

# Chapter 13

## Morphosyntactic variation in Old Swahili

Lutz Marten<sup>a</sup>, Hannah Gibson<sup>b</sup>, Rozenn Guérois<sup>c,d</sup> & Kyle Jerro<sup>b</sup>

<sup>a</sup>SOAS University of London <sup>b</sup>University of Essex <sup>c</sup>LLACAN CNRS <sup>d</sup>University of KwaZulu-Natal

The comparative and historical study of Bantu and other African languages is often based on contemporary, synchronic data since many African languages do not have a long-written record. In contrast, for Swahili such a record exists in the form of an extensive tradition of written poetic texts. This study presents a comparison of the language used in these texts with present-day Standard Swahili, focusing on morphosyntactic variation. Harnessing the morphosyntactic parameters of Guérois et al. (2017), we show that present-day Swahili differs from Old Swahili in terms of loss of variability and loss of morphosyntactic forms, with only limited cases of innovation. We also show that compared to a sample of 18 neighbouring East African Bantu languages, Standard Swahili shows less similarity to these neighbouring languages than Old Swahili. We propose that these differences are related to the sociolinguistic development of Swahili as a language of wider communication, and the processes of standardisation and regularisation this involved.

### 1 Introduction

Comparative research on Bantu languages has often focused on lexical and phonological data, or on specific morphosyntactic construction types. These studies are also mainly based on synchronic data. The present study develops a novel approach to the examination of morphosyntactic variation in Bantu, by including historical data from classical Swahili poetry and by adopting both qualitative



and quantitative methods of comparison.<sup>1</sup> The study shows that the relation between “Old Swahili” and Standard Swahili is characterised by a loss of variability and processes of regularisation. We propose that this is related, at least in part, to processes of language planning and standardisation which Swahili underwent from the twentieth century onwards. The study also shows that Old Swahili is more similar than Standard Swahili to neighbouring Bantu languages in terms of the morphosyntactic parameters adopted in the study. We propose that this is similarly related to the process of standardisation. Our results present a new perspective on the investigation of morphosyntactic variation as they show the effect of standardisation and a particular trajectory of morphosyntactic development. They also show the benefit of combining qualitative and quantitative methods in the study of morphosyntactic variation.

The majority of comparative and typological studies of Bantu languages are based on synchronic material and draw on contemporary data. This is partly due to methodological reasons, since rich and varied contemporary data are easier to find, include negative evidence, and can in principle be replicated, thus making results more reliable. But, in part, it is also the result of an absence of historical data for most Bantu languages. However, as will be shown in the present study, there are exceptions to this latter challenge.

On the one hand, there are linguistic descriptions of many Bantu languages dating back to the late nineteenth and early twentieth century, and in some cases considerably earlier than this. These can be used to develop diachronic studies and trace language change across a trajectory of several generations. For example, Balestrieri (2017) compares data from three Tanzanian Bantu languages – Haya, Nyamwezi, and Shambala – from different historical periods from the late nineteenth century onwards. An even longer documentary history exists for languages of the Kongo Basin. For example, Dom & Bostoen (2015) use early Kikongo sources to build a diachronic corpus stretching over several centuries.

Beyond linguistic descriptions of languages, there are various written texts that can be analysed to give clues to the linguistic structure of earlier stages of particular languages. One of the most rewarding languages for this kind of diachronic study is Swahili, for which a large body of historical literature exists in the form of a collection of religious, poetic texts written in Arabic script and dating from the nineteenth and early twentieth century. Considerable work has been devoted to the collection, translation, and analysis of the texts from early

---

<sup>1</sup>While historical work has a long tradition in research on other language families and regions, such work is limited in Bantu, for reasons we discuss below. Our approach here is novel in the context of Bantu linguistics, where there has been little work that makes use of historical data (with notable exceptions such as Balestrieri 2017 and Dom & Bostoen 2015).

work such as Taylor (1891) to recent work of, for example Vierke (2011), and – despite the variation in geographic origin, genre, and style of the texts – they provide a good basis for the study of language change and grammaticalisation. We will harness aspects of the language of these texts for the present study. We refer to these data as “Old Swahili” and discuss relevant complexities in more detail below.

## 2 Methodological background

### 2.1 Methodological approach

Our methodological approach is based on recent work by Guérois et al. (2017), which investigates typological, diachronic-historical, and contact-related aspects of morphosyntactic variation in Bantu languages, based on a set of 142 parameters or variables. These parameters reflect salient and well-described aspects of Bantu grammar and are used for the establishment of a large-scale comparative database, the Bantu morphosyntactic variation (BMV) database (Marten et al. 2018). The database contains data from more than forty Bantu languages, eighteen of which are spoken in Eastern Africa and are included in the language sample we will use in our comparative study below. The 142 parameters are divided into twelve thematic groups, as shown in Table 1.

Table 1: Thematic grouping of parameters in Guérois et al. (2017)

1.	Nouns and pronouns (14)
2.	Noun modifiers (11)
3.	Nominal derivation (4)
4.	Lexicon (6)
5.	Verbal derivation (13)
6.	Verbal inflection (38)
7.	Relative clauses, clefts and questions (15)
8.	Verbless clauses (3)
9.	Simple clauses (6)
10.	Constituent order (14)
11.	Complex sentences (15)
12.	Expression of focus (3)

Data in the database come from published sources such as descriptive grammars or more specialised studies focusing on specific grammatical aspects of a

given language or languages. For some languages data come from fieldwork conducted to investigate some or all of the parameters in Guérois et al. (2017); for example, the chapters in Shinagawa & Abe (2019) use the parameters for the study of East African Bantu languages. As discussed further below, for Old Swahili, we rely on Mieke (1979), and for Standard Swahili on Ashton (1947) supplemented by data from contemporary consultants. The data we have available for Old Swahili address only some of the thematic areas shown in Table 1. As a broad observation, we have a good amount of data for the more morphological variables, such as nouns and pronouns, and nominal and verbal derivation and inflection, but less data for the parameters which relate to syntax and information structure. For Standard Swahili our data are complete with respect to the 142 parameters.

## **2.2 Old Swahili**

Evidence of older forms of Swahili comes from a body of texts of religious poetry, written in Arabic script. These texts reflect the literary culture on the East African coast, which was influenced by the introduction and adaptation of Islamic thought – and correspondingly, language contact between Swahili and Arabic – from the ninth century onwards (e.g. Whiteley 1969, Mbaabu 1978, Mugane 2015). Whilst there is little doubt that there was significant interaction between speakers of Arabic and Swahili, the nature of the contact warrants an additional note here. Swahili was used as an important lingua franca throughout the area and became the language of trade, including being used by traders from the Arabian Peninsula. However, it is likely that levels of bilingualism were often asymmetric and restricted in domain. Studies examining Arabic borrowings into Swahili (e.g. Krumm 1940, Lodhi 2000, Baldi 2012, Mwaliwa 2018), for example, indicate a high degree of lexical borrowing across nouns, verbs and grammatical markers (primarily prepositions and temporal adverbs). However, there is little evidence of structural influence from Arabic, and despite prolonged societal bilingualism, the structure of Swahili remains largely similar to neighbouring Bantu languages (see discussion below).

The language of the texts shows variation which can be related to both time and space (Mieke 1979). Although the actual manuscripts largely date from the twentieth century, the language contained in them is likely to cover a longer period of several hundred years and reflect the language of several artistic and cultural centres along the coast. Despite this, the majority of texts are written in northern Swahili dialects, and in particular in Kiamu, the language of Lamu Island, which can be regarded as one of the main centres of Classical Swahili literary production. Given the comparative heterogeneity of the corpus, and the

predominance of northern varieties of Swahili, it is clear that there is no direct line from an idealised “Old Swahili” to (to some extent similarly idealised) modern Standard Swahili, which is to a large extent based on Southern Swahili varieties, in particular the variety spoken on Zanzibar Island. Furthermore, a number of features of the texts which appear to be archaic or to have disappeared in Modern Swahili are often found in present-day dialects. On the other hand, even though the texts were produced at different times and in different places, they can be said to represent a distinct form of Swahili, defined by its specific genre of religious poetry, its historical extension – the majority of texts were produced before the mid-twentieth century and the rise of Standard Swahili – and by drawing primarily on northern varieties of Swahili spoken at the time. It is in this sense that we compare the languages of these older texts, which we refer to collectively as “Old Swahili” with so-called Modern Swahili. However, we acknowledge that the comparison is to some extent dialectal (and to some degree, artificial), and related to the specific genre of the texts, rather than representing a solely diachronic investigation.

Classical Swahili poetry has attracted scholarly attention for more than a century (Taylor 1891, Harries 1962, Mieke 1979, Mulokozi & Sengo 1995, Bertoncini Zúbková et al. 2009, Vierke 2011), and substantive collections of Old Swahili texts are held in different archives and libraries, and a number of which have been edited and analysed. The present study draws in particular on the work of Mieke (1979), which provides a linguistic analysis of the different grammatical – mainly morphological – features found in the Old Swahili texts. As noted above, we also follow Mieke (1979) in treating the language of the texts as one variety – or genre – of Swahili, even though there is considerable internal variation.

### 2.3 Standard Swahili

Swahili has a long history of use as lingua franca in East Africa (cf. Whiteley 1969, Mbaabu 1978, Blommaert 2014, Mugane 2015). It has been used as a language of commerce, education, and intellectual exchange along the East African coast for most of the last millennium. From the 19th century onwards, Swahili was increasingly used in the East African mainland, following the growth in trading activities from the coast. With the onset of European colonialism, the language became a state-sponsored administrative language under both German and British colonial rule. After independence, Swahili was strongly supported as an official and national language. In Tanzania, Swahili was promoted across all public domains, and while there was still a role for English, the space for community languages

has become very restricted. In both Tanzania and Kenya, Swahili plays a central role in the linguistic ecology.

The colonial authorities, as well as associated missionaries, played a major part in language planning and the standardisation of Swahili. Early on, a Latin script-based orthography was developed, as this was seen as more suitable for the use of the language as an administrative language in the European-controlled territories, as well as for the use as a language of promoting Christianity, since the Arabic-based Ajami writing system of classical Swahili was seen as being associated too closely with Islam (Whiteley 1969). From the mid-nineteenth century onwards, there were attempts at developing a standard version of Swahili, which was more or less established by the turn of the twentieth century, and later further developed by the Inter-territorial Swahili Committee which was set up by the British colonial authorities in the 1930s. After independence, Swahili occupied a major role in language policy and planning in East Africa, and in particular in Tanzania and Kenya, where the language has been consistently promoted across a wide range of public domains and supported by institutional infrastructure (Mugane 2015).

Standard Swahili, or *Kiswahili Sanifu*, has several key characteristics which set it apart from earlier and other contemporary varieties of the language. As noted above, Standard Swahili is written in Latin script, thus breaking with the writing tradition of classical Swahili in Arabic script. Secondly, Standard Swahili is based on southern Swahili dialects, in particular on Kiunguja, the dialect of Zanzibar, while classical Swahili was largely based on northern dialects, such as the more literary dialects Kiamu, spoken in Lamu, or Kimvita, spoken in Mombasa. Thirdly, as can be expected from a standardised variety, Standard Swahili is more homogenous, regularised, and has less internal variation than is found in Old Swahili. Fourthly, Standard Swahili has undergone major influence from non-first language speakers – the main foundational works of Standard Swahili were written by non-native Swahili speakers. For example, works by linguists and speakers of Swahili as another language, including foreigners such as Steere (1870) and Ashton (1947) who have been highly influential in the formation of Standard Swahili, and for a large number of speakers and writers past and present, Swahili is used in addition to one or more community and/or European languages.

While there certainly exists variation within Standard Swahili, this has not been investigated in detail so far. As with Old Swahili, we are assuming here an artificially homogeneous version of Standard Swahili and the data we use are mainly based on Ashton (1947), which remains one of the most comprehensive descriptions of Swahili grammar to date, Schadeberg (1992), and on contemporary native speaker judgements.

In the following sections, we provide an analysis of Old Swahili with respect to the parameters of morphosyntactic variation developed in Guérois et al. (2017), focusing on the difference between Old Swahili and Standard Swahili. We provide detailed discussion of relevant parameters in §3 and present a wider comparative analysis and synthesis in §4.

### 3 Parameters of variation

In this section we discuss the differences between Old Swahili and Standard Swahili in terms of the morphosyntactic parameters developed in Guérois et al. (2017). We have data from both Old and Standard Swahili for 61 out of 142 parameters. This is due to the limited data available for Old Swahili (data for Standard Swahili are available for all 142 parameters), especially in the area of syntax and information structure, as noted above. The available data are thus mainly focused on parameters addressing morphological properties.

Specifically, we focus our discussion on the eight parameters of variation shown in Table 2, for which we have data and for which there is variation between the Old Swahili and Standard Swahili. We are aware that concepts such as “typically” in parameter P073 are somewhat subjective and may be difficult to determine, and the issue is discussed in further detail in §3.7. For this study we have adopted this question from the parameters as formulated in Guérois et al. (2017) where such caution was deemed to be necessary, particularly in the case of less

Table 2: Parameters of variation for Old and Standard Swahili

P018	Are there specific pronominal forms for different kinds of possession?
P020	Are there morphological divisions in the system of demonstratives?
P028	Does suffixation of the agentive marker <i>-i</i> occur as a verb-to-noun derivational process?
P038	How is the agent noun phrase in passives introduced?
P058	Is the negative imperative formally distinct from the negative subjunctive?
P068	Is there a tense/aspect suffix <i>-ile</i> or a similar form?
P073	Is preverbal marking of tense/aspect/mood typically restricted to one slot?
P075	Are there object markers on the verb?

well described languages where saying “always” or “in all constructions” may be difficult to prove.

We will discuss the differences between Old Swahili and Standard Swahili with respect to these parameters in more detail in the present section, and then turn to wider comparative analyses in §4.

### 3.1 The coding of alienable and inalienable possession (Parameter 18)

The first parameter in which Old Swahili and Standard Swahili differ relates to the formal distinction between alienable and inalienable possession in possessive pronouns. The relevant parameter and its possible values are detailed below (P018):

- (1) Parameter 18: Kinds of possession: Are there specific possessive pronominal forms for different kinds of possession?  
null unknown  
n.a. there are no possessive pronouns (e.g. only connective constructions)  
no possessive pronouns do not display variation  
yes specify which kind(s) of possession (inalienable/kinship terms/“community”)

The value for this parameter for Standard Swahili is “yes”, since there are specific possessives for kinship terms, while for Old Swahili, the answer is “no” – since although there is variation between different pronominal forms, these are not systematically related to different kinds of possession.

Standard Swahili has two types of possession constructions: one is a class of possessive pronominal stems which are generalized across all types of possessive relation, and the other is a class of possessive suffixes which are restricted to (extended) kinship relations. With respect to the first construction type, a series of six possessive pronominal stems makes a distinction between person (first, second, or third person) and number (singular or plural). Distinct from some other Bantu languages, there are no dedicated possessive forms for different classes, with the third-person forms being used across all classes.

- (2) Standard Swahili possessive pronominal stems (Ashton 1947: 55)

	Singular	Plural
1 <sup>st</sup> person	-angu	-etu
2 <sup>nd</sup> person	-ako	-enu
3 <sup>rd</sup> person	-ake	-ao



These pronominal stems mark agreement in noun class with the possessee, as shown in (3).<sup>2</sup>

(3) Standard Swahili possessive pronouns

- a. nyumba y-ake  
9.house 9-POSS3SG  
'her/his house'
- b. wa-toto w-ao  
2-child 2-POSS3PL  
'their children'
- c. m-pango w-angu  
3-plan 3-POSS1SG  
'my plan'

The data in (3) show examples of different pronominal stems – *-ake* (3a), *-ao* (3b) and *-angu* (3c) – combined with agreement prefixes of classes 9, 2, and 3 respectively. This type of possessive construction is not restricted to any particular noun classes, nor is it restricted to a particular kind of possession, possessor type, or possessive relation.

In addition to these full, analytic possessive pronouns, there exists a class of suffixed forms in which the possessive stem is suffixed to the possessee without any inflecting agreement prefix (Ashton 1947: 56, Schadeberg 1992: 20):

(4) Standard Swahili possessive suffixes

- a. dada-ke  
9.sister-POSS3SG  
'her/his sister'
- b. mw-enz-angu  
1-friend-POSS1SG  
'my friend'

This second means of expressing possession, which is illustrated in (4), is only available with (extended) kinship terms. These kinship terms are found in several noun classes – for example *dada* 'sister' in (4a) is a class 9 noun, while *mwenzi* 'friend' in (4b) is in class 1. There are therefore two ways of expressing possessive

---

<sup>2</sup>Unless otherwise indicated, Standard Swahili examples are our own. We are grateful to Ida Hadjivayanis for discussing relevant Swahili examples with us.

relations in Standard Swahili – an analytic one available for all possessive relations, and a synthetic one employing possessive suffixes which is only available for kinship terms, irrespective of their noun class.

In contrast to Standard Swahili, Old Swahili does not distinguish between kinship and non-kinship possessive relations. As in Standard Swahili, there are analytic and synthetic forms, but crucially, both forms – including the synthetic forms – can be used with either kinship or non-kinship terms.

Analytic forms are very similar in form and function to Standard Swahili. Miehe (1979: 166) calls these forms “disjunct” forms:<sup>3</sup>

- (5) Old Swahili possessive pronouns  
sifa            z-akwe  
10.qualities 10-POSS3SG  
'her good qualities' (Miehe 1979: 166)

In contrast to these pronominal, disjunct forms, there are suffixed forms which Miehe (1979: 159) calls “conjunct” forms:<sup>4</sup>

- (6) Old Swahili possessive suffixes
- a. wa-na-w-e  
2-child-2-POSS1  
'her sons' (Miehe 1979: 162)
  - b. rafiki-o                    (< rafiki-(y)-o)  
9.friend-9.POSS2SG    9.friend-9-POSS2SG  
'your friend' (Miehe 1979: 161)
  - c. mu-lango-w-o  
3-door-3-POSS2SG  
'your door' (Miehe 1979: 161)
  - d. mahali-p-e  
16.place-16-POSS1  
'his position' (Miehe 1979: 162)

The conjunct forms illustrated in (6) are suffixed forms, similar to Standard Swahili suffixed forms like those illustrated in (4); however, they differ from

---

<sup>3</sup>Old Swahili examples are taken from Miehe (1979). We have added glosses and provided English translations for translations given in Dutch, French, or German in the original.

<sup>4</sup>Unfortunately we do not have enough data to present a full paradigm of these forms, and the examples provided in (6) thus serve merely to illustrate the contrast with Standard Swahili.

Standard Swahili in that they contain morphological marking of agreement with the possessee, e.g. class 16 *p-* in (6d), and because the pronominal stem is contracted, e.g. third singular *-akwe* in (5) becomes *-e* in (6a, d). In contrast to Standard Swahili possessive forms, Old Swahili possessive suffixes can be used to express possession other than kinship terms, for example, as shown in (6), with *mulango* ‘door’ (6c) or *mahali* ‘place/position’ (6d) as possessee. Conjunct forms as shown in (6) are no longer in used in Standard Swahili, although some lexicalised forms are still used, e.g. *mwenzio* ‘your friend’. Despite the morphological differences between Old and Standard Swahili possessive suffixes, Mieke (1979) assumes that the two forms indicate the same semantic possessive relation, and she notes their difference in distribution: “In the texts, the conjunct form is not only used for kinship terms – as in Standard Swahili – but with nouns with a range of meanings” (Mieke 1979: 168). If the two forms mark comparable meanings, then we can conclude that Standard Swahili has innovated a restriction in the use of suffixed forms to indicate the expression of possession with kinship terms. It is this difference which is reflected in the distinct values for Parameter 18.

### 3.2 Demonstrative morphology (Parameter 20)

Another difference between Old and Standard Swahili is related to demonstrative morphology. Parameter 20 distinguishes between different demonstrative systems according to the number of morphological distinctions in the system. This division often relates to the distance from the speaker and/or the deictic centre, or to the visibility of the referent. The parameter identifies systems with two-way, three-way, four-way, and five-way (or more) distinctions.

- (7) Parameter 20: Demonstrative morphology: Are there morphological divisions in the system of demonstratives? (e.g. in terms of spatial and temporal deixis and/or visibility)
- |      |  |
|------|--|
| null | unknown  |
| no   | no distinction                                 |
| 1    | yes, there is a two-way distinction            |
| 2    | yes, there is a three-way distinction          |
| 3    | yes, there is a four-way distinction           |
| 4    | yes, there is a five-way (or more) distinction |

We show that Old Swahili has a larger inventory of demonstratives (value 3 for the parameter) than Standard Swahili (which has value 2).

Standard Swahili has a three-way distinction between distal, proximal and referential demonstratives. The last group is used for entities to which reference has already been made, or those which are available in the context. The three forms are based on the noun class concord morphology, and so the demonstrative forms agree with their head noun. The forms can be schematically described as in (8) (Schadeberg 1992: 18).

- (8) Standard Swahili demonstratives:
- a. proximal:  $h + V(\text{owel}) + C(\text{oncor})d$
  - b. distal:  $Cd + le$
  - c. referential:  $h + V + Cd + o$

The proximal form is based on a demonstrative formative  $h + V-$ , where  $V$  stands for a vowel copied from the concord vowel, to which the concord is suffixed – so, for example, for the class 1 concord  $-yu$ , the proximal demonstrative is  $huyu$ , as in (9a):

- (9) Standard Swahili proximal demonstratives:  $h + V + Cd$
- a.  $m-tu \quad hu-yu$   
1-person DEM-CD1  
'this person'
  - b.  $ma-ji \quad ha-ya$   
6-water DEM-CD6  
'this water'
  - c.  $vi-ti \quad hi-vi$   
8-chair DEM-CD8  
'these chairs'

The distal demonstrative form is built from the concord and a demonstrative formative  $-le$ , so for class 1, the demonstrative form is  $yule$ :

- (10) Standard Swahili distal demonstratives:  $Cd + le$
- a.  $m-tu \quad yu-le$   
1-person CD1-DEM  
'that person'
  - b.  $ma-ji \quad ya-le$   
6-water CD6-DEM  
'that water'

- c. vi-ti vi-le  
8-chair CD8-DEM  
'those chairs'

The final Standard Swahili demonstrative form is the so-called referential form, which is used for entities to which reference has already been made, for example in the preceding discourse. The form is based on the proximal demonstrative, but with the vowel of the concord replaced by the formative *-o* (with some phonological effects observable with some of the concords). The same form of the concord with a final *-o* vowel – “the *o* of reference” in Ashton (1947) – is found in other parts of the grammar of Standard Swahili, for example in relative clause formation.

(11) Standard Swahili referential demonstratives: *h + V + Cd + o*

- a. m-tu hu-y-o  
1-person DEM-CD1-DEM  
'this (aforementioned) person'
- b. ma-ji ha-y-o  
6-water DEM-CD6-DEM  
'this (aforementioned) water'
- c. vi-ti hi-vy-o  
8-chair DEM-CD8-DEM  
'these (aforementioned) chairs'

In addition, demonstrative forms can be reduplicated to encode emphasis. In (12), the reduplicated distal demonstrative form has a reading which means something like ‘the very same’:

- (12) vi-ti vi-le-vile  
8-chair 8-DEM-RED  
'these very chairs'

Forms like the one illustrated in (12) could arguably be analysed as constituting a separate morphological class of demonstratives. However, we do not assume such an analysis here, and so consider Standard Swahili to show a three-way distinction between proximal, referential, and distal demonstratives.

In contrast to Standard Swahili, Old Swahili has not only three formatives participating in demonstrative expressions, but four, which can be used in a range of combinations. In fact, as Miehe (1979: 137) observes, there is a high degree of

variation in the data, and it is sometimes difficult to identify consistent patterns or paradigms because different formatives can be combined with each other in different ways. However, allowing for a certain amount of variation, at least four main demonstrative paradigms can be distinguished. Three of these have a corresponding paradigm in Standard Swahili, even though there is some phonological variation. The fourth one, however, based on a formative *-no*, is not found in Standard Swahili.

- (13) Old Swahili demonstratives: *s/h + V + Cd*
- a. proximal (Type 1): *s/h + V + Cd*
  - b. distal: *Cd + le*
  - c. referential: *(s/h) + (Cd) + o*
  - d. proximal (Type 2): *(-s/h) + (V) + Cd + no*

Proximal demonstratives of Type 1 are expressed with a formative *h + V-* or *s + V-* plus the relevant concord, where V is a copy of the concord vowel:<sup>5</sup>

- (14) Old Swahili proximal demonstratives (Type 1): *s/h + V + Cd*
- a. hu-yu binti  
DEM-CD1 1.daughter  
'this daughter' (Miehe 1979: 143)
  - b. ngamia su-yu  
9.camel DEM-CD1  
'this camel' (Miehe 1979: 142)

This is quite similar to Standard Swahili, except that in Old Swahili there is variation between /h/ and /s/ in the formative,<sup>6</sup> while in Standard Swahili it is uniformly /h/.

Distal demonstratives in Old Swahili are formed as in Standard Swahili with a demonstrative formative *-le* and the relevant concord. Lengthened forms, as in (15b), with a long vowel /e/, might have been emphatic forms (Miehe p.c.).

---

<sup>5</sup>In Old Swahili, like in Standard Swahili, the demonstrative can follow or precede the head noun (cf. Rugemalira 2007, Van de Velde 2005).

<sup>6</sup>It is not clear to us at present whether the different kinds of variation described for Old Swahili are dialectal or free variation in the speech of a single speaker/writer.

(15) Old Swahili distal demonstratives: Cd + *le*

- a. za-le      zi-tunda  
 CD8-DEM 8-fruit  
 ‘those fruits’ (Miehe 1979: 138)
- b. u-lee      isilamia  
 CD1-DEM 1.muslim  
 ‘that Muslim’ (Miehe 1979: 138)

Referential demonstratives are formed as in Standard Swahili with a formative *-o*.

(16) Old Swahili referential demonstratives: (*s/h*) + (Cd) + *o*

- a. dini,      ni-i-fuweṭe-yo      na-we  
 9.religion SM1SG-OM9-follow-PFV-REL9 CONJ-PRON2SG  
 u-fuwat-e      i-yo  
 SM2SG-follow-SBJV CD9-DEM  
 ‘the faith I followed, you also follow it’ (Miehe 1979: 140)
- b. s-u-yo      yatima  
 DEM-CD1-DEM 9.orphan  
 ‘this orphan’ (Miehe 1979: 144)

While in Standard Swahili referential demonstratives are built on the proximal demonstrative form, in Old Swahili there is variation as to the elements involved – other than the referential *-o*. In (16a), for example, the referential form *iyō* is simply based on the concord, without the use of the formative *h-/s-* which is found in the proximal, but in other examples such as (16b) *h-/s-* is found in referential demonstrative forms as well.

A final Old Swahili demonstrative form is based on a formative *-no* (cf. Nicolle 2012). This can be suffixed to proximal forms, to form another proximal form. Although it is not fully clear from the descriptions, it is possible that while the normal proximal forms contrast with the distal forms, the proximal form with *-no* refers specifically to speaker proximity, which is the function of the proximal demonstrative form *\*-nóo* reconstructed for Proto-Bantu in Meeussen (1967: 107), of which the Old Swahili form *-no* is likely to be a reflex.

(17) Old Swahili proximal demonstratives (Type 2): (*-s/h*) + (V) + Cd + *no*

- Hu-yu-no      si      malaika  
 DEM-CD1-DEM NEG.COP 1.angel  
 ‘whether this is (not) an angel at all’ (lit. ‘this one is not an angel’) (Miehe 1979: 146)

Abstracting away from variation in Old Swahili, the two different paradigms of demonstratives can be summarised as in Table 3, which shows how the three forms of Standard Swahili contrast with the four forms of Old Swahili.

Table 3: Demonstrative forms in Standard Swahili and Old Swahili

	Standard Swahili	Old Swahili
Proximal	$h + V + Cd$	$s/h + V + Cd$ $(-s/h) + (V) + Cd + no$
Distal	$Cd + le$	$Cd + le$
Referential	$h + V + Cd + o$	$(s/h) + (Cd) + o$

As can be seen from Table 3, the Old Swahili demonstrative system is very similar to Standard Swahili with respect to three forms, but differs through the presence of an additional formative *-no*. As already noted, a demonstrative form in *-no* is reconstructed for Proto-Bantu and is found in other present-day Bantu languages, but is not found in Standard Swahili. It seems that Old Swahili has maintained the form, but it was lost in Standard Swahili. It is also noteworthy that the Standard Swahili system appears to be more regular, with three distinct paradigms, while the Old Swahili system is more complex and irregular: “In the texts we encounter an extraordinary diversity in the forms of demonstratives” in contrast to “the comparatively simple three-way distinction of Standard Swahili” (Miehe 1979: 137). This difference is perhaps reflective of efforts of standardisation in the development of Standard Swahili where pre-existing variation in patterns has been minimised, possibly in a bid to facilitate learning and adoption but also with a view of seeking some more standard “norm” (as discussed in further detail in §4.1).

### 3.3 The use of the agentive suffix *-i* (Parameter 28)

The next difference between Old Swahili and Standard Swahili concerns the use of the deverbial agentive nominalising suffix *-i*. The form is usually part of a wider set of nominal suffixes which can be used with verbal or adjectival bases. Agentive *-i* has been reconstructed for Proto-Bantu, as well as derived forms such as *mu-ĩbĩ* ‘thief’ from *ĩb* ‘steal’ (Meeussen 1967: 93; cf. Standard Swahili *mwizi* ‘thief’). The situation with respect to the use of the form is complex, but we assume that it is found productively in Old Swahili, but only with limited productivity in Standard Swahili. The relevant parameter is shown in (18):



- (18) Parameter 28: Agentive suffix *-i*: Does suffixation of the agentive marker *-i* occur as a verb-to-noun derivational process (possibly in addition to classes 1/2 prefixes)?
- null unknown
  - n.a. there is no agentive noun derivation in the language
  - no this derivational process is not attested in the language, but there are other suffixes
  - 1 yes, it is used productively
  - 2 yes, but it is no longer productive (e.g. there might be frozen forms)

The parameter value for Old Swahili is “1 – yes, it is used productively”, while the value for Standard Swahili is “2 – yes, but no longer productive”.

The older Swahili texts contain numerous examples of the suffix, including those shown in (19):

- (19) Old Swahili agentive forms in *-i* (Miehe 1979: 78)
- a. muwumbi ‘creator’ (< umba ‘create’)
  - b. mpai, mpayi ‘giver’ (< pa ‘give’)
  - c. msomi ‘reader’ (< soma ‘read’)

In addition, the agentive suffix *-aji* is also found, which appears to be an innovation in Swahili. Schadeberg (1992: 11) suggests that the new form results from the combination of the habitual suffix *-ag* and the agentive suffix *-i*, resulting in a new, innovated agentive suffix:

- (20) Old Swahili agentive forms in *-a(j)i* (Miehe 1979: 78)
- muumbai, muwumbaji ‘creator’ (< umba ‘create’)

As the example in (20) shows, there is variation in the form (*-ai*, *-aji*), and both the older form in *-i* (19a), as well as the newer form in *-aji* (20), can be found with the same verbal stem. Miehe (p.c.) suggests that there might have been a semantic difference, whereby *-i* encoded professional occupation, while *-aji* encoded habitual activity.

In Standard Swahili, while there are many examples of agentive nouns in *-i*, the productive method of agentive derivation is with the suffix *-aji*. Miehe (1979: 79) observes: “The formation of the first group [in *-i*] is very rare in Standard Swahili and typically the second group [in *-aji*] is used”. Examples are provided in (21).

(21) Standard Swahili agentive forms in *-aji*

- a. msomaji 'reader' (< soma 'read')
- b. mwuzaji 'seller' (< uza 'sell')
- c. mwendeshaji 'driver' (< endesha 'drive, make go')
- d. mchoraji 'painter/artist' (< chora 'draw, paint')

In (21) agentive nouns are derived with the suffix *-aji* from *soma* 'read' (21a), *uza* 'sell' (21b) and *endesha* 'drive, make go' (21c). There are also alternative ways of creating agentive nouns, for example through borrowing (22) (Krumm 1940, Lodhi 2000, Zawawi 1979) or through nominalisation (23) (both of which were also available in Old Swahili):

(22) Standard Swahili agentive borrowings

- a. dereva 'driver' (< English *driver*)
- b. spika 'speaker' (< English *speaker*)
- c. mwalimu 'teacher' (< Arabic *mu'allim*)
- d. katibu 'clerk' (< Arabic *kātib*)
- e. waziri 'minister, secretary' (< Arabic, Persian *wazīr*)

In (22a) *dereva* is a loan from English *driver*, so creating a (near) lexical doublet: *dereva* and *mwendeshaji* (21c). The following example, *spika* (22b) is also borrowed from English, while *mwalimu* (22c) and *katibu* (22d) are loans from Arabic and the last example, *waziri* (22e) is a loan from Arabic via Persian (Lodhi 2000: 222).

(23) Standard Swahili agentive nominalisation in *-a*

- a. mwuza samaki 'fish monger' (uza 'sell' + samaki 'fish')
- b. mshona viatu 'cobbler' (shona 'sew' + viatu 'shoes')
- c. mfua fedha 'silver smith' (fua 'forge, hammer' + fedha 'silver')

In (23), agentive derivation derives class 1/2 nouns and retains the *-a* suffix of the verb stem. The nominalisation includes an object noun and the resulting form denotes a professional artisan or trader.

In addition to the processes discussed so far, Standard Swahili has also agentive nouns derived with an *-i* suffix.

- (24) Standard Swahili agentive nominalisation in *-i*
- a. *msomi* ‘scholar’ (< *soma* ‘read’)
  - b. *mcheshi* ‘joker, jester’ (< *cheka* ‘laugh’)
  - c. *fundi* ‘technician, artisan, expert’ (historically from *°funda* ‘learn’)

As in Old Swahili, for a number of verbal bases there are two agentive derivations such as *msomaji* ‘reader’ (21a) and *msomi* ‘scholar’ (24a) and often the difference between the two forms relates to habitual activity (*-aji*) as opposed to professional occupation (*-i*), similar to what might have been the case in Old Swahili.

However, for many agentive nouns in *-aji*, there is no corresponding form in *-i*.

- (25) Putative Standard Swahili agentive nominalisation in *-i*
- a. \**mchori* (cf. *mchoraji* ‘painter/artist’)
  - b. \**mwendeshi* (cf. *mwendeshaji* ‘driver’, and also *dereva* ‘driver’)

Furthermore, there are agent nominals in *-aji* which denote professions, so the interpretation of ‘someone doing X habitually’ appears to arise mainly in contrast with another form, in *-i* or a loanword.<sup>7</sup>

- (26) Standard Swahili agentive nominalisation in *-aji* denoting professions
- a. *mchekeshaji* ‘comedian’ (*cheka* ‘laugh’)
  - b. *mtungaji* ‘designer’ (*tunga* ‘compose, design’)
  - c. *mchezaji mpira* ‘footballer’ (*cheza mpira* ‘play football’)
  - d. *mshonaji* ‘tailor’ (*shona* ‘sew’)

Finally, forms in *-aji* are also found in verb-object nominalisations:

- (27) Standard Swahili agentive nominalisation in *-aji* in verb + object nominalisations
- a. *mwuzaji kompyuta* ‘computer salesperson’
  - b. *mwuzaji bima* ‘insurance salesperson’

<sup>7</sup>A *Swahili Times* headline reads: *Wachekeshaji 10 waliolipwa zaidi 2018* ‘The 10 highest paid comedians in 2018’ (*Swahili Times*, 18/08/19, <https://twitter.com/swahilitimes/status/1162958909814059008>)

In sum, while the situation is complex, and to some extent involves semantic distinctions, there is some evidence that in Standard Swahili agent derivations in *-i* are more lexicalised and less productive than agent derivations in *-aji*. The examples show that in Old Swahili, agent derivation with the suffix *-i* is common and more frequent than agentive derivation with *-aji* (which also occurs). Since it is difficult to assess productivity in a closed corpus, we take frequency, and the rarity of forms in *-aji* without a similar form in *-i* as a proxy for productivity. In contrast, forms in *-i* are less frequent in Standard Swahili, where agent derivation is typically achieved with the more productive suffix *-aji*, or by other means such as borrowing or other derivational processes.

### 3.4 The coding of the agent phrase in passives (Parameter 38)

This difference relates to the coding of the agent phrase in passives, where a number of different strategies can be distinguished across the Bantu family (Fleisch 2005, Guérois forthcoming). The relevant parameter is given in (28):

- (28) Parameter 38: Agent noun phrase: How is the agent noun phrase (when present) introduced?  
 null unknown  
 no an agent noun phrase cannot be added to a passive construction  
 1 by a comitative or instrumental preposition (e.g. *na*)  
 2 by class 17 locative morphology (e.g. *ku-* or *kwa-*)  
 3 by another preposition  
 4 by a copula  
 5 there is no overt marker used to introduce the agent noun phrase  
 6 using two (or more) of the above strategies

While for Standard Swahili the value of the parameter is “1 – by the comitative (*na*)”, for Old Swahili it is “6 – using two strategies” (*na* and the copula *ni*).

In Standard Swahili agents of passives are introduced by the comitative preposition *na*:

- (29) a. Kesi hi-yo i-me-fungul-i-w-a na m-ke w-ake  
 9.case DEM-CD9 SM9-PERF-open-APPL-PASS-FV COM 1-wife 1-POSS3SG  
 ‘The case was opened by his wife.’  
 b. Wa-me-shik-w-a na njaa  
 SM2-PERF-hold-PASS-FV COM 9.hunger  
 ‘They were grabbed by hunger.’

The examples in (29) show the use of *na* with both a human agent (29a), and an abstract agent (29b).

In Old Swahili, agents of passive constructions can be expressed by either the comitative preposition *na* (30a) as in Standard Swahili, but also by the copula *ni* (30b, c). From the available examples, it is not clear whether there are any semantic or functional differences between the use of the two forms.

- (30) a. mw-ema, m-za-w-a      **na** w-ema  
 1-good    1-bear-PASS-FV COM 14-good  
 ‘the good one, born in goodness’ (Miehe 1979: 196)
- b. mahari            a-l-o-pa-w-a                      **ni** Jabiri  
 9.bride\_price SM1-PST-REL-give-PASS-FV COP Jabiri  
 ‘the bride-price set by him by Jabir’ (Miehe 1979: 196)
- c. me-zing-iw-a                      **ni** mal’una  
 SM1.PERF-surround-PASS-FV COP 10.cursed  
 ‘He was surrounded by the cursed.’ (Miehe 1979: 196)

In contrast to Old Swahili, the use of *ni* to introduce the passive agent is not found in Standard Swahili. The difference between Old Swahili and Standard Swahili is noted by Miehe (1979: 197): “Frequently the copula *ni* is used instead of *na* which is used in Standard Swahili to express the agent of the action”. Meeussen (1967: 116) proposes that in Proto-Bantu agents in passives were introduced by *na*, and so the use of the copula *ni* in Old Swahili (and other Bantu languages such as, for example, Chichewa, Digo or Gikuyu) would be an innovation, which, however, was then no longer used in Standard Swahili.

### 3.5 Negative imperatives (Parameter 58)

Negative imperatives in Bantu are often formed in a manner identical to negative subjunctives, but there are also languages which employ a distinct form for negative imperatives. This observation is investigated in Parameter 58, which is presented in (31) below:

- (31) Parameter 58: Negative imperative: Is there a negative imperative which is formally distinct from the negative subjunctive?  
 null unknown  
 n.a. there is no negation (or means to express negation) in the language  
 no  
 yes

In Standard Swahili, negative imperatives are formally identical to negative subjunctives, so the setting for Parameter 58 is “no”. Negative imperatives and negative subjunctives are formed with a negative marker *si-* and a final vowel *-e*:

- (32) U-si-ondo-e                      vy-ombo  
SM2SG-NEG-remove-SBJV 8-dish  
‘Don’t take the things away.’ (Ashton 1947: 119)

The subject marker can sometimes be omitted in negative imperatives, although this is not common and mainly found in formal or written language. Ashton (1947: 119) notes that forms without subject markers are often found in proverbs, which might be an indication that this reflects past usage:

- (33) Si-pig-e  
NEG-beat-SBJV  
‘Don’t beat.’ (Ashton 1947: 119)

Both the forms in (32) and (33) have the same negative marker and final vowel *-e*. The final vowel *-e* is also found in affirmative subjunctives, but not in affirmative imperatives, which end in *-a* (or the “original” vowel in verbs that end in vowels other than *-a*). Since the optional drop of the subject marker in subjunctives is restricted to formal and written registers, we assume that Standard Swahili is a language where the negative imperative is formally identical to the negative subjunctive.

In Old Swahili, negative imperatives are typically identical to negative subjunctives, like in Standard Swahili (Miehe 1979: 249), and also the formal markers employed in the construction are the same: a negation marker *si-* and a final vowel *-e*. The example in (33) also shows that the subject marker can be omitted in Old Swahili negative subjunctives, in the same way we noted for Standard Swahili in (34).<sup>8</sup>

- (34) Inuk-a, si-keti tena  
rise-FV NEG-sit again  
‘Get up and stop sitting.’ (Miehe 1979: 249)

---

<sup>8</sup>The final vowel in *-keti* is lexically determined and does not change in the subjunctive. Unfortunately, it is the only example provided in Miehe (1979). The German translation is ‘*Steh auf und bleibe nicht länger sitzen*’ (1979: 249) which means ‘stand up and do not sit any longer’ (translation our own).

However, Miehe (1979: 251) notes an alternative way of expressing negative imperatives, observed by Krapf (1850: 56), and confirmed by Whiteley (1955) as still being heard in Mombasa in the 1950s, although rarely. In these forms, there is a negative marker *si-*, but the final vowel is *-a*, not *-e*.

- (35) a. *si-pend-a*  
 NEG-like-FV  
 ‘Don’t like.’ (Miehe 1979: 251)
- b. *si-pend-a-ni*  
 NEG-like-FV-PL  
 ‘Don’t like (pl).’ (Miehe 1979: 251)

The forms can be seen as direct negative counterparts of affirmative imperatives, which similarly (typically) have a final vowel *-a*. Through the difference in final vowel, they are distinct from negative subjunctives, and so for Old Swahili, the value of Parameter 58 is “yes”.

### 3.6 The formation of the perfect (Parameter 68)

A well-known development in the history of Swahili is the development of different perfect markers, each involving a grammaticalisation cycle of a verb meaning ‘finish’ (e.g. Heine & Reh 1984, Marten 1998, Drolc 2000). The oldest of these cycles predates Swahili and has been located at a pre-Bantu stage (Voeltz 1980). It involves a reconstructed verbal form *\*-gid* ‘finish’ which developed into the widespread Bantu perfect marker *-ile*. Parameter 68 is concerned with the presence of this form:

- (36) Parameter 68: Suffix *-ile*: Is there a tense/aspect suffix *-ile* or a similar form (as a reflex of *\*-ide*)?  
 null unknown  
 no indicate how perfect/perfective verb forms are formed  
 yes

The common Bantu perfect form *-ile* is still found in Old Swahili, but it has disappeared in Standard Swahili. Both Old Swahili and Standard Swahili also have an additional perfect marker *me-* resulting from a grammaticalisation process of *mala* ‘finish’ (a verb form now obsolete in Standard Swahili but whose root survives in the historical causative form *maliza* ‘finish’). Standard Swahili has, in addition, a perfect marker based on a more recent grammaticalisation process, namely the emerging perfect marker *sha-* from *isha* ‘finish’. While the

beginning of this process can be seen in Old Swahili, the form has become more widely accepted only recently (see Marten 1998). The overall situation indicates that with respect to Parameter 68, the value for Old Swahili is “yes”, since there is evidence of the use of *-ile*, while for Standard Swahili, it is “no”, since *-ile* is no longer used.<sup>9,10</sup>

Several examples of the use of *-ile* are found in Old Swahili:

(37) Old Swahili perfects in *-ile*

- a. ni-kom-ile            ku-kutubu  
SM1SG-finish-PERF 15-write  
‘I have finished writing.’ (Miehe 1979: 179)
- b. u-tu-p-ile            kuwwa  
SM1-OM1PL-give-PERF 9.power  
‘He has given us power.’ (Miehe 1979: 180)
- c. na    ratabu u-ni-pee  
CONJ dates SM1-OM1SG-give.PERF  
‘and dates he gave me’ (Miehe 1979: 178)
- d. Athumani ondosh-ile    ...  
Athumani SM1.leave-PERF  
‘Othman went ...’ (Chuo cha Herekali, Knappert 1967: 159)

Miehe notes that *-ile* in Old Swahili is “only partly productive” (1979: 178), and that there is already evidence for the development of perfects in *me-*, which is the form which has taken over the function of *-ile* in Standard Swahili (Miehe 1979: 178).<sup>11</sup> There is also some evidence of the incipient development of a completive or perfect marker from *isha* ‘finish’ in Old Swahili, although it is much less widespread than in Standard Swahili and seems to be restricted to temporally underspecified contexts (Marten 1998).

<sup>9</sup>Miehe (p.c.) notes that the use in Standard Swahili of a verb form ending in *-e* after the preposition *tangu* ‘since’ is likely to reflect an old perfect form rather than a subjunctive which it is often interpreted as.

<sup>10</sup>As helpfully pointed out by an anonymous reviewer, it is important to note that in example (37c) the long vowel *-ee* incorporates an allomorph of the *-ile* suffix, meaning that this example also supports the presence of this marker in Old Swahili.

<sup>11</sup>Space prevents discussion of the details of the perfect grammaticalisation processes in Swahili. The loss of the suffix *-ile* may in part have been the result of morphological pressures to mark tense and aspect in pre-verbal position, in part the result of phonological loss or reduction due to the loss of intervocalic /l/, and in part related to wider grammaticalisation paths involving ‘finish’, completive, perfectives, perfects and pasts (Heine & Kuteva 2002: 134–138, 231). For the development of *me-* and *sha-*, see Marten (1998).



Standard Swahili examples of perfects in *me-* (38) and *sha-* (39) are shown below:

- (38) Standard Swahili perfect in *me-*  
 Wa-tu wa-**me**-fik-a  
 2-person SM2-PERF-arrive-FV  
 ‘People have arrived.’
- (39) Standard/Colloquial Swahili perfects in *sha-*
- a. A-**me**-kwisha (ku-)imb-a  
 SM1-PERF-finish (INF-)sing-FV  
 ‘S/he has finished singing ~ has already sung.’
  - b. A-**me**-**sha**-imb-a  
 SM1-PERF-COMPL-sing-FV  
 ‘S/he has already sung.’
  - c. Ni-**sha**-fahamu  
 SM1SG-PERF-understand  
 ‘I have (already) understood.’

The use of *me-* as shown in (38) is the standard way of expressing perfect in Standard Swahili. The use of *sha-* (39) is more recent, and examples are more common in spoken than in written language. While forms like those seen in (39a) and (39b) are more widely accepted, the form in (39c) is still stylistically very restricted. Semantically, *sha-* contains an element of both completion and counter-expectation and is thus semantically more complex than the pure perfect in (38). The specific semantic contribution of *sha-* can be seen in the common co-occurrence of *sha-* with the older perfect *me-* (39b), where it is typically translated as ‘already’.

As discussed in this section, Standard Swahili has two perfect markers (*me-* and *sha-*) but no reflex of the Proto-Bantu perfect markers \*-*ide*. In contrast, perfect forms with *-ile* are found throughout the Old Swahili texts, accounting for the difference in parameter setting between the two varieties.

### 3.7 Preverbal TAM slots (Parameter 73)

Another difference between Old Swahili and Standard Swahili in relation to tense-aspect-mood marking concerns the number of preverbal morphological slots available for TAM marking. Parameter 73 distinguishes between languages with more than one slot and languages with typically only one slot.

- (40) Parameter 73: TAM slots: In an inflected verb form, is preverbal marking of tense/aspect/mood typically restricted to one slot?  
 null unknown  
 n.a. there are no tense/aspect/mood prefixes in the language  
 no there are two or more preverbal slots for tense/aspect/mood marking  
 yes there is typically only one preverbal slot for tense/aspect/mood marking

There is some flexibility in the interpretation of this parameter, as the issue is a question of degree. Since the “yes” value includes “typically”, it can be true even if there are isolated instances of more than one slot being used to mark TAM distinctions. This situation is true of Standard Swahili, where typically, preverbal TAM marking is restricted to one position, although there are exceptions. In contrast, in Old Swahili, there are more instances of two TAM slots, although even in Old Swahili, this is not freely available.

In Old Swahili, the past marker *ali-* (itself developed from a tense marker *a-* and a copula verb *li*, cf. Nurse & Hinnebusch 1993: 412, 443, 455, 502, 505; Nurse 2008: 83) can be combined with either the perfect *me-*, already noted above, or with the situative marker *ki-* (Miehe 1979: 219–220):<sup>12</sup>

- (41) Old Swahili combination of (*a*)*li-* and *me-*
- a. **ali-me-itind-a** ndia  
 PAST-PERF-block-FV 9.road  
 ‘(S/he) blocked the way.’ (Miehe 1979: 219)
- b. **a-li me-keti nyumba-ni**  
 SM1-PAST PERF-sit house-LOC  
 ‘He was sitting in his house.’ (Miehe 1979: 219)
- (42) Old Swahili combination of (*a*)*li-* and *ki-*
- N-ali-ki-kw-evuz-a** mno  
 SM1SG-PAST-SIT-OM2SG-search-FV very  
 ‘I was looking for you a lot.’ (Miehe 1979: 220)

<sup>12</sup>The analysis of the TAM forms in (41) when used with a class 1 subject marker *a-* is complex, especially as the subject marker can be omitted in certain contexts. A form such as *ali* can thus be analysed as either subject marker *a-* plus TAM form *ali*, with vowel shortening of the two adjacent /a/ vowels, or as TAM form *ali* without subject marker. The analyses in (41) follow Miehe (1979: 219), who translates (41a) as ‘(er) versperrte den Weg’.

The past tense marker *ali-* develops in Standard Swahili into the past tense marker *li-* (meaning that historically the past tense marker in Standard Swahili is derived from the copula form *li*, so historically a copula form), which cannot be combined with other TAM forms. In Old Swahili, *ali-* can be found in combination with *me-* (41) or *ki-* (42). Although the examples illustrate two TAM slots between subject marker and object marker as seen in (42), orthographic variants may indicate the ambiguous status of the form, as they are frequently written disjunctively, as in (41b), and can even be divided by intervening material – in which case *li* must still have been analysed as a separate predicate. The forms are bound up in the grammaticalisation process of past marking, and indeed in that of the grammaticalisation of the perfect with *me-* from the verb stem *mala* ‘finish’, already noted above. However, at least some of the examples in the text, such as (41a) and (42), provide evidence of two TAM slots (cf. also Nurse & Hinnebusch 1993: 443, 459).

In Standard Swahili, TAM marking is typically restricted to one marker per verb, and TAM markers are typically monosyllabic (cf. Schadeberg 1992).

Table 4: Standard Swahili TAM markers

Affirmative		Negative	
a-	General present	-i	Present
na-	Progressive present		
li-	Past	ku-	Past
ta-	Future		
me-	Perfect	ja-	Perfect
mesha-	Unexpected perfect		
ki-	Situational		
nge-	Present conditional		
ngali-	Past conditional		
ka-	Subsecutive		
hu-	Habitual		

Table 4 shows Standard Swahili TAM markers. Typically, only one marker can be used on an inflected verb form, and unlike in Old Swahili, the past tense marker *li-* cannot be combined with other TAM markers. There are two polysyllabic markers in Table 4. The conditional marker *ngali-* is diachronically complex but is synchronically better analysed as monomorphemic. The case of perfect *mesha-* has been mentioned above: it is part of the grammaticalisation of the verb

stem *isha* ‘finish’ into a perfect marker, and in addition to *mesha-* other intermediately grammaticalised forms of the process can be found, as seen in (39), above. As an on-going grammaticalisation process, some intermediate forms combine the older TAM marker *me-* (and in some varieties the situational marker *ki-*) with the newly emerging marker *sha-*, but there is strong pressure in the system to reduce the form to a monomorphemic (*mesha-*) and ultimately monosyllabic (*sha-*) marker.

In terms of the question of the number of verbal TAM slots, we analyse Standard Swahili as having only one slot, while for Old Swahili we propose that two TAM slots are more regularly available, although, as we noted, these are also related to on-going grammaticalisation processes.

### 3.8 Pre-verbal and post-verbal object marking (Parameter 75)

A final difference between Old Swahili and Standard Swahili discussed here is related to object marking. Variation in object marking across Bantu languages is well attested (cf. Beaudoin-Lietz et al. 2004, Marten & Kula 2012, Marlo 2015). A common cross-linguistic difference is the presence of pre-verbal (or pre-stem) and/or post-verbal object markers, and this is expressed in Parameter 75:

- (43) Parameter 75: Object marking: Are there object markers on the verb (excluding locative object markers)?
- null unknown
  - no there is no slot for object marking in the language (i.e. only independent pronouns)
  - 1 yes, there are only pre-stem object markers
  - 2 yes, there are only post-verbal object markers (enclitics)
  - 3 yes, there are both pre-stem and post-verbal object markers

Bantu languages vary as having only pre-verbal, only post-verbal, or both pre- and post-verbal object markers. While Standard Swahili has only pre-verbal object markers (so the value for Parameter 75 is “1”), in Old Swahili we find both pre-verbal and post-verbal object markers (value “3”) (see also Gibson et al. 2019).

Like many Bantu languages, Standard Swahili only allows pre-stem object markers, and only one object marker at a time.

- (44) Standard Swahili object marking (cf. Marten et al. 2007: 263/4)
- a. ni-li-m-p-a  
SM1SG-PAST-OM1-give-FV  
‘I gave him/her.’

- b. ni-li-m-p-a                      hi-zi  
SM1SG-PAST-OM1-give-FV DEM-CD10  
'I gave them (to) him/her.'
- c. \* ni-li-zi-m-p-a  
SM1SG-PAST-OM10-OM1-give-FV  
Intended: 'I gave them (to) him/her.'
- d. \* ni-li-m-p-a-zi/-zo  
SM1SG-PAST-OM1-give-FV-OM10  
Intended: 'I gave them (to) him/her.'

The examples show that one object marker is acceptable, even with a seemingly ditransitive verb such as *pa* 'give', which allows object drop in a context for which the referent of the (object) nominal being referred to is recoverable (44a). If a second pronominal object is licensed, it will be expressed by using a full pronominal form (44b). A second object cannot be expressed by a second pre-stem object marker (44c), nor by post-verbal object marker (44d) (there are post-verbal locative clitics, which we ignore here).

Like Standard Swahili, Old Swahili does not allow multiple pre-verbal object markers. However, there are examples of both pre-verbal and post-verbal object markers, in very specific circumstances. In (45), the pre- and post-verbal object markers co-occur, and they crucially refer to the same participant in the event.<sup>13, 14</sup>

(45) Old Swahili emphatic object marking

- a. na u-me-n-amkuw-a-mi  
and SM2SG-PERF-OM1SG-call-FV-OM1SG  
'then you called me' (Miehe 1979: 99)

<sup>13</sup>There are only a few examples of these constructions in the literature, and more empirical evidence would be needed to further explore this typologically unusual pattern.

<sup>14</sup>An anonymous reviewer notes that a comparable effect can be seen in some relative clause constructions in Standard Swahili, where in so-called "tenseless relatives" (cf. Schadeberg 1989) a relative marker is suffixed to the verb. In object relatives such as (i), the object marker and the relative marker are co-referential and so resemble the double marking discussed here for Old Swahili.

- (i) U-ki-nunua-cho                      ni nini?  
SM2SG-OM7-buy-FV-REL7 COP what  
'What is it that you buy?'

- b. ni-mu-dhamini-ye            jaza  
SM1SG-OM1-guarantee-OM1 reward  
'I will guarantee him reward.' (Miehe 1979: 100)
- c. a-ka-zi-angush-a-zo  
SM1-CONS-OM10-throw.down-FV-OM10  
'and he threw them down' (Steere 1884: 108, in Miehe 1979: 101)

Miehe (1979: 101) suggests that the combination of pre- and post-verbal markers might have emphatic function. The morphological shape of the post-verbal markers differs from the pre-verbal ones. For discourse participants (1st and 2nd person) (45a), these object markers seem to be shortened pronominal forms, while for classes such as class 1 (45b) and class 10 (45c) the forms are “bound substitutives” (Schadeberg 1992) which are also used, for example, in demonstratives and relatives. Miehe notes the difference between Old Swahili and Standard Swahili in this respect: “mention should be made of the additional suffix with presumably emphatic function, which is not (no longer?) used in this function in Standard Swahili” (Miehe 1979: 101).<sup>15</sup> However, according to Steere (1884: 108), the post-verbal object markers are not used in the dialect of Zanzibar – indicating a dialectal, as well as or in addition to a diachronic analysis.

### 3.9 Summary of comparative results

When comparing Old Swahili and Standard Swahili, the values for 53 of the 61 parameters are the same, but for 8 parameters, the two varieties differ. In terms of shared values, the two varieties thus show 87% similarity. The eight parameters which differ between the two varieties are summarised in Table 5.

## 4 Old Swahili in the context of the development of Swahili and of wider Bantu variation

The previous sections have shown differences between Old Swahili and Standard Swahili related to the eight parameters in which the two varieties differ and have presented a detailed discussion of the specific forms and structures involved. In the present section, we discuss the differences in a wider context and develop

<sup>15</sup>The use of a post-verbal formative *-ni*, often analysed as pluralising, in the formation of 2nd plural object marking could be seen as a historical remnant of the Old Swahili system: *ni-na-ku-ambi-e-ni*, SM1SG-PRS-OM2SG-tell-FV-PL, or *ni-na-wa-ambi-e-ni*, SM1SG-PRS-OM2-tell-FV-PL, both meaning ‘I am telling you (pl.)’ (cf. Gibson et al. 2019).

Table 5: Parameters of variation for Old and Standard Swahili

Parameter	Old Swahili	Standard Swahili
P018: Are there specific pronominal forms for different kinds of possession?	No	Yes
P020: Are there morphological divisions in the system of demonstratives?	There is a four-way distinction	There is a three-way distinction
P028: Does suffixation of the agentive marker <i>-i</i> occur as a verb-to-noun derivational process?	It is used productively	It is no longer productive
P038: How is the agent noun phrase in passives introduced?	Using two (or more) strategies	By the comitative or instrumental (e.g. <i>na</i> )
P058: Is the negative imperative formally distinct from the negative subjunctive?	Yes	No
P068: Is there a tense/aspect suffix <i>-ile</i> or a similar form?	Yes	No
P073: Is preverbal marking of tense/aspect /mood typically restricted to one slot?	No	Yes
P075: Are there object markers on the verb?	There are pre-stem and post-verbal object markers	There are only pre-stem object markers

qualitative and quantitative approaches towards a better understanding of the patterns observed. We show that, on the one hand, the overall difference between Old Swahili and Standard Swahili is related to innovation and loss, but also to the processes of standardisation which have resulted in Standard Swahili, and that, on the other hand, this process has also resulted in a development which sets Standard Swahili more clearly apart, in terms of morphosyntactic structure, from neighbouring Bantu languages than Old Swahili.

#### 4.1 Qualitative differences and the standardisation of Swahili

The differences between Old Swahili and Standard Swahili discussed above can be related to three different processes: Loss and innovation on the one hand, and standardisation on the other. The first two are well-established processes of language change, while the third one reflects the particular socio-historical development of Swahili, and provides the context in which these processes of change have taken place. As has been noted in previous literature, certain sociolinguistic situations may affect the rate of language change; for example, societal multilingualism has been argued to have the effect of accelerating processes of language change (Kusters 2003, Trudgill 2009, 2011, McWhorter 2011). We contend here that the standardisation of Swahili may have served as an accelerating, or in this case regularising, process of both loss and innovation in the language as well as reducing optionality and variability. We discuss each of these three processes in turn.

The majority of differences between Old Swahili and Standard Swahili can be seen as instances of loss, where Standard Swahili appears to have lost forms or structures which were still available in Old Swahili. The most well-known example of this is probably the loss of the perfect in *-ile*. We also noted the loss of the fourth demonstrative formative *-no*. In both cases there is evidence of these forms in Old Swahili, whilst they are not found in Standard Swahili. Furthermore, both forms are well-attested across Bantu and have been reconstructed for Proto-Bantu.

Other examples of loss include: 1) the loss of negative imperatives as distinct from negative subjunctives; 2) the use of the copula *ni* for introducing the agent in passives which is no longer possible in Standard Swahili, and 3) the agentive derivational suffix *-i*, which was fully productive in Old Swahili but is no longer fully productive in Standard Swahili. The latter process shows that change is gradual, as the form is found in both varieties, but the change relates to the distribution of the two agentive forms in the two varieties of Swahili and the frequency in their use. A final example is the use of post-verbal object markers, which is found in Old Swahili for emphatic purposes, but which is not possible in Standard Swahili.

As noted at the outset of the paper, there is no unambiguous direct diachronic line from what we here call “Old Swahili” to Standard Swahili. This means that the case for analysing the differences discussed here as loss differs from example to example. The most robust examples are those where there is a clear Proto-Bantu reconstruction – such as the perfect *-ile*, the demonstrative *-no*, and the agentive *-i* – since it is fair to assume that these forms existed in some earlier form



of Swahili. On the other hand, structures like the double object marking found in Old Swahili, which are not attested widely across Bantu and not reconstructed for Proto-Bantu, may always have been restricted to only specific varieties of Swahili (e.g. Northern dialects), and so have not, strictly speaking, been lost in Standard Swahili.

In addition to processes of loss there are processes of innovation. However, in our study, there are far fewer examples of innovation than of loss. The main example concerns the development of perfect markers. Corresponding to the loss of the perfect marker *-ile*, two new perfect markers have developed. The perfect marker *me-* is already attested in Old Swahili but has become the main perfect marker in Standard Swahili. Furthermore, the more recent perfect marker *sha-* is only found in Standard Swahili, even though evidence for initial stages of the grammaticalisation process can already be seen in Old Swahili. The markers *me-* and *sha-* are claimed to have grammaticalised from *mala* ‘finish’ and *kwisha* ‘finish’ respectively (Schadeberg 1990, Muzale 1998, Marten 1998, Nurse 2008). The second example of innovation is the development of the agentive derivational suffix *-aji*, which is used more productively in Standard Swahili than the older suffix *-i*. However, despite these examples, the overall relation between Old Swahili and Standard Swahili is characterised by loss rather than by innovation.

A third dimension of change observable in the data is related to standardisation, and the loss, not of forms and structures as such, but of variability and optionality. Miehle (1979) comments on this point in relation to different developments, for example, as noted above, in relation to the demonstrative system. While the difference in the demonstrative system is in part related to a loss of a specific formative (the morpheme *-no*), it also undergoes a process of regularisation. While in Old Swahili a variety of structures can be built from the basic four formatives, so that it is difficult to distinguish or enumerate distinct paradigms, in Standard Swahili three discrete and invariable demonstrative paradigms can be identified. In the marking of agents in passives, the option to use the copula *ni* is lost (even though the copula as a form survives), and so the paradigm becomes simplified, involving only the form *na* ‘and, with’. Similarly, in Old Swahili, two negative imperatives could be formed: one identical to the negative subjunctive, with a final vowel *-e*, and one dedicated negative form with a final vowel *-a*. With the loss of the second option, the grammar shows less variation in this regard and the end result is a loss of a category distinction (for a negative imperative meaning) in Standard Swahili, since only one form is used for the function. The loss of post-verbal object markers could similarly be seen as a regularisation of the object marking paradigm, which now only includes pre-verbal object markers. A final example of increased regularisation involves possessive marking. As noted

above, there are two kinds of possessive markers in both Old Swahili and Standard Swahili – independent forms and possessive suffixes. However, while in Old Swahili there was functional overlap, and hence variation between the forms, in Standard Swahili the difference in form has been interpreted as a difference in function, related to the semantic criterion of kinship possessors, and so as more regular.

In addition to loss of forms and functions, increased regularisation and loss of variability is a second major factor in the relation between Old Swahili and Standard Swahili. Here, as well, the differences between our corpora have to be kept in mind. Our Old Swahili data come from texts produced at different times, in different places, and by different authors. In contrast, our data for Standard Swahili come mainly from two linguistic works, Ashton (1947) and Schadeberg (1992). In some regards, the Old Swahili corpus is broader since it reflects different time periods and different contributors. However, this corpus is based on an almost exclusively literary or poetic register. In contrast, although the Standard Swahili data come from two primary sources, these two both draw on a larger body of contemporary data and can be assumed to be much wider and representative in terms of genre. The difference in variability is therefore to some extent unsurprising. However, we believe that this is not the only explanation, and that the increase in regularity and the decrease in variability in Standard Swahili is a consequence of the process of standardisation. In part, it reflects the involvement of choice in relation to the creation of a standard version of the language, but it is likely that it also partly reflects the agency of second-language speakers in the standardisation of Swahili (cf. Whiteley 1969, Mlacha 1995, Mazrui 2007, Blommaert 2014). Variability and variation, which were possibly linked to sociolinguistic or register variables, were difficult for early students of Swahili to grasp, and even more difficult to represent as part of descriptive or pedagogical works. It would have been much easier to reduce variability, or to imbue variant forms with more tangible, referential-semantic differences, as in the case of the kinship relations in possessives.

In summary, when comparing Old Swahili and Standard Swahili the main differences are related to the loss of forms, or the loss of function of a given form in a specific context, and to regularisation and loss of variability. In contrast, innovation of forms or structures plays a less important role. In part, these two factors are related to the differences between the two corpora we compare – differences in terms of age, genre, authorship, dialects, heterogeneity, and other factors. However, we have argued that, to some extent, the differences reflect the process of standardisation which Swahili has undergone over the past century. We have proposed that loss of variability is an integral part of standardisation,

but that, in addition, the specific history of standardisation of Swahili, which involved many second-language speakers, plays a role in this as well. The differences between Old Swahili and Standard Swahili in terms of morphosyntax are thus related to, and provide further evidence for, the particular historical trajectory of the language. In the next section, we relate this finding to the wider comparative Bantu context.

#### 4.2 Quantitative differences and comparative Bantu contexts

As noted above, the comparison between Old Swahili and Standard Swahili presented here is embedded in a wider project on morphosyntactic variation, following Guérois et al. (2017) and Marten et al. (2018). In this section we draw on these wider data, and compare Old Swahili and Standard Swahili with the 18 Eastern African Bantu languages included in our corpus, which are spoken in Kenya and Tanzania, around the Great Lakes, and in Mozambique: Nyolo (E35), Gikuyu (E51), Rombo (E623), Digo (E73), Bende (F12), Rangì (F33), “Normal” Mbugu (G221KK), Chindamba (G52), Kinyarwanda (JD61), Kifuliiru (JD63), Ha (JD66), Nyoro (JE11), Luganda (JE15), Matengo (N13), Sena (N44), Yao (P21), Makhuwa (P31) and Cuwabo (P34). The languages of the sample have been chosen to provide the comparative context for our comparison of Old Swahili and Standard Swahili. They are all spoken in the East African region, and they all belong to the Eastern or Southeastern group of Bantu languages (cf. Grollemund et al. 2015) and include languages from all six of Guthrie’s (1967–1971) East African zones (J, E, F, G, N and P). They thus provide a balanced, if somewhat selective and unsystematic, snapshot of the linguistic context in which Swahili is used and as such an appropriate background for comparison in geographic and genetic-linguistic terms.<sup>16</sup>

The comparison of the languages is based on the comparative Bantu Morphosyntactic Variation (BMV) database (Marten et al. 2018) and includes values for up to 142 parameters for the twenty languages of the sample (although for most languages of the sample we do not have a complete data set).

A summary of all twenty languages of our sample, including Old Swahili and Standard Swahili, is provided in Table 6.

---

<sup>16</sup>However, we have not taken into account differences in the sociolinguistic profiles of the languages of the sample, e.g. the use as cross-border languages, as regional lingua francas, use in education or wider public domains, or levels of language shift and endangerment. Since our findings in part reflect the sociolinguistic history of Swahili, taking into account the sociolinguistic histories of the other languages of the sample would provide a promising avenue for further research.

Table 6: Languages of the quantitative comparison

Language name	Language code	Main location of use
Nyole	E35	Kenya, Uganda
Gikuyu	E51	Kenya
Digo	E73	Kenya
Rombo	E623	Tanzania
Bende	F12	Tanzania
Rangi	F33	Tanzania
Mbugu	G221KK	Tanzania
Standard Swahili	G42	Kenya, Tanzania
Old Swahili	G42_Old	Kenya, Tanzania
Chindamba	G52	Tanzania
Kinyarwanda	JD61	Rwanda
Kifuliiru	JD63	DRC
Ha	JD66	Tanzania
Nyoro	JE11	Uganda
Luganda	JE15	Uganda
Matengo	N13	Tanzania
Sena	N44	Mozambique
Yao	P21	Malawi, Mozambique, Tanzania
Makhuwa	P31	Mozambique
Cuwabo	P34	Mozambique

For our comparative analysis we used a report available in the database which calculates the pairwise similarity between the languages, so that for each language pair, we have the percentages of shared parameter values. This is a measure of how similar two languages are, given by the percentage of parameters for which the two languages have the same value. It is based on the method used in lexicostatistics to measure the percentage correspondence of lexical cognates between two languages (Swadesh 1952). The results of the comparison are provided in Table 7. The shared percentages are based on the available data for each language pair, so that only parameters are taken into account for which we have values for both languages of the pair. The percentage calculated for each language pair then reflects the number of parameters with the same value out of all parameters with values for both languages.

Table 7: Pairwise comparison of Old Swahili and Standard Swahili in the context of 18 selected Eastern African Bantu languages (20 languages in total)

	E35	E51	E623	E73	F12	F33	G221KK	G42-Ash	G42_Old	G52	JD61	JD63	JD66	JE11	JE15	N13	N44	P21	P31
E35																			
E51	63%																		
E623	61%	65%																	
E73	57%	58%	66%																
F12	68%	60%	66%	62%															
F33	61%	64%	74%	65%	70%														
G221KK	65%	67%	71%	62%	70%	65%													
G42-Ash	57%	58%	58%	65%	65%	64%	66%												
G42_Old	63%	67%	67%	85%	69%	70%	70%	87%											
G52	59%	58%	62%	58%	63%	71%	65%	67%	62%										
JD61	66%	55%	54%	52%	61%	57%	58%	53%	56%	57%									
JD63	64%	52%	51%	58%	66%	61%	56%	53%	53%	56%	59%								
JD66	54%	55%	62%	68%	65%	63%	57%	58%	66%	60%	58%	60%							
JE11	73%	59%	58%	60%	70%	64%	65%	57%	67%	64%	67%	56%	61%						
JE15	74%	61%	57%	56%	61%	57%	61%	57%	62%	60%	57%	58%	54%	72%					
N13	64%	51%	62%	55%	62%	57%	51%	49%	48%	55%	54%	53%	49%	58%	53%				
N44	51%	45%	51%	55%	59%	56%	56%	62%	66%	59%	51%	53%	62%	50%	54%	48%			
P21	56%	60%	57%	59%	67%	66%	60%	62%	66%	61%	52%	54%	56%	62%	55%	63%	53%		
P31	54%	53%	55%	58%	61%	54%	62%	64%	67%	51%	53%	53%	56%	58%	56%	52%	63%	54%	
P34	47%	48%	47%	56%	57%	50%	57%	60%	68%	49%	46%	50%	58%	54%	52%	49%	62%	56%	80%

The comparative data show that shared parameters range from the highest similarity of 87% (between Old Swahili and Standard Swahili) to the lowest similarity of 45%, between Gikuyu (E51) and Sena (N44). We have noted the 87% similarity between Old and Standard Swahili before, but the figure can now be seen in a wider comparative Bantu context. Given the overall typological similarity between Eastern Bantu languages – most languages of the sample have 50% or more shared values – the high amount of shared values between Old and Standard Swahili confirms their status as very closely related varieties.

The data also show a clear difference between Old Swahili and Standard Swahili with respect to the other languages of the sample. Overall, Old Swahili is more similar than Standard Swahili to the other languages of the sample, with respect to the morphosyntactic parameters. The values of the pairings involving Old Swahili and Standard Swahili are summarised in Table 8. The data show that, typically, the shared value for Old Swahili and another language is higher than the shared value of Standard Swahili with the same language.

The relevant difference can be seen, for example, with Nyolo (E35) which shares 63% of the parameter values with Old Swahili, but only 57% with Standard Swahili, a difference of 6%. In fact, it is true for 15 out of the 18 pairings that the shared value with Old Swahili is higher than the shared value with Standard Swahili, and only in three cases does this not hold. In one pairing, the values are the same: Both Old Swahili and Standard Swahili share 53% of value with Kifuliiru (JD63). In two pairings, the percentage for Standard Swahili is higher than the percentage for Old Swahili: With Chindamba (G52), Old Swahili shares 62% of parameter values, but Standard Swahili shares 67%. With Matengo (N13), Old Swahili shares 48% of values, but Standard Swahili shares 49%. A possible explanation for this difference is that both Chindamba and Matengo are Tanzanian community languages which have been shown to have been heavily influenced by Swahili, particularly in more recent years (see Yoneda 2010, Kutsukake & Yoneda 2019 for Matengo and Edelsten & Lijongwa 2010 for Chindamba). Given the prevalence of (Standard) Swahili in the areas where these two languages are spoken, the higher percentage can be seen to reflect a higher level of language contact and multilingualism in these areas in the present day, and the resulting convergence effects.

In some cases, the difference in shared values is comparatively small, e.g. 56% vs. 53% in the case of Kinyarwanda (JD61), while in others it is quite considerable. The biggest difference is found with Digo, with 85% vs. 65%. The case of Digo is interesting, as the data show that the similarity between Digo and Old Swahili (85%) is about the same as the similarity between Old and Standard Swahili (87%). Digo and Swahili are closely related – both are members of the Eastern-Bantu

Table 8: Pairwise similarity of Old Swahili and Standard Swahili with 18 East African Bantu languages

Language code	Language name	% shared with Old Swahili	% shared with Standard Swahili	Difference
G52	Chindamba	62%	67%	+5%
N13	Matengo	48%	49%	+1%
JD63	Kifuliiru	53%	53%	0%
JD61	Kinyarwanda	56%	53%	-3%
P31	Makhuwa	67%	64%	-3%
F12	Bende	69%	65%	-4%
G221KK	Mbugu	70%	66%	-4%
N44	Sena	66%	62%	-4%
P21	Yao	66%	62%	-4%
JE15	Ganda	62%	57%	-5%
E35	Nyolo	63%	57%	-6%
F33	Rangi	70%	64%	-6%
P34	Cuwabo	68%	60%	-8%
JD66	Ha	66%	58%	-8%
E51	Gikuyu	67%	58%	-9%
E623	Rombo	67%	58%	-9%
JE11	Nyoro	67%	57%	-10%
E73	Digo	85%	65%	-20%

Sabaki sub-group – and Digo is the closest relative to Swahili in our sample. The comparison shows that there is a very close morphosyntactic resemblance between Old Swahili and Digo, but that the resemblance is much less close with Standard Swahili. As noted above, there are at least two relevant explanations for this difference. Firstly, Northern dialects of Swahili, which had a stronger influence on Old Swahili than on Standard Swahili, are likely to be more similar to Digo, spoken in Kenya, than Southern dialects of Swahili, and so the difference reflects the difference in Swahili dialects. Secondly, the difference is also likely to be an effect of standardisation, which resulted in a development away from other Bantu languages overall, and in particular in changes away from historically closely related languages like Digo.

The difference in similarity between Old Swahili and Standard Swahili with respect to neighbouring languages can also be seen from the weighted average

of similarities. The weighted average match percentage of a given language is the average match percentage of that language compared with all the other languages of the sample, weighted by their respective numbers of common parameters. In other words, this value calculates all similarity values for each language, resulting in one overall value, and the higher the value, the more similar the language is to the rest of the sample. The relevant data are summarised in Table 9.

Table 9: Weighted averages of Old Swahili and Standard Swahili in the context of 18 East African Bantu languages

Language code	Language name	Weighted average
G42_Old	Old Swahili	66%
F12	Bende	64%
F33	Rangi	63%
G221KK	Mbugu	62%
JE11	Nyoro	62%
E35	Nyolo	61%
G42-Ash	Standard Swahili	61%
E623	Rombo	60%
E73	Digo	60%
G52	Chindamba	60%
JD66	Ha	59%
P21	Yao	59%
JE15	Ganda	58%
P31	Makhuwa	58%
E51	Gikuyu	57%
JD61	Kinyarwanda	56%
JD63	Kifuliiru	56%
N13	Matengo	55%
N44	Sena	55%
P34	Cuwabo	55%

The data in Table 9 show that values for weighted average are distributed quite narrowly, ranging from 55% to 66%. When interpreting the data, this has to be kept in mind, and probably not too much should be read into very small differences in percentage points between different languages. However, against this backdrop, the data show that in terms of the morphosyntactic parameters assumed in this study, Old Swahili has the highest weighted average with 66%,



while Standard Swahili is found lower in the table, with 61%. Data from weighted average thus show (in a different way than data from pairwise comparison) that Old Swahili is morphosyntactically more similar to the East African Bantu languages of the sample than Standard Swahili is.

This finding dovetails with previous work on grammatical complexity in Standard Swahili. Specifically, Jerro (2018) compares Standard Swahili to five East African Bantu languages (Kinyarwanda, Gikuyu, Lingala, Haya, and Luganda) in their morphophonological complexity, measured mostly by phonological and morphological inventory sizes (cf. Kusters 2003, McWhorter 2011). The conclusion of the study is that while Standard Swahili differs in many ways from other Bantu languages, there is no evidence that it exhibits less morphological or phonological complexity than the other Bantu languages spoken in the area. While Jerro (2018) looks at the role of bilingualism between Swahili and Arabic as a potential explanation for the divergence of Standard Swahili from other Bantu languages (cf. Trudgill 2009, 2011), the effect of bilingualism would have been present through both Old and Standard Swahili, and therefore cannot be driving the differences between them. Combining the findings of that work and the present paper, we see that the diachronic changes that give way to Standard Swahili from Proto-Bantu are an admixture of language contact/bilingualism, standardisation, loss and innovation.

## 5 Conclusions

The study of language change has always played an important role in Bantu linguistics, and there is a long history of comparative-historical work (cf. van der Spuy forthcoming). However, this work has often focused on lexical and phonological data, and on synchronic evidence due to the (perceived) absence of historical data for Bantu languages. The current study extends the debate, by using morphosyntactic data from historical texts, and by adopting both qualitative and quantitative methods of comparison. The study has focused on Old Swahili – the language used in classical Swahili poetry of the twentieth and earlier centuries – and compared selected morphosyntactic features of Old Swahili with Standard Swahili, and with a sample of 18 neighbouring East African Bantu languages. The methodology adopted for the comparison is based on the Bantu morphosyntactic parameters developed in Guérois et al. (2017) and uses the associated database of Marten et al. (2018).

There are inescapable restrictions in the study of historical texts, and Swahili is no exception. Our corpus of Old Swahili is based on a single genre – religious

poetry – and as a result is very limited in terms of text types and genres. On the other hand, it includes texts from a variety of writers, places, and times, and so is, in these respects, heterogeneous. Furthermore, there is no straightforward diachronic relation between Old Swahili and Standard Swahili, as the former is largely formed of northern dialects of Swahili, while the latter is mainly based on the southern dialect of Zanzibar, Kiunguja. For the purpose of comparison, we have assumed an idealised version of Old Swahili, based in Mieke's (1979) work on the language of classical Swahili poetry, and have likewise based our analysis of Standard Swahili on descriptive works such as Ashton (1947) and Schadeberg (1992), which was supplemented by native speaker judgements.

The starting point of our analysis was the comparison of Old Swahili and Standard Swahili with respect to the parameters of Guérois et al. (2017), and we have shown that out of the 61 parameters with data in both varieties, 53 are shared and 8 differ, resulting in 87% similarity. When looking at the differences in more detail, we have shown that most of them result from loss of either form or function, while there are comparatively few innovations. In addition, there are several instances of regularisation of functions of paradigms, which we have attributed, at least in part, to the process of standardisation which led to the development of Standard Swahili from the early 20<sup>th</sup> century onwards, and to the effect of second-language speaker agency in the process.

We then turned to a quantitative analysis, where we compared the difference between Old Swahili and Standard Swahili in the wider context of East African Bantu languages, based on a sample of 18 East African Bantu languages. A pairwise comparison of Old Swahili and Standard Swahili with the languages of the sample showed that, overall, Old Swahili is more similar to neighbouring languages. In 15 out of the 18 pairings, the shared values of the relevant language with Old Swahili are higher than the values shared with Standard Swahili. The difference was particularly notable in relation to Digo, a closely related Mijikenda language of Kenya, which shows 85% similarity with Old Swahili, but only 65% with Standard Swahili. We have suggested that the difference illustrates the trajectory of Swahili from its closest neighbours to a standardised language of wider communication.

A second set of data was shown to illustrate the same point in a slightly different way. We constructed weighted averages across all pairwise values for a given language to provide an indication of the overall similarity of each language to all other languages in the sample. While the range of values for the sample was narrow (ranging from 55% to 66%), we noted that Old Swahili had a higher score (66%) than Standard Swahili (61%). We have proposed that this distribution shows that through processes of regularisation and standardisation,

Standard Swahili has developed away from neighbouring Bantu languages in terms of morphosyntax.

Results of the study show significant differences between the two varieties. In particular, it shows that the relation between Old Swahili and Standard Swahili is characterised by a loss of variability. This is most likely related to processes of language change, but also more specifically to the processes of language planning and standardisation in the formation of Standard Swahili. The results of the study provide a good demonstration of these effects with respect to morphosyntax.

The findings of the study shed new light on morphosyntactic variation since they show the effect of standardisation and a particular trajectory of morphosyntactic development. They also show the strength of combining qualitative and quantitative methods in the study of morphosyntactic variation. For the examination of Bantu languages and the associated morphosyntactic variation, the study points to the importance of the development of languages of wider communication. There are several Bantu languages which have developed to become national or regional *lingua francas*, and their relation to neighbouring Bantu languages may have been affected similarly to what we have shown for Swahili. It is certainly a factor which should be kept in mind in future comparative studies.

Finally, the study also provides a meaningful background for the development of non-standard varieties of Swahili (and other *lingua francas*) such as youth languages like Sheng, but also colloquial varieties such as Kenyan or Mainland Tanzania varieties of Swahili. In these, we can often see processes which are the inverse to the regularisation effects observed here – including increase of variability and the re-introduction of morphosyntactic features often through contact with neighbouring Bantu languages which have maintained these features. A more detailed investigation of these varieties along the lines of the current study would be very likely to yield interesting results.

## **Acknowledgements**

Parts of the research presented in this chapter were supported by a Leverhulme Trust Grant (RPG-2014-208) for the project “Morphosyntactic variation in Bantu: typology, contact and change” (Lutz Marten), a British Academy Postdoctoral Research Fellowship for the project “Pathways of change at the northern Bantu borderlands” (Hannah Gibson), and a Leverhulme Trust Grant (RPG-2021-248) for the project “Grammatical variation in Swahili: contact, change and identity” (Hannah Gibson), all of which are hereby gratefully acknowledged. We are grateful to Ida Hadjivayanis, Gudrun Mieke, Muaadh Salih, and the late Dr Abel Mreta,

as well as audiences at Beijing Foreign Studies University, Shanghai International Studies University, SOAS, the seventh international Bantu conference (Cape Town, July 2018), and Syntax of the World's Languages 6 (Paris, September 2018) for helpful comments and suggestions on the findings of this chapter.

## Abbreviations

Glossing conventions follow the Leipzig Glossing Rules with the following additions:

1, 2, 3 etc.	noun class number	PLUR	pluractional
AUG	augment	PREP	preposition
CD	concord	PRO	pronoun
CONN	connective	RED	reduplication
CONJ	conjunction	REF	referential
FV	final vowel	REL	relative
INC	inceptive	SBJV	subjunctive
INT	intensive	SG	singular
OM	object marker	SIT	situative
PERS	persistent	SM	subject marker
PLA	plural addressee		

## References

- Ashton, Ethel O. 1947. *Swahili grammar*. 2nd edn. London: Longman.
- Baldi, Sergio. 2012. Arabic loans in East African languages through Swahili: A survey. *Folia Orientalia* 49. 37–52.
- Balestrieri, Fiona Agnes. 2017. *An analysis of Swahili lexical borrowings into Shambala, Nyamwezi, and Haya: 1880s to present*. SOAS University of London. (MA thesis).
- Beaudoin-Lietz, Christa, Derek Nurse & Sarah Rose. 2004. Pronominal object marking in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 175–188. Cologne: Köppe.
- Bertoncini Zúbková, Elena, Mikhail Gromov, Said Khamis & Kyallo Wamitila. 2009. *Outline of Swahili literature*. 2nd edn. Leiden: Brill.
- Blommaert, Jan. 2014. *State ideology and language in Tanzania*. 2nd edn. Edinburgh: Edinburgh University Press.

- Dom, Sebastian & Koen Bostoen. 2015. Examining variation in the expression of tense/aspect to classify the Kikongo language cluster. *Africana Linguistica* 21. 163–211.
- Drolc, Ursula. 2000. Zur Typology des Perfekts (am Beispiel des Swahili). In Werner Breu (ed.), *Probleme der Interaktion von Lexik und Aspekt (ILA)*, 91–112. Tübingen: Niemeyer.
- Edelsten, Peter & Chiku Lijongwa. 2010. *A grammatical sketch of Chindamba: A Bantu language (G52) of Tanzania*. Cologne: Köppe.
- Fleisch, Axel. 2005. Agent phrases in Bantu passives. In Erhard Voeltz (ed.), *Studies in African linguistic typology*, 93–111. Amsterdam: John Benjamins.
- Gibson, Hannah, Gastor Mapunda, Lutz Marten, Sheena Shah & Julius Taji. 2019. Variation in double object marking in Swahili. In Daisuke Shinagawa & Nico Nassenstein (eds.), *Swahili Forum 26, Special Issue "Variation in Swahili"*, 142–165.
- Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences* 112(43). 13296–13301.
- Guérois, Rozenn. Forthcoming. Passive constructions. In Lutz Marten, Nancy Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages*. Oxford: Oxford University Press.
- Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list*. Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.
- Guthrie, Malcolm. 1967–1971. *Comparative Bantu: An introduction to the comparative and prehistory of the Bantu languages*, vol. 1–4. London: Gregg International.
- Harries, Lyndon. 1962. *Swahili poetry*. Oxford: Clarendon Press.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.
- Jerro, Kyle. 2018. Linguistic complexity: A case study from Swahili. In Jason Kandybowicz, Travis Major & Harold Torrence (eds.), *The Selected Proceedings of the 45th Annual Conference on African Linguistics: African Linguistics on the Prairie*, 3–20. Berlin: Language Science Press.
- Knappert, Jan. 1967. *Traditional Swahili poetry: An investigation into the concepts of East African Islam as reflected in the Utenzi literature*. Leiden: Brill.

- Krapf, Johann Ludwig. 1850. *Outline of the elements of the Kisuáheli language with special reference to the Kinika dialect*. Tübingen: L.F.Fues.
- Krumm, Bernard. 1940. *Words of oriental origin in Swahili*. London: Sheldon Press.
- Kusters, Wouter. 2003. The fate of complex languages: Classical Arabic and Old Norse in the age of globalisation. *Nordlyd* 31. 275–289.
- Kutsukake, Sayaka & Nobuko Yoneda. 2019. Contact-induced language divergence and convergence in Tanzania: Forming new varieties as language maintenance. *Swahili Forum* 26. 181–204.
- Lodhi, Abdulaziz Y. 2000. *Oriental influences in Swahili: A study in language and culture contacts*. Gothenburg: Acta Universitatis Gothoburgensis.
- Marlo, Michael R. 2015. On the number of object markers in Bantu languages. *Journal of African Languages and Linguistics* 36(1). 1–65. DOI: 10.1515/jall-2015-0001.
- Marten, Lutz. 1998. Swahili *-kwisha*: Sketching the path of grammaticalization. *SOAS Working Papers in Linguistics and Phonetics* 8. 141–163.
- Marten, Lutz, Peter Edelsten, Hannah Gibson & Rozenn Guérois. 2018. *Bantu Morphosyntactic Variation (BMV) database*. London: SOAS University of London.
- Marten, Lutz & Nancy Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30(2). 237–253.
- Marten, Lutz, Nancy Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105. 253–338.
- Mazrui, Alamin. 2007. *Swahili beyond the boundaries*. Athens: Ohio University Press.
- Mbaabu, Ireri. 1978. *Kiswahili: Lugha ya taifa*. Nairobi: Kenya Literature Bureau.
- McWhorter, John. 2011. *Linguistic simplicity and complexity: Why do languages undress?* Berlin: Mouton de Gruyter.
- Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
- Miehe, Gudrun. 1979. *Die Sprache der älteren Swahili-Dichtung: Phonologie und Morphologie*. Berlin: Reimer.
- Mlacha, S. A. K (ed.). 1995. *Kiswahili katika Kanda ya Afrika Mashariki*. Dar es Salaam: Taasisi ya Uchunguzi wa Kiswahili.
- Mugane, John M. 2015. *The story of Swahili*. Athens, OH: Ohio University Press.
- Mulokozi, Mugyabuso M. & Tigiti S. Y. Sengo. 1995. *History of Kiswahili poetry: A. D. 1000–2000*. Dar es Salaam: Institute of Swahili Research.
- Muzale, Henry R. T. 1998. *A reconstruction of the Proto-Rutaran tense/aspect system*. Memorial University of Newfoundland. (Doctoral dissertation).

- Mwaliwa, Hanah Chaga. 2018. Modern Swahili: The integration of Arabic culture into Swahili literature. *Tydskrif vir Letterkunde* 55(2). 120–133.
- Nicolle, Steve. 2012. Semantic-pragmatic change in Bantu *-no* demonstrative forms. *Africana Linguistica* 18. 193–234.
- Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford: Oxford University Press.
- Nurse, Derek & Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history*. Berkeley: University of California Press.
- Rugemalira, Josephat M. 2007. The structure of the Bantu noun phrase. *SOAS Working Papers in Linguistics* 15. 135–148.
- Schadeberg, Thilo C. 1989. The three relative constructions in Swahili (Kisanifu). In Marie-Françoise Rombi (ed.), *Le swahili et ses limites: Ambiguïté des notions reçues*, 33–40. Paris: Editions Recherche sur les Civilisations.
- Schadeberg, Thilo C. 1990. Schon – noch – nicht – mehr: Das Unerwartete als grammatische Kategorie im KiSwahili. *Frankfurter Afrikanistische Blätter* 2. 1–15.
- Schadeberg, Thilo C. 1992. *A sketch of Swahili morphology*. 3rd edn. Cologne: Köppe.
- Shinagawa, Daisuke & Yuko Abe (eds.). 2019. *Descriptive materials of morphosyntactic microvariation in Bantu*. Tokyo: Research Institute for Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.
- Steere, Edward. 1870. *A handbook of the Swahili language as spoken at Zanzibar*. London: Bell & Daldy.
- Steere, Edward. 1884. *A handbook of the Swahili language as spoken at Zanzibar*. 3rd edn. London: Society for the Promotion of Christian Knowledge.
- Swadesh, Morris. 1952. Lexico-statistical dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society* 96. 452–463.
- Taylor, William Ernest. 1891. *African aphorisms or, saws from Swahili-land*. Collected, translated and annotated by the Rev. W. E. Taylor. London: Society for Promoting Christian Knowledge.
- Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an eng variable*, vol. i, 98–109. Oxford: Oxford University Press.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics with special focus on Bantu and Mande*, 425–441. Tervuren: Royal Museum for Central Africa.

- van der Spuy, Andrew. Forthcoming. A history of research into Bantu linguistics. In Lutz Marten, Nancy Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages*. Oxford: Oxford University Press.
- Vierke, Clarissa. 2011. *On the poetics of the Utendi: A critical edition of the nineteenth-century Swahili poem “utendi wa haudaji” together with a stylistic analysis*. Zurich: LIT.
- Voeltz, F. K. Erhard. 1980. The etymology of the Bantu perfect. In Luc Bouquiaux (ed.), *L’expansion bantue*, 487–492. Paris: Centre National de la Recherche Scientifique.
- Whiteley, Wilfred Howell. 1955. Kimvita: An inquiry into dialectal status and characteristics. *Journal of the East African Swahili Committee* 25. 10–39.
- Whiteley, Wilfred Howell. 1969. *Swahili: The rise of a national language*. London: Methuen.
- Yoneda, Nobuko. 2010. “Swahilization” of ethnic languages in Tanzania: The case of Matengo. *African Study Monographs* 31(3). 139–148.
- Zawawi, Sharifa. 1979. *Loan words and their effect on the classification of Swahili nominals*. Leiden: Brill.