda* ‘like.INF.NEG’, *kutopenda* ‘like.INF.NEG’ (see e.g. Steere 1884: 154–155, Delaunay 1885: 56, Nurse & Hinnebüsch 1993: 417).

affects form, not function, and offers a potential example of speakers having identified and redistributed phonological substrings within inflected wordforms.

| | pre-SM | subject marker | secondary negative | tense/aspect | object marker | lexical root | final vowel | post-FV |
|-----|-----------|----------------|--------------------|--------------|---------------|--------------|-------------|---------|
| 1SG | | <i>ni</i> | | <i>ta</i> | | <i>fany</i> | <i>a</i> | |
| 2SG | | <i>u</i> | | <i>ta</i> | | <i>fany</i> | <i>a</i> | |
| 3SG | | <i>a</i> | | <i>ta</i> | | <i>fany</i> | <i>a</i> | |
| 1PL | | <i>tu</i> | | <i>ta</i> | | <i>fany</i> | <i>a</i> | |
| 2PL | | <i>m</i> | | <i>ta</i> | | <i>fany</i> | <i>a</i> | |
| 3PL | | <i>wa</i> | | <i>ta</i> | | <i>fany</i> | <i>a</i> | |
| 1SG | | <i>si</i> | | <i>ta/to</i> | | <i>fany</i> | <i>a</i> | |
| 2SG | <i>h</i> | <i>u</i> | | <i>ta/to</i> | | <i>fany</i> | <i>a</i> | |
| 3SG | <i>h</i> | <i>a</i> | | <i>ta/to</i> | | <i>fany</i> | <i>a</i> | |
| 1PL | <i>ha</i> | <i>tu</i> | | <i>ta/to</i> | | <i>fany</i> | <i>a</i> | |
| 2PL | <i>ha</i> | <i>m</i> | | <i>ta/to</i> | | <i>fany</i> | <i>a</i> | |
| 3PL | <i>ha</i> | <i>wa</i> | | <i>ta/to</i> | | <i>fany</i> | <i>a</i> | |

Table 4. Future affirmative and negative forms for *-fanya* ‘do, make’.

3.2 Underdetermination and contrasts of shape

Canonically, it is expected that each morphological property should be associated with an overt exponent, but, even in Swahili, this is not always the case. In practice, certain inflectional forms are identified by a distinctive combination of filled and unfilled slots. The most striking example of this in Swahili is the affirmative imperative, shown in Table 5 for the lexeme *-soma* [‘soma] ‘read’, which in the second person singular consists simply of a bare root with optional object prefix (e.g. *soma!* ‘read.IMPV.AFF.2SG’, *kisoma!* ‘read it!’ [where the direct object is a singular of the KI/VI class, such as *kitabu* ‘book’]). For this affirmative imperative form, there is thus no distinctive exponent of subject or of mood. What is distinctive is instead the overall shape of the form, in contrast to all other forms within the inflectional system (or inflectional paradigm, if such a structure is recognised): this insight is most plausibly captured within a descriptive framework which recognises full wordforms as units in their own right.⁴

Second person plural affirmative imperatives exhibit a slightly different pattern, characterised by final vowel [ɛ] and a post-final-vowel exponent [ni], but neither of these exponents is unique to the affirmative imperative: [ni] is traditionally analysed as a plural marker since it can also occur with certain plural object exponents (Mpiranya 2015: 85), while [ɛ] is also found throughout the subjunctive and negative imperative. Thus, a key distinguishing characteristic of the second person plural affirmative imperative is, again, its overall shape, with no prefixes other than an optional object exponent preceding the lexical root. This point is particularly salient for the inflectional class of verbs with invariant final vowels (see §4.1 below), e.g. *sahau* ‘forget.IMPV.AFF.2SG’, *sahauni* ‘forget.IMPV.AFF.2PL’.

⁴ The imperative might be classed as a case of ZERO EXPONENCE, which is itself an intrinsically relational concept: by definition, the putative ‘absence’ of an exponent is not directly observable or identifiable for a single inflected form in isolation, but can only be inferred via comparison between multiple inflected forms.

| | pre- SM | subject marker | secondary negative | tense/aspect | object marker | lexical root | final vowel | post- FV |
|-----|------------|-------------------|-----------------------|--------------|------------------|-----------------|----------------|-------------|
| 2SG | | | | | | <i>som</i> | <i>a</i> | |
| 2PL | | | | | | <i>som</i> | <i>e</i> | <i>ni</i> |
| 2SG | | <i>u</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |
| 2PL | | <i>m</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |

Table 5. Imperative affirmative and negative forms for *-soma* ‘read’.

In the negative imperative, contrast of shape combines with contrast in exponents. For the past negative and future negative forms illustrated in §3.1, exponents of negation occur in initial, Pre-SM position, preceding the subject exponent, and are principally of the form *h(a)-*, whereas for the negative imperative, exponents of negation occur in the secondary negative slot, following the subject exponent, and are consistently of the form *si-*. Negation with *si-* also occurs throughout the subjunctive (Table 6); thus, if feature values are to be assigned to *si-*, they must include values for mood as well as polarity, in another instance of cumulative exponence.

| | pre- SM | subject marker | secondary negative | tense/aspect | object marker | lexical root | final vowel | post- FV |
|-----|------------|-------------------|-----------------------|--------------|------------------|-----------------|----------------|-------------|
| 1SG | | <i>ni</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |
| 2SG | | <i>u</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |
| 3SG | | <i>a</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |
| 1PL | | <i>tu</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |
| 2PL | | <i>m</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |
| 3PL | | <i>wa</i> | <i>si</i> | | | <i>som</i> | <i>e</i> | |

Table 6. Negative subjunctive forms for *-soma* ‘read’.

A final related example is shown in Table 7, comparing the present progressive affirmative and present negative. Whereas in the present progressive affirmative, the tense/aspect slot is standardly occupied by *-na-*, in the present negative there is no exponent corresponding to this slot. Instead, the present negative is characterised both by the distinctive pattern of filled and unfilled slots – the lexical root preceded only by negative and subject exponents, and an optional object exponent – and by the final vowel /i/. Here, the final vowel exhibits cumulative exponence of tense and polarity, while polarity once again receives extended exponence. While in this example it is still possible to associate individual feature values with individual exponents, such that no feature value is entirely without an exponent, it is hard to avoid the sense that such an approach is missing a generalisation: the identification of a form such as *sifanyi* [si'fani] ‘do.PRS.NEG.1SG’ as present negative is based in part on its *inventory* of exponents and overall *shape*, rather than simply the content of the exponents present.

| | pre-SM | subject marker | secondary negative | tense/aspect | object marker | lexical root | final vowel | post-FV |
|-----|-----------|----------------|--------------------|--------------|---------------|--------------|-------------|---------|
| 1SG | | <i>ni</i> | | <i>na</i> | | <i>fany</i> | <i>a</i> | |
| 2SG | | <i>u</i> | | <i>na</i> | | <i>fany</i> | <i>a</i> | |
| 3SG | | <i>a</i> | | <i>na</i> | | <i>fany</i> | <i>a</i> | |
| 1PL | | <i>tu</i> | | <i>na</i> | | <i>fany</i> | <i>a</i> | |
| 2PL | | <i>m</i> | | <i>na</i> | | <i>fany</i> | <i>a</i> | |
| 3PL | | <i>wa</i> | | <i>na</i> | | <i>fany</i> | <i>a</i> | |
| 1SG | | <i>si</i> | | | | <i>fany</i> | <i>i</i> | |
| 2SG | <i>h</i> | <i>u</i> | | | | <i>fany</i> | <i>i</i> | |
| 3SG | <i>h</i> | <i>a</i> | | | | <i>fany</i> | <i>i</i> | |
| 1PL | <i>ha</i> | <i>tu</i> | | | | <i>fany</i> | <i>i</i> | |
| 2PL | <i>ha</i> | <i>m</i> | | | | <i>fany</i> | <i>i</i> | |
| 3PL | <i>ha</i> | <i>wa</i> | | | | <i>fany</i> | <i>i</i> | |

Table 7. Present progressive affirmative and present negative forms for *-fanya* ‘do, make’.

3.3 Syncretism and suppletion

Three lexemes in Swahili are described as having ‘irregular’ imperatives: *-leta* ‘bring’, *-enda* ‘go’, and *kuja* ‘come’, though in each case the distribution and/or the exponent affected is different.

In the case of *-leta*, the irregularity concerns the distribution of final vowels: the second person singular form is not *leta* as expected, but *lete* with final vowel [ɛ], even in the absence of an object exponent; the second person plural form is expected *leteni*. In the case of *-enda*, both singular and plural imperative forms are characterised by initial [n], contrasting with the forms *-kwenda* and *-enda* found in other inflected forms (see §4.2 for additional discussion of this verb): *nenda* ‘go.IMPV.AFF.2SG’, *nendeni* ‘go.IMPV.AFF.2PL’. Finally, in the case of *kuja* ‘come’, the imperative forms are suppletive: *njoo* ‘come.IMPV.AFF.2SG’, *njooni* ‘come.IMPV.AFF.2PL’,⁵ and resist analysis distinguishing the lexical root from the final vowel.

The negative imperative exhibits complete syncretism (Baerman et al. 2005: 59) with the negative subjunctive: i.e. the inflected forms corresponding to both functions are phonologically identical, for any given lexeme. This generalisation remains exceptionless even for verbs with irregular affirmative imperatives, e.g. *usije* ‘come.IMPV.NEG.2SG, come.SBJV.NEG.2SG’, *msije* ‘come.IMPV.NEG.2PL, come.SBJV.NEG.2PL’.

3.4 Negative and subject markers

An obvious problem in the negative forms shown in Tables 3, 4, 6 and 7 concerns the first person singular. Instead of forms with initial **hani-*, as might be expected based on the usual series of exponents occupying the subject marker slot, all negative first person singular forms outside the

⁵ For Mpiranya (2015: 91), the irregular forms are pragmatic alternatives to ‘regular’ *leta*, *kwenda*, *kwendeni*, *kuja*, *kujani*, which are described as “directive and almost rude”; in which case, these three verbs also present overabundance (Thornton 2011).

subjunctive have initial *si-*. The exponent *si-* constitutes a ‘portmanteau morph’, or instance of cumulative exponence, since it is associated with values of multiple distinct features: negative polarity, first person, and singular number. Stump (1993: 144–145) analyses *si-* as instantiating a ‘portmanteau position class’: as *si-* stands in opposition to sequences of exponents occupying the pre-SM and subject marker slots respectively, but cannot be further segmented or definitively assigned to either slot, it is deemed to occupy both slots at once. Within an analysis where the exponents associated with each slot are described by realisational rules exhaustively ordered into discrete blocks, accounting for a portmanteau such as *si-* further requires the introduction of an exceptional type of rule block, spanning the two ordinary rule blocks corresponding to the pre-SM and subject marker slots. The necessity of positing a specific type of position class and rule block is thus driven by theory-internal strict initial assumptions that “ordering relations among rules partition a rule set into totally ordered subsets” (Blevins 2016: 148); in place of special mechanisms to deal with exceptional cases, Blevins argues for a reassessment of the initial assumptions in the light of empirical reality (2016: 146–149).

Based on comparison with negative subjunctive forms (Table 6),⁶ it is tempting to ask whether *si-* in the first person singular is an instance of the ‘secondary negative’ and fills this separate slot instead; but an arbitrary contrast still remains to be accounted for, in that *si-* is always preceded by *ni-* in first person singular negative subjunctive forms, e.g. *nisisome* ‘read.SBJV.NEG.1SG’, whereas in other tense-aspect-mood categories, first person singular negative forms admit only *si-*, without preceding *ni-*, e.g. *sisomi* ‘read.PRS.NEG.1SG’, *sikusoma* ‘read.PST.NEG.1SG’.

Moreover, first person singular forms are not alone in presenting realisations which are not immediately predictable from straightforward concatenation of morphs. Table 8 shows present progressive forms and present negative forms of *-potea* ‘become lost’ for all singular persons and noun classes, with an approximate division into the relevant position slots. Although the subject marker for class 3 and 11 noun subjects is phonologically identical to the subject marker for second person singular subjects, second person singular negative forms have initial *hu-* [hu], while negative forms for subjects of class 3 and 11 have initial *hau-* [ha.u] instead. A similar contrast, albeit less strictly parallel, is observed for the other vowel-initial subject markers: third person singular forms of class 1 have initial *ha-* [ha] in negative forms, while forms with class 9 subjects have initial *hai-* [ha.i]. The series of realisations *si-*, *hu-*, *ha-*, *hau-*, *hali-*, *haki-*, *hai-* occurs systematically across tense-aspect-mood categories (except those in which all persons have the exponent *-si-* in the secondary negative position): it is a recurrent, consistent property of the inflectional system.

⁶ Stump (1993) illustrates the theoretical formalisation of a data set given in a linguistics textbook; as negative subjunctive forms are not included in the initial data set, these do not receive discussion.

| | | PRS.PROG.AFF | | | PRS.NEG | | |
|-----|------|----------------|--------------|--------------|-----------|----------------|--------------|
| | | subject marker | tense/aspect | root+FV | pre-SM | subject marker | root+FV |
| 1SG | | <i>ni</i> | <i>na</i> | <i>potea</i> | | <i>si</i> | <i>potei</i> |
| 2SG | | <i>u</i> | <i>na</i> | <i>potea</i> | <i>h</i> | <i>u</i> | <i>potei</i> |
| 3SG | 1 | <i>a</i> | <i>na</i> | <i>potea</i> | <i>h</i> | <i>a</i> | <i>potei</i> |
| | 3,11 | <i>u</i> | <i>na</i> | <i>potea</i> | <i>ha</i> | <i>u</i> | <i>potei</i> |
| | 5 | <i>li</i> | <i>na</i> | <i>potea</i> | <i>ha</i> | <i>li</i> | <i>potei</i> |
| | 7 | <i>ki</i> | <i>na</i> | <i>potea</i> | <i>ha</i> | <i>ki</i> | <i>potei</i> |
| | 9 | <i>i</i> | <i>na</i> | <i>potea</i> | <i>ha</i> | <i>i</i> | <i>potei</i> |

Table 8. Singular present progressive affirmative and present negative forms of *-potea* ‘become lost’.

Both Ashton (1944), and Stump (2016: 40, discussing Ashton’s data), appear to assume that the second and third person singular forms *hu-*, *ha-* are phonological reductions of the sequences *ha-u-*, *ha-a-* respectively. If only personal forms are considered, a phonological generalisation might be made, to the effect that negative *ha-* remains [ha] before a consonant-initial subject exponent, e.g. *hatupotei* ‘become_lost.PRS.NEG.1PL’, *hampotei* ‘become_lost.PRS.NEG.2PL’, *hawapotei* ‘become_lost.PRS.NEG.3PL’, and reduces to [h] before a vowel-initial subject exponent. Yet the evidence of forms with class 3 and 11 noun subjects shows that the sequence /hau/ is phonologically acceptable in the language. Thus, if *hu-* and *ha-* are to be attributed to a phonological process, it must be admitted that phonology is sensitive to morphosyntactic properties such as noun class agreement, a suboptimal solution. While a concatenative, morpheme-based approach encounters difficulties in accounting for the contrasts among these realisations – seemingly arbitrary in terms of their phonological context, yet entirely systematic and predictable in terms of their paradigmatic distribution – the problem dissolves in an abstractive approach, where such contrasts can be recognised as irreducibly morphological properties of the inflectional system.

3.5 Interim summary

The data discussed in §3.1–§3.3 illustrate how, in Swahili, it is relatively straightforward to divide inflectional forms into substrings and to associate these substrings with position slots (though see §3.4 for more complex and arbitrary phenomena). However, attempts to assign meaning to substrings and to position slots reveal that non-canonical exponence is more pervasive than the traditional ‘agglutinating’ label suggests: many-to-one or one-to-many relationships are common, the inventory of meanings associated with a given substring may appear to shift according to the other substrings with which it occurs in a given inflectional form, and different feature combinations arbitrarily require different combinations of filled and unfilled slots. These properties of the inflectional system indicate that the substrings at stake are units of form rather than genuine morphemes associating form and meaning.

Evidence that substrings are reliably and consistently identifiable without reference to meaning is provided by Goldsmith & Mpiranya’s (forthcoming) experiments in segmentation of Swahili verb forms. These authors describe a linguistics classroom situation in which students were presented with verb forms, but, unlike in typical exercises (Spencer 1991: 231, Gleason 1955 cited in Stump 1993: 129, Aronoff & Fudeman 2005: 29–30), were not provided with translations:

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

We begin by putting a word on the board: *ninasema* [...] Then we ask everyone else to divide it into morphemes. There is silence, of course, because the students think they have no idea what the right answer is. Then we ask them to guess how many morphemes there are here: one? two? more? Students guess there are at least two morphemes, and if pressed, typically offer a cut into *nina-sema*. Then we write *unasema*, and ask them if this allows them to change their minds. Everyone with an opinion opines that the correct cuts are *ni-nasema* and *u-nasema*. [...] Then we consider a third word, *anasema*, and the students feel confirmed in their judgment after the second word, since they can easily extend their hypothesis to *ni/u/a + nasema*. [...] The next word is *ninaona*, and the students easily conclude that there is a break after *nina* (comparing *ninasema* and *ninaona*) and furthermore, the word should be divided up as *ni-na-ona*. The next two words we offer are *ninampiga* and *tunasema*. The first they break up as *ni-na-mpiga*, and the second as *tu-na-sema*. How about *ninawapiga*? That must be *ni-na-wa-piga*, and then they realize we must go back and reanalyze *ninampiga* as *ni-na-m-piga*. [...] The point to bear in mind is that the students have done this without being told what the Swahili words mean in English. At some point we explain that in other linguistics courses, the teacher gives their students the same words along with their translations in to English, but we tell them that we do not think it is necessary to know the meanings of the words to find the morphemes, and that the external form (which is to say, the spelling) is enough to discover the right morphological structure. By the end of the class, we have analyzed about 30 Swahili words, and found the right structure (Goldsmith & Mpiranya forthcoming).

The wordforms⁷ in this exercise are, of course, carefully chosen by the instructors, and presented in a specific order to facilitate comparison by the students. Yet the form of the experiment is significant: the students arrive at the same segmentation that linguists and grammarians would suggest for the relevant forms, without knowledge of the meanings, or indeed any attempt to assign meanings; they do so via a process of abstractive generalisation in which they compare full inflected wordforms and observe recurrent substrings and recurrent patterns of ordering of those substrings. The substrings themselves are never presented in isolation, or as free-standing entities; rather, they are perceived only as generalisations across multiple wordforms (Blevins 2016: 45–47). Although Goldsmith & Mpiranya’s controlled classroom situation is not to be directly compared with native speaker cognition, the processes of abstractive reasoning consciously applied by the students do recall the inferences which native speakers are assumed to rely on in producing inflected forms (Ackerman et al. 2009). Most importantly, it is clear that the ‘morphemes’ identified by the students cannot be understood as morphemes in the classical sense (for which see Anderson 2015), since the segmentation task is entirely divorced from any consideration of meaning. A more accurate characterisation of the task is instead “finding the proper analysis of a word into consecutive morphs” (Goldsmith & Mpiranya forthcoming); the authors subsequently develop a machine-learning algorithm for the same task, capable of producing an analysis close to that of the standard grammatical descriptions (Goldsmith & Mpiranya forthcoming).

4. Arbitrary distributions

The previous sections observe non-canonical properties of Swahili verb inflection at the level of individual exponents illustrating non-canonical relationships between form and function. In this section,

⁷ The items cited here are, in order: *ninasema* ‘say.PRS.PROG.1SG’; *unasema* ‘say.PRS.PROG.2SG’; *anasema* ‘say.PRS.PROG.3SG’; *ninaona* ‘see.PRS.PROG.1SG’; *ninampiga* ‘hit.PRS.PROG.1SG[SBJ].3SG[OBJ]’; *tunasema* ‘say.PRS.PROG.1PL’, *ninawapiga* ‘hit.PRS.PROG.1SG[SBJ].3PL[OBJ]’. Note that the cited items do not illustrate the non-canonical phenomena discussed in §3, or morphomic structure of the type discussed in §4 below.

I turn instead to a wider perspective exploring the systematic distributional patterns observed in verb inflection: paradigmatic distribution patterns, or metamorphemes in the terms of Round (2015), and inflectional classes (RHIZOMORPHOMES in the terms of Round 2015). Arbitrary, stable and conventionalised, such patterns resist motivation in terms of extramorphological factors, yet represent fundamental generalisations, productively exploited by speakers, about the organisation of the inflectional system. As essentially relational objects, their observed behaviour is most accurately modelled within paradigm-based approaches, namely those approaches which recognise implicative relationships of mutual predictability between inflected forms (Blevins 2016: 163–173, Maiden 2018: 9–24).

4.1 Final vowels: metamorphemes and inflectional classes

As seen in the examples discussed in §3, all inflected verb forms have a final vowel, which for the majority of lexemes is selected from the inventory [a], [ɛ], [i]. For verbs such as *soma* ‘read’, *fanya* ‘want’, exemplified above, the paradigmatic distribution of these three vowels is as follows. The vowel [i] occurs only in the present negative indicative. The vowel [ɛ] occurs throughout the subjunctive (affirmative and negative), the negative imperative (all persons), and the second person plural affirmative imperative.⁸ The vowel [a] is a default occurring in all other forms. This pattern of distribution is stable, and may be considered conventionalised; in terms of alignment with morphosyntactic and/or morphosemantic feature values, it is largely arbitrary and does not correspond to natural class distinctions.

The stability of the final vowel contrast is also striking in the light of its very low functional load. In almost all inflected verb forms, final vowels participate in extended exponence. The only contrast of TAM categories borne entirely by the final vowel is that between the aspect-neutral present (Ashton 1944: 36–38) and the affirmative subjunctive, for certain subject person/number/gender combinations: for example, in the third person singular the only formal contrast concerns the final vowel, e.g. *wasoma* ‘read.PRS.AFF.3PL’ vs. *wasome* ‘read.SBJV.AFF.3PL’, whereas in the first person singular the initial syllable also contrasts, e.g. *nasoma* ‘read.PRS.AFF.1SG’ vs. *nisome* ‘read.SBJV.AFF.1SG’. Despite the low functional load of the final vowel contrasts, there is no evidence for elimination of the alternation pattern, or for alignment of the formal contrasts with functional contrasts.

The arbitrary, conventionalised and stable nature of this distributional pattern recalls the morphomic patterns of stem distribution (METAMORPHOMES, in the terms of Round 2015) identified for Indo-European languages, particularly Romance; the parallel becomes clearer if a verb such as *soma* ‘read’ is considered to have three different stems, *-soma*, *-some*, and *-somi*, each distinguished by its final vowel and with a consistent distribution across the array of all inflected verb forms for this lexeme.⁹

The Swahili system is less complex in that each lexeme with final vowel alternation has the same inventory of three final vowel variants, as well as the respective distribution of each variant

⁸ Also some indicative forms with object exponent *ku-* or *wa-* (Mpiranya 2015: 84). Detailed consideration of this point is beyond the scope of the present study.

⁹ Note that in traditional analyses, the final vowel and lexical root are separated because additional material realising contrasts of valency and argument structure can appear between them. Traditionally termed DERIVATIONAL EXTENSIONS, these exponents occur with the default order CAUSATIVE, APPLICATIVE, RECIPROCAL, PASSIVE (CARP TEMPLATE, Hyman 2003: 249). Within a paradigmatic analysis, forms with derivational extensions may potentially be treated either as separate, derivationally related lexemes, or as distinct stems; though consideration of such forms falls outside the scope of the present study.

remaining the same from lexeme to lexeme; by contrast, in Romance, only the distributional patterns recur across lexemes, with the phonological shape of stem allomorphs differing considerably across lexemes (Maiden 2018). Thus, while the near-agglutinating system of Swahili does not preclude the presence of metamorphemes, it is associated with a much simpler manifestation of such patterns.

It is also noteworthy that a subset of verb lexemes have invariant final vowels. Examples of such verbs include *-ishi* ‘live, inhabit’, *-safiri* ‘travel’, *-hitaji* ‘need’, *-subiri* ‘wait’, *-jaribu* ‘try’, *-shukuru* ‘be grateful’, *-ruhusu* ‘permit’, *-samehe* ‘forgive’, *-wemo* ‘include’, *-dharau* ‘despise’, *-sahau* ‘forget’ (for further examples, see e.g. Mpiranya 2015: 83–84, 133–135, Lodhi 2000: 115–120, Schwarz 2004 cited in Versteegh 2014). Table 9 compares forms of verb lexemes with invariant final vowels and forms of verb lexemes with the majority pattern of contrast between [a], [ɛ], and [i]: all these lexemes behave identically with respect to prefixal material, but verbs with invariant final vowels in effect have only a single stem. The distinction between the two types of verbs is arbitrary and lexically specific; it embodies, in fact, another type of morphomic patterning, to wit, an inflectional class distinction.

Verbs with invariant final vowels are sometimes described as ‘Arabic verbs’, reflecting the fact that many such verbs originated as loan-words from Arabic; more recently, loans from English may follow this pattern, such as *-rekodi* ‘record’, *-promoti* ‘promote’, and *-dili* ‘handle the ball’ < *deal* (Dzahene-Quarshie 2012). The emergence of a novel inflectional class for loan verbs is not without precedent in the morphological literature: see e.g. Luís (2011) for the example of verbs of substrate and adstrate origin in Indo-Portuguese.

| | <i>-soma</i> ‘read’ | <i>-fanya</i> ‘do’ | <i>-subiri</i> ‘wait’ | <i>-jaribu</i> ‘try’ | <i>-samehe</i> ‘forgive’ | <i>-sahau</i> ‘forget’ |
|------------------|------------------------|-----------------------|--------------------------|-------------------------|-----------------------------|---------------------------|
| INF.AFF | <i>kusoma</i> | <i>kufanya</i> | <i>kusubiri</i> | <i>kujaribu</i> | <i>kusamehe</i> | <i>kusahau</i> |
| PRS.PROG.AFF.1PL | <i>tunasoma</i> | <i>tunafanya</i> | <i>tunasubiri</i> | <i>tunajaribu</i> | <i>tunasamehe</i> | <i>tunasahau</i> |
| PRS.NEG.1PL | <i>hatusomi</i> | <i>hatufanyi</i> | <i>hatusubiri</i> | <i>hatujaribu</i> | <i>hatusamehe</i> | <i>hatusahau</i> |
| PRS.SBJV.AFF.1PL | <i>tusome</i> | <i>tufanye</i> | <i>tusubiri</i> | <i>tujaribu</i> | <i>tusamehe</i> | <i>tusahau</i> |
| PRS.SBJV.NEG.1PL | <i>tusisome</i> | <i>tusifanye</i> | <i>tusisubiri</i> | <i>tusijaribu</i> | <i>tusisamehe</i> | <i>tusisahau</i> |
| IMPV.AFF.2SG | <i>soma</i> | <i>fanya</i> | <i>subiri</i> | <i>jaribu</i> | <i>samehe</i> | <i>sahau</i> |
| IMPV.AFF.2PL | <i>someni</i> | <i>fanyeni</i> | <i>subirini</i> | <i>jaribuni</i> | <i>sameheni</i> | <i>sahauni</i> |

Table 9. Comparison of verbs with final vowel alternation and verbs with invariant final vowels.

4.2 ‘Short’ or ‘monosyllabic’ verbs

A further inflectional class distinction in Swahili concerns the group of so-called ‘short’ or ‘monosyllabic’ verbs (for which see e.g. Mpiranya 2015: 94–100). This group comprises nine lexemes with monosyllabic stems, such as *-ja* [ja] ‘come’, *-wa* [wa] ‘be’, *-nywa* [ɲwa] ‘drink’; verbs of the ‘short’ class exhibit the same set of final vowel alternations as the majority *-fanya* ‘do’ type discussed in §4.1, but are distinguished by a differing distribution of the prefixal exponent *ku-*, as shown in Table 10. Two additional lexemes, *-enda* ‘go’ and *-isha* ‘finish’, can optionally inflect according to the pattern of ‘short’ verbs; all other lexemes with vowel-initial roots, such as *-ona* ‘see’ and *-ishi* ‘live’, inflect according to the ‘long’ pattern.

| | -soma 'read' | -fanya 'do' | -nywa 'drink' | -wa 'be' | -enda 'go' |
|------------------|-------------------|--------------------|---------------------|-------------------|-----------------------|
| INF.AFF | <i>kusoma</i> | <i>kufanya</i> | <i>kunywa</i> | <i>kuwa</i> | <i>kwenda</i> |
| INF.NEG | <i>kutosoma</i> | <i>kutofanya</i> | <i>kutokunywa</i> | <i>kutokuwa</i> | <i>kuto(kw)enda</i> |
| PRS.PROG.AFF.1PL | <i>tunasoma</i> | <i>tunafanya</i> | <i>tunakunywa</i> | <i>tunakuwa</i> | <i>tuna(kw)enda</i> |
| PRS.NEG.1PL | <i>hatusomi</i> | <i>hatufanyi</i> | <i>hatunywi</i> | <i>hatuwi</i> | <i>hatuendi</i> |
| PRS.SBJV.AFF.1PL | <i>tusome</i> | <i>tufanye</i> | <i>tunywe</i> | <i>tuwe</i> | <i>twende</i> |
| PRS.SBJV.NEG.1PL | <i>tusisome</i> | <i>tusifanye</i> | <i>tusinywe</i> | <i>tusiwe</i> | <i>tusiende</i> |
| PST.AFF.1PL | <i>tulisoma</i> | <i>tulifanya</i> | <i>tulikunywa</i> | <i>tulikuwa</i> | <i>tuli(kw)enda</i> |
| PST.NEG.1PL | <i>hatukusoma</i> | <i>hatukufanya</i> | <i>hatukunywa</i> | <i>hatukuwa</i> | <i>hatu(kw)enda</i> |
| FUT.AFF.1PL | <i>tutasoma</i> | <i>tutafanya</i> | <i>tutakunywa</i> | <i>tutakuwa</i> | <i>tuta(kw)enda</i> |
| FUT.NEG.1PL | <i>hatutasoma</i> | <i>hatutafanya</i> | <i>hatutakunywa</i> | <i>hatutakuwa</i> | <i>hatuta(kw)enda</i> |
| IMPV.AFF.2SG | <i>soma</i> | <i>fanya</i> | <i>kunywa</i> | <i>kuwa</i> | <i>nenda</i> |
| IMPV.AFF.2PL | <i>someni</i> | <i>fanyeni</i> | <i>kunyweni</i> | <i>kuweni</i> | <i>nendeni</i> |

Table 10. Comparison of “long” and “short” verbs.

For both the class of ‘long’ verbs and the class of ‘short’ verbs, the paradigmatic distribution of the element *ku-* is consistent and conventional; however, the distribution of this element differs between the two groups. A third group, comprising the two items *kwenda* ‘go’ and *kwisha* ‘finish’, displays variation between the two distributional patterns (‘long’ and ‘short’).

In diachrony, there is some evidence that *ku-* is undergoing progressive eviction: eighteenth-century attestations of ‘long’ verbs exhibit *ku-* in positions characteristic of ‘short’ verbs (Nurse & Hinnebusch 1993: 369). A plausible interpretation is one of syncope (for which there is considerable precedent in Swahili, Nurse & Hinnebusch 1993), favoured in verbs with roots of greater phonological substance; ‘short’ verbs are conservative in retaining *ku-* in many forms, while *kwenda* and *kwisha* are intermediate between conservative and innovative patterns. In synchrony, Mpiranya (2015: 98–99) characterises the class of ‘short’ verbs exclusively in terms of phonology, motivating the appearance of *ku-* in specific forms based on an Optimality Theory analysis proposed by Park (1997), according to which Swahili verb stems (i.e. the complex {lexical root, extension, final vowel}) are subject to a ‘disyllabic minimality’ requirement. Yet, under this analysis, the present negative affirmative, affirmative subjunctive and negative subjunctive of ‘short’ verbs all contravene the disyllabic minimality requirement; it is not clear why these forms would not require *ku-*. Furthermore, it remains to be explained why *kwenda* and *kwisha* retain initial *ku-/kw-*, while vowel-initial ‘long’ verbs such as *-ona* ‘see’, *-ishi* ‘live’, of apparently similar phonological shape, do not.

The behaviour of these forms, together with the variable inflection of *kwenda* and *kwisha*, points to an irreducibly arbitrary, morphological element in the distribution of *ku-*: a lexically specific inflectional class phenomenon, originally and still partially correlated with phonological properties (see Bach 2018 for further examples), and a recurrent, arbitrary distribution across inflectional forms of a given lexeme. As in the case of final vowels, alternation between forms with and without *ku-* in a given lexeme can potentially be conceptualised in terms of stem alternation, with each stem having a specific and consistent distribution.

4.3 Interim summary

The phenomena discussed in §4 illustrate two types of purely morphological, morphomic structure present in the Swahili inflectional system: metamorphomes and inflectional classes. These distributional patterns are irreducibly morphological properties of the system, constituting a conventionalised and stable organisational principle. As morphomic phenomena go, the Swahili examples are neither rich nor complex compared to the diverse, cross-cutting patterns described for ‘fusional’ systems, particularly Romance. Nevertheless, it is significant that they exist. The Swahili data indicate that morphomic phenomena are not a priori excluded from occurrence in ‘agglutinating’ inflectional systems, though they may of necessity be more circumscribed where there are few instances of alternation between lexemes or between exponents: the fewer formal contrasts occur between inflected forms or substrings of forms in a given lexeme, and the fewer dimensions of contrast exist between lexemes, the fewer opportunities there are for speakers to infer and perpetuate distinctive patterns organising the distribution of the respective alternants. Morphomic patterning in creoles has already been documented (Luis 2011, Henri 2021), showing that relatively small inventories of inflected wordforms may limit the incidence of purely morphological structure, but do not preclude its development; the Swahili case is complementary to these studies, illustrating a further situation in which incidence is limited (in this instance, due to a large inventory with low incidence of alternation). Together, these studies point to the potential for morphomic phenomena such as metamorphomes and inflectional classes to arise in any inflectional system: inflectional typology and paradigm size are not intrinsically correlated with the presence or absence of such structure, only with the degree to which it manifests.

The existence of morphomic structure (and of non-canonical exponence more generally) in Swahili further bears on the relationship between inflectional typology and the cognitive units which speakers manipulate in inflection, as discussed in §2.2 above. The examples discussed in §4 all illustrate inflectional phenomena which are ordinarily considered to be more accurately represented within a paradigm-based approach, capable of capturing distributional regularities even when these do not instantiate canonical form-function relationships. In order to account for non-canonical exponence and purely morphological structure in Swahili, it is thus desirable to adopt a paradigm-based approach in which morphological cognition manipulates full inflected wordforms. Developing a paradigm-based account of Swahili inflection, comparable to the paradigm-based accounts of ‘fusional’ languages, further resolves the conceptual problem (raised in §2.2) of a sharp theoretical cut-off between analyses which treat sub-word units as cognitive primitives in inflection, and analyses in which sub-word units exist only as abstract generalisations: inflection in Swahili (and, by extension, typologically similar languages) can be assumed to involve similar cognitive structures to most other languages, reflecting the continuous and gradient nature of inflectional typology.¹⁰

Adopting this view is not to deny the empirical, typological observation of greater ease and certainty in segmenting Swahili verb forms into recurrent substrings, compared to languages with ‘fusional’ inflection such as the Romance varieties studied by Maiden (2018) or the Estonian examples discussed by Blevins (2016: 28–29). Rather, this typological contrast is understood as relating purely to form (the ANALYSABILITY of a form, i.e. how readily demarcated ‘recurrent partials’ are; Blevins 2016: 74), instead of to the relationship between form and meaning (association of individual feature

¹⁰ By contrast, attempting to resolve the problem by applying a morpheme-based framework to all languages results in serious descriptive shortcomings, such as the inability to capture structural generalisations about morphomic patterning.

values with individual exponents, or association of holistic morphological property sets with whole words). In the ‘fusional’ systems, where comparison of inflected forms reveals multiple points of similarity and contrast (e.g. patterns of stem vowel alternation may be orthogonal to patterns of stem-final consonant alternation), imposing any single segmentation “forces a choice between [multiple] convergent analyses”, and “treats one pattern as significant and disrupts the other[s]” (Blevins 2016: 28–29). By contrast, in the Swahili system, where context-sensitive allomorphy and cross-cutting¹¹ alternation patterns are rare, comparison of inflected forms will tend to identify recurrent partials which are clearly demarcated and of consistent phonological shape: segmentation into recurrent partials does not disrupt or obscure the formal patterns of inflectional contrast. Thus, the typological difference can be captured with reference to segmentation into units of form, as opposed to segmentation into morphemes associating form with meaning (see also Blevins 2016: 225–227, Beniamine et al. 2021): correspondingly, an ‘agglutinating’ system is understood as one in which the realised phonological forms associated with inflectional paradigm cells are of high analysability into recurrent partials, rather than as one in which inflectional forms constitute strings of morphemes.

5. Conclusions

The traditional view of Swahili verb inflection in the morphological literature is as a system of highly canonical, agglutinating, type, and as such suited to description within a morpheme-based analysis. Even in the context of paradigm-based frameworks (Stump 2001, 2016) or frameworks which admit both paradigm-based and segmentation-based analyses (Crysmann & Bonami 2017), paradigm-based or abstractive analyses of Swahili are presented as optional, affording little or no improvement in descriptive adequacy compared to morpheme-based or constructive analyses.

Yet a range of non-canonical phenomena are attested for Swahili verb inflection in both theoretical and descriptive literature. This paper collates and reviews examples of non-canonicity in synthetic verb forms, including examples of morphomic structure, arguing that such examples are too numerous and systematic to be dismissed as marginal within the inflectional system.

The examples analysed illustrate the tasks of segmenting Swahili wordforms into component substrings, assigning these substrings to slots within the general position-class template set out for Bantu languages by Nurse (2008), and associating each substring with morphological feature values. This procedure confirms that Swahili verb forms are readily, near-exhaustively, segmentable into substrings; furthermore, the segmentation part of the task can be achieved without knowledge of function, simply by aligning inflected forms and observing systematic correspondences or contrasts of form between them. Most substrings can be associated with a particular position-class slot, and in many cases, it is also possible to identify one or more feature values associated with an individual substring. However, it is not consistently possible to isolate biunique relationships either between an individual feature value and an individual substring, or between an individual feature and an individual position-class slot: while the meaning of the carefully chosen example *alikusukuma* is “compositionally computable from its form” (Stump 2016; and indeed its form is compositionally computable from its meaning), for many Swahili verb forms no such direct associations exist. Reliable and systematic relationships exist between forms and meanings at the word level; between inflectional forms,

¹¹ Such patterns are rarer in Swahili than in most other Bantu languages, in part due to the fixed stress assignment and the absence of tone, which mean that segmental alternations do not cross-cut prosodic alternations. However, even in Swahili, some examples do exist: consider, for instance, the orthogonal distributions of final vowels and of stems with and without *ku-*.

independently of meaning; but not between substrings or positions and meaning. Furthermore, the paper finds some evidence for the existence of genuinely morphomic phenomena in Swahili.

These observations constitute a non-trivial obstacle for constructive accounts, in which the selection and arrangement of morphemes is exclusively and exhaustively determined by feature content: the set of morphemes required abounds with arbitrary redundancies, co-occurrence restrictions, and morphological zeros (compare Spencer 1991: 208–214). By contrast, an abstractive, paradigm-based account can readily describe non-canonical exponence and capture the consistency of implicational relationships between forms in the overall organisation of the Swahili inflectional system. The fact that such accounts prove empirically plausible and useful in describing inflectional systems traditionally held up as examples of near-canonicity, such as Swahili, is striking: if the paradigm-based approach is not only possible but plausible and even advantageous for such systems, there can be few real-life inflectional systems which resist description in paradigm-based terms, and conceivably few for which a constructive approach is demonstrably superior.

Via exploration of the morphological structures identifiable for Swahili verb inflection, this paper also addresses a wider theoretical question about the relationship between inflectional typology, inflectional theory and morphological cognition. In much morphological literature, proceeding from the apparently innocuous view that different theoretical models of morphology may be appropriate for languages with typologically different inflection, paradigm-based models are argued to afford descriptive advantage for ‘fusional’ systems, while ‘agglutinative’ systems are considered to be tractable within morpheme-based analyses. Yet these differing theories also embody incompatible claims about the basic cognitive units of morphology manipulated by speakers: full inflected wordforms in an implicational network, or morphemes which individually contribute both form and function as speakers combine them to form words. Treating ‘fusional’ inflection as paradigm-based and ‘agglutinating’ inflection as morpheme-based thus leads to the assumption of a stark cognitive contrast between speakers of languages with ‘fusional’ and ‘agglutinative’ inflection respectively. Such an assumption cannot be sustained in the light of empirical observations that contrasts of inflectional typology are quantitative, relative and gradient. Instead, this paper argues for a single, paradigm-based model encompassing diverse inflectional systems, with typological contrasts understood as relating to analysability of morphological form independently of function.

Acknowledgements

This paper was completed during a Visiting Fellowship at Trinity College, University of Oxford, funded by the CNRS Institut des Sciences Humaines et Sociales via the grant scheme ‘Soutien à la mobilité internationale 2022’. The research reported here was initially presented under the title ‘Morphological units in Swahili verb inflection’ as an invited talk at the Inflectional Reflections workshop (Paris, 15 December 2021) organised by the LabEx Empirical Foundations of Linguistics workgroup GL4 ‘Typology of inflectional systems with non-canonical inflection’ (part of Strand 3 ‘Typology and dynamics of linguistic systems’). As such, this work has benefited from the support of the Agence Nationale de la Recherche via the ‘Investissements d’Avenir’ scheme, reference ANR-10-LABX-0083, and contributes to the IdEx Université de Paris – ANR-18-IDEX-0001. Subsequent versions were presented at the Llacan seminar (Paris, 26 January 2022), and at the 13th Mediterranean Morphology Meeting (Rhodes, 19–22 May 2022) under the title ‘The descriptive advantage of word-based models extends to agglutinative inflection’. I thank colleagues at these events, and the two anonymous Word Structure reviewers, for many stimulating and constructive comments on the paper.

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

References

- Ackerman, Farrell, Blevins, James & Malouf, Robert. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In Blevins, James & Blevins, Juliette (eds.), *Analogy in grammar*, 54–81. Oxford: Oxford University Press.
- Anderson, Stephen R. 2015. The morpheme: Its nature and use. In Baerman, Matthew (ed.), *The Oxford handbook of inflection*, 11–33. Oxford: Oxford University Press.
- Aronoff, Mark. 1994. *Morphology by itself*. Cambridge, M.A: MIT Press.
- Aronoff, Mark & Fudeman, Kirsten. 2005. *What is morphology?* Oxford: Blackwell.
- Ashton, Eric Ormerod. 1947. *Swahili grammar*. London: Longman.
- Bach, Xavier. 2018. *The origins of inflectional classes*. Oxford: University of Oxford. (Doctoral dissertation).
- Baerman, Matthew & Brown, Dunstan & Corbett, Greville G. 2005. *Syncretism*. Cambridge: Cambridge University Press.
- Beniamine, Sacha. 2021. One lexeme, many classes: inflection class systems as lattices. In Crysmann, Berthold & Sailer, Manfred (eds.), *One-to-many relations in morphology, syntax, and semantics*, 23–51. Berlin: Language Science Press.
- Beniamine, Sacha & Bonami, Olivier & McDonough, Joyce. 2017. When segmentation helps. Implicative structure and morph boundaries in the Navajo verb. (Paper presented at *First International Symposium on Morphology*, Lille, 13–15 December 2017).
- Beniamine, Sacha & Bonami, Olivier & Luís, Ana R. 2021. The fine implicative structure of European Portuguese conjugation. *Isogloss* 7. 1–35.
- Blevins, James. 2006. Word-based morphology. *Journal of Linguistics* 42. 531–573.
- Blevins, James. 2016. *Word-and-paradigm morphology*. Cambridge: Cambridge University Press.
- Bonami, Olivier. 2014. *La structure fine des paradigmes de flexion*. Paris: Université Paris Diderot. (Habilitation à diriger des recherches).
- Brown, Dunstan. 2010. Morphological typology. In Song, Jae Jung (ed.), *Handbook of linguistic typology*, 487–503. Oxford: Oxford University Press.
- Corbett, Greville G. 2009. Canonical inflectional classes. In Montermini, Fabio & Boyé, Gilles & Tseng, Jesse (eds.), *Selected proceedings of the 6th Décembrettes*, 1–11. Somerville, MA: Cascadilla Proceedings Project.
- Cruschina, Silvio, Martin Maiden & John Charles Smith (eds.). 2013. *The boundaries of pure morphology*. Oxford: Oxford University Press.
- Crysmann, Berthold & Bonami, Olivier. 2017. *Atomistic and holistic exponence in Underspecified Realisational Morphology*. In Müller, Stefan (ed.), *24th International Conference on Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications.
- Deen, Kamil Ud & Hyams, Nina. 2006. The morphosyntax of mood in early grammar with special reference to Swahili. *First Language* 26. 67–102.
- Delaunay, P. 1885. *Grammaire kiswahili*. Paris: Levé.
- Dzahene-Quarshie, Josephine. 2012. English loans in Swahili newspaper football language. *Ghana Journal of Linguistics* 1(1). 35–56.
- Enger, Hans-Olav. 2021. Meta-morphomic patterns in North Germanic. *Nordic Journal of Linguistics*, First View, 1–19. Published online 18 October 2021.

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

- Evans, Nicholas. 2019. Waiting for the word: Distributed deponency and the semantic interpretation of number in the Nen verb. In Baerman, Matthew & Bond, Oliver & Hippisley, Andrew (eds.), *Morphological perspectives: Papers in honour of Greville G. Corbett*, 100–123. Edinburgh: Edinburgh University Press.
- Feist, Timothy & Palancar, Enrique. 2021. Paradigmatic restructuring and the diachrony of stem alternations in Chichimec. *Language* 97(1). 1–41.
- Gleason, Henry A. Jr. 1955. *Workbook in descriptive linguistics*. New York: Holt, Reinhart & Winston.
- Goldsmith, John & Fidèle Mpiranya. forthcoming. Learning Swahili morphology. In Sibanda, Galen & Ngonyani, Deo & Choti, Jonathan & Biersteker, Ann (eds.), *Descriptive and theoretical approaches to African linguistics: Selected papers from the 49th Annual Conference on African Linguistics*. Berlin: Language Science Press.
- Henri, Fabiola. 2021. Morphomic structure in Mauritian Kreol: On change, complexity and creolization. *Morphology* 31. 447–489.
- Herce, Borja. 2020. *A typological approach to the morpheme*. Guildford: University of Surrey. (Doctoral dissertation).
- Hyman, Larry. 2003. Suffix ordering in Bantu: A morphocentric approach. In Booij, Geert E. & van Marle, Jaap (eds.), *Yearbook of morphology 2002*, 245–281. Dordrecht: Kluwer.
- Katamba, Francis. 1978. How agglutinating is Bantu morphology? *Linguistics* 210. 77–84.
- Lodhi, Abdulaziz. 2000. *Oriental influences in Swahili. A study in language and culture contacts*. Göteborg: University of Göteborg. (Doctoral dissertation).
- Luís, Ana R. 2011. Morphomic structure and loan-verb integration: evidence from Lusophone creoles. In Maiden, Martin & Smith, John Charles & Hinzelin, Marc-Olivier & Goldbach, Maria (eds.), 2011. *Morphological autonomy*, 235–254. Oxford: Oxford University Press.
- Maiden, Martin & Smith, John Charles & Hinzelin, Marc-Olivier & Goldbach, Maria (eds.). 2011. *Morphological autonomy*. Oxford: Oxford University Press.
- Maiden, Martin. 2018. *The Romance verb. Morphomic structure and diachrony*. Oxford: Oxford University Press.
- Matthews, Peter. 1974. *Morphology*. Cambridge: Cambridge University Press.
- Meadows, Tom. 2020. Clause structure and the morphosyntax of Swahili verbs. (Paper presented at SyntaxLab, Cambridge, 21 January 2020).
- Meeussen, A. E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 61(3). 79–121.
- Mpiranya, Fidèle. 2015. *Swahili grammar and workbook*. London: Routledge.
- Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford: Oxford University Press.
- Nurse, Derek & Hinnebüsch, Thomas. 1993. *Swahili and Sabaki. A linguistic history*. Berkeley: University of California Press.
- O’Neill, Paul. 2014. The morpheme in constructive and abstractive models of morphology. *Morphology* 24. 25–70.
- Palancar, Enrique. 2021. Paradigmatic structure in the tonal inflection of Amuzgo. *Morphology* 31. 45–82.
- Park, Jae-Ick. 1997. Disyllabic requirement in Swahili morphology. *University of Pennsylvania Working Papers in Linguistics* 4(2). 245–259.
- Robins, Robert H. 1959. In defence of WP. *Transactions of the Philological Society* 58. 116–144.
- Round, Erich R. 2013. *Kayardild morphology and syntax*. Oxford: Oxford University Press.

Esher, Louise. 2022. Swahili verbs and the value of abstractive accounts for agglutinating inflection. *Word Structure* 15(2). 148–173. <https://doi.org/10.3366/word.2022.0204>.
Author accepted version. The Version of Record published by Edinburgh University Press (July 2022) is available at: <https://www.eupublishing.com/doi/abs/10.3366/word.2022.0204>

- Round, Erich R. 2015. Rhizomorphemes, metamorphemes and meromorphemes. In Baerman, Matthew & Brown, Dunstan & Corbett, Greville G. (eds.), *Understanding and measuring morphological complexity*, 29–52. Oxford: Oxford University Press.
- Schwarz, N. 2004. *Arabic loan verbs in Swahili: A study of the history of language contact on the East African coast and an analysis of Swahili verbs of Arabic origin*. Nijmegen: University of Nijmegen. (Masters thesis).
- Spencer, Andrew. 1991. *Morphological theory*. Oxford: Blackwell.
- Steere, Edward. 1884. *A handbook of the Swahili language as spoken at Zanzibar*. London: Society for promoting Christian knowledge.
- Stump, Gregory T. 1993. Position classes and morphological theory. In Booij, Geert E. & van Marle, Jaap (eds.), *Yearbook of morphology 1992*, 129–180. Dordrecht: Kluwer.
- Stump, Gregory T. 2001. *Inflectional morphology*. Cambridge: Cambridge University Press.
- Stump, Gregory T. 2016. *Inflectional paradigms*. Cambridge: Cambridge University Press.
- Stump, Gregory T. & Finkel, Rafael A. 2013. *Morphological typology. From word to paradigm*. Cambridge: Cambridge University Press.
- Thornton, Anna. 2011. Overabundance (multiple forms realizing the same cell): A non-canonical phenomenon in Italian verb morphology). In Maiden, Martin & Smith, John Charles & Hinzelin, Marc-Olivier & Goldbach, Maria (eds.), 2011. *Morphological autonomy*, 358–381. Oxford: Oxford University Press.
- Versteegh, Kees. 2014. Pidgin verbs: Infinitives or imperatives? In Buchstaller, Isabelle & Holmberg, Anders & Almoaily, Mohammad (eds.), *Pidgins and creoles beyond Africa-Europe encounters*, 141–170. Amsterdam: Benjamins.
- Vitale, Anthony. 1985. Aspects of Kiswahili stress. *Kiswahili: Journal of the Institute of Kiswahili Research* 52. 106–113.