Grammaticalization of tense and aspect in Mbugwe: a preliminary investigation

VERA WILHELMSEN

_____

_____

EL Publishing

Grammaticalization of tense and aspect in Mbugwe: a preliminary investigation

Vera Wilhelmsen

1. Introduction[1]

Mbugwe is a relatively small Bantu language spoken in Northern Tanzania, between the cities of Arusha and Babati. The formal classification is Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Narrow Bantu, Central, F, Nyilamba-Langi (F.30) (Lewis 2009). It is of particular interest to linguists because it is isolated from the other Bantu languages, surrounded by languages from other families, such as the Cushitic languages Iraqw and Gorowa to the west and south respectively, and the Nilotic Maasai language to the east. The Khoisan/ Khoe[2] language Sandawe is also spoken in the area, as well as the language isolate Hadza. Derek Nurse (1999: 11) states: 'the West Tanzania area is the only place in Africa where representatives of the continent's four language families are still spoken'. The aim of this article is to study a small part of the tense and aspect system of Mbugwe, specifically the forms with past time reference, and to present a hypothesis of how they are being grammaticalized. Hopefully the paper will give some new insight as to how this grammaticalization process takes place in Bantu languages.

Mbugwe is classified as F.34 under the Guthrie referential classification system for Bantu languages (1948). The nearest Bantu relative is Rangi, F.33, with which it has 72% lexical similarity, according to Masele & Nurse (2003: 121). However, Masele & Nurse (2003: 122) conclude in their chapter about the Bantu zone F that 'F33 and F34 are close to each other, but not particularly close to the whole group'. As is the case for many other minority languages in the country, Mbugwe is under heavy pressure from Swahili (classified by Guthrie as G42), which is the national language and the only

[2] The relatedness of the Khoe family to the Khoisan language family is uncertain (see Güldemann & Elderkin 2003).

language of instruction in Tanzanian schools. Significant influence from Swahili has already been observed in the language, especially in the vocabulary of younger speakers (Mous 2004: 1).

Mbugwe may be considered as endangered according to UNESCO's list of factors conditioning endangerment, as it has no official status, no written language, and has not been adequately described in academic literature (UNESCO 2003). The Mbugwe are also a minority of Tanzania, and the language is not used in any official domain. The precise number of Mbugwe speakers is not known: the *Ethnologue* (Lewis 2009) suggests 24,000, but the number appears to be lower than this, according to survey work done by SIL personnel in the area (Julia Larsen, p.c.). What is clear, however, is that Mbugwe speakers are a small minority in the country, which has a total population of about 42 million. Without a description and a standardized written form of the language, language shift towards Swahili is probable.

Previous linguistic work on Mbugwe is limited, and mostly dated. Baumann (1894) has a list of numerals, Seidel (1898) lists a few words, Struck (1909) a few notes, and Dempwolff (1915–1916) has written a short grammatical description of the language. Newer works include some words in Kesby (1986) and a word list in Masele (2001). The best and most recent grammatical sketch was written by Maarten Mous (2004). He emphasizes that the work is tentative, based on a very short period of fieldwork with just a couple of informants, and that the sketch is incomplete in many respects. Nevertheless, it gives a good overview of the basic structure of the language, and I have mainly based my analysis on his data.

Two linguists from SIL International, Viggo and Julia Larsen, are working with the Mbugwe people and language. They have started to do language analysis and orthography development, and have conducted several orthography workshops in the community. They have generously shared their FLEx (FieldWorks Language Explorer) database, containing about 1,300 lexical items and several stories, which I have used as a reference during the present analysis.

This paper is structured as follows: first I present a general introduction to tense and aspect in Bantu, in order to provide some background to the investigation of the Mbugwe system. I will then explore Mous' analysis of the verbal system of Mbugwe, focusing on forms with past time reference, and comment on his labels. Then I will offer an alternative analysis of Mous' data, based on an analysis by Nurse (2008), and drawing from literature on related languages and some of the Larsens' findings. Finally, I will offer a new proposal, which is a synthesis of the two different analyses. I propose that Mbugwe is in the process of reanalyzing the perfect forms as referring to different degrees of past, following a grammaticalization path which takes the verbal system from an aspectual system to a fully developed metric tense system.

2. Tense and aspect in Bantu languages

According to Comrie (1985: 9), tense is a 'grammaticalized expression of location in time', whereas aspects are grammaticalized expressions of the 'different ways of viewing the internal temporal constituency of a situation' (Comrie 1976: 3). The literature on tense and aspect is vast, as they are important but elusive grammatical categories, playing an important role in the grammar of most languages of the world. I follow Comrie in definitions and the differentiation between them, but as we shall see, I find that the categories are not always easy to keep apart.

Nurse (2008: 13) holds that every tense-aspect form has a specific meaning, which is different from other forms. They make up a coherent system, he claims, and therefore it is not difficult to separate tense from aspect, although the system is dynamic and flexible. As the investigation will show, it will not always be easy to decide if a morpheme represents tense or aspect, but like Nurse, I do assume that the morphemes make up a coherent, interlocking system.

Bantu languages are well known for having a rich tense and aspect system (see e.g. Dahl 1985: 85, 175). They have many affixes which may attach to the verb stem, as well as periphrastic constructions which encode different tense and aspect categories. In addition, other grammatical categories are also expressed on the verb, such as subject, object, mood, negation, relativity, condition, focus and also verbal derivations (often valency-changing, such as passive and applicative). This makes the verbs rather complicated and flexible, and Nurse identifies 11 traditional slots on the verb stem (Nurse 2008: 31ff.). He argues for a slight modification of the traditional system with a collapse of a couple of the slots, and proposes the following template for Bantu verbs:

Pre-SM + SM + NEG2 + TA + OM + root + extension + FV + post-FV[3]

Tense and aspect are normally encoded in the TA slot, but also in the FV slot. Some languages allow several morphemes in the TA slot, but most do not (Nurse 2008: 34). Nurse (2008: 14) also notes that 'most Bantu languages encode tense on the left and aspect to the right', whether or not they both appear before the stem. This is, however, not universal.

When it comes to the encoding of tenses in Bantu, Nurse (2008: 80ff) looks at typical or widespread markers for the past, present and future tenses.

---

[3] FV= final vowel, NEG2= negation 2 (negation 1 is marked in the Pre-SM slot), OM= object marker, SM = subject marker, TA= tense/aspect marker, The 'final vowel' is sometimes longer, e.g. *-ile.* In indicative, the unmarked final vowel usually is *-a.*

In his sample of 100 'matrix' languages, 78% had a morpheme *a* with a past time reference, most commonly in the TA slot of the verb. In 59% of the languages the final vowel (FV) was also *-a* in the past tense, but for 43% the suffix was *-ile*. A large percentage also had *-ile* as a suffix but zero marking at the TA slot. *-ile* is however also associated with the anterior aspect, as well as near past. This is an important point for my investigation, which we will look at in more detail later in the paper. Some languages have a null-past, which is quite interesting typologically, but is not relevant for this discussion.

In addition to their rich verbal morphology, Bantu languages are well known for the remoteness degrees of their tense system (see for instance Dahl 1985: 120–21). This is also found in other places of the world, such as in New Guinea, South America, and a few instances in North America and Australia (Dahl & Velupillai 2011b). This phenomenon is termed TEMPORAL DISTANCE by Dahl (1985: 120). It divides past (and sometimes future) time into two or more time frames, ranked from remote to close to the time of speaking (the present). Nurse reports that some Bantu languages have up to five degrees of past and future, although the average number is two or three (Nurse 2008: 89). 80% of the languages in Nurse's sample had more than one division in past. For future, the number of languages was much lower (41% had two or three futures) (Nurse 2008: 22). Dahl (1985: 121) notes that this seems to be a universal tendency.

In general, there is often a cut-off point between 'today' and 'before today', called 'hodiernal' and 'hesternal' (Dahl 1985: 125). The hesternal category is often subdivided into several periods (for multiple remoteness distinctions). The cut-off point for the hodiernal tense tends to be the evening the day before, even for Bantu languages (Östen Dahl, p.c.), but the exact time is often not given in reference grammars and is therefore uncertain. The cut-off points for other remoteness categories are even harder to determine. Also, the system can be flexible or rigid (subjective or objective) (Dahl & Velupillai 2011b, Nurse 2008: 22). Nurse does however note that for languages with three pasts, the most common categorization is hodiernal, hesternal and earlier than hesternal (2008: 91). Some languages also have an immediate past, which refers to something that just happened (Nurse 2008: 91).

When it comes to languages with several futures, they seem to be less rigid than the past systems, and some authors would say that the remote future is simply less certain than the near future, indicating that it could be a modal category rather than a temporal category (Nurse 2008: 92 discusses this). The morphemes which mark future time are less uniform than the past tense morphemes, although *ka-* appears in 29% of the sample languages (Nurse 2008: 85ff). Another form is *la(a-)* or *ra(a)-,* which are probably allomorphs or reflexes of the same morpheme. This occurs in 12% of the languages in

Nurse's sample. Another source for the future tense is the verb 'come', which occurs all over the Bantu area, in 15% of the languages.

Another form that is typical of Bantu languages in particular is the narrative tense (Dahl 1985: 113–14; Nurse 2008: 120). It is characterized by the fact that a string of situations, which happen in sequence, is marked by the same form, indicating that they are part of the same discourse. Often, the first verb in the string gives the tense reference, and the consecutive ones share the same tense. Sometimes, however, the tense is not given, though it is clear from the context or otherwise clear to the hearers (Nurse 2008: 120). The most common narrative marker is *-ka* or a reflex of it (30%) (Nurse 2008: 122).

The perfect or anterior aspect is very common in Bantu languages. 81% of the languages in Nurse's survey are marked for anterior. There are three major ways of encoding anterior aspect, according to Nurse (2008: 157): 29% of the languages had *Ø-...-ile,* 14% had *a-...-ile*, whereas 19% had *a-...-a*. The *Ø- ...-ile* pattern is predominantly used to encode a present anterior, and in languages in which also *a-...-ile* occurs, it indicates a further removed time. The *a-* element is therefore probably added in order to signal past time. Where the pattern *a-...-a* occurs as an anterior, it usually refers to present anterior.

In relation to this, Nurse (2008: 94ff, 155ff) discusses a problem which is relevant to the current investigation: how to distinguish the near past from the anterior aspect. One of the difficulties he mentions is the nature of anteriors; a definition of anterior aspect is 'a situation that started in the past but continues into the present' or with 'continuing present relevance of a previous situation' (Comrie 1976: 52). Recent events are more likely to be of relevance to the present than earlier events. An option could be to call this 'near past', if the focus is on the time, and not on the aspectual meaning of the verb.

In several Romance languages and some German dialects, the perfect has taken over for simple past as the main form for reference to past tense, and according to Comrie (1976: 11), there has been 'a gradual relaxation of the requirement of present relevance in the use of the Perfect to refer to a past situation'. He calls one of the types of perfect 'perfect of recent past', as it seems that the recentness itself might be enough to fulfill the requirement of present relevance.

There are several ways of testing whether a form is temporal or aspectual, if they indeed are separate, and Nurse suggests four ways of testing whether a form is anterior or near past (Nurse 2008: 95–99):

The first test is systemic. In some languages, aspect is always expressed in the post-stem (FV) position, and tense always in the TA pre-stem position, and there is no ambiguity between them. However, this does not hold for all

languages (e.g. Sukuma, Swahili), and it appears to be difficult to apply to a language before the whole system is well known and understood.

A second test is to look at compound constructions. In some languages, the first word in a compound verb construction contains the auxiliary, such as 'to be', whereas the second contains the lexical verb. In many languages, the auxiliary can have a tense marker or an aspect marker, but only a small set of aspect markers may appear on the second (lexical) verb. Therefore, the second verb can never express tense, only aspect (e.g. Swahili).

A third way of testing is to check how anteriors behave with different kinds of verbs, such as stative/inchoative vs. dynamic. Nurse (2008: 97–98) states that 'anteriors used with stative verbs represent the state resulting from the action, and the translation into English is present': Swahili *wa-me-lala* 'they are sleeping', *tu-me-shiba* 'we are satisfied', *i-me-vimba* 'it is swollen', etc. In languages with several pasts, however, it is possible to have several anteriors as well, and it is possible that 'a current state might result from situations at different past points' (Nurse 2008: 98).

A fourth test is the range of reference of the morpheme. Does it refer to only today, or also previous events which either have some relevance to the present, or continue a state entered in the past, or do they appear recent in relation to some other (distant) past? In Giryama, *dza* is both a recent past and an anterior, according to Nurse (2008: 98). Östen Dahl (p.c) suggest that this is a lot more common than Nurse seems to believe.

Nurse also observes that many languages indicate different degrees of past with a length or tone distinction involving *a,* both in the TA (pre-stem) position, and in the FV (post-stem) position (2008: 107). 22% of languages investigated by Nurse & Phillipson (2006) had contrastive *a*, but in very few cases was a difference in length or tone on the *a* the only difference between two forms. In Nurse's data of 53 languages with sufficient information on the tones, he found that 76% of the languages have *a-...-a* encoding either near past or anterior, the first *a-* being either high or low (Nurse 2008: 111). Other typical pasts are *a-...-ile,* with varying tones on both affixes. A long *aa* usually encodes other pasts, or even future, and very rarely near past or anterior. He concludes that the *a-...-a* forms seems to have originally referred to general past, but then being restricted to near past or anterior, whereas *a-...-ile,* originally encoding (past) anterior aspect, came to refer to a more remote past.

It is noteworthy that Proto-Bantu is reconstructed as being 'poor in tenses and rich in aspects' (Nurse 2008: 281). The (often multiple) tense distinctions have arisen at a later stage. There is therefore reason to assume that the temporal categories, and especially the different degrees of temporal distance, are younger than the aspectual categories. Bybee, Perkins & Pagliuca (1994:

51ff) show that it is not uncommon for anteriors to develop into past markers, for instance, and that such grammaticalization paths are quite common (see also Nurse & Philippson 2006: 181 and Dahl & Velupillai 2011a).

3. The verbal system of Mbugwe

We now turn to the verbal system of Mbugwe, and the data at hand. My starting point is Mous (2004), as I have limited data from elsewhere.

3.1. Mous' analysis

For the verbal system, Mous emphasizes the tentative nature of his analysis. He holds that it is incomplete, and probably needs to be revised. Still, his analysis gives an impression of a rich verbal system (Table 1, below).

3.2. Past time reference

Mous (2004: 11) states: 'the perfect refers to the past with present relevance, the past perfect to the past and the far past to anything earlier than the day of reference'. This statement appears a bit vague, and the difference between the past perfect and the far past is not clear. Let us take a closer look at the verb forms he refers to, and investigate their semantics.

Past time reference in Mbugwe seems to be made by the prefix *á(a)-,* which is a high short or long *a-*[4]. Mous refers to the form which includes this morpheme, and a high tone on the final vowel, as far past.

---

[4] The tones, transcriptions and glosses are Mous' in all the examples, but some of the labels are changed in accordance with the Leipzig Glossing Rules (http://www.eva.mpg.de/lingua/resources/glossing-rules.php).

Only sur¬face tone and phonetic sounds are marked, not phonemes. A phonological analysis is still in progress. An acute accent on vowels denotes a high tone; low tone is not marked. Abbreviations used in the tables and examples are: √ = root, 1 = first person, 2 = second person, 3 = third person, COND = conditional, CSEC = consecutive, FUT = future, H = high tone inserted in the root, HAB = habitual (Mous used both HAB and GEN, but I use only HAB), INF = infinitive, IRR = irrealis, NEG = negative, OPT = optative, PL = plural, PRES = present, PRF = perfect (Mous had PF), PROG = progressive, PST = past, SC = subject concord prefix, SG = singular. Numbers which are not followed by SG or PL refer to noun classes.

Table 1: Verbal conjugations (Adapted from Mous 2004:7)

form	tense marker(s)	final vowel	tense
SC-*á*(*a*)-√-*á*	PST	PST	far past
SC-√-*íye*	PRF	PRF	perfect
SC-*áa*-√-*íye*	PST	PRF	past perfect
SC-*ka*-√	CSEC	—	consecutive
SC-*ké*-H-√	COND	—	conditional
SC-*ándá*-√	HAB	—	habitual
SC-*ende*-√	PROG	—	progressive
SC-*jé*-√	FUT	—	future
SC-ka-H-√-*íye*	IRR	PRF	irrealis
SC-H-√-*e*	—	OPT	optative
NEG-SC-*ándá*-H-√	HAB	—	not yet
NEG-SC-*áa-ré*-H-√	PST-*be*	—	negative past imperfective
NEG-SC-*á*-√-*íye*	PST	PRF	negative past perfect[5]
NEG-SC-*jé~já*-H-√	NEG	—	general negative
NEG-SC-*jé*-H-√-*íye*	NEG	PRF	negative past
NEG-SC-*ká*-√-*íye*	IRR	PRF	negative irrealis
SC-*ki-sé*-√ ~	COND-NEG	—	prohibitive
SC-*káy-sé*-H-√			
INF + SC-*kénde*	PRES.PROG		present progressive[6]
INF + SC-*je*	*come*:OPT		future
INF + SC-*anda*	HAB		habitual
INF + SC-*áa-re*	PST-*be*		past-imperfective

---

[5] Mous had 'negative far past', but agreed in p.c. that it probably is negative past perfect according to his analysis.

[6] In Mbugwe periphrastic verb forms, the infinitive precedes the auxiliary. This is atypical of SVO languages, and indeed for Bantu, but is a shared feature with Rangi (F.33). See Mous (2000). Hannah Gibson at SOAS is writing her PhD dissertation on the parallel construction in Rangi.

(1)   *n- áa-rem-á*   *yonda*   *r-áne*   *wiki*   *aloká*
        1SG-PST-cultivate- PST   5.farm   5-my   week   past
        'I cultivated my farm last week.' (Mous 2004: 8)

Mous reports this form to be used referring to last week, as in example (1), but the Larsens have found that it is mostly used for events which took place longer than a month ago. The exact time frame is to be investigated further, but there is little doubt that this form refers to a remote past time.

Another form which includes *á(a)-* is what Mous calls past perfect, which also has the (typically) anterior suffix *-íye*[7]:

(2)   *n-áa-rem-íye*   *yonda*   *r-áne*   *néejo*
        1SG-PST-cultivate-PRF   5.farm   5-my   yesterday
        'I cultivated my farm yesterday.' (Mous 2004: 8)

Mous writes that it is used for past events, without any further qualifications. What is common for his data and the Larsens' findings, however, is that they indicate that it is used for events that happened yesterday, and up to a week or a month ago. Therefore it is natural to question whether it has primarily an aspectual (anterior) meaning, or if it refers mainly to a time frame, such as middle past, even if its suffix is a typical Bantu anterior marker. In the typical past perfect meaning, a past state is related to an even earlier situation (Comrie 1976: 53), but in this example there is only one event, and it is not related to any other event.

The form also seems to retain an aspectual reading, in certain cases such as in (3), where a lion is returning after a week, in order to see whether a child he has been promised is big enough for him to eat:

(3)   *o-kúndia*   *mwána*   *w-a-kúr-íye*
        INF-meet   1.child   3SG-PST-grow-PRF
        'Finding that the child had become big.' (Mous, 2004: 27)

The translation given is from Mous (2004: 33). Literally it is closer to 'the child had grown'. The form fulfills the conditions for a typical past perfect, as it refers to a past state which is related to an even earlier situation (Comrie 1976: 53). It might however also be interpreted as a middle past, as it happened within a week from the time of the event of returning. This example

---

[7] This is a reflex of the more common Bantu suffix *-ile,* discussed in the previous section.

shows that the form can function as a past perfect in some cases, but in other cases it must have a purely temporal reading. Note also that in this example the tone and length of the past marker (*a-*) are not consistent with Mous' generalization of the form. This is something that needs to be checked.

There is also a form which Mous calls perfect, which includes the typically anterior suffix, but no past tense marker:

(4)   *ndem-íye*   *yonda*   *r-áne* *énsíkʊ* *na*   *mʊtóndo*
     1SG:cultivate-PRF   5.farm 5-my 9.today with 3.morning
     'I cultivated my farm this morning' (Mous 2004: 3)

Again, the question is whether the form has primarily past time reference, or is aspectual in meaning (anterior). In this example there is no apparent present relevance of the event, other than that it happened not long ago. It therefore looks like it is near past, but it is difficult to draw a conclusion based on only this example.

In narratives, this form is sometimes used with what might very well be a perfect reading, as in example (5), but as noted above, these forms are hard to interpret, and further research is needed in order to confidently identify the form as aspectual or temporal, if that is even possible. In this case the free translation includes the word 'already', and it also needs to be investigated whether that sense is part of the meaning of the verb form. Note that the English translation uses the past perfect, and not the present perfect.

(5)   *Mwána*   *kómbá*   *o-pampá-ye*
     1.child   well   1-clever-PRF
     'The child, however, had already become smart.' (Mous 2004: 27)

3.3. Future time reference

Before I proceed into the alternative analysis in more detail, I will look briefly at the future forms in Mbugwe, in order to check whether there could be several degrees of future time reference, as there might be in the past tense. The future marker which Mous found to be *jé-*, which is also a well-known Bantu future form, probably grammaticalized from the verb for 'to come', *ja* (Mous 2004: 7).

(6)   *n-jé-mu-vekéra*
     1SG-FUT-3SG-dress
     'I will dress her/him.'
     (Mous 2004: 10)

In the context of the story in which this form is found, a woman tells the lion that she will dress her son in a white sheet that night, so that he will know which one he is. It is therefore plausible to assume that this form indicates near future. This is also supported by the data collected by the Larsens, and it is the label they have given to this form.

Future can also be expressed with an infinitive plus the optative of 'to come':

(7)   *o-ra*   *ko-je*   *mohogo*
        INF-eat   1PL-come:OPT   cassava
        'We will eat cassava.' (Mous 2004: 11)

In a story there is an example of this form used for an action expected to happen the day after:

(8)   *o-móna*   *ó-je*   *na*   *ka-mwaulá*   *k-eéru*
        INF-1:see   2SG-FUT   with   13-3.necklace   13-white
        'You will see him with the white necklace.' (Mous 2004: 28)

The data is not sufficient at this point to determine whether there is a difference in meaning between the inflectional and the periphrastic form. This might be an earlier stage of the grammaticalization process of *-ja* coming to denote future tense, but it could also be a middle future.

The Larsens have reported a distant future form, which is also periphrastic, and in addition to the verb *ja* 'to come' (not in the optative form in this case), has the prefix *ka-*. It is interesting to note that the same marker is used in the irrealis form, which fits well with the idea of something which takes place in a distant future. More research is needed in order to determine whether this form has an irrealis sense, as the prefix *ka-* is very common in Bantu verb morphology (Botne 1999: 473). It is also interesting to note that Botne (1999: 492) gives an example of an F zone language (Sumbwe, F.23) as a language which has the *-ka* prefix in a remote future form. I have refrained from giving the *-ka* morpheme a label, until more is known about the nature of this form.

(9)   *o-k-wera*   *a-ka-ja*
        INF-2SG-tell   3SG-*ka*-come
        'He will tell you.' (Larsen & Larsen 2011, my translation)

3.4. An alternative analysis

In his Appendix 2 Nurse (2008) presents an analysis of the tense and aspect system of Mbugwe, shown here in Table 2.

Table 2: Nurse's analysis of Mbugwe past reference forms

Alternative 1:

Past 3 (remote)	*-á-...-á*
Past 2(middle)	*-á-...-íe*
Past 1(near)	*-Ø-...-íe*

Alternative 2:

Past perfective	*-á-...-á*
Past anterior	*-á-...-íe*
Present anterior	*-Ø-...-íe*

It is evident that Alternative 2 is more similar to the system Mous suggested, whereas Alternative 1 is in line with the alternative analysis I hinted at above, with multiple degrees of past. Nurse seems to base his analysis on Mous' data, as he uses examples from Mous' sketch to illustrate his verb forms. Nurse does not take a stand as to which of the two analyses is correct, and based on the data presented in this paper, it seems that they both are possible for Mbugwe.

The alternative analysis then, is to call the verb form with the suffix *-íye* near past (alternatively hodiernal past), mostly referring to situations which happened on the day of speaking, and to call the verb form with both the past tense marker *á(a)-* and the suffix *-íye* middle past (alternatively hesternal past), mostly referring to situations which happened the day before speaking and up to a week or a month earlier. The fact that there seem to be at least two degrees of the future tense also supports the theory that there are several degrees of past, although it does not prove it, as one anonymous reviewer pointed out.[8]

---

[8] A reviewer pointed out that even though there is a strong typological tendency that the past tense should display at least as many, if not more, distinctive metric forms than the future, there is at least one other language which displays a metric distinction in the future but not in the past, namely the Austronesian language Seediq spoken in Taiwan.

Regarding this analysis it is relevant to notice Masele and Nurse's claim concerning a 'significant innovation in the tense-aspect pattern', which according to them is restricted to F.21, the Dakama dialect of F.22, F.24 and F.32 (Masele & Nurse 2003: 127). They explain this innovation in the following way (Masele & Nurse 2003: 128):

> The original Bantu system was likely one in which pre-stem markers (here /aa/ or /a/) indicated tense whereas suffixes marked aspect. In this new system the pre-stem marker simply represents past, while the aspectual suffixes (with some minimal help from tone) have been recycled to refer to different degrees of past.

According to the alternative analysis, this 'recycling' of markers might have happened or indeed seems to be happening right now in Mbugwe.

Interestingly, Stegen (2006) asserts that Rangi (F.33), which is the closest language to Mbugwe, has also undergone such a semantic shift, and some of the forms are still being redefined (the distant past is used by some speakers as a past habitual). Comparing the Rangi system with the core Bantu F group, Stegen (2006) presents the analysis given here in Table 3.

Table 3: Comparison of past tenses between Rangi and other Bantu F languages (from Stegen 2006)

	Core Bantu F (Masele & Nurse 2003)	*Rangi*	*with semantic shift*
Far past	S-*àa*-V-*a*	S-*á*-V-*á*(*g*)*a*	past progressive
Hesternal past	S-*á*-V-*íle*	S-*a*-H-V -*á*	recent past
Hodiernal past	S-*á*-V-*aga*	S-*a*-H-V-*ire*	past hab./dist. past
Immediate past	S-*á*-V-*a*	S-*áá*-V-*a*	intermediate past

Regarding the tests that Nurse proposed in order to check whether a form is temporal or aspectual, the first, systemic test is not available to us at this point, as the heart of the matter is whether the suffix, which often is an aspect marker in Bantu, can have a temporal reading in addition to the aspectual one. This test is a bit circular, as one must know how the system works before one can determine if the tense and aspect markers are always separated. Until more is known about the Mbugwe tense and aspect system, this test is not possible.

The test concerning compound constructions does not work on Mbugwe either, as the periphrastic forms are very different form the typical Bantu pattern. No tense or aspect is marked on the main verb, which is in infinitive form, and the auxiliary seems to be able to mark both tense and aspect. This

also needs further investigation in order to be sure if the markers are purely aspectual or temporal.

The third test, checking how anteriors behave with different kinds of verbs, such as stative/inchoative vs. dynamic, is possible for Mbugwe. Mous does say that the perfect form is used to express such a state, which would indicate that at least in this context it functions as an anterior:

(10)   *n-kat-íye*
        1SG-fatigue-PRF
        'I am tired.' (Mous 2004: 11)

As noted above, however, it is possible to have several anteriors in languages with several pasts. More research is needed in order to determine what role *aktionsart* plays in the Mbugwe verbal system.

The final test is the range of reference of the morpheme. From the analysis it is clear that the form which is called perfect, or near past, refers mainly to the day of speaking, but also to previous events which either have some relevance to the present, or a continuation of a state entered in the past. Nurse mentions that Giryama *dza* is both a recent past and an anterior, and it appears that the same is true for the Mbugwe verb form with just the suffix *-íye*, and possibly for the form with *áa-…-íye* as well (which might function as both a middle past and a perfect).

3.5. An integrated analysis

Östen Dahl (p.c.), when presented with some of the data in this paper, suggested that Nurse's Alternative 2 (with aspectual readings) might be an older system, which is being replaced with a system of multiple degrees of past time reference. When considering it as a case of grammaticalization of the anterior into different degrees of past, it is not surprising that the recent past is hard to distinguish from the anterior aspect. It could in fact be that this form in particular can function as both a hodiernal (recent) past and as anterior, and the differentiation we want to make between them might not exist in the language at this stage. The same can be said for the past perfect/middle past, as the data clearly shows that sometimes the form has a purely temporal reading, and sometimes a past perfect reading.

My proposal then, is that Mbugwe is in the process of reanalyzing the perfect and past perfect forms as different degrees of past. At the present stage both a metric, purely temporal reading and an aspectual, perfect reading are possible for the two forms which have the perfect suffix. As such, Nurse's two alternative analyses give a good picture of the situation, and emphasize the ambiguity of the forms. Whether this process necessarily will lead to a loss of

the possible perfect reading, or whether they will continue to co-exist, is still an open question.

The new, integrated analysis, based on the present discussion, and including some additional information from my SIL colleagues, is presented in Table 4. The negative forms need to be investigated independently, but are labelled in analogy with the positive forms for now.

Table 4: New categorization of Mbugwe verbs (Adapted from Mous 2004)

form	*tense marker(s)*	*final vowel*	*tense*
SC-*jé*-√	FUT	—	near future
SC-*á(a)*-√-*á*	PST	PST	far past
SC-*áa*-√-*íye*	PST	PRF	middle past/past perfect
SC-√-*íye*	PRF	PRF	near past/perfect
SC-*ándá*-√	HAB	—	habitual
SC-*ké*-H-√	COND	—	conditional
SC-*ká*-H-√-*íye*	IRR	PRF	irrealis
SC-H-√-*e*	—	OPT	optative/subjunctive
SC-*ka*-√	CSEC	—	consecutive
NEG-SC-*ándá*-H-√	HAB	—	not yet
NEG-SC-*áa-ré*-H-√	PST-*be*	—	negative past imperfective
NEG-SC-*á*-√-*á*	PST	PST	negative far past
NEG-SC-*á*-√-*íye*	PST	PRF	neg. mid. past/past perfect
NEG-SC-√-*íye*	—	PRF	neg. near past/perfect
NEG-SC-*jé~já*-H-√	NEG	—	general negative
NEG-SC-*jé*-H-√-*íye*	NEG	PRF	negative past
NEG-SC-*ká*-√-*íye*	IRR	PRF	negative irrealis
SC-*ki-sé*-√~SC-*káy-sé*-H-√	COND-NEG	—	prohibitive
INF + SC-*je*	*come:*OPT		near future
INF + SC-*ka-ja*	*ka-come*		distant future
INF + SC-*áa-re*	PST-*be*		past imperfective
INF + SC-*anda*	HAB		present habitual
INF + SC-*kénde*	PRES.PROG		present progressive

4. Conclusion and future research

The data presented in this paper suggest that Mbugwe is in the midst of a process of reinterpreting the past verb forms as different degrees of past, instead of having a primarily aspectual function. Hopefully, the discussion sheds some light on the development of metric tense systems from tense/aspect systems, as one reviewer remarked. This paper is the result of an initial survey only, and should be taken as such. Clearly a lot more data and a deeper understanding of the language are necessary in order to confirm the hypothesis put forth here. Of particular interest is to test whether it is possible to assign discrete meaning to each tense/aspect marker and to each verb form, and to gain more insight into the process of grammaticalization which seems to be happening in Mbugwe. Further testing of the range of reference of the verb forms is required, as well as a full overview of the whole verbal system, including tense, aspect, mood and perhaps evidentiality.

It is also important to keep in mind that language is never static, but that it keeps changing. As mentioned previously, there is a high rate of bilingualism with Swahili, and any influence from Swahili or other neighbouring languages needs to be taken into account when analyzing Mbugwe grammar.

During my PhD project I plan to spend two or three extended periods in the Mbugwe area for data collection, language learning and analysis. The value of consulting with native speakers and also gaining a certain level of command of the language cannot be overestimated. Another valuable asset is the input from my SIL colleagues, who live and work in the area and will therefore know the language to a much higher degree than I will. My first fieldtrip will take place during autumn of 2011, and I am looking forward to gaining more insight into this fascinating language.

References

Baumann, Oskar. 1894. *Durch Massailand zur Nilquelle: Reisen und Forschungen der Massai-Expedition des deutschen Antisklaverei-Komite in den Jahren 1891–1893.* Berlin: D. Reimer.

Botne, Robert. 1999. Future and distal-ka-s: Proto-Bantu or nascent form(s). In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics,* 473–515. Stanford: CSLI.

Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *Evolution of grammar: tense, aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.

Comrie, Bernard. 1976. *Apect.* Cambridge: Cambridge University Press.

Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.

Dahl, Östen. 1985. *Tense and aspect systems.* Oxford: Blackwell.

Dahl, Östen & Viveka Velupillai. 2011a. Tense and Aspect. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World atlas of language structures online.* Munich: Max Planck Digital Library, supplement 7. Available online at http://wals.info/supplement/7 (30 May 2011).

Dahl, Östen & Viveka Velupillai. 2011b. The Past Tense. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World atlas of language structures online.* Munich: Max Planck Digital Library, chapter 66. Available online at http://wals.info/chapter/66 (30 May 2011).

Dempwolff, Otto. 1915–1916. Beiträge zur Kenntnis der Sprachen in Deutsch-Ostafrika, 7. Buwe. *Zeitschrift für Kolonialsprachen* 6, 1–27.

Güldemann, Tom & Edward D. Elderkin. 2003. On external genealogical relationships of the Khoe family. In Matthias Brenzinger & Christa König (eds.), *Khoisan languages and linguistics: the Riezlern symposium 2003* (*Quellen zur Khoisan-Forschung* 24), 1–27. Köln: Rüdiger Köppe.

Guthrie, Malcolm. 1948. *The classification of the Bantu languages.* London: Oxford University Press for the International African Institute.

Kesby, John D. 1986. *Rangi natural history: the taxonomic procedures of an African people.* New Haven: Human Relations Area Files.

Larsen, Julia & Viggo Larsen. 2011. *Paradigms of Mbugwe.* Unpublished manuscript.

Lewis, M. Paul (ed.), 2009. *Ethnologue: languages of the World,* 16th edn. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com.

Masele, Balla F.Y.P. 2001. *The linguistic history of Sisuumbwa, Kisukuma and Kinyamweezi in Bantu zone F.* St. John's: Memorial University of Newfoundland.

Masele, Balla F.Y.P. & Derek Nurse. 2003. Stratigraphy and prehistory: Bantu zone F. In Henning Andersen (ed.), *Language contacts in prehistory: studies in stratigraphy,* 115–134. Amsterdam: John Benjamins.

Mous, Maarten. 2000. Counter-universal rise of infinitive-auxiliary order in Mbugwe (Tanzania, Bantu F 34). In Reiner Voßen, Angelika Mietzner & Antje Meißner (eds.), *Mehr als nur Worte…: afrikanistische Beiträge zum 65. Geburtstag von Franz Rottland,* 469–481. Köln: Rüdiger Köppe.

Mous, Maarten. 2004. *A grammatical sketch of Mbugwe: Bantu F34, Tanzania.* Köln: Köppe.

Nurse, Derek. 1999. Towards a historical classification of East African Bantu Languages. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics*, 1–41. Stanford: CSLI.

Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford: Oxford University Press.

Nurse, Derek & Gérard Philippson. 2006. Common tense-aspect markers in Bantu. *Journal of African languages and linguistics* 27, 155–196.

Seidel, Alfred. 1898 "Grammatik der Sprache von Irangi" In Waldemar Werther (ed.), *Die mittleren Hochländer des nördlichen Deutsch-Ost-Afrika*, 387–434. Berlin: Hermann Paetel.

Stegen, Oliver. 2006. Participatory research. TAM functions, and dialect survey among the Rangi: a reply to Dunham (2004–2006). Handout, 36th CALL.

Struck, Bernard. 1909. Mbugwe. *Mitteilungen aus den deutschen Schutzgebieten,* Ergänzungsheft 2, 62–63.

UNESCO (=UNESCO Ad Hoc Expert Group on Endangered Languages). 2003. *Language vitality and endangerment. Document submitted to the International expert meeting on UNESCO Programme Safeguarding of endangered language. Paris, 10–12 March 2003*. Paris. (manuscript.)