▣ AN ANALYSIS OF THE VERBAL MARKER *TSA* IN LUGURU

Malin Petzell

University of Gothenburg

This paper deals with a morphosyntactic phenomenon found in the under-described Bantu language Luguru, spoken in central Tanzania: the verbal marker *tsa*. This marker encodes shared knowledge or shared reference. The meanings conveyed by the marker stretch from 'at a specific time' or 'at that place' to 'as we know', or even 'for that reason'. In Mkude's grammatical description of Luguru from 1974, there is a mention of a marker (*zaa*) signalling what he calls "recollected reference", which restricts the event to one specific moment in the past; this marker is believed to have developed into today's *tsa*.

1. INTRODUCTION

This paper explores the verbal marker *tsa* in the under-described Bantu language Luguru, classified as G35 in Guthrie's (1967–1971) standard classification (ISO 639-3: ruf).[1] Luguru is spoken by 403,602 people in the Morogoro region of central Tanzania (Languages of Tanzania Project 2009) and is the major mother tongue in the region. Swahili is the official language and almost all Luguru speakers are bilingual (Petzell & Khül 2017: 36). Below is a map of the languages of the Morogoro region (Figure 1), in which Luguru is marked with green hexagons.

The verbal marker *tsa*, it will be argued, encodes shared knowledge or shared reference. The meanings conveyed by the marker range from 'at a specific time' or 'at that place' to 'as we know', or even 'for that reason'. Typically, it is used to express determinate moments in the past. Compare *nikala* 'I sat' without the marker, in (1), with **tsa**nikala 'I sat at that point in time', with the marker prefixed, in (2). Apart from adding a temporal (or even explanatory) deixis, the *tsa* form in (2) also implies that the event is not ongoing (i.e. I am not sitting anymore). The marker itself is most likely a grammaticalization of the verb *kutsa* 'to come'. In contexts where the marker occurs, it is always optional, although there are some settings where it is unidiomatic; see Section 3.3.

Finnish Oriental Society

Figure 1 Map of "linguistic centres" in the Morogoro region, Tanzania.[2]
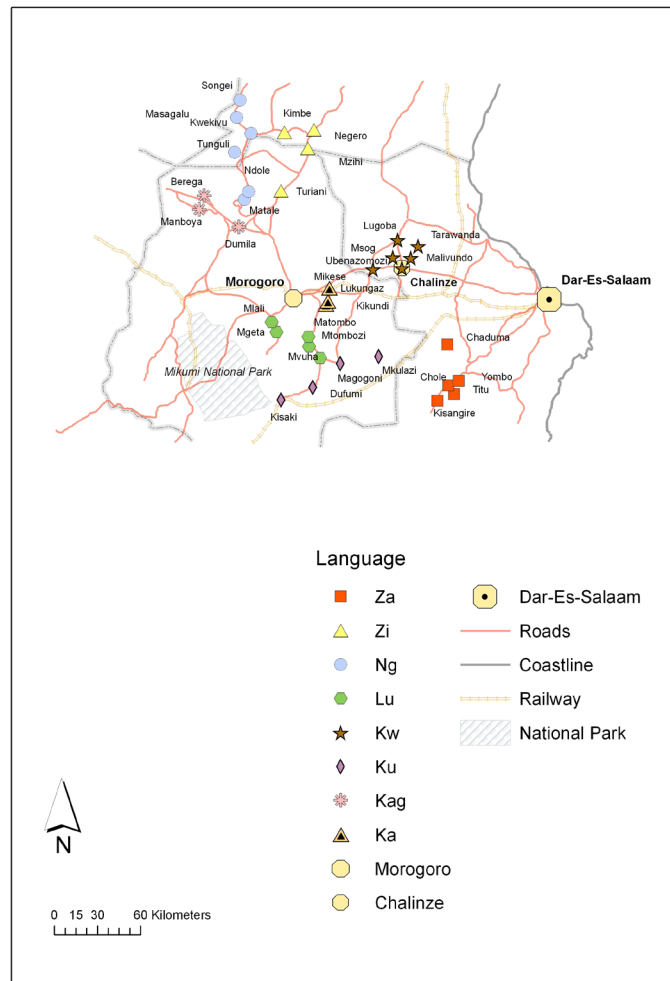
(1)   *Ni-kal-a.*
      1SG.SM-sit-FV

      'I sat.' (It is not important when, where, or why the person sat.)

(2)   ***Tsa**-ni-kal-a.*
      TSA-1SG.SM-sit-FV

      'I sat then (or 'for that reason').'

Given that the marker *tsa* is used in the verb phrase, this paper will begin by providing a brief sketch of the Luguru verb (Section 2). Additionally, since there are no grammatical descriptions of modern Luguru, this account of the verb will contribute to our knowledge of the Luguru spoken today. Following the sketch, the marker itself is explored in Section 3 – its form, usage and non-usage, frequency and distribution, and grammaticalization. The final section in this

---

2 The map was created by the cartographer, Ulf Ernstson (from the Department of Human and Economic Geography, University of Gothenburg), and Malin Petzell (from the Department of Languages and Literatures, University of Gothenburg). Languages are abbreviated to the two (or three) first letters in the name: Zalamo (G33, ISO 639-3: zaj), Zigua (G31, ISO 639-3: ziw), Nguu (G34, ISO 639-3: ngp), Luguru (G35, ISO 639-3: ruf), Kwere (G32, ISO 639-3: cwe) Kutu (G37, ISO 639-3: kdc), Kagulu (G12, ISO 639-3: kki), and Kami (G36, ISO 639-3: kcu).

paper, Section 4, concludes that the marker *tsa* is used to establish rapport by functioning as an anchor to set the time and the place, and that it is only infelicitous together with vague references to indefinite time.

The Luguru data that form the basis for this paper were collected by the author during field trips to the Morogoro region in 2014 and 2016–2019. The data are supplemented by email conversations that took place between 2017 and 2020. All examples in this paper, apart from the Bible passages, come from the author's fieldwork. The data consist mostly of structured interviews and elicitation, including translations of wordlists, sentences, and stories. The 1184 token sentences alone form a database which is tagged for tense, aspect, and negation, among other features (see Jordan & Petzell, in press). In addition to this database, other sources are the recently published translation of the New Testament into Luguru (available at <world-bibles.org/language_detail/eng/ruf/Luguru>), pre-publication drafts of the same (Pioneer Bible Translators, pers. comm.), and a small booklet with four traditional stories (Lukanza et al. 2001). The different versions of the New Testament, edited over a number of years, turned out to be highly interesting, since some instances of *tsa* occur in the draft but not in the printed version, and others occur in the printed version but not in the draft. This is discussed further in Section 3.2. Other sources of Luguru include a grammar of Luguru (Mkude 1974) and a dated grammatical sketch containing some elicited sentences (Seidel 1898). The works of Johnston (1922) and Guthrie (1948) also include some Luguru data. More recent sources discussing specific phenomena include, for instance, Marten & Ramadhani (2001) and Marten (2003), while comparative works include Petzell (2012b) and Petzell & Hammarström (2013). None of these sources, apart from Mkude (1974), mentions the marker *tsa*.

The informants are all mother tongue speakers of Luguru and about a third of them work as Bible translators. All of them were born in, and all but one still live in the Morogoro region. The Luguru speakers are bilingual in Swahili, which is the national language of Tanzania and is prevalent in all public settings, including school and church (Petzell 2012a). Swahili was used as a vehicular language during fieldwork, and occasionally English, with the few informants who can speak it. There is always the risk that the vehicular language affects the data collection, and even more so when the vehicular language is dominant in the region and code switching between the two is the rule rather than the exception. In order to avoid influence from the metalanguage (or metalanguages), transformational elicitation (Kibrik 1977: 60; Mosel 2011: 84) was often employed. In transformational elicitation, the informant is asked to transform rather than translate a grammatical construction. For instance, the informant can be given a sentence in the past tense and asked to transform it into the present tense, or to change an ongoing action into a completed one (Bloom Ström & Petzell, in press). That is, instead of asking for a translation of a sentence where the informants are "provided" with the tense/aspect, such as the Simple Past or Perfect in Swahili (*-li-* or *-me-*), the informants were given the sentence in the present, such as *Tunaenda Dar es Salaam* 'We are going to Dar es Salaam' and asked to say it as if the event took place this morning, yesterday, or ten years ago, or as if it had taken place already.

2. THE VERB IN LUGURU

Relative to most other Bantu languages, the Luguru verbal morphology is reduced and there are exceptionally few tense, aspect, and mood (TAM) markers. My data also show that Luguru has no tonal contrast, either lexical or grammatical (see also Guthrie 1948: 50). The minimal shape for the verb is the root plus a final vowel, but the structure is generally more morphologically

complex and includes at least one other affix. The agglutinating structure is illustrated in the template below in Table 1. The template and the labels are based on Meeussen (1967: 108–111) and Nurse (2008a), with my addition of a pre-pre-initial slot containing the marker *tsa* ahead of the first slot in their templates. The reason for dividing the template into slots is simply to show how the morphemes concatenate (see also Maho 2007). The ordering of morphemes is very strict in Bantu, while the ordering of syntactic elements is less restricted.

Table 1 Template of the Luguru verbal slots

Slot	Marker
1. Pre-pre-initial	*tsa*
2. Pre-initial	Relative, negation
3. Initial	Subject marker
4. Formative	First tense/aspect/mood marker
5. Infix	Object marker
6. Radical	Verb root
7. Suffix	Extensions
8. Pre-final	Second tense/aspect marker
9. Final	Final vowel
10. Post-final	Plurality

It should be noted that Meeussen describes how, infrequently, two different types of formatives may occupy his first slot (Meeussen 1967: 108), but he does not divide the slot into two. This additional pre-pre-initial slot also exists in neighbouring Kagulu G12 (Petzell 2008: 98–100). In neither Meeussen's data nor in Kagulu is the slot occupied by an element that is assumed to be a grammaticalized verb such as *tsa*. Usually a relative marker, a conditional/temporal marker, or negation is found in this initial position.

The first slot can contain the marker *tsa*, which is the topic of this paper. The second slot, that is the pre-initial, may host a relative marker, the conditional/temporal marker 'if/when', or a negation marker. Slot 3 can only hold the subject marker and slot 4 hosts the TAM formatives. Slot 5 is restricted to holding the object marker or a reflexive marker, while slot 6 constitutes the root of the verb. The verbal extensions are in slot 7, while slot 8 again takes TAM markers; two markers alternate in this slot, namely Imperfective *-ag-* and Perfective *-ile* (which takes up slot 9 as well). The final vowel (FV) is in slot 9 and the post-final plurality marker (plural of addressees) is in slot 10. The FV is inflectional, like the other TAM markers, but it is not optional. The "default" FV is the indicative *-a* (see Rose, Beaudoin-Lietz & Nurse 2002: 32). The only slots that are obligatorily filled are slot 6 (the root) and slot 9 (the FV), which combined represent the Imperative. The majority of the slots can be filled at the same time; the restrictions lie in which morphemes can co-occur. Some of the morphemes may have merged – the present tense, for instance, is marked by a merger of the morphemes in slots 3 (the subject marker) and 4 (the TAM marker), and as such, it may be difficult to tell what slots they actually occupy. As mentioned above, the linear ordering of the morphemes is more important than the division into slots. In Table 2 below, the Luguru sentence in example (3) is fitted into the linear template. Notice that the Past has null realization in slot 4.

(3) *Tsa-ni-mw-ambik-il-a   i-mw-ana   pfi-moka.*
 TSA-1SG.SM-OM1-cook-APPL-FV AUP-1-child 8-potato

 'I cooked the child potatoes.'

Table 2 Linear template containing an example sentence

1	2	3	4	5	6	7	8	9
Pre-pre-In	**PreIn**	**In**	**Fo**	**Infx**	**Rad**	**Sfx**	**PreF**	**F**
tsa		*ni*		*mw*	*ambik*	*il*		*a*

2.1 Tense, aspect and mood

There are only two pre-root tense markers, and one conditional marker, that can occupy slot 4. The tense markers are the present -*o*-, which merges with the preceding subject marker, and the Future, which is marked with -*tso*- (4). There is also the null realization here, which is used for the Past (both Perfective and Imperfective) and for the Anterior. Mkude (1974: 319) calls the null realization a Static Tense, or Perfect Definite. Note that the future marker in (4) may very rarely carry a modal meaning:

(4) *tu-**tso**-gend-a*
 1PL.SM-**FUT**-go-FV

 'we will go' (or rarely 'lest we go' in certain contexts)

The present marker -*o*- and the future -*tso*- are both subject to vowel height harmony and are realized as -*a*- and -*tsa*-[3] respectively when followed by the infinitive (class 15) -*ku*- or the reflexive object marker -*i*- (Mkude 1974: 96). This is mostly corroborated by my data, where -*a*- is used in compound verb forms, that is, when followed by -*ku*- or (-*kw*-) (5), although there are no occurrences of reflexives in the future tense in my data.

(5) *Esta ka-**tsa-ku**-w-a ka-many-a.*
 PN   SM1-**FUT-15**-be-FV SM1-know-FV

 'Esther will have known.'

There is a third marker, a Conditional (-*ng'a*-),[4] that can also occur in slot 4, as seen in (6).

(6) *u-**ng'a**-ts-a*
 2SG.SM-**COND**-come-FV

 'if you come [...]'

Apart from the pre-root TAM markers mentioned, there are also two post-root markers. As is common in Bantu, the pre-root markers usually pertain to tense and the post-root markers usually pertain to aspect (Nurse & Devos 2019: 206). The post-root markers in Luguru are the imperfective -*ag*- in slot 8, which is in complementary distribution with the perfective marker -*ile* (intermittently -*ire*) in the merged slots 8 and 9; see (7). Although -*ile* probably used to

---

3 Note that this realization of the future is not immediately related to the marker *tsa* which is the topic of this paper, even though they may originally have stemmed from the same verb 'come'.
4 <ng'> is the grapheme for a velar nasal.

mark the Perfective (Nurse 2008a: 264), in present day Luguru it is only used in conditional/ temporal constructions, negatives and relatives. This is substantiated by Mkude (1974: 11), who states that *-ile* "is only used in dependent and in negative clauses", and by Guthrie (1948: 49), who notes that there are some G30 languages where *-ile* behaves curiously and does not occur in "regular" affirmative sentences.

2.2 Negation

The negative morpheme is *si-* (7) for the first person singular and *ha-* (8) (also realized as *ng'a-*) in all other persons and classes. Apart from these there is the invariable negative *ba(y) e* (8), also described by Mkude (1974: 100). The invariable negative cannot stand alone but requires morphological negation. Both *ba(y)e* and morphological negation can be combined with the marker *tsa*, although this construction is not very common and occurs only rarely in my data (see (7) and (9)).

(7) *M-ji-ni*   ***tsa-si*-gend-ire.**
     3-town-LOC   **TSA-1SG.SM.NEG**-go-PFV

     'I did not go to town (then/at that time).'

(8) ***Ha*-wa-m-many-ile**   *baye*   *i-pinga.*
     NEG-SM2-OM1-know-PFV   NEG   AUP-woman

     'They do not know the woman.'

(9) ***Tsa-ng'a*-wa-pfikits-e**   *bae.*
     TSA-NEG-SM2-think-FV   NEG

     'They did not think (but they were expected to do so).'

3. THE MARKER *TSA*

This section will explore the marker *tsa*, including its meaning, usage, distribution, non-usage, and finally its form and possible grammaticalization. The marker is used in the past and it conveys various meanings similar to 'at a specific time', a locative reference 'at that place', or even 'for a reason'. Its main function is to evoke rapport, signalling some form of shared knowledge or reference. It can also be used as a marker of expectation; see (9). It is challenging to produce a context where *tsa* would be obligatory, nor is there any context where it is disallowed, only dispreferred.

Usually, *tsa* refers to an already established point of reference which is more often than not temporal, but may also be locational. Compare *nigenda* 'I went' (10) with *tsanigenda* '(at that time) I went' (11).

(10) *Ni-gend-a*   *m-ji-ni.*
      1SG.SM-go-FV 3-town-LOC

      'I went to town.'

(11) ***Tsa*-ni-gend-a**   *m-ji-ni.*
      TSA-1SG.SM-go-FV 3-town-LOC

      'I went to town (then).'

The same sentence, *tsanigenda*, may also mean 'because I went'. It is not possible to say *tsanigenda* out of context − it needs to be anchored somehow. Below are examples where the first sentence (12) contains a specific temporal reference and the second (13) depends on prior knowledge or expectation (similar to presupposition). You expect the person to have seen the glasses in (13).

(12) ***Tsa**-ni-gend-a   m-ji-ni   nemitondo.*
    TSA-1SG.SM-go-FV 3-town-LOC this_morning

    'I went to town this morning.'

(13) ***Tsa**-kw-ion-a   mi-wani?*
    TSA-2SG.SM-OM4-see-FV 4-glasses

    'Did you see the glasses?' or '(Was there any time when) you saw the glasses?'

As mentioned, the marker *tsa* is rarely infelicitous, although it may be unidiomatic. The following sentence would not be spoken at the beginning of a conversation or just after greetings if two people meet.

(Context: Husband comes home and says to wife: "Hello, how are you? What did you do today?")

(14) #***Tsa**-ni-gend-a   m-ji-ni   na ni-tingan-a   na   wa-nu.*
    TSA-1SG.SM-go-FV 3-town-LOC and 1SG.SM-meet-FV with 2-person

    'I went to town and I met people.'

On the other hand, in a setting where there is some sort of expectation, the usage of *tsa* in a sentence like the one above is appropriate. One informant offered the following context where the sentence above (14) would be felicitous: "If the husband comes home and expects dinner to be ready but finds that it is not, he may ask the wife what she did today (implicit: instead of making dinner) and then her answer would be: *Tsang'enda mjini na nitingana na wanu.* 'I went to town and I met people (implicit: which kept me busy so I did not have time to cook)'."

The marker bears some similarity to the Swedish multifunctional particle *ju* 'as you know', which may express "what the speaker thinks is shared knowledge" (Aijmer 1997: 221). Like *tsa*, the particle *ju* is often not translated into English (or Swahili) – Aijmer shows that *ju* is only translated into English 20 percent of the time (Aijmer 1997: 417). But while *ju* can function in other ways, such as being contrastive and/or emphatic, that is not the case for *tsa*.

The marker *tsa* is also realized with voicing (*dza*, as seen in (15), (16) and (18)), depending on the dialect of the informant.[5] It is usually written conjointly with the verb by the informants, although on rare occasions it may be written separately (15). There is no other disjoint writing of verbal markers in Luguru. This means that it is presumably undergoing a process of grammaticalization; see Section 3.4. The marker is placed at the beginning of the verb phrase but may be preceded by adverbials (16).

(15) ***Dza** ka-u-pfik-a   u-sage   ako*
    TSA SM1-OM14-find-FV 14-flour DEM

    'S/he found it (the flour) there.'

---

5 Luguru is divided into dialects that differ mainly in phonology and lexicon (Moses 2018).

(16)　*Jana*　　　***dza**-ka-som-a*　　*i-chi-tabu.*
　　　　yesterday　**TSA**-SM1-read-FV　AUP-7-book

　　　'Yesterday s/he read the book (that I had told her/him to read).'

Since the Bantu language family constitutes a fairly coherent group of language varieties (Philippson & Grollemund 2019: 341), many grammatical markers can be traced back to Proto-Bantu, which is not the case for *tsa*. Moreover, there is, to my knowledge, no analogous marker in the neighbouring Bantu languages. The closest in form that I have found is the Shangaci (P321) purpose marker, which is derived from 'come' (Devos 2014: 307). In terms of function, there is a substitutive marker of class 14 in Nyakyusa (M31), *bo*, which can have "an aspectual function of establishing or reintroducing a temporal anchor" (Persohn 2017: 45), analogous to the function of *tsa*.

3.1 Aspect and mood (or lack thereof)

Tsa can occur in most aspects and moods, although usually only in the past.[6] Also, in elicited sentences, *tsa* only occurs where the consultants were given the simple past marker *-li-* in Swahili. Likewise, when the consultants were asked to give a back translation into Swahili from Luguru, they offered only the simple past marker *-li-*. The Swahili perfect marker *-me-* was never used. When asked to give a back translation into Swahili, the second most common way was to use the Swahili construction in example (17), which is usually translated as a Simple Past, as seen in *alikuwa amelala* 's/he was asleep' (Ashton 1944: 249), or as a Past Perfect (Nurse & Hinnebusch 1993: 382).

(17)　SM-*li-kuwa*　SM-*me*-VERB
　　　SM-PST-be　　SM-PFV-VERB

Moreover, even though *tsa* rarely occurs together with the habitual marker *-ag-*, it may do so in some very specific contexts, such as (18). The background here is that someone wants to know for sure (or with some sort of emphasis) that someone else used to do something in the past.

(18)　*Ku-shule*　　*dza a-kari*　*a-gend-ag-a?*
　　　17-school　　TSA SM1-PER SM1-go-HAB-FV

　　　'Did s/he used to go to school anyway/by the way?' or 'Did s/he really used to go to school at all?'

In Mkude's grammatical description of Luguru, there is a comparable marker (*zaa* (19)) restricting the event to one determinate moment in the past (Mkude 1974: 94–95). Mkude calls the marker "'recollected' reference" (Mkude 1974: 103) or "'recollected' time" (Mkude 1974: 104), which is very much in line with the analysis of *tsa* in this paper.

(19)　***Za-a***　　*ka-mu-lag-a.*　　　　　　　　　　　(Mkude 1974: 95)
　　　TSA-ADV?　SM1-OM1-hit-FV

　　　'At one time, s/he hit her/him.'

6　Mkude (1974: 105) gives only two examples of *tsa* used in non-past contexts, only one of which is corroborated by one of our informants: *tsa alime* 's/he should dig later'. There are no other occurrences of non-past *tsa* in my corpus.

Mkude's marker is not included in Seidel's (1898) brief grammar of Luguru, and occurs only a handful of times in his own grammatical description, where he tentatively assumes that it has the same origin as the future tense allomorph *-tsa-* (Mkude 1974: 104). The working hypothesis is that Mkude's form *zaa* has developed into today's *tsa*, losing its geminate vowel. However, it is hard to explain the scarcity or even absence of the marker in older sources of Luguru compared to its frequent use in today's oral and written stories, conversations, elicited sentences, and Bible translations.

3.2 Frequency and distribution

Although frequency itself does not say much about the validity of a claim or analysis, it may add to our understanding of the usage of a grammatical feature. Not until a frequency count was made did it become clear just how frequent *tsa* is in contemporary Luguru texts, and to some extent also in spoken language. This marker is very rarely found in old data, but now occurs in some 43 percent of the past tense sentences (298 occurrences) in my database. It should be mentioned, however, that many sentences contain several demonstratives such as 's/he found it/them there' (20), which may partly explain the high frequency of a marker signalling "shared knowledge or reference", at least compared to other sentences in isolation.

(20) ***Tsa**-ka-tu-pfik-a*   *ako.*
    TSA-SM1-1PL.OM-find-FV   DEM

    'S/he found us there.'

Likewise, in the newly published Luguru Bible, *tsa* occurs frequently. Apart from the online version, available at <worldbibles.org/language_detail/eng/ruf/Luguru>, I also received a pre-publication draft dated 2013 from the Pioneer Bible Translators. In the entire Book of Luke (22,000 words), there are 102 occurrences in the draft and 149 in the published version, which means that 46 percent of the markers were added by the translators in the reviewing process.[7] Since this is unannotated text, it is not feasible to know the total number of verbs and thus get a percentage for the frequency of verbs occurring with *tsa* compared to verbs occurring without the marker, but the high percentage of added markers is striking. The motivation behind the adding of *tsa* is an issue for further research. When asked why they use *tsa*, one informant, who is not a Bible translator, says that it is "light" without it, and another that it is more proper to use it and that without it "you're cutting corners". Example (21) below is from the Bible draft dated June 2013; *tsa* is added in example (22), which is from the online published version, dated 2017.

(21) *I-wa-nu*   *wa-ingi wa-ndus-a*   *u-kw-andik-a* […]   (Luke 1:1, draft 2013)
    AUP-2-person   2-many SM2-begin-FV   AUP -15-write-FV

    'Many have undertaken to write […]'

(22) *I-wa-nu*   *wa-ingi* ***tsa**-wa-ndus-a*   *u-kw-andik-a* […]   (Luke 1:1, 2017)
    AUP-2-person   2-many **TSA**-SM2-begin-FV AUP-15-write-FV

    'Many have undertaken to write […]'

---

7 Searching for the string 'tsa' over-generated hits, so to narrow it down, I searched for all verbs containing *tsa* and the subject markers for classes 1 and 2 (i.e. third person singular and plural) *tsaka* and *tsawa*.

This adding of a morpheme in different drafts of the same Luguru Bible excerpt is similar to a process observed for a nominal morpheme in Luguru – the augment prefix.[8] The augment prefix is a nominal marker related to specificity; it can be determinative, and it is used as a marker for reference (Petzell & Kühl 2017: 42). Analogous to *tsa*, the Bible translators *added* the augment prefix during their editing process in order to make the text more "authentically Luguru" (Petzell & Kühl 2017: 44). Petzell & Kühl analyse this overuse of the augment prefix as *stability despite contact* due to *covert prestige*. It appears that there is an analogous process of adding *tsa* – not just by the Bible translators but other informants as well. When asked to repeat a sentence the same speaker may add *tsa* the second time. I tentatively draw a similar conclusion for the usage of *tsa* as for the augment prefix, although in this case it is more likely *divergence despite contact*, again due to covert prestige, seeing that this is a novel feature that does not exist in neighbouring languages nor in Swahili. The augment prefix, on the other hand, occurs in neighbouring languages and can be reconstructed for Proto-Bantu, which is why it is analysed as *stability* and not *divergence*. Furthermore, it should also be noted that such behaviour in general, that is, speakers editing their texts, is not uncommon during data collection (Marten & Petzell 2016).

In the above mentioned booklet of short stories (Lukanza et al. 2001), *tsa* is rather infrequent. It usually occurs in the first sentence; it may then be reiterated in the second sentence, but rarely after that. Correspondingly, in these stories, the marker seems to establish rapport by functioning as an anchor setting the time and the place. The material is very small though – four stories of about 900 words in sum and a total of only seven occurrences of *tsa*. It is clear both in my collected data (oral and written) and in the Bible that *tsa* is used more frequently today. The reasons for its low frequency in the traditional stories and in sources such as Mkude's grammar (1974), and its total omission in Seidel's grammatical sketch (1898), is peculiar. Either we assume that *tsa* was present in the Luguru language but ignored by linguists, or that rapid language change has taken place, which is what I propose. This is in itself not unusual – smaller language communities go through faster language change, and the evolution of linguistically marked constructions (such as *tsa*) are more likely to take place in smaller communities (Nettle 1999: 119). Finally, a similarly speedy change in Luguru has been documented for the nominal augment prefix (Petzell & Kühl 2017: 40).

3.3 Non-usage

Given that *tsa* may be used in many different settings, for instance in both punctual and extended temporal references, and that there is no setting where it is obligatory, the focus will now be on contexts where *tsa* is *less* felicitous. *Tsa* would be dispreferred with its intended meaning in (23) and (26), since it does not combine with adverbials such as 'recently' (23) or 'long ago' (26), which do not denote a specific time. Nevertheless, if *tsa* is used with *sambi* 'now/ recently', as in (24), the meaning becomes determinate, as in 'just now' or 'at this moment'. The data was verified with back translation into English (Swahili does not make the distinction between 'now/recently') to avoid tautology.

(23) *Ka-uk-a   sambi.*
SM1-leave-FV   now

'S/he left now/recently.'

---

8 Also referred to as "pre-prefix", "initial vowel", or simply "augment".

(24) ***Tsa*-*ka-uk-a***   *sambi.*
      TSA-SM1-leave-FV   now

      'S/he left (just) now.'

(25) ***Tsa*-*tu-long-a***   *naye*   *mw-aka*   *gu-bit-ile.*
      TSA-1PL.SM-speak-FV   with_her/him   3-year   SM3-pass-FV

      'We spoke with her/him last year.'

(26) *Tu-long-a*   *naye*   *mwande.*
      1PL.SM-speak-FV   with_her/him   long_ago

      'We spoke with her/him a long time ago.'

(27) #***Tsa*-*tu-long-a naye mwande.***

      Intended: 'We spoke with her/him a long time ago.'

(28) #***Zsa* *ka-mu-lag-a mwande.***   (Mkude 1974: 95)

      Intended: 'At one time, s/he hit her/him long ago.'

The reason that the marker is not used (26) or is infelicitous (27)–(28) in combination with 'long ago' is that the adverbial is indeterminate and *tsa* requires a specific (or in Mkude's terminology "recollected") moment. Notice that *tsa* can occur with other temporal adverbials if they have a specific temporal reference such as 'in the year which has passed' (25). This pattern emerged clearly during elicitation, where *tsa* was often used together with several adverbials denoting specific times, such as 'this morning', 'last week', 'last month', and 'last year' (25), while it was not used together with the unspecific 'a long time ago' (26).

Interestingly, when the unspecific *mwande* 'long ago' is used with a demonstrative and thus "specified" as a determinate time (although long ago), acquiring the narrative meaning 'once upon a time', *tsa* is commonly used, as in (29).

(29) *Aho mwande,* ***tsa*-*ku-kal-a***   *na*   *mu-nu.*
      DEM long_ago **TSA-**SM17-remain-FV   with   1-person

      'Once upon a time there was a person'.

3.4 Grammaticalization

The etymology of the marker *tsa* is probably a grammaticalization of the verb *kutsa* 'to come',[9] from the Proto-Bantu reconstruction **jìja* 'come'. The current form of *kutsa* is consistent with other phonological changes in Luguru. The verb *kutsa* 'to come' is still used as a main verb in Luguru; see (30).

(30) *Kuya* *tsa-ka-**ts**-a*   *mu-nu*   *yu-ngi.*
      DEM   TSA-SM1-**come**-FV   1-person   1-another

      'Then came another person.'

---

9 The meaning could speculatively stem from 'the time/thing that came to happen' or similar.

What is more, the future marker *-tso-* in slot 4[10] is presumably also a grammaticalization of 'come'. I make use of Hopper & Traugott's (2003: 18) definition of grammaticalization: "the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and once grammaticalized, continue to develop new grammatical functions". *Tsa* is an illustration of the latter part of the definition, referring to the continued development of new grammatical functions. Not only is the function of *tsa* new in Luguru (it is not attested in older sources), but it also appears to be highly unusual in other Bantu languages.

Overall, *tsa* follows the pattern usually described for grammaticalization, that is a content item becoming a grammatical word (see Hopper & Traugott 2003: 7). The marker still displays the full phonological shape of the original verb stem, although it cannot take any affixes itself. Bernander (2017) lists parameters of grammaticalization in a schema (Table 3). According to this schema, the *tsa* marker has undergone the conceptual change and begun the process of decategorialization, that is the loss of the source verb's morphosyntactic characteristics. The reason for claiming that it has not yet finished the process is that the marker does not exhibit the property of phonological dependence, as clitics do, and that it has not yet moved closer to the verb root. Such affixation is one of the formal processes identified for grammaticalization (Bybee, Perkins & Pagliuca 1994: 106).

Table 3 The parameters of grammaticalization (Bernander 2017: 120)

	Parameters	Linguistic domain	Type of change	Direction
1.	Extension	Pragmatics	Conceptual change	
2.	Desemanticalization	Semantics		
3.	Decategorialization	Morpho-syntax	Formal change	
4.	Erosion	Phonetics		

As mentioned, the original verb *kutsa* 'to come' still functions as a regular verb as well, and can occur in double verb constructions, as seen in (31). Nevertheless, here it does not appear to be grammaticalized, since no desemanticalization has taken place. The main verb in the constructions retains the lexical meaning 'to come'. It is followed by the infinitive, which is in alignment with one grammaticalization path for Bantu: "the incorporation of original strings of inflected auxiliary plus infinitive ([tense + AUX] + infinitive)" (Nurse 2008a: 291).

(31)   *tsa-we-**ts**-a*   *ku-mw-ing-its-a*   *li-kumbi*   (Luke 1:59)

      TSA-SM2-**come**-FV   15-OM1-enter-CAUS-FV   5-circumcision

      'they came to circumcise (the child)'

It is not uncommon for a future marker to derive from 'come' in Bantu (Nurse 2008a: 298), whereas 'come' grammaticalizing into a marker only used with the past appears to be very rare. Nurse (2008a: 307) points out a few examples but they are realized in the negative, with only one exception (Giryama E72). In other language families, however, 'come' has occasionally developed into an anterior (Bybee, Perkins & Pagliuca 1994: 56). The Luguru *tsa* seems to be a grammaticalization of 'come' used in the past. That said, the marker does not necessarily carry an inherent past meaning. Moreover, Persohn (2018: 106) describes a metaphorical extension of 'come' markers where there is a shift "from denoting translational motion towards the deictic centre to denoting that the subject reaches, achieves or is led to a specific condition". This transfer from non-deictic motion could have taken place in the grammaticalization of *tsa* too.

---

10 "the most productive cradle of new verb morphology in Bantu" (Güldemann 2003: 185).

As for the syntagmatic position of *tsa* in the verb, it is regular in the sense that recently grammaticalized items start in the periphery and work their way to the core. The first slot is where new material often becomes grammaticalized (Nurse 2003: 91).

4. CONCLUDING DISCUSSION

This paper has briefly sketched the structure of the verb in Luguru to set the background for a discussion of the marker *tsa*. The usage, form, and frequency of the marker have been outlined, together with its function as a marker of shared knowledge, reference, or even expectation. The marker is never obligatory, it is not found in older Luguru sources, and there is no equivalent in other Bantu languages. The settings where the marker is unidiomatic have been described, along with its lack of grammatical aspect. My data demonstrate that the marker *tsa* is an indicator of what the language user thinks is shared knowledge in a broad sense. It refers to something like a "definite span" of time or space, or to more abstract notions such as reasons and expectations. *Tsa* is a marker that cannot be used in out-of-the-blue contexts without some sort of "anchor". This is corroborated by the infelicitous sentences combining *tsa* with an "indefinite" adverbial. That *tsa* is dispreferred in combination with unspecific or vague references to indefinite time is in alignment with Besha's description of the tenses in another Tanzanian Bantu language, Shambala (G23). Besha (1989: 188–190) divides the TAM markers into dependent and independent formatives, whereby the dependent ones can only co-occur with a "definite" determiner such as 'this year', while the independent ones can only co-occur with "indefinite" ones such as 'a long time ago'. *Tsa* in Luguru, like the independent formatives in Shambala, is infelicitous with the indefinite adverbial *mwande* 'long ago' (see example (26)); however, the more "definite" use of the same adverbial, that is, combined with a demonstrative, is felicitous in (32) (repeated from (29)).

(32) *Aho   mwande, **tsa**-ku-kal-a   na   mu-nu.*
      DEM   long_ago **TSA**-SM17-remain-FV   with   1-person

   'Once upon a time there was a person'.

In languages with so few tense, aspect, and mood markers, do markers such as *tsa* surface because of that scarcity? Nurse (2008b: 170) makes a general claim for the languages in the Morogoro region, stating that they compensate for the simplicity of tense reference by making use of auxiliaries such as 'come' and 'go'. Since I assume that *tsa* has grammaticalized from 'come', it is possible that this innovation happened because there is so little TAM morphology. That said, the neighbouring language, Kami (G36), has even fewer markers, but no parallel innovation (Petzell & Aunio 2019). Even so, I pose the question of whether innovations like this one are more likely to take place in languages that have little tense, aspect, and mood morphology.

ABBREVIATIONS

Numbers refer to person (first 1, second 2, third 3) or to noun (class) prefixes.

ADV	adverb		OM	object marker
APPL	applicative		PER	persistive
AUP	augment prefix		PFV	perfective
CAUS	causative		PL	plural
DEM	demonstrative		PN	personal name
FUT	future		PST	past
FV	final vowel		SG	singular
HAB	habitual		SM	subject marker
LOC	locative		TAM	tense, aspect, mood
NEG	negative		TSA	the marker *tsa*

REFERENCES

ASHTON, Edith 1944. *Swahili Grammar* (*including intonation*). London: Longmans.

AIJMER, Karin 1997. Epistemic Modality as a Discourse Phenomenon: A Swedish–English Cross-Language Perspective. In: U. FRIES, V. MÜLLER & P. SCHNEIDER (eds), *From Ælfric to the New York Times: Studies in English Corpus Linguistics*: 215–227. Amsterdam: Rodopi.

BERNANDER, RASMUS 2017. *Grammar and Grammaticalization in Manda: An Analysis of the Wider TAM Domain in a Tanzanian Bantu Language*. PhD dissertation, University of Gothenburg.

BESHA, Ruth Mfumbwa 1989. *A Study of Tense and Aspect in Kishambala*. Berlin: Dietrich Reimer.

BLOOM STRÖM, Eva-Marie & Malin PETZELL *in press*. Micro-Variation Approaches to Bantu Language Varieties. In: L. MARTEN, E. HURST, N. KULA & J. ZELLER (eds), *The Oxford Guide to the Bantu Languages*. Oxford: OUP.

BYBEE, Joan L., Revere D. PERKINS & William PAGLIUCA 1994. *Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: UCP.

DEVOS, Maud 2014. Motion Verbs in Shangaci: Lexical Semantics and Discourse Functions. In: M. DEVOS & J. VAN DER WAL (eds), *COME and GO off the Beaten Grammaticalization Path*: 281–317. Berlin: De Gruyter Mouton.

GÜLDEMANN, Tom 2003. Grammaticalization. In: D. NURSE & G. PHILIPPSON (eds): 182–194.

GUTHRIE, Malcolm 1948. *The Classification of the Bantu Languages*. London: OUP for the International African Institute.

GUTHRIE, Malcolm 1967–1971. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages*. 4 vols. London: Gregg International.

HOPPER, Paul J. & Elizabeth Closs TRAUGOTT 2003. *Grammaticalization*. 2nd edn. Cambridge: CUP.

JOHNSTON, Sir Harry Hamilton 1922. *A Comparative Study of the Bantu and Semi-Bantu Languages*, II. Oxford: Clarendon Press.

JORDAN, Caspar & Malin PETZELL *in press*. Evaluation of Two Methods of Data Processing in Linguistics. In: A. BEREZ-KROEKER, B. MCDONNELL, E. KOLLER & L. COLLISTER (eds), *Open Handbook of Linguistic Data Management*. Cambrigdge, MA: MIT Press Open.

KIBRIK, Aleksandr Evgenevic 1977. *The Methodology of Field Investigations in Linguistics:* (*Setting up the Problem*). The Hague: Mouton.

Languages of Tanzania Project 2009. *Atlasi ya Lugha za Tanzania*. Dar es Salaam: Mradi wa Lugha za Tanzania, Chuo Kikuu cha Dar es Salaam.

LUKANZA, Ramadhani, Christuan SENGO, Nitu Mkkulu KUMWAO & Rashidi KIBEKU 2001. *Hadithi za Kilughuru* [Luguru stories]. Dallas & Kleinmond: Pioneer Bible Translators & Word for the World.

MAHO, Jouni 2007. The Linear Ordering of TAM/NEG Markers in the Bantu languages. *SOAS Working Papers in Linguistics* 15: 213–225.

MARTEN, Lutz 2003. Dynamic and Pragmatic Partial Agreement in Luguru. *Typologie des langues d'Afrique et universaux de la grammaire* 1: 113–139.

MARTEN, Lutz & Malin PETZELL 2016. Linguistic Variation and the Dynamics of Language Documentation: Editing in "pure" Kagulu. *Language Documentation and Conservation* 10: 105–129.

MARTEN, Lutz & Deograsia RAMADHANI 2001. An Overview of Object Marking in Kiluguru. *SOAS Working Papers in Linguistics and Phonetics* 11: 259–275.

MEEUSSEN, Achille E. 1967. Bantu Grammatical Reconstructions. *Africana Linguistica* 3: 79–121.

MKUDE, Daniel 1974. *A Study of Kiluguru Syntax with Special Reference to the Transformational History of Sentences with Permuted Subject and Object.* PhD dissertation, University of London.

MOSEL, Ulrike 2011. Morphosyntactic Analysis in the Field: A Guide to the Guides. In: N. THIEBERGER (ed.), *The Oxford Handbook of Linguistic Fieldwork*. Oxford: OUP.

MOSES, Godian 2018. *Are Highland Luguru and Lowland Luguru Distinct Dialects?* Unpublished MA thesis, University of Dar es Salaam.

NETTLE, David 1999. Is the Rate of Linguistic Change Constant? *Lingua* 108(2–3): 119–136.

NURSE, Derek 2003. Aspect and Tense in Bantu Languages. In: D. NURSE & G. PHILIPPSON (eds): 90–102.

NURSE, Derek 2008a. *Tense and Aspect in Bantu*. Oxford: OUP.

NURSE, Derek 2008b. Tense and Aspect in Bantu: Appendices. Available at <faculty.mun.ca/dnurse/Tabantu/>, accessed May 2018.

NURSE, Derek & Maud DEVOS 2019. Aspect, Tense and Mood. In: M. VAN DE VELDE, K. BOSTOEN, D. NURSE & G. PHILIPPSON (eds): 173–203.

NURSE, Derek & Thomas J. HINNEBUSCH 1993. *Swahili and Sabaki: A Linguistic History*. Ed. Thomas J. Hinnebusch, with a special addendum by Gérard Philippson. Berkeley: UC Press.

NURSE, Derek & Gérard PHILIPPSON (eds) 2003. *The Bantu Languages*. London: Routledge.

PERSOHN, Bastian 2017. *The Verb in Nyakyusa: A Focus on Tense, Aspect, and Modality*. Berlin: Language Science Press.

PERSOHN, Bastian 2018. Basic Motion Verbs in Nyakyusa: Lexical Semantics and Associated Motion. *Studies in African Linguistics* 47: 101–127.

PETZELL, Malin 2008. *The Kagulu Language of Tanzania: Grammar, Texts and Vocabulary*. Cologne: Rüdiger Köppe.

PETZELL, Malin 2012a. The Linguistic Situation in Tanzania. *Moderna språk* 106(1): 136–144.

PETZELL, Malin 2012b. The Under-Described Languages of Morogoro: A Sociolinguistic Survey. *South African Journal of African Languages* 32(1): 17–26.

PETZELL, Malin & Lotta AUNIO 2019. Kami G36. In: K. BOESTON, M. VAN DE VELDE, D. NURSE & G. PHILIPPSON (eds): 563–590. London: Routledge.

PETZELL, Malin & Harald HAMMARSTRÖM 2013. Grammatical and Lexical Comparison of the Greater Ruvu Bantu Languages. *Nordic Journal of African Studies* 22(3): 129–157.

PETZELL, Malin & Karoline KHÜL 2017. The Influence of Non-Linguistic Factors on the Usage of the Pre-Prefix in Luguru. *Linguistic Discovery* 15(1): 35–49.

PHILIPPSON, Gerard & Rebecca GROLLEMUND 2019. Classifying Bantu Languages. In: M. VAN DE VELDE, K. BOSTOEN, D. NURSE & G. PHILIPPSON (eds): 335–354.

ROSE, Sarah, Christa BEAUDOIN-LIETZ & Derek NURSE 2002. *A Glossary of Terms for Bantu Verbal Categories, with Special Emphasis on Tense and Aspect*. Munich: Lincom Europa.

SEIDEL, August 1898. Grundriss der Wa-Ruguru-Sprache. In: C.W. WERTHER (ed.), *Die mittleren Hochländer des nordlischen Deutsch-Ostafrika*: 436–455. Berlin: Verlag von Hermann Paetel.

VAN DE VELDE, Mark, Koen BOSTOEN, Derek NURSE & Gérard PHILIPPSON (eds) 2019. *The Bantu Languages*. 2nd edn. London: Routledge.