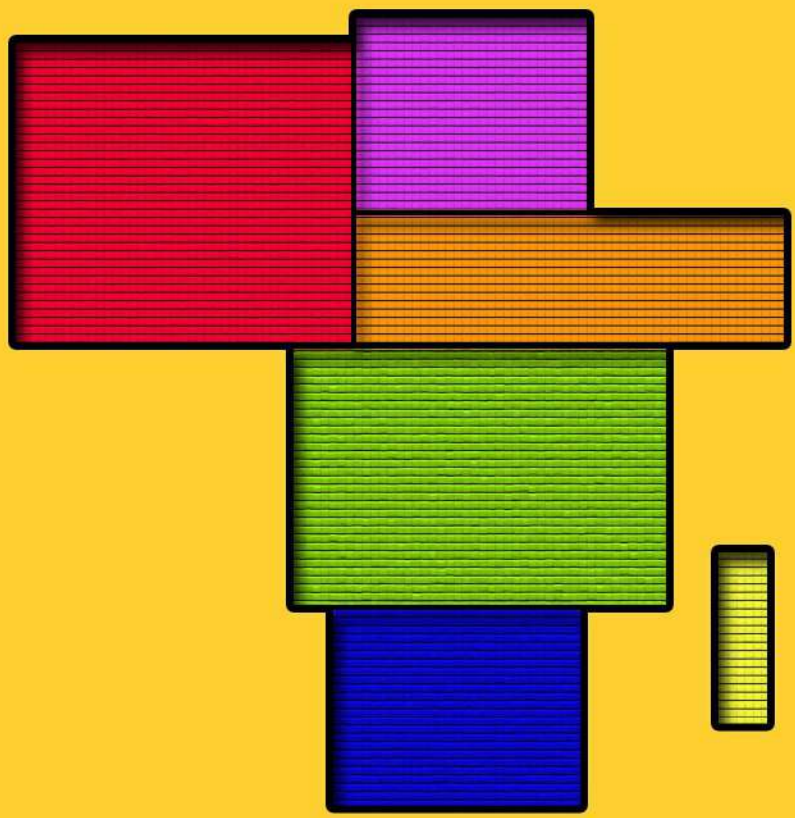
3/2022



JOURNAL OF AFRICAN LANGUAGES AND LITERATURES

UP

3/2022

JOURNAL OF AFRICAN LANGUAGES AND LITERATURES

Department of Asian, African, and Mediterranean Studies
University of Naples "L'Orientale"
jalalit@unior.it

ISSN 2723-9764

UniorPress

Table of Contents

JOURNAL OF AFRICAN LANGUAGES AND LITERATURES
3/2022, 1-21

Revisiting the English-Swahili debate on Tanzania's medium of instruction policy at secondary and post-secondary levels of education

GASTOR C. MAPUNDA

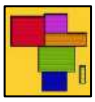
University of Dar es Saalam
mapundag@gmail.com

ABSTRACT

Like in other African countries, in Tanzania the debate on the medium of instruction has focused on the use of either English or Swahili in secondary and post-secondary education. During British colonialism, the focus of the debate was on ethnic languages, Swahili and English at primary level of education. Swahili was used in lower primary education and English in upper primary, middle, and in secondary education. After independence, pedagogical-cum-nationalist opinions wanted a complete changeover from English to Swahili. In 1967 Swahili replaced English in primary education, and speaking English in public was frowned on. Consequently, mastery of English declined. Swahili was also to replace English in secondary and post-secondary education, but it has not happened until now. Whilst it is true that most children have not mastered English to be able to use it comfortably in their studies, similar problems apply to children in remote rural Tanzania who have not mastered Swahili, especially in beginner classes at the primary level of education. Yet, the problem with Swahili and ethnic languages is never seriously debated. Nevertheless, English still commands symbolic and material value. Using a translanguaging perspective, we find that the merits of English in education outshine its demerits. It is recommended that the debate take a pragmatic-cum-utilitarian angle that multilingualism can unlock opportunities for learners.

KEY WORDS: medium of instruction policy, Tanzania, Swahili, trans-languaging perspective, monolingual practices

DOI https://doi.org/10.6093/jalalit.v3i3.9141
ISSN 2723-9764

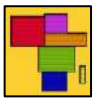


1. Introduction

The decision as to which language should be used in a country's education system should not be based solely on past histories, nationalist or patriotic aspirations, and pronouncements. The debate on whether or not Swahili should replace English in the Tanzania's secondary and post-secondary education is a clear illustration of medium of instruction (MoI) decision debates. Debates on MoI decisions are also common in other parts of the post-colonial nations, including South Africa, Zimbabwe, etc. (HUNGWE 2007, PROBYN 2009, BRISTOWE, OOSTENDORP and ANTHONISSEN 2014). This article revisits the debate on whether Swahili should replace English in Tanzania's secondary and post-secondary education system, a debate that has been there for more than five decades now. Nevertheless, it is important to still talk about it now because the pressure to change from English to Swahili medium is resurfacing, now that Swahili is gaining more international acknowledgement, such as by UNESCO (2021) (and by the African Union). The study sets out to address two main questions: (i) Why has there been a feeling that Swahili needs to replace English in education? (ii) Why is changeover from English to Swahili taking so long? The rest of the chapter is devoted to answering these two questions in detail.

As early as 1907, when what is now Tanzania was part of *Deutsch-Ostafrika* (German East Africa), the colonial authority was under pressure by the Asian community who wanted only Arabic and English to be used as media of instruction (MoI) in education (CAMERON and DODD 1970: 75). The Government did not yield to the pressure, and opted for Swahili instead. At that time, the suggestion to use English was facilitated by the presence of a sizeable number of Asians. The debate re-emerged also during the British colonial rule, as to which among the ethnic languages, Swahili or English should be used as the MoI (CAMERON and DODD 1970, BARRETT 2014). The Phelps-Stokes Commission's report of 1924 recommended the use of ethnic languages in primary school because education was to be adapted to the context and needs of the communities it served (CAMERON and DODD 1970: 134). The argument was that most communities did not know Swahili or English, but only their own ethnic language. Nevertheless, the British continued to use Swahili in the first five years of primary education as the MoI, and switched to English in the last three years, and throughout secondary education (BARRETT 2014: 4).

Those who have generally taken part in the Tanzania's language-in-education debate on Tanzania (including linguists, educationists, politicians, and the general public) have based their arguments on the relative benefits of using either English or Swahili in the country's education system (e.g. MEKACHA 1994, Roy-CAMPBELL and QORRO 1997, KAPOLI 1998, RUBAGUMYA 1999, KADEGHE, 2003).

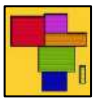


Conversely, those who are opposed to the use of English have taken the angle that English is colonial, imperialist, and thus incapable of serving as a viable vehicle for delivering education in a meaningful way (e.g. MOCHIWA 2001, MASELE 2005). This group holds that, foreign as it is, English is elitist because it caters for only a small fraction of the citizenry, mainly serving the interests of its native speakers. On another note, those who are pro-English feel that English is a useful MoI, and suggest that it has the potential of facilitating education transactions (e.g. KAPOLI 1998, KADEGHE 2003). This group has viewed English as a global language, and paid little attention to the argument that it serves the interests of particular nations. These two camps have also been surfacing in both formal and non-formal dialogues among staff at the University of Dar es Salaam from time to time.

From the early days of independence, the Government of Tanzania preferred the use of Swahili in government activities, including education (specifically in primary education). Even before independence, the Tanganyika African National Union (TANU) made a proposition in 1954 that Swahili be made the second language in the Legislative Council's meetings (TANU 1954: 5). Furthermore, in the early 1960s when Tanganyika became independent, one urgent issue was to change the MoI from English to Swahili. Consequently, starting from late 1960s to mid-1970s, the use of English in public spaces was looked at with disdain, and it was even referred to as *kasumba ya kikoloni* (colonial hangover) (TRAPPES-LOMAX 1985: 12; MAPUNDA 2015: 37). For example, in 1964 the then Second Vice-president, Rashidi Kawawa, instructed public offices not to use English, if it was not necessary. In 1967, Kawawa also revisited his instruction:

> "The Government has made a decision that as from now Swahili shall be used in all government transactions, and the practice of using English or any other foreign language should cease forthwith, except if unavoidable." (*BAKITA* 1970, 1; our translation)

More recently, language activists, such as some academics, politicians and partisans, have taken the lead in seeing to it that the changeover from English to Swahili becomes a reality; and whoever is pro-English is viewed as an apostate. However, almost six decades have passed since the Government of Tanzania wanted a radical change in the use of Swahili as a MoI, and yet the Government is still indecisive and ambivalent on a clear and fine-grained language-in-education policy for all levels of education. It is within this context this indecisiveness and ambivalence that the question of which language should be used as the MoI in Tanzania is being revisited now.



2. Overview of the use of Swahili and English

Swahili and English languages have co-existed in Tanzania's education since the onset of British colonialism. Since part of this story has already been explained in the introduction section, in this part we will talk briefly about their ecological co-existence. To begin with, Swahili was preferred by the Germans, and then by the British, and later on by the independence government (CAMERON and DODD 1970, BARRETT 2014). After independence in 1961, more efforts were made by the Government. For example, according to HILL (1980: 368-369), Swahili was used as the MoI for adult education even in areas where the language was not spoken by the community. Many educational books were published in Swahili; and the Institute of Kiswahili[1] Research and the Tanzania Council for Kiswahili worked hard to coin vocabulary for educational purposes. Educational radio programmes were also aired in Swahili.

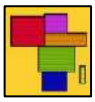
The Government was aware that Swahili was not known throughout the country, and so adult education was also used by the Government as a forum for teaching it. Writing for UNESCO, VISCUSI (1971: 10) observes that the Tanzanian Government made efforts which were commended:

> "For those rural Tanzanians who have never attended school, the government looks to adult literacy classes as a means of helping to establish Swahili as a true national tongue. For some of these people, therefore, becoming literate also means acquiring or perfecting a second language (which is Swahili)."

English started to be used in the education system after 1918 when the Germans lost the First World War to the British. It started to be used as the MoI in the last three years of primary education and throughout secondary education (Barrett 2014: 4). After independence, secondary and post-secondary education still retained English as the MoI, even though attempts were underway to eventually replace it by Swahili. For example, in 1969 the Ministry of National Education sent a circular to all headmasters and headmistresses of all secondary schools in Tanzania regarding the possibility of introducing Swahili as the MoI in secondary schools in at least some subjects. The circular suggested that political education could be taught in Swahili starting from 1969/70, and domestic science in 1970.

Other actors also made several attempts to discourage the use of English, and the National Council for Kiswahili (1970) was among them. In 1967 the Government adopted the use of Swahili as an exclusive MoI in primary education. Following this movement to uphold Swahili, it became apparent, in the words of those

[1] The prefix *ki-* is often used in referring to the language; but in this article we prefer to call the language Swahili.



whom I may call language purists, the standard[2] of English in the country deteriorated. Remarking on his personal experience, ISHUMI (2004: 13), for example, observes that

> "The English spoken today by our students, and hence by their teachers, is no longer the impeccable English that was practiced in the classrooms and in the corridors of the institutions of the early years of the independence period. The same can be said of the wider population outside the walls of institutions."

An often quoted report on the situation of the English language in Tanzania by CRIPER and DODD (1984) also noted that the teachers themselves, both degree and diploma graduates, had problems with the English language. They noted that the use of English in the schools was very minimal, something that needed redressing (CRIPER and DODD 1984: 36).

To date, there is still no indication that Tanzania will do away with English as a MoI in secondary and post-secondary education. In 1969, Harries noted the ambivalence in the language policy. He said, "There is no active opposition towards the use of English by the authorities, but rather a positive commitment to promote the use of Swahili as widely as possible. No special or extraordinary steps are taken to promote the use of English, but the further use of Swahili is encouraged" (HARRIES 1969: 278). This endless ambivalence has given an indecisive decision in regard to the Swahili-English debate. Nyerere, president of Tanzania (1962-1985), defended the use of English in education, arguing that Tanzania still needed the support of the rest of the world due to its economic woes (RUBAGUMYA 1990: 27, BARRETT 2014: 7). Government ambivalence was also shown in the 1980s when Nyerere, the president of Tanzania then, set a presidential commission to study the state of education in Tanzania, headed by Jackson Makweta. When their report came out in February 1982, they recommended a changeover of the MoI from English to Swahili at secondary and later post-secondary levels.

Interestingly, in 1984 the Ministry of Education issued a statement, saying that both English and Swahili would still be used as a MoI, and that the teaching of English would be further improved. Additionally, the President made it even more explicit that English was still important and Tanzanians needed it. In a statement quoted in *Mzalendo* newspaper (owned and run by the ruling party – *Chama cha Mapinduzi*) on 28 October 1984, Nyerere used very strong words, "It is wrong to leave English to die. To reject English is foolishness, not patriotism"

[2] This deterioration is in comparison to the level of mastery during the days when English was taught by native speakers before independence.

and so "English will be the medium of instruction in secondary schools and institutions of higher education, because if it is left as only a normal subject it may die". This background information suggests that there is a need to look into the MoI debate in a better way.

3. Arguments for and against the use of English

The feeling that Swahili should replace English emerges mostly within findings from the library research which we conducted. There are several arguments in the literature that propose the replacement of English by Swahili in the Tanzania's education system. The arguments can be categorized as weaknesses in the use of English, nationalist arguments, and monolingual tendencies in education. These are presented in the next sub-sections.

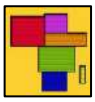
Weaknesses in the use of English

What is apparent in Tanzania's educational institutions is a general weakness in the mastery of the English language by students, but also by teachers. In universities, for example, a study by MAPUNDA and MAFU (2014) which involved 450 first year students from the Universities of Dar es Salaam (UDSM), Sokoine (SUA) and Tumaini - Iringa (TUICO), in which students were required to write essays on a common topic, identified a number of consistent weaknesses in the mastery of English. The errors included transfer of concepts from Swahili, wrong choice of words and translating directly from Swahili, rampant spelling errors, rule overgeneralization, wrong use of adjectives, and vague expressions.

Also, from the mass media, a number of complaints have been written about. One example is by MAKUMBA (6 August 2010) who reported in the *Business Times* about a concern by a Tigo Human Resource Officer, who complained about the weaknesses of Tanzanian graduates who applied to Tigo for jobs. She noted that Tanzanian graduates (both secondary and post-secondary) have poor communication skills, especially when it comes to the use of the English language, which was evident in almost every interview in which she participated.

Nationalist arguments

Many arguments against the use of English as the MoI are based on nationalist arguments, even though on the surface they may sound pedagogical. MAZRUI (1992) makes an important remark as to why English is rejected in post-colonial Africa, including Tanzania. He argues that the rejection of English can be an expression of nationalist sentiments. The fact that the language was associated with the brutality associated with colonialism and enslavement, and also that in



the beginning it was accessible only to a selected few, has engendered ill-feeling against it. Thus, the need to replace English by Swahili comes about because of

> "[…] (Europeans') brutal and humiliating experiences of enslavement and colonization. As a result, the nationalist sentiments towards Kiswahili were bound to involve a rejection of any seemingly negative projection of the language which may have been engendered by colonial discourse." (MAZRUI 1992: 89).

ROY-CAMPBELL and QORRO (1997) are among the scholars who have strongly argued against the use of English in education in Tanzania. In their book titled "Language crisis in Tanzania", the authors argue against the use of a *foreign* language in Tanzania's education. A quote from their book suggests that a foreign language should not be used in Tanzania's education system:

> "Why should cultural and moral values, customs and traditions be learnt through the medium of English, a foreign language? This is tantamount to suggesting to the students that their background is inadequate and of little value." (ROY-CAMPBELL and QORRO 1997: 78-79)

MZEE (1995) also suggests that using a foreign language is like nurturing colonialism. In which case, the problem may not be that the language cannot fulfil certain sociopolitical functions, but that because of its colonial background it is not acceptable.

> "In an attempt to counter campaigns, run by foreign newspapers against efforts by the Tanzanian people to liberate themselves from colonialism, and finding their own development, the TANU Party (at that time) decided to launch its own newspaper called *Sauti ya TANU* (Voice of TANU) which was published in Swahili; it did a lot to educate the citizenry about TANU politics." (Our translation from the original Swahili version)

Furthermore, other scholars have linked the continued use of English to class formation in society. One such example is MALEKELA (2003: 110) who argues that "The continued use of English as a medium of instruction in post-primary education is meant to boost the status or prestige of a few who have inferiority complex towards our national language, Kiswahili". In a further criticism of English, CHACHAGE (2004) has likened the use of English in Tanzania to the propagation of cultures of other nations, and particularly those of the economically powerful nations.

Monolingual tendencies

It has been observed that teachers and students who have been codeswitching between English and Swahili are regarded as doing the wrong thing. They have been viewed as incapable of sustaining the process of teaching/learning in English. An example of this is from ROY-CAMPBELL and QORRO (1997: 61), who hold that, "If more than half of the respondents admit that they sometimes ask the teacher questions in Kiswahili when in class, then the students' commitment to English can be called into question". For them, it is a weakness of the students to ask questions in Swahili.

This view is also held by MOCHIWA (2001), MALEKELA (2005), and MASELE (2005). MOCHIWA (2001:5), for example, has argued that the use of English in education is a hindrance: "English continues to impede communication between teachers and students, so the development of the child's brain does not get an opportunity to develop". Thus, for these people, the immediate solution is for English to be replaced.

4. Theoretical basis

We approach the debate on the MoI in Tanzania from a translanguaging perspective as an alternative lens which can help iron out the dilemma of which language should be used or which language should replace another language in an education system. For GARCÍA and WEI (2014: 2) translanguaging is

> "an approach to the use of language, bilingualism and the education of bilinguals that considers the language practices of bilinguals not as two autonomous languages [or more] systems as has been traditionally the case, but as one linguistic repertoire with features that have been societally constructed as belonging to two separate languages."

They also view translanguaging as being capable of looking "at linguistic realities of the world today and how human beings use their linguistic knowledge holistically to function as language users and social actors", not as language users who operate according to some policy imperative.

The history of translanguaging can be traced back to Welsh classrooms in the 1980s. It was found that when learners were allowed to freely interchange between English and Welsh (WEI 2018: 15) their learning experiences became even more successful.

The ideas that underlie translanguaging also find support in the works of LANZA (2007), MATRAS (2009), and LÜPKE and STORCH (2013), among others, who prefer

the view of language repertoires, as opposed to language systems. For example, MATRAS (2009: 4) argues that bilingual and multilingual speakers

> "have a complex repertoire of linguistic structures at their disposal. This repertoire is not organized in the form of 'language' or 'language systems'. Rather, elements of the repertoire (word-form, phonological rules, constructions, and so on) gradually become associated through a process of linguistic socialization, with a range of social activities."

Besides, LÜPKE and STORCH (2013: 2) posit that, it is counterproductive to think of 'fixed languages or fixed linguistic identities' in multilingual settings. For them, speakers use repertoires; and repertoire choices are determined by 'domains, contexts, addresses, and many other factors'. This is not the same as forcing language users to adopt one particular language and leave out another one. They are not comfortable with the view that one 'named' language operates separately, and so should replace another 'named' language. MATRAS (2009: 9) who is even more critical of the view of languages as linguistic systems, suggests an abandonment of the idea of linguistic systems in favour of linguistic repertoires which consider languages mainly as communication.

Coming back to translanguaging, GARCÍA (2017: 258) holds that the perspective is aware of the linguistic potential that students come into the school with. This potential links them to their past, and at them same time, influences their progression towards acquiring the dominant standard languages, and also makes a way for them to move toward creative languaging that opens up further possibilities of knowledge creation. The main claim of translanguaging is thus the use of two or more languages in multilingual educational settings enables teachers and leaners to transcend language boundaries since multilingual speakers do not operate each language separately; all the codes operate as a unified whole complementing each other.

Another proponent of translanguaging is CANAGARAJAH (2011) who, like the others, is opposed to monolingual tendencies in multilingual settings and feels that there is a need for an "affirmative action" to undo such tendencies. His position is based on the fact that many of the discussions that prevailed in the past about language use in multilingual classrooms focused on the view that the use of more than one language is a weakness. GARCÍA and WEI (2014: 137) argue that "through multiple discursive practices that constitute the language users' linguistic repertoires, translanguaging makes visible the different histories, identities, heritages and ideologies of multilingual language users". However, looking at classroom research on multilingual settings, one almost always encounters arguments that discourage the use of more than one language in the

classroom (e.g. SENKORO 2004: 53; QORRO 2004: 109); thus arguing for one 'named' language replacing another 'named' language. PROBYN (2009) has even shown ironically that bringing into the classroom a language that is not designated by policies to be used as such as tantamount to 'smuggling' into the classroom an illegal language.

It is therefore our conviction that translanguaging can be a feasible alternative to classroom discourses in multilingual contexts. As such, using more than one language in an education system of a multilingual nation concurrently can be well accommodated, and the dilemma of whether or not one language should replace another language will no longer resurface.

5. Methods

In this study, three methods of data collection were used, namely library research, questionnaire, and interview. Through the library research, we consulted journal articles, books, and newspaper articles. The questionnaire was administered to 467 students from four secondary schools in Songea Town, Tanzania.[3] These students are approximately 17-21 years old. For ethical reasons the schools are named here as School1 (in which 162 students participated), School2 (where 81 students participated), School3 (where 92 students participated), and School4 (where 132 students participated). The administration of the questionnaire was meant to collect data on students' language practices, hypothesizing that language practices may provide information on whether or not government's (which are also schools') policies on the MoI align with such practices.

As for the interviews method, two categories of participants were involved. Besides, two reasons prompted their use. For the students' interviews the goal was to get descriptive information on the responses from the questionnaire. This first category involved 12 advanced level[4] secondary school students from three schools in Songea Town (School1, School2, and School3). These same students had participated in filling the questionnaire, which had also been filled by the 467 students. The second category involved ten (10)[5] high ranking executives

[3] This was part of a larger study on language attitudes and youth identities.

[4] According to Tanzania's education system, primary school takes seven years, general certificate level of secondary education four years, and advanced level of secondary education takes two years. At the advanced level, students specialise in specific subject combinations according to their pass grades at the general certificate level.

[5] These are part of a larger project commissioned by the Tanzania Commission for Universities (TCU). I am grateful to the TCU for allowing me to use these data.

from Government ministries (Foreign Affairs, Works, and Labour), four university vice-chancellors (from two private, and two public universities), a country-level official in a multinational communication and financial company, a top official from the Tanzania Private Sector Forum, and an executive from the Public Service Commission. The reason for holding interviews with these officials is a utilitarian one: to get their opinions on whether or not it is important to make a changeover from English to Swahili MoI in Tanzania's secondary and post-secondary levels of education.

Interviews with students were held in three public secondary schools located in Songea Town, in south-western Tanzania, in vacant classrooms which were made available to us. These were from School1 (which admits only boys); School2 (which admits only girls) and School3 (which admits both boys and girls at the general certificate level, and only girls at the advanced level). Interviews with the various senior officials were held in their offices, some in Dar es Salaam Region and some in Morogoro Region.

6. Findings

In this section we present findings based on each of the two research questions which the study set out to address. The data are therefore presented according to the research questions, but also in line with particular data collection methods.

6.1 Derailment of changeover from English to Swahili

It has now become apparent that a changeover from English to Swahili is taking a long time, and maybe it may take even longer. This will be shown in the following sub-sections with data from questionnaires and interviews.

Findings from the questionnaire

We administered a questionnaire to 467 secondary school students, probing into matters of students' language practices, focusing on English, Swahili and ethnic language in the Ruvuma Region, Tanzania. Secondary school students are normally in the range of 14-17 years for the general certificate level, and 18-20 for those at the advanced level. In this study we are reporting findings on three questions, namely (i) *Which language do you speak with your teachers at the school, but outside the classroom?* (ii) *Do you think that you will work in Tanzania or abroad in the future?* and (iii) *Which language do you think will give you the best working opportunities?* In answering question (i) (which language they usually speak with their teachers at the school, but outside the classroom), we found that 329 (70.4%) speak Swahili, 124 (26.6%) speak English, and only 14

(3%) speak an ethnic language. However, even though Swahili was the language most spoken with teachers outside the classroom, their attitude to the language was quite the opposite. Regarding the second question (ii), that is, where they would like to work in the future, out of the 467 participants, 313 (67%) will work outside Tanzania, while only 154 (33%) will work locally. As to question (iii) (which language they think would give them the best working opportunities in the future), 83.9 per cent chose English, while only 15 per cent chose Swahili and 0.2 chose ethnic languages. Thus, their imagined future work place, and which language would facilitate their imagination match. These findings suggest that even some of those who wish to work in Tanzania still think that English is more promising than Swahili.

Findings from students' interviews

The findings from the interviews are from 12 advanced level secondary school students from School1, School2 and School3. The main interview question asked to the students was about the relative amount of time they spent learning the languages (English, Swahili, or an ethnic language) and their reasons for doing so. First, out of the 12 students, nine spend more time learning English because English carries more material and symbolic value for their future, while the remaining three spend more time learning Swahili. Those who attach more value to English have friends outside Tanzania, want to learn fashion from the internet, which is in English, want to become lawyers and soldiers; and some want to live outside Tanzania, and particularly in Uganda, South Africa and the USA.

Another reason for attaching more value to English is that examinations in Tanzanian secondary schools are given in English; and if they want to do well, then they need to master English well. Some of them want to pursue further education abroad, and scholarships are given to those with good examination grades; and these will be those who have mastered the English language. Yet for some, speaking English is prestigious in their communities. Amina[6] (female, School1, studies arts subjects, Form VI) will be self-employed, and for her, Swahili may help her; but it is English that would provide more opportunities for her. She said, "I want to become an interpreter-cum-translator, but what I need more is English. It is English which will give me another language to become what I want". Also, Rose (female, School3, studies science subjects, Form VI) works hard to master English because she does not see the necessity of spending time learning Swahili, which she already knows: "I already know Swahili, so I need to learn English. After all, Swahili is useful only for communication with the

[6] All the names used here are pseudonyms for ethical reasons.

local people. For me, what is more important and useful is English". Rose also feels that employment is mostly provided by foreign investors, for whom, knowledge of English is more valued than Swahili. For Teddy (female, School3, studies arts subjects, Form VI), Swahili is useful only to a small extent: "I think Swahili may help me, to a small extent, but I am not even sure how. I want to become a university lecturer, and I think it is English which will help me". Peter (male, School 2, studies arts, Form V) wants to become a teacher in an English medium school, and thinks that he needs to become better in English.

Interestingly, three out of the nine interviewees think that for them Swahili carries their future. One of them is John (male, School2, studies science subjects, Form V) who wants to become a Swahili language teacher, and spends a lot of time learning Swahili. However, he also thinks that English is more important, "One cannot become a teacher with only one language". As such, he needs to spend more time learning English. In the future, he also wants to write Swahili books and sell them; and yet he invests more efforts in learning English. Also, Mary (female, School 3, studies arts subjects, Form V) is the only participant who does not subscribe to the view that English will help her; and for her, Swahili is more important, "I want to become a Swahili teacher or professor, and there aren't many experts in Swahili. I want to become one; and Swahili is my future hope". For Mary, the usefulness of English is only when she comes across those who do not speak Swahili. Also, for the other students who have more regard for Swahili, their reasons include the fact that they have no intention to go and work abroad; Swahili was becoming more and more popular within and outside Tanzania, so there is assurance of employment; and they did not have an opportunity to learn English well while still in primary school.

Findings from senior officials' interviews

Interviews with high level officials in governmental and non-governmental sectors focused on the usefulness of English to their organization, and what would be the consequences of the Government adopting a Swahili-only policy in the education system.

All their responses were related in many ways. For example, Participant1 (works in the Ministry of Works, Transport and Communication) is of the view that, the kind of university graduates whom they need, should have English language skills, particularly interview skills, self-expression, organization of thoughts, report writing, and communication skills in general. This was also a concern by Participant2 (works for a multinational communication and financial company)

who also valued the skills of organization, marketing, and sustaining a point of view.

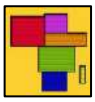
In the Ministry of Foreign Affairs, Participant3 was of the view that language skills, both English and Swahili, are highly needed in their Ministry because they interact with internationals on a daily basis. She also argued that those with good mastery of English are advantaged, "Those who studied abroad, or in English medium primary schools are in a better position in most cases. During interviews, most of them cannot express themselves, and they cannot write reports well". In addition, all the university vice-chancellors held the same view that the English language is important at the university level. To be more specific, Participant5 (works with University1, Public) even recommends that there should be a compulsory English language course throughout university schooling:

> "English language skills are highly needed at this level. More importantly, the soft skills are lacking. In fact, there should be a compulsory credited English course for all students all the way from the first year. The nation needs English to reach the middle income level status."

Participant6 (University2, Private) lamented that students at his university do not value English, which he sees as being very important. He saw that there was a growing tendency at his university to view English as imposed and colonial. He proposed that there be a sensitization campaign for students to be willing, proud, and motivated to learn and master English as a tool for employment. There is a need for continuous English courses for all students. Participant7 (University3, Private) underscored the relevance of English but lamented that mastery of the language by the public is very poor, and that, when they introduced an entry examination at his university, 200 of the applicants failed.

7. Discussion of the findings

In this section we discuss the findings in the order of the research questions. The first question asks why there has been a feeling that Swahili should replace English. From the findings, it has been observed that those who have taken part in the debate have tended to associate nationalist feelings and the thinking that monolingual policies in education work better. Deep down, they uphold the liberation role played by Swahili during the fight against colonial oppression, and ignore the usefulness of the English language in productive sectors, including education. MAZRUI (1992: 88) made an important remark on this issue:



> "Many Africans developed strong nationalistic sentiments towards Kiswahili, seeing it as a language of national sovereignty, as a possible symbol of transnational and continental unity, and as a reminder of the common origins of African descent now scattered throughout the globe [world]."

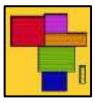
Also, looking at the current situation through the lens of Translanguaging Theory, we may need to think about classroom languaging not on the basis of its history in Tanzanian schools, or even whether English was a foreign or imperial language, but rather, in terms of knowledge construction and sharing. We should go beyond thinking of languages as separate codes. WEI and LIN (2019: 211-212) have argued that:

> "When we talk about the classroom, we tend to have an immediate image of a confined physical space with specified and often hierarchical role sets and planned learning objectives and tasks. Translanguaging classroom discourse is not only about encouraging fluid multilingual practices within the limits and boundaries set up by these role sets, objectives and tasks, but to aim at challenging and transforming them."

As such, when classrooms allow the use of multiple cultural resources, learning becomes active, interesting, and transformative. It is at this point that GARCÍA and WEI (2014) and LANZA (2007), among others, are opposed to segregating linguistic systems because bilinguals/multilinguals automatically transcend linguistic system boundaries when they communicate. When we want to create global citizens, we should not think in terms of fragmenting our education systems in terms of one language replacing another language, but rather different linguistic and extra-linguistic resources working together to produce graduates who can appreciate diversity.

Arguments like the colonial past and foreignness given in support of eventual replacement of English by Swahili in the education system are not sufficiently convincing. It is not foreignness or even colonial background that matter, but the promise that a language carries with it. Chinese is a more recent example of a foreign language which is currently gaining students' attention in Tanzanian. SCHMIED (1986) was surprised to see that, in the Department of Foreign Languages and Linguistics (University of Dar es Salaam), it was French and not English, which attracted more attention among students. But there was a material incentive for this:

> "Another, perhaps surprising, fact is that among students at the Department of Foreign Languages and Linguistics French seems to be a more popular subject than English. This can be attributed to dreams about a career in the diplomatic service or to the study term spent abroad (formerly in France, now "only" in Burundi or Madagascar)." (SCHMIED 1986: 89)



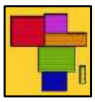
While emphasis has been on Swahili replacing English, some scholars are aware of possible consequences of overhauling the MoI status quo. ROY-CAMPBELL and QORRO (1997) and RUBAGUMYA (1999) are among them. For example, RUBAGUMYA (1999: 138-139) is critical of the desire for a radical overhaul, but at the same time maintains that an overhaul should have been the more logical option. This dilemma is partly due to the thinking that one language should replace another one, a situation that will almost always trigger a dilemma.

Indeed, what is apparent in this discussion is that it is not the English language itself which is the problem; the problem lies beyond the bounds of the language factor. Even the often-quoted Makweta's report (*Jamhuri Ya Muungano Wa Tanzania (JMT)*, 1982), which recommended a changeover from English to Swahili, did not see the problem as being exclusively the use of English. In fact, the report observes that there are problems with the teaching of both English and Swahili. To be exact, this is what the report says:

> "Students at various levels of education are unable to advance themselves in English or Swahili. This problem has been caused by scarcity of facilities and insufficiency of good language teachers, students not being encouraged to learn foreign languages, and not getting sufficient time to practice speaking foreign languages because of fear that results from fear that they would be seen as having *kasumba ya ukoloni* (colonial hangover)." (Our translation from Swahili)

Some of the arguments provided even by those who support the eventual replacement of English, have talked about Swahili as a commodity, that it should also be used across Africa, and in other parts of the world. There is, in addition, the question of fluency in Swahili in remote rural Tanzania. They do not even talk about the denial of children of the benefits of learning in their own ethnic languages in education in their early school years (MAPUNDA 2013). They have from time to time argued that Swahili is a Bantu language, thus it is known within a short time. However, the truth is that many children in remote rural Tanzania do not speak Swahili before they start school (VISCUSI 1971: 10, MSANJILA 1990: 53), and is not being discussed as a more serious problem.

It is possibly Ngugi who convincingly summarises the debate by calling for a triple language policy with a view to decolonising the African mind, "Mother tongue, *lingua franca*, and then French or English or whatever. That is how we are going to decolonise Africa because that also creates an attitude" (KHELEF 2018).



8. Conclusion

The chapter has shown that the debate on whether or not Swahili should replace English as a MoI in Tanzania's secondary and post-secondary education after independence is based on both its colonial history and its pedagogical potential (or lack of it). The literature on this question describes the situation of how English was before and immediately after independence. We also know that, to date, those schools which have sufficiently invested in teaching English and other subjects, have been able to make their students and teachers able to use English in the classrooms.

While it is true that English was introduced during colonialism, this colonial relationship does not remain as such forever. The colonial masters have gone, but the English language has remained, and Tanzania should make use of it. The importance of having English well mastered is in the interest of the nations which use it. It is against this background that even some countries which have never been colonized by Britain, such as Ethiopia and Rwanda, have willingly opted to use English in their education systems.

A further conclusion relates to the reason why the changeover from English to Swahili has not yet taken place. Speaking from a pragmatic point of view, the youth have heard of the woes of colonial brutality, but they do not have first-hand feeling and experience. It is logical for them to take an interest in their future, and not cling so much to their country's pasts, which they have not even lived or experienced. For the youth, it is not the English language that hinders their development, but how resources are managed and shared. For many, and particularly those who will be employed, their future hope lies mostly in the hands of the private sector for employment, which is mostly dominated by foreign companies. It is more plausible for them to think that English is among the tools that might help them hope to realize their imagined futures.

On another note, even staunch supporters of a Swahili-only policy covertly attach a lot of value to the English language as an important MoI. For example, it is known that many of such people have themselves sent their children, grandchildren or children of their siblings to English medium primary schools in order to prepare them well for secondary education. It is because of this reality that there is a lot of competition among parents for privately owned English medium primary schools where English is taught well.

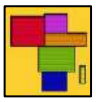
In view of what we have argued, it appears doubtful that Swahili shall eventually replace English as the MoI in the near future. This doubt seems plausible because of the fact that now both Swahili and English have become even more relevant, and stakeholders see value in both. In 2021 UNESCO proclaimed Swahili one of

the world languages, and has also become one of the languages of the African Union, and of the East African Community as well. Many countries in Africa and outside Africa have shown interest in learning it. As for English, most job opportunities are through the private sector; and mostly foreign. For these reasons, students and parents have divided opinions about the value of these languages. Among those who have expressed doubt about a complete changeover from English to Swahili is TRAPPES-LOMAX (1985: 13), who did it more than 40 years ago. Furthermore, more than 30 years ago, RUBAGUMYA (1998: 88) wrote in his two-page article that changeover from English to Swahili "might not happen for another 5, 10, 20 years (?) depending on how policy makers are determined to delay it". We suppose the question mark meant much longer than 20 years. To be pragmatic we do not see any changeover happening in the near future, because of the symbolic value attached to the English language. As long as Tanzania's economy is tightly intertwined with the economies of the rest of the world, English is there to stay as a MoI. Our prediction is that, if this changeover is forced, what is likely to happen is that, privately owned schools and institutions which will still use English will have more clients than ever before. What is likely to happen if English is completely removed from the education system, schools in neighboring countries will have an assured market from Tanzanian students.

Finally, with the passing of time, we should not be worried about one language replacing the other. A monolingual approach to education in multilingual situations should be seen as a thing of the past, because it denies learners their desired goal of becoming global citizens. Let two or more languages be allowed to co-exist in the education system; this has the potential of further strengthening education in Tanzania. The focus should now be on proper teacher training and resource mobilization for educational institutions.

Acknowledgements

I am thankful to the anonymous reviewers and the editorial team for their constructive comments, which have enriched the current outlook of the article. I also wish to acknowledge my colleague, Tove Rosendal, and my research assistant Edna James whom I worked with on the Youth Language Attitudes in Songea Town, Tanzania. I also acknowledge the Tanzania Commission for Universities (TCU) for allowing me to use some findings from a study which they commissioned in 2016.



References

BARRETT, Judith (2014). "Why English is still the medium of education in Tanzanian secondary schools", *Language, Culture and Curriculum* (7) 1: 3-16.

BRISTOWE, Anthea, Marcelyn OOSTENDORP and Christine ANTHONISSEN (2014). Language and youth identity in a multilingual setting: A multimodal repertoire approach. *Southern African Linguistics and Applied Language Studies,* 2014, 32/2: 229–245

CAMERON, John and William DODD (1970). *Society, schools and progress in Tanzania*, Oxford: Perganon Press Limited.

CANAGARAJAH, Suresh (2011). "Translanguaging in the classroom: Emerging issues for research and pedagogy". *Applied Linguistics Review* 2: 1-27.

CHACHAGE, Seith L.C. (2004). *Kiswahili katika muktadha wa utandawazi. Katika Baraza la Kiswahili na utandawazi*, Dar es Salaam: BAKITA.

CRIPER, Clive and William A. DODD (1984). *Report on the Teaching of the English Language and it's Use as a Medium in Education,* Dar es Salaam: The British Council.

García, Ofelia (2017). "Translanguaging in schools: Subiendo y bajando, bajando y subiendo as afterword", *Journal of Language, Identity & Education* 16 (4): 256-263.

GARCÍA, Ofelia and Li WEI (2014). *Translanguaging: Language, Bilingualism and Education*, Basingstoke: Palgrave Macmillan.

GARDNER, Robert and Wallace LAMBERT (1972). *Attitudes and motivation in second-language learning*, Massachusetts: Newbury House Publishers.

HARRIES, Lyndon (1969). "Language policy in Tanzania", *Africa: Journal of the International African Institute* 39 (3): 275-280.

HUNGWE, Kedmon (2007). "Language policy in Zimbabwean education: historical antecedents and contemporary issues", *Compare* 37 (2): 135-149.

HILL, Charles P. (1980). "Some developments in language in Tanzania since 1969", in Edgar C. POLOMÉ and Charles P. HILL (eds.), *Language in Tanzania*, Oxford: Oxford University Press, 362-420.

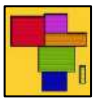
ISHUMI, Abel G.M. (2004). "The place of research in a myriad of competing knowledge bases: Selected cases of dilemmas in development arising from a relegation of research-based knowledge", *Papers in Education and Development* 24: 1-24.

JMT (1982). *Ripoti juu ya mapendekezo ya elimu nchini Tanzania*. (*Makweta's Presidential Commission's Repor*t), Dar-es-Salaam: Ministry of Education.

KADEGHE, Michael (2003). "In defence of continued use of English as the language of instruction in secondary and tertiary education in Tanzania", In Birgit BROCK-UTNE, Zubeida DESAI, and Martha QORRO (eds.), *Language of instruction in Tanzania and South Africa (LOITASA)*, Dar es Salaam: E&D Limited, 170-186.

KAPOLI, Irenius J. (1998). "Teacher control and language input in ESL writing: an exploration of classroom discourse and writing in secondary schools", *Journal of Linguistics and Language in Education* 4: 53-74.

KHELEF, Mohammed (2018). "African languages need to talk to each other", *Deutsche Welle*, accessed 6 July 2021, <https://www.dw.com/en/ngugi-wa-thiongo-african-languages-need-to-talk-to-each-other/a-44297656>



LANZA, Elizabeth (2007). "Multilingualism and the family", in Peter AUER and Li WEI (eds). *Handbook of multilingualism and multilingual communication*, Berlin: Mouton de Gruyter, 45-67.

LÜPKE, Friederike and Anne STORCH (2013). *Repertoires and choices in African languages*, Boston & Berlin: De Gruyter Mouton

MAKUMBA, Mwemezi (2010). "Is Tanzanians' lack of employability a challenge or a threat?", *Business Times Limited* (Dar es Saalam), 6 Aug. 2010.

MALEKELA, George A. (2003). "English as a medium of instruction in post-primary education in Tanzania. Is it a fair policy to the learners?", in Birgit BROCK-UTNE, Zubeida DESAI, and Martha QORRO (eds), *Language of instruction in Tanzania and South Africa (LOITASA)*, Dar es Salaam: E&D Limited, 102-112.

MALEKELA, George A. (2005). Secondary school students' views on the medium of instruction in Tanzania. In Birgit Brock-Utne, Zubeida Desai, and Martha Qorro (eds.), *LOITASA research in progress*, Dar es Salaam: KAD Associate, 125-138.

MAPUNDA, Gastor (2013). "Ngoni people's attitudes towards the use of Kingoni in beginner classes", *Journal of Linguistics and Language in Language in Education* 7 (1): 74-92.

MAPUNDA, Gastor and Safari T.A. MAFU (2014). "Corpus-based analysis of academic writing errors by first year Tanzanian university students: Cases from UDSM, SUA and TUICO", *Journal of Linguistics and Language in Education* 8 (2): 46-65.

MAPUNDA, Gastor (2015). "The changing roles of English in Tanzania: Past and present", in Martha MICHIEKA and Yousif ELHINDI (eds), *The Changing Roles of English in Eastern Africa*, Illinois: Common Ground Publishing LLC., 34-57.

MASELE, Bala F.Y.P. (2005). "Kiswahili au Kiingereza? Siasa na lugha muafaka ya kufundishia Tanzania", *Kioo cha Lugha* 3: 31-50.

MATRAS Yaron (2009). *Language contact*, Cambridge: Cambridge University Press.

MAZRUI, Alamin (1992). "Roots of Kiswahili: Colonialism and the dual heritage", *Ufahamu: A Journal of African Studies* 20 (3): 88-100.

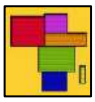
MEKACHA, Rugatiri D.K. (1994). "Language death: Conceptions and misconceptions", *Journal of Pragmatics* 21(1): 101-116.

MOCHIWA, Zacharia S.M. (2001). "Elimu au maigizo? Uhakiki wa sera ya lugha katika elimu Tanzania", *Kioo cha Lugha* 2: 41-58.

MSANJILA, Yohana P. (1990). "The role and status of ethnic languages in Tanzania", *Kiswahili* 57: 46-54.

MZEE, H. (1995). *Uhuru na Mzalendo katika kukuza Kiswahili. Katika Mlacha, S.A.K. (Mh). Kiswahili na Vyombo vya Habari*, Dar es Salaam: TUKI, 79-83.

QORRO, Martha A.S. (2004). "Popularising Kiswahili as the language of instruction through the media in Tanzania", in Birgit BROCK-UTNE, Zubeida DESAI, and Martha QORRO (eds.), *Researching the language of instruction in Tanzania and South Africa*, Cape Town: African Minds, 93-116.



PROBYN, Margie (2009). "Smuggling the vernacular into the classroom': conflicts and tensions in classroom codeswitching in township/rural schools in South Africa", *International Journal of Bilingual Education and Bilingualism* 12 (2): 123-136.

ROY-CAMPBELL, Zaline M. and Martha A. QORRO (1997). *Language crisis in Tanzania: The myth of English versus education*, Dar es Salaam: Mkuki na Nyota.

RUBAGUMYA, Casmir M. (1990). *Language in education in Africa: A Tanzanian perspective*, Clevedon: Multilingual Matters.

RUBAGUMYA, Casmir M. (1998). "Disconnecting education: A response to RM's rejoinder", *Journal of Linguistics and Language in Education* 4: 87-88.

RUBAGUMYA, Casmir M. (1999). "Choosing the language of instruction in post-colonial Africa: Lessons from Tanzania", *Papers in Education and Development* 20: 125-145.

SCHMIED, Josef J. (1986). "English in Tanzanian education", *Bayreuth African studies series 5: Language and education in Africa* 5: 63-114

SENKORO, Fikeni E.M.K. (2004). "Research and approaches to the medium of instruction in Tanzania: perspectives, directions and challenges", Birgit BROCK-UTNE, Zubeida DESAI, and Martha QORRO (eds.), *Researching the language of instruction in Tanzania and South Africa*, Cape Town: African Minds, 42-56.

TANU (1954). *The first conference of Tanganyika African National Union*, Dar es Salaam: Unpublished Typescript.

TRAPPES-LOMAX, Hugh R. (1985). "English language teaching in Tanzania: A colloquium". *Utafiti* 7 (11): 11-26.

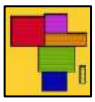
UNESCO (2021). "World Kiswahili Language Day" (UNESCO General Conference, 41st, 5 November 2021), accessed 8 March 2021,
<https://unesdoc.unesco.org/ark:/48223/pf0000379702>

VISCUSI, Margo (1971). *Literacy for working functional literacy in rural Tanzania*, Paris: UNESCO.

WEI, Li (2018). "Translanguaging as a practical theory of language", *Applied Linguistics* 39 (1): 9-30.

WEI, Li and Angel M.Y. LIN (2019). "Translanguaging classroom discourse: pushing limits, breaking boundaries", *Classroom Discourse* 10 (3-4): 209-215.

JOURNAL OF AFRICAN LANGUAGES AND LITERATURES
3/2022, 22-85



Modal auxiliary verb constructions in East African Bantu languages

RASMUS BERNANDER [1] / MAUD DEVOS [2] / HANNAH GIBSON [3]

[1] University of Helsinki
rasmus.bernander@helsinki.fi

[2] Royal Museum for Central Africa / Ghent University
maud.devos@africamuseum.be

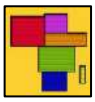
[3] University of Essex
h.gibson@essex.ac.uk

ABSTRACT

In this article we offer an overview of the use of modal auxiliary verb constructions in East African Bantu (encompassing languages spoken from eastern Congo in the north-west to northern Mozambique in the south-east; viz. Guthrie zones JD, JE, E, F, G, M, N and P). Modality, here conceptualized as a semantic space comprising different subcategories (or flavors) of possibility and necessity, has traditionally been a neglected category within Bantu linguistics, which has tended to focus instead on the more grammatical(ized) categories of tense, aspect and to a lesser extent mood. Nonetheless, our survey shows that there exists a rich number of different verbs with specialized modal functions in East African Bantu. Moreover, when comparing the variety of modal verbs in East African Bantu and the wider constructions in which they operate, many similar patterns arise. In some cases, different languages make use of cognate verbs for expressing similar modal concepts, in other cases divergent verbs, but with essentially the same source meaning(s), are employed. In addition, both Bantu-internal and Bantu-external contact have played a key role in the formation of several of the languages' inventories of modal verbs. A typologically significant feature recurrently discovered among the languages surveyed is the tendency of structural manipulations of the same verb base to indicate semantic shift from participant-internal to participant-imposed modal flavors.

KEY WORDS: East Africa, Bantu, modality, auxiliary, verb

DOI https://doi.org/10.6093/jalalit.v3i3.9153
ISSN 2723-9764



RASMUS BERNANDER / MAUD DEVOS / HANNAH GIBSON

Modal auxiliary verb constructions in East African Bantu languages

1. Introduction

This study surveys modal auxiliary verb constructions as they occur in the Bantu language varieties of East Africa (EA). Based on HEINE (1993) and ANDERSON (2006, 2011: 2), we define an auxiliary as "a verbal element on a diachronic form-function continuum standing between a fully lexical verb and a bound grammatical affix" and an auxiliary verb construction as mono-clausal complex constructions (minimally) consisting of an auxiliary verb operating on a lexical predicate verb. Thus, in this conceptualization any verb-verb construction consisting of an (erstwhile) verb which contribute some functional quality to a lexical verb is counted as an auxiliary verb construction, regardless of the functional or formal specificities of either the auxiliary verb itself or the construction as a whole. We are aware that this is a broader conceptualization than in many other (traditional) definitions of auxiliaries (cf. HEINE 1993: Ch. 1).[1] We also acknowledge the fact that many of the modal verbs discussed in this paper arguably are at the very initial stage of the verb-to-affix continuum, i.e. they exhibit few formal signs of grammaticalization (as will be further described below). We treat modality as a grammatical domain which fundamentally encodes expressions of possibility and necessity (see e.g. VAN DER AUWERA and PLUNGIAN 1998 and further discussion in §2). Consequently, a modal auxiliary verb construction is a linguistic composition consisting of a verb expressing either possibility or necessity. An example of a possibility auxiliary is **-wez-** 'can' in Swahili in (1), and an example of a necessity auxiliary is **-xoy-** 'ought' in Lusaamia in (2).[2]

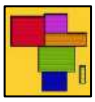
(1) Standard Swahili (G42d; ASHTON 1947: 276)

siku	**y-o**	**y-ote**	**wa-wez-a**	**ku-wa-on-a**	**p-o**	**p-ote**
9.day	9-of	9-all	SM2SG.PRS-able-FV	INF-OM2-see-FV	16-of	16-all

'Any day you can see them anywhere'

[1] For example, as noted by one of the reviewers, our conceptualization differs from the tradition in Southern Bantu studies of distinguishing a more narrow category of "auxiliaries" in contrast to so called "deficient verbs". Together with ANDERSON (2006: 11), we believe that such a systematization leads to a too restrictive categorization of auxiliaries.

[2] Cited languages are given with their "Guthrie code" as commonly used when referring to Bantu languages, based on the alpha-numeric coding system introduced by GUTHRIE (1948, 1967-1971) and later updated by MAHO (2003, 2009); see also HAMMARSTRÖM (2019). In terms of transcription, we have represented the examples as in the original versions, but we have taken the liberty to add or alter the glosses as to fit with conventionalized abbreviations within linguistic typology and comparative Bantu linguistics.



(2) Lusaamia (JE34; BOTNE et al. 2006: 60)
a-xóy-eré **a-mény-é** **anó**
SM1-ought_to-PRF SM1-live-SBJV here
'S/he must live here'

As ANDERSON (2006: 30) notes: "Indeed, virtually every non-nominal (person, number, class) category described as 'inflectional' [...] can or has been encoded through an (erstwhile or present) AVC [Auxiliary Verb Construction]." To be sure, the functional category of modality is often treated as one of the primary categories expressed with auxiliaries, not least due to the fact that they form a salient category in English and many other Indo-European languages (see e.g. ZIEGELER 2011). Nonetheless, modal auxiliaries have traditionally gone under the radar in Bantu studies (cf. NURSE and DEVOS 2019). Modal auxiliaries are not included in NURSE (2008), to date the most thorough cross-Bantu work on verbal constructions. ANDERSON's (2011) imposing investigation of auxiliary verb constructions in Africa (including a sample of 100 Bantu languages) only mentions a single Bantu modal verb. We can think of several possible reasons for this fact. Firstly, there is a general understanding that modality is mainly marked verb-internally in Bantu through the use of the subjunctive suffix **-e** (cf. NURSE and DEVOS 2019), which, in turn, is connected to the traditional focus on the already very intricate verb-internal Bantu morphology. When complex verbal constructions are addressed in work on Bantu languages, the description is often exclusively dedicated to copula compounds and/or markers of Tense-Aspect.[3] Modal verbs are also generally considered as semantically "more weighty", i.e. they pattern more with lexical verbs and typically show less signs of grammatical status than auxiliaries expressing e.g. tense and aspect (see e.g. KRUG 2011). This is a tendency which would also seem to be borne out in Bantu and which may have led to modal auxiliary verbs to be misleadingly classified as lexical verbs, with the result being that they have been excluded from grammars. The "lexicality" of modal verbs arguably also makes them more easily exposed to contact-induced change, a fact which may have caused modal verbs to be omitted from linguistic descriptions (e.g. in puristic attempts of editing out 'borrowed' terms; cf. MOUS 2019, BERNANDER, forthcoming).

[3] A poignant example is the grammars written by the "Hamburg School" of German colonial linguists whose grammatical descriptions tend to include a chapter on "Hilfsverba 'Sein und Haben'" while rarely addressing any other auxiliary verbs. In some cases, copula verbs inflected in the subjunctive in Bantu can have functions that can be argued to be modal, in line with the definition of modality by some scholars. However, such constructions do not fall within our delimitation of modality, as will be made further clear in §2).

However, since the article by DEVOS (2008a), which constitutes a systematic account of the modal system of Shangaci (P312), recent years have witnessed an increasing interest in modality in Bantu, including both language-specific studies and work addressing a smaller set of Bantu varieties. Still, no overarching comparative study of a larger group of Bantu languages has been executed. The purpose of this paper is to offer such a comprehensive overview of modal auxiliary verb constructions, with a focus on East African Bantu. It surveys the various parts of Bantu-speaking East Africa (see §3.1 for discussion on this delimitation) to produce an inventory of modal verbs. Based on a comparison of the collected data, we offer some initial conclusions on general functional and morphosyntactic traits of modal auxiliary verb constructions in Bantu languages and their overall systematization within the domain of modality. The study also sets out to trace source constructions and to comment on pathways of change, as well as any cases of areal spread of certain construction types, thus considering the role of both internal and external (i.e. contact-induced) change in the development of Bantu modal auxiliation.

The remainder of this paper is organized as followings. Section 2 provides a background to the concept of modality and its internal taxonomy of expressing subcategories of possibility and necessity. Section 3 surveys the use of auxiliary verb constructions for expressing modality in Bantu languages spoken across the East African area. Based on the comparative data presented in Section 3, Section 4 offers a generalized account of the organization and development of the modal domain. Section 5 represents a summary, offering some final conclusions while also pointing to possible fruitful avenues for further research.

2. On modality

2.1 Introduction and setting the scope

In this paper, we treat modality as a part of the wider domain of TAM (Tense-Aspect-Mood) categories (cf. NUYTS 2006, 2016). We chose to delimit this onomasiological grid as encompassing expressions of possibility (capability, ability, potentiality) and necessity (need and necessities). This delimitation is based on the generally held understanding of modality in functional-typological studies (e.g. BYBEE et al. 1994, VAN DER AUWERA and PLUNGIAN 1998, NUYTS 2006, NAUZE 2008) including those targeting Bantu varieties. This entails, for example, that we exclude so called "volitive" or "bouletic" modality from our conceptualization of modality. See e.g. VAN DER AUWERA and PLUNGIAN (1998) for the motivation of such a position. It is our feeling that volitional modality is often straightforwardly expressed with a volitional verb in East African Bantu, viz. a

verb meaning 'look for, want, need, love' or similar. In some languages (e.g. Manda N11; see BERNANDER 2017: 220-224, 292), one and the same verb *-lond-* expresses both 'want' and 'need' which has the result that this verb (in an analogous construction) is used as a marker of both volitional modality as well as participant-internal modality. In Shangaci (P312; DEVOS 2008a), *-sakh-* 'want' is used as a marker of both epistemic possibility and necessity. Thus, essentially, we do not include volitive modality, except when the same auxiliary is also used for expressions of possibility/necessity. Similarly, although we discuss the "input" to the modal domain, viz. the lexical source verbs, we do not offer any detailed discussion of the "output" or the further expansion of erstwhile modal auxiliaries into marking other "postmodal" functional categories, such as future tense, imperatives and various subordinate clauses (cf. VAN DER AUWERA and PLUNGIAN 1998).[4]

2.2 The taxonomy of the modal domain

The modal categories, or modal forces (Kratzer 1981, 1991) of possibility and necessity stand in a scalar relationship to each other, with expressions of possibility encompassing a weaker force vis-à-vis necessity (cf. Coates 1983, Verstrate 2005). The notions of possibility and necessity may be further divided into various subcategories, or flavors (Kratzer 1981, 1991), generating particular modal meanings. Authors who otherwise agree in delimiting the modal domain to necessity and possibility as paradigmatic variants, tend to differ on the details of their further subcategorization as regards semantic cut-off points and labeling. We have chosen to operate with the following subcategories and associated terms: participant-internal (PI), participant-external (PE), deontic and epistemic modality.

Although some scholars prefer to merge the subtype of PE modality with that of deontic modality, we are inclined to agree with NUYTS (2006: 35) that it make sense to keep deontic modality apart from other types of participant-external modality. This is because deontic modality together with epistemic modality convey the speaker's own viewpoint rather than describing the state of affairs more objectively. That is to say, they encompass more subjective/subjectified concepts, coding "attitudinal" notions of whether the state of affairs "exists or not, or is morally acceptable or not, or is 'agreeable' or not" (NUYTS 2015: 109). What

[4] We also exclude any discussion about the conceptually very closely related linguistic category of evidentiality but refer the readers to the recent overviews of BOTNE (2020) and CRANE et al. (forthcoming) on this category in Bantu languages. As implied in these works, evidentiality is rarely, if ever, co-expressed with modal auxiliary verbs.

this further implies is that the various modal subcategories are associated with a progression of increased grammaticality as well as subjectivity, viz. "the speaker's subjective belief state / attitude toward the proposition" (NUYTS 2006, 2015, 2016: 35). In this taxonomy, "dynamic" PI and PE possibility/necessity are in both synchronic and diachronic perspectives taken as representing less mature, more lexical-like "pre-core" modal categories while deontic and particularly epistemic possibility/necessity are treated as representing a more mature stage of "core modals" (cf. TRAUGOTT and DASHER 2002, MUNRO 2006).

The following examples from different East African Bantu languages are meant to illustrate the subtypes within the onomasiological grid of modality, constituted by possibility and necessity further divided into its various subcategories. In this way, we also have a chance to briefly present the semantic particularities of each type.

Participant-Internal modality refers to properties inherent to the first participant engaged in the State of Affairs (SoA); either abilities or capacities in expressions of Participant-Internal possibility (3) or needs or necessities in expressions of Participant-Internal necessity (4).

(3) Luguru (G35; NYINONDI and LUSEKELO 2020: 75)

mai	**ko-dah-a**	**ku-kal-a**	**ghoya**	**na**	**i-wa-ana**
mother	SM1.PRS-POSB-FV	INF-stay-FV	better	COM	AUG-2-child

ku-bit-a **Mwenda**
INF-surpass-FV Mwenda
'The mother can cope better with children than Mwenda'

(4) Shangaci (P312; DEVOS 2008a: 15)
ki-ná **y'oo-ráfúun-a**
SM1SG-NEC CONN9'INF-chew-FV
'I need to eat'

Participant-External modality indicates that there are circumstances (partly) beyond the control of the first participant which either enable the SoA in expressions of Participant-External possibility (5) or makes them necessary in expressions of Participant-External neccesity (6).

(5) Manda (N11; BERNANDER 2017: 279)
n-a-wonang-íni **na** **vá-ndu** **bahápa** **pa-Ludéwa** **gólo**
SM1SG-PST-meet-REC.PFV COM 2-people EMPH.LOC LOC16-Ludewa yesterday
ndáya **y-áke** **n-a-hotw-íli** **ku-kín-a** **m-píla**
reason 9-POSS3SG SM1SG-PST-POSB-PFV INF-play-FV 3-football
'I met with some people here in Ludewa yesterday (and) because of that I could play football'

(6) Rundi (JD62; MBERAMIHIGO 2014: 137)
ku-gum-a **u-´-ri** **mu-zima** **mu** **ki-he** **ki-óóse** **u-fit-´-ye**
INF-stay-FV SM2SG-CJ-be 1-alive LOC18 7-time 7-all SM2SG-have-REL-FV
i-N-bányi **u-tégerezw-a** **ku-nyó-a** **i-bi-rahure** **u-mu-naáni**
AUG9-9-pregnancy SM2SG-NEC-FV 15-drink-FV AUG8-8-glass AUG3-3-eight
bya **a-ma-əzi** **cánké** **bya** **i-mi-tóbe**
CONN8 AUG6-6-water or CONN8 AUG4-4-juice
'In order to stay in good health during your entire pregnancy, you should drink eight glasses of water or juice'

Deontic modality can be seen as an outlier of Participant-external modality which is subjectively construed by the speaker and where the external circumstances affecting the first participant are socially created, e.g. through personal authority (usually the speaker) or a societal norm. While Deontic possibility expresses permission (7), Deontic necessity expresses obligation (8).

(7) Lusoga (JE16; KAWALYA et al. 2019: 18)
é-N-sáwó **yo** **o-sóból-á** **ó-kú-gi-rek-á**
AUG9-9-bag 9.POSS2SG SM2SG.PRS-POSB-IPFV AUG15-INF-OM9-leave-FV
wa-nó **obá** **o-kú-gy-á** **ná-yo**
16-DEM or AUG15-INF-go-FV with-SBST9
'Your bag, you can either leave it here or go with it'

(8) Lusaamia (BOTNE et al. 2006: 62)
ómu-xasí **a-xóy-eré** **a-mány'-e** **nga** **a-koy-á**
1-woman SM1-NEC-PFV SM1-know-SBJV now SM1-brew-FV
áma-lwa **kééne** **kanó**
6-beer 6-kind 6.this
'A woman ought to know how to brew this particular kind of beer'

Epistemic modality, finally, refers to the speaker's subjective judgements or estimations of the relative certainty of the truth of the conditions encompassed in the state of affairs. As seen in (9), epistemic possibility refers to a relative high degree of uncertainty whereas epistemic necessity refers to a relatively high degree of certainty (10).

(9) Shangaci (P312; DEVOS 2008a: 15)
a-ń-sákh-a **o-fiy-á** **á-tthú** **e-énkeénye**
2-PRS-POSB-FV INF-arrive-FV 2-person 2-many
'A lot of people might arrive'

(10) Luganda (JE15; KAWALYA et al. 2019: 367)
obá **mugánda** **w-affe** **Napóólíyâni** **a-yogéddé**
if 1.brother 1-POSS1PL Napoleon SM1.PRS-talk.PFV
bw'âtyo **erá** **kí-téékw-á** **kú-bá** **ki-túufú**
like=1-that then SM7.PRS-NEC-IPFV INF-be 7-true
'If our brother Napoleon has talked like that, then it must be true'

2.3 Overlap in modal expressions

Often there is overlap in the reading of a modal expression between various flavors. An inferred reading can be conventionalized thus expanding and consolidating the semantic scope of a particular construction. This is exemplified by (11) from Nyamwezi.

(11) Nyamwezi (F22; JONSSON 1954: 60)
Enea **w-a-dul-a** **ku-buk-a** **haha**
Enea SM2SG-?-POSB-FV INF-rise-FV now
'Eneas kan stiga upp nu (Eneas can/might get up now)'

This example is ambiguous. In the absence of further context and with its original Swedish translation, where the possibility auxiliary 'can' may be used for all types of flavors of possibility unlike English where the epistemic use tends to be restricted to questions and negative contexts (see e.g. VAN DER AUWERA and PLUNGIAN 1998, NARROG 2005) the utterance can be associated with more or less any subtype of possibility, depending on how one wishes to interpret the proposition. That is, as participant-internal or -external possibility if the sentence is taken to denote that Eneas is able to rise now due to some internal or circumstantial reason, deontic possibility if the speaker uttering the sentence or some other normative force gives him permission to rise but also epistemic possibility if the speaker wishes to convey that there is a chance that Eneas will get up.

Just as a modal verb can expand between various flavors and from expressing pre- to core modal notions, it may also expand between the two modal forces, particularly within the deontic subdomain (cf. VAN DER AUWERA and PLUNGIAN 1998). Similarly, a modal source does not necessarily have to start out marking PI modality but can be recruited directly into any of the subcategories of

modality, e.g. like English 'be supposed (to)' which went directly into the modal system as a marker of PE.

With this in mind, a caveat is also needed. Namely, with a few notable examples, our surveyed sources on East African Bantu languages seldom offer any comprehensive overviews of the modal system in the given language, nor any detailed information on which exact subcategories of modality a specific auxiliary construction covers or does not cover (or offer enough contextualized examples for us to judge for ourselves). Our findings are necessarily therefore determined by the descriptive status of the languages under examination. As such there is a risk that some of the data presented and analyzed in the current work is inadequate or non-comprehensive.

3. Surveying Bantu modal auxiliary verb constructions across Eastern Africa

3.1 Introduction

In this section we offer a survey of auxiliary verbs expressing modality found in East African Bantu languages. The data comes both from general grammatical descriptions and articles specifically dedicated to modal expressions in East African Bantu. The latter are far fewer in number but much more detailed in their description of the modal system of a given language variety. It should be mentioned at the outset that with some notable exceptions, the sources surveyed seldom provide any complete descriptions of the inventory of modal verbs in a given language nor any detailed accounts of the various modal flavors any given auxiliary may express. Plenty more linguistic descriptions than those cited in this work have been perused in the hunt for modal auxiliaries, but where no information on the matter could be obtained.

The survey is geographically based in its scope and delimited to East Africa, which in this study is conceptualized as covering an area which reaches from Uganda and the Democratic Republic of the Congo (DRC) in the northwest to the coast of the Indian Ocean and further southward to Northern Mozambique, Malawi and Zambia. The languages considered are undeniably all also closely genetically related Eastern Bantu languages (cf. GROLLEMUND et al. 2015). By and large, the pool of languages (and subgroupings) equals those discussed in the genealogically oriented work of NURSE (1999) or EHRET's (1998, 1999) grouping of "Mashariki-Kaskazi" under the "Eastern-Savannah subgroup" of "Savannah Bantu". Since most genealogical (sub-)groupings also are geographically coherent (and indeed often carry a geographically based designation) we have

taken the liberty of using terminology perhaps mainly associated with genealogically-based studies when helpful.

3.2 Great Lakes

We start with the Great Lakes (zone J) languages in the north-western part of our designated area (cf. SCHOENBRUN 1990, 1994, NURSE 1999, BASTIN 2003), where the most substantial set of well-studied research on modal auxiliary verb constructions to date is to be found. Consequently, this section will also by far constitute the most elaborate and longest section. For ease of reference, we will further subdivide this section using the genealogical subgroups for Great Lakes established by SCHOENBRUN (1994, 1997). (Recall, however, that overall our study is geographically rather than genealogically organized).

3.2.1 Western Lakes

The interest in detailed research on modality in the Great Lakes Bantu was instigated by the diachronically detailed, corpus-driven investigations of modality markers in Rundi, a member of the Western Lakes subgroup of Great Lakes (JD62; BOSTOEN et al. 2012, MBERAMIHIGO 2014). In BOSTOEN et al. (2012), which focuses on auxiliary verbs expressing possibility, it is shown that **-shóbor-** is the most prominent modal possibility verb in Rundi, since it both transverses all subcategories of possibility and is most frequently attested in the corpus (see also Mberamihigo 2014: 80).[5]

Still, the verb **-shóbor-** is mostly confined to pre-core modal or "dynamic" senses, as in the participant-external construction in (12). Although it may also be used for core modal flavors as the deontic construction in (13) shows.

(12) Rundi (JD62; Mberamihigo et al. 2012: 16)

n-ra-shóbor-a	**ku-siinziir-a**	**ha-ri-hó**	**i-ki-tănda**
SM1SG-PRS.DISJ-POSB-FV	INF-sleep-FV	SM16-be-LOC16	AUG7-7-bed

'I can sleep (because) there is a bed'

[5] We exclude here from further discussion the verb **-shóbok-** which is an anti-causative or intrasitivizing derivation of **-shóbor-**. See BOSTOEN et al. (2012) for a detailed account and the further discussion in § 4.1.4.

(13) Rundi (JD62; MBERAMIHIGO et al. 2012: 16)
a-ba-ntu **ba-ó** **mu** **mi-ryango** **i-óóse**
AUG2-2-person 2-CONN LOC18 4-clan 4-all
ba-á-ra-shóbor-a **ku-bandw-a**
SM2-REM-DISJ-POSB-FV INF-participate.in.the.Kiranga.cult-FV
'People of all clans were allowed to participate in the Kiranga cult'

The verb **-shóbor-** may also be used for epistemic possibility in which case it has to occur in a more complex construction which includes the copula verb **-b-** in the infinitive and the semantic main verb in a type of dependent verb form, the so called "*conjunctif*". This verb form is characterized by an initial high tone and is used as a special verb form of unmarked subordinate clauses but also after auxiliaries, as in this case (see MEEUSSEN 1967: 113, NURSE 2008: 309; see also BOSTOEN et al. 2012: 17 for Rundi-specific references).

(14) Rundi (JD62; MBERAMIHIGO et al. 2012: 17)
u-u-jîsh-a **i-bi-ziriko** **a-shóbor-a** **ku-bá**
REL1-SM1-plait-FV AUG8-8-rope SM1-POSB-FV INF-be
´-a-a-ra-bi-i-têr-ir-ye
CJ-SM1-REM-DJ-OM8-REFL-plant-APPL-PFV
'Someone who plaits ropes may have planted them (for) himself'

The verb **-shóbor-** can be traced to a verb stem ***-còbʊd-** 'be able' reconstructed for Great Lakes Bantu (BASTIN et al. 2002). Cognates to Rundi **-shóbor-** also appear in several other closely related West Lakes languages like Ha (JD66 HARJULA 2004: 148), Fuliiro (JD63; VAN OTTERLOO 2011) and Vinza (JD67; Ko 2014), as well as other Great Lakes Bantu languages like Luganda, the latter two languages are discussed later in this section.

In addition to **-shóbor-**, Rundi also makes use of another possibility verb, namely **-bâsh-** with the source meaning 'be active', used in particular for marking pre-core modal participant-internal and -external modality (BOSTOEN et al. 2012: 16, MBERAMIHIGO 2014: 100-106). MBERAMIHIGO (2014: ch.3) investigates an additional set of other more peripheral possibility auxiliaries and in MBERAMIHIGO (2014: ch.4) several necessity auxiliaries are described. The most prominent necessity verb in terms of frequency is **-tégerezw-** 'have to, be obliged to'. This verb is particularly used for the expression of deontic necessity (but has been attested for the other subcategories too). The verb can be traced to the source verb **-tég-**, a reflex of Proto-Bantu *** -tég-** '(set) trap' with similar lexical meanings in Rundi (MBERAMIHIGO 2014: 122) and further derived with the suffix

phraseme **-erez-** (consisting of a non-compositional sequence of two applicatives and a causative) and the common Bantu passive extension.[6]

Both Fuliiro (JD63) and Shi (JD53) – as described by VAN OTTERLOO (2011: 265-266) and POLAK-BYNON (1975), respectively – make use of a possibility verb **-hash-**, presumably related to **-bâsh-** previously discussed. As mentioned, Fuliiro also makes use of **-shóbol-**. In addition, the verbs **-lóng-** 'get' and **-zìgir-** 'be able' are employed as markers of PI possibility.

Another description of a West Lake language which includes a (brief) discussion of modal verbs is the aforementioned study on Vinza (JD67). KO (2014) shows that Vinza too makes use of a cognate of the Rundi **-shóbor-**, **-yoboor-** which seems to be to the most prominent (if only?) possibility auxiliary in the language. Furthermore, Ko (2014) describes a spectacular combination of the verbs **-aang-** (< 'refuse') and **-kuund-** (< 'like') - used to express deontic necessity 'must' in the language, as in (15). Since the semantic main verb is inflected in the (finite) subjunctive verb form and not in the infinitive, however, it is difficult to be sure that we are really dealing with a mono-clausal (double) auxiliary construction in this case (although the translation would imply that this is indeed the case).

(15) Vinza (JD67; KO 2014: 97-98)

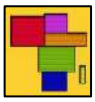
ya-aang-ye	**a-kuund-ye**	**a-garuk-e**	**i-muhira**	**ntyanyene**
SM1-refuse-PRF	SM1-like-PRF	SM1-return-SBJV	23-home	now

'He/She must return home now (lit. he refuses or likes, he returns home now)'

3.2.2 West Nyanza

The possibility verbs discussed for Rundi also show up in the West Nyanza languages (JE10, JE21-24), mostly spoken in Uganda and across the border into Tanzania. KAWALYA et al. (2014) is also a corpus-driven investigation on **-sóból-** in Luganda, a reflex of ***-còbud-** and cognate to the Rundi (and other Western Lakes) forms above. However, unlike its Rundi cognate, **-sóból-** in Luganda is not associated with epistemic possibility. In KAWALYA et al. (2021), it is furthermore shown that **-sóból-** stands in competition with the verb **-yînz-** (from a lexical source 'be powerful', 'overcome', 'manage', 'control', claimed by the authors to be obsolete as a lexical verb in present-day Luganda). Unlike **-sóból-, -yînz**- covers the entire range of possibility flavors, including the expression of the subjective core modal category of epistemic possibility.

[6] For the use of the term 'suffix phraseme' see BOSTOEN and GUÉROIS (forthcoming).



KAWALYA et al. (2019), in turn, investigates the modal auxiliary verbs **-téekw-** and **-lina** that are used (alongside other markers) in coding the necessity domain in Luganda. The modal verb **-téekw-** derived from a passivized form of the lexical verb **-téek-** 'make a law, bind by law, edict'. Typically combined with an infinitive predicate verb, **-téekw-** is employed to express different subtypes of necessity, although it is primarily used with a deontic flavor, as in (16). In addition, it may not be employed for expressing PI necessity.

(16) Luganda (JE16; KAWALYA et al. 2019: 365)
ssemwandu **a-teekw-a** **o-ku-kungubag-a** **o-ku-mal-a**
1.widower SM1.PRS-NEC-IPFV AUG15-INF-mourn-FV AUG15-INF-complete-FV
o-mu-aka **mu-lamba**
AUG3-3-year 3-full
'(And it is our custom,) a widower must mourn for a full year'

The possessive 'have'-verb **-lina**, stems from the lexicalization (formal univerbation and functional re-analysis) of what was originally a copula verb **-li** and the comitative **na**. As a modal auxiliary verb, **-lina** too is mainly employed as marking deontic necessity as seen in (17).

(17) Luganda (JE16; KAWALYA et al. 2019: 372)
o-lina **o-ku-bu-tuukiriz-a**
SM2SG.PRS-NEC AUG15-INF-OM14-fulfil-FV
'(Walukagga, that responsibility of making Bemba think that he is winning the battle, while he is being defeated,) you have to fulfil it…'

The modal system described for closely related Ruruuli-Lunyala (JE103) in NAMYALO et al. (2021) appears to be identical to that of Luganda.

KAWALYA et al. (2018), in turn, broaden the scope of research to include the expression of possibility in all of Luganda's closest relatives of the West Nyanza branch of Great Lakes. As seen in Table 1, a simplified version of Table 6 in KAWALYA et al. (2018: 21), a relatively scattered picture appears when comparing these otherwise tightly interconnected languages. Thus, no modal verb is shared by all member languages of the West Nyanza branch. Some closely affiliated varieties to Luganda share its use of **-sóból** and **-yînz-** as possibility verbs, with roughly the same functional range. However, in the other languages **-yînz-** does not exist as a modal verb whereas **-sóból-**, where it does exist, differs in its semantic range (with more flavors of possibility covered in Rutooro, and less flavors covered in Haya and Nyamba compared to Luganda).

LANGUAGE / AUXILIARY	**-sóból-**	**-yînz-**	**-báas-**	**-kuhích-**	**(-ezy-)**[7]
Luganda (JE15)	YES	YES	NO	NO	NO
Lusoga (JE16)	YES	YES	NO	NO	NO
Lugwere (JE17)	YES	YES	NO	NO	YES
Runyoro-Rutooro (JE11)	YES	NO	NO	NO	NO
Runyankore-Rukiga (JE13/14)	NO	NO	YES	NO	NO
Haya (JE22)	YES	NO	YES	NO	NO
Nyambo (JE21)	(YES)	NO	YES	NO	NO
Kerewe (JE24)	NO	NO	NO	YES	NO
Sinza (JE23)	NO	NO	NO	YES	NO

Table 1 – Possibility auxiliaries in West Nyanza Bantu languages (KAWALYA et al. 2018: 21)

Runyankore, Haya and Nyambo make use of a modal verb **-báas-**, cognate to Rundi **-bâsh-** presented above (and just as **-sobol-** a reflex of a regional reconstruction ***-báac-**; BASTIN et al. 2002). In Runyankore (JE13), **-baas-** is the only possibility auxiliary employed, a fact also pointed out by NYINONDI and LUSEKELO (2020). Example (18) illustrates the use of **-baas-** as a deontic marker in the language.

(18) Runyankore (JE13; NYINONDI and LUSEKELO 2020: 69)
ka-ku-mar-a **ku-kór-a** **e-mi-rimo=ye** **n-a-báas- a**
if-INF-finish-FV INF-do-FV AUG-4-work=POSS3SG PROG-SM3SG-POSB-FV
ku-zan-a
INF-play-FV
'Once he is done with his work, he may play (on the street)'

NYINONDI and LUSEKELO (2020: 72-74) also report the use of the necessity auxiliary **-téekw-** 'must/ought to', which they link to a passivized variant of an original root **-téek-** 'put', but which most likely is cognate with the Luganda form previously discussed.

Those JE20-varieties spoken at the eastern shores of Lake Victoria which are still considered to form a part of the West Nyanza subgroup, viz. Kerewe (JE24) and Zinza (JE23), make exclusive use of a completely different possibility verb, namely **-kuhich-**, which appears to be of unknown origin. These languages are

[7] KAWALYA et al. (2018) does not provide any etymology for **-ezy-**, only used as a PI possibility auxiliary in Lugwere. The authors furthermore note that "the verb **-ezy-** seems to be highly threatened by **-sobol-** in this language. The former is only provided as a synonym of the latter, but without clear contexts of usage" (ibid: 21). Maybe it is cognate with Ishenyi (JE45) **-esh-** (cf. §3.2.4).

known to be in heavy contact with Suguti languages like Jita (JE25; cf. §3.2.4 below). However, we could not find any trace of this verb in Jita either.

3.2.3 Greater Luhya

There are also a substantial set of studies on modality from the eastward parts of Great Lakes Bantu, where we can start with zooming in to the Luhya branch (the JE30 languages + JE41). GLUCKMAN et al. (2017) have described the modal system, including several auxiliary verbs, in various Luhya varieties as summarized in Table 2 below. In working out the semantic differences between modal expressions in Luhya these authors have operated with an extra layer of scalar relationship of "weak" versus "strong" necessity. Their conceptualization of weak necessity relates to expressions such as English 'should' and refers to instances where "the source of the obligation or requirement is relative to something "less strict" than in instances of strong necessity modality […]" (GLUCKMAN and BOWLER 2020: 224). In addition to this study, the same authors (GLUCKMAN and BOWLER 2020) offer a careful semantic investigation of the modal system in JE41 alone, whereas BOTNE et al. (2006: 60-62) address modal verbs in Lusaamia (JE34). Note that BOTNE et al.'s (2006) description of Lusaamia differs from that in Table 2 in that **-khoy-** (or **-xoy-** in their spelling) is formed without an additional applicative.

MODAL USE	possibility	necessity (weak)	necessity (strong)
NON-MODAL USE	'manage'	'want'	'arrive/reach'
Llogoori (JE41)	**-nyal-**	**-eny-**	**-duk-**
Lubukusu (JE31c)	**-nyal-**	**-eny-**	**-enyekh-/-khoy-**
Lunyore (JE33)	**-nyal-**	**-eny-**	**-ol-**
Lusaamia (JE34)	**-nyal-**	**-eny-/-dakh-**	**-khoyer-**
Lutiriki (JE413)	**-nyal-**	**-eny-**	**-tukh-**
Luwanga (JE32a)	**-nyal-**	**-eny-**	**-la-**

Table 2 – Modal auxiliary verbs within the Luhya cluster (GLUCKMAN et al. 2017: 4)

As seen in Table 2, all Luhya varieties make use of the cognate verb **-nyal-** for expressing possibility. As further discussed in §4.1.4, this verb is borrowed from Luo. It is used to cover all types of flavors within the modal force of possibility, from participant-internal to epistemic possibility. Example (19) from Lunyore is an illustration of the latter function.

(19) Lunyore (JE33; GLUCKMAN and BOWLER 2017: 6)
professor **a-nyal-a** **okhu-its-a** **ng'ondi**
1.professor SM1-POSB-FV INF-come-FV today
'The professor might come today'

Expressions of weak necessity are also formed with reflexes of the same verb **-eny-** (cognate with the necessity verb which occurs in some JE40 varieties discussed below).[8] This weak necessity verb is reportedly rarely used with epistemic flavors (with the exception of Lubukusu). The verb, derived from a lexical source meaning 'want', occurs in its original active form for the expression of PI necessity alone. For expressions of PE and deontic necessity, the verb is additionally derived with a subject-demoting extension, in this case with the anti-causative and intransitivizing neuter suffix, as in (20) from Logoori. The predicate verb, in this case **-sav-** 'ask', appears in the subjunctive.

(20) Logoori (JE41; GLUCKMAN and BOWLER 2020: 229)
Sira **y-enyek-a** **a-sav-e** **amu-aavo**
1.Sira SM1-NEC-FV SM1-ask-SBJV 1-brother
'Sira should ask his brother'

The auxiliary verbs expressing "stronger" necessity in Luhya are all derived from lexical verbs meaning 'arrive/reach'. Thus, although these modal verbs are not cognates, they (more or less) seem to share the same source semantics. The 'arrive'-verbs are used to express all types of flavors of necessity, e.g. PE and epistemic necessity, as in (21). As indicated in these examples, however, they occur in different constructions. In (21a), the predicative verb is inflected in the subjunctive whereas in (21b) it is in the infinitive. Furthermore, in (21a), the verb is additionally derived with an applicative extension, which is not the case in the epistemic example in (21b).

(21) Lubukusu (JE31c; GLUCKMAN et al. 2017: 8)
a. **o-khoy-el-a** **o-bukul-ɛ** **endeke** **yi-no**
SM2SG-NEC-APPL-SBJV SM2SG-take-SBJV 9.flight 9-DEM
'You have to take this flight (to get to Kisumu)'

[8] The only exception in the table is Lusaamia for which the authors claim that **-dakh-**, also meaning 'want', can be used in free variation with **-eny-**. This verb is not discussed by BOTNE et al. (2006) as a modal verb, however (although it is attested in their appended vocabulary as marking 'want (great need or desire)' (ibid: 131, 171). Can it be a Swahili loan from **-tak-** further integrated in the language through the productive process of Dahl's Law (where voiceless stops are voiced when the succeeding syllable also consists of a voiceless consonant; cf. DAVY and NURSE 1982)? Standard Swahili **-tak-** is a fortified variant of ***-càk-** (NURSE and HINNEBUSCH 1993: 581), a type of sound law not characteristic of Luhya as far as we know.

b. **e-fula** **a-khoy-a** **khu-b-a** **nekwa**
9-rain SM9-NEC-FV INF-be-FV fall
'It must be raining'

Note finally for Luhya, that the verbs used for expression of necessity are more often not part of an auxiliary construction but in different types of impersonal constructions, either inflected with a dummy morph from a certain class (22)[9] and further derived with valency-decreasing suffixes, namely associative suffix **-an-**, or with the verb in the infinitive (23).[10]

(22) Llogori (JE41; GLUCKMAN et al. 2017: 14)
ga-duk-an-a **ndee** **u-zi-ɛ** **m-skolu**
SM6-NEC-ASSOC-FV that SM2SG-go-SBJV 18-9.school
'You must go to school'

(23) Llogori (JE41; GLUCKMAN et al. 2017: 16)
ku-duk-a **Sira** **a-z-ɛ** **Nairobi**
INF-NEC-FV 1.Sira SM1-go-SBJV Nairobi
'Sira must go to Nairobi'

Although the infinitival use seems to be uncommon, the use of one and the same verb in both straightforward mono-clausal modal auxiliary constructions and in constructions with expletive subject markers is attested in many other eastern Bantu modal systems too. This fact will become clearer as we move along the other parts of Bantu-speaking East Africa. In §4.1.4 we return with a general discussion of such constructions and how some of them pose a problem in formal terms in relation to our delimitation of auxiliary verb constructions.

3.2.4 East Nyanza

Finally, for East Nyanza, Aunio et al. (2022) describes the modal system in some of the Mara languages, specifically those of the Western Serengeti branch of South Mara, i.e. Ikoma, Ishenyi, Nata (all three categorized as JE45) and Ngoreme (JE401). As for many of the other Great Lakes Bantu languages discussed above, these Mara languages also have one prominent possibility auxiliary employed

[9] Interestingly, Luhya varieties seem to employ expletives from different noun classes for different functions, even in terms of modal flavor (see §4.1.4).

[10] It may need to be stressed that the presentation of the data and discussion in GLUCKMAN and BOWLER (2020) seem to suggest that these verbs indeed stand in the infinitive (/deverbal noun) form and thus with the nominal class prefix of noun class 15, and not the subject marker of the same class (or the homomorphic locative class 17) as an expletive.

throughout the whole spectrum of this modal force from PI (24) to epistemic possibility (25).[11]

(24) Ikoma (JE45; AUNIO et al. 2022)
omo-kúngu **n-a-agho-tór-a** **ku-ghĕgh-a** **amá-nche** **má-áru**
1-mother FOC-SM1-PRS-POSB-FV INF-carry-FV 6-water 6-plenty
'Mother is able to carry a lot of water'

(25) Ngoreme (JE401; AUNIO et al. 2022)
umu-kufu **go** **Neema** **gu-ra-tor-a** **ku-bh-a** **go-sir-re**
3-necklace CONN3 Neema 3-IPFV-POSB-FV INF-be(come)-FV 3-be_lost-PFV
'Neema's necklace might be lost'

The verb in question is **-tor-** (or occasionally **-tur-** due to processes of vowel harmony within the verbal word). All reflexes of the verb are shown to ultimately go back to a reflex of ***-túd-** 'pierce' (cf. BASTIN et al. 2002),[12] a curious etymology further discussed in §4.1.1.

Along with **-tor-**, Ishenyi also makes use of the verb **-esh-** 'become fitting', as illustrated in (26).

(26) Ishenyi (JE45; Aunio et al. 2022)
omó-óna **n-eesh-íre** **a-sháág-e**
1-child FOC-SM1.POSB-PFV SM1-swim-SBJV
'The child might be swimming'

As implied in the gloss, **-esh-** is a change-of-state verb and thus inflected in the perfective in present temporal contexts.[13] In addition, the second verb appears in the subjunctive. This verb is only used for marking core modal flavors, i.e. the subject-oriented sub-categories of deontic and in particular epistemic possibility, as in (26) above. The source 'be(come) fitting' is a common source meaning for

[11] The pre-initial prefix **N-** found in the Ikoma-example stems from the common focus marker ***ní-** which has petrified on some verb forms in some of the Mara-varieties. Tone is not indicated on the Ngoreme-examples as recordings for this particular set of data were lacking.

[12] Nurse and ROTTLAND (1991: 201), EHRET (1998: 316) suggest that what appears to be a cognate verb, **-tol-** 'bore' in Sonjo (E46) and in other E50 languages, is borrowed from Cushitic. In light of the spread of ***túd** 'pierce' (3 PB Regions (NW Ce NE) and 6 zones B D E J M N (BASTIN et al. 2002)), we would be more inclined to believe the Mara modal verb to be an inherited verb (although attestations of Cushitic influence on Mara exists, see e.g. SHETLER (2003: 11-14, 288)).

[13] In Bantu studies, the term change-of-state verbs are commonly referred to "those verbs whose aspectual potential includes a transition from one state to another on the part of the subject" (CRANE and PERSOHN 2019: 311), viz. 'become X'. Such verbs gain a (present) resultative reading when inflected with the perfective, as is also the case here.

modal auxiliaries cross-linguistically and in Bantu. However, it is more often associated with necessity (e.g. BYBEE et al. 1994 'be fitting' associated with obligation and thus necessity) than possibility as in this case. Based on this typological pattern one might be tempted to infer that the **-esh-** has expanded from an erstwhile necessity marker to a possibility marker (in line with the description in §2.3 above). In the absence of any concrete evidence, however, we have to refrain from drawing any such conlusions.

For expressing necessity, the JE45 language varieties use **-end-** 'look for, want, need, wish, desire'. As seen in (27), Ngoreme makes use of a different non-cognate source verb **-tun-**, which, however, has a partly parallel etymology, namely that of 'need'. As also described for many of the languages above, in their unaltered active forms **-end-** and **-tun-** are used to indicate the flavor of PI nessecity. For PE and deontic necessity, **-end-** and **-tun-** is further derived with a participant-demoting passive extension.

(27) Ngoreme (JE401; AUNIO et al. 2022)
u-ra-tunu **gw-eghi** **ebh-into,**
SM2SG-IPFV-NEC(<want-PASS) INF-wash 8-utensils
gasi **tare** **go** **che-nguru**
but NEG.COP CONN17 10-force
'You should (**must) do the dishes, but you are not obliged to'

Both **-end-** and **-tun-** + PASSIVE VOICE are only used to express weak obligation, similar to JE30 **-eny-** above. This is visible in (27) above, where a more forceful deontic reading of **-tun-**, in this case, is cancelled by the following adversative clause.

To express a more forceful, strong obligation, the verb **-hatek-** can be used in Ikoma and Nata (at least), either derived in the passive form (28) in line with the common pattern of non-PI necessity verbs. Alternatively, the affected participant is demoted to an object and marked as such on the modal verb (29).

(28) Ikoma (JE45; AUNIO et al. 2022)
n-o-oku-háteku **u-mar-é** **a-shúúre**
FOC-SM2SG-PRS-NEC SM2SG-finish-SBJV 9-school
o-keeré **ghu-tĕ́m-a** **omo-bhííra**
SM2SG-before INF-hit-FV 3-football
'You must finish school before playing football'

(29) Ikoma (JE45; AUNIO et al. 2022)
n-e-egho-kú-hatek-a **u-mar-é** **a-shúúre**
FOC-9-PRS-OM2SG-NEC-FV SM2SG-finish-SBJV 9-school

o-keeré **ghu-tḗm-a** **omo-bhííra**
SM2SG-before INF-hit-FV 3-football
'You must finish school before playing football'

The verb can be suggested to be linked with ***-pát-** 'hold' > ***pát-ık-** (J,L) 'wedge in', so in the passive form 'be wedged in' (cf. BASTIN et al. 2002).[14] Notice, that such a meaning is close to that of 'be made to fit in' ~ 'be fitting' and thus that the semasiological background of **-hatek-** could be argued to parallel other necessity source verbs related to meanings discussed for G42, N11 and P22 below.

Other peripheral modal verbs (i.e. verbs with fewer attestations and with a less extensive semantic range not including core modality) in the Mara languages, include PI possibility markers **-many-** from 'know' in Ikoma and **-nagy-** in Ngoreme (probably with the etymology 'have strength, power' + the causative extension, also present in other closely related languages such as Zanaki [JE44]).

3.3 Central Kenya

Moving further eastwards to Central Kenya (Guthrie's E50), far less detailed information on modal auxiliary verb constructions is available, both with regard to their specific semantics and the range of flavors they cover within the modal system, as well as to their semasiological background.

For Kamba (E55), LINDBLOM (1926: 59) describes two auxiliaries, which, given the examples provided, at least seem to express PI possibility. These are **-tonĩ̯-** 'be able (to)' (see also **-tony-** 'be able' in WHITELEY and MULI (1962: 123)), as in (30), and **-manĩ̯-** know' (see P22, JE40 and further discussion in §4.1.1) as in (31).

(30) Kamba (E55; LINDBLOM 1926: 59)
n̥dı-tonĩ̯-a **u-sæmb-a** **mĩtuk̲ı**
SM1SG-POSB-FV INF-run-FV quickly
'I am not able to run quickly'

(31) **w-ĩ-manĩ̯-a** **ku-u-a** **lĩ-ú**
SM2SG-?-POSB-FV INF-cook-FV 5.food
'Can you cook food?'

Other auxiliary verbs with modal meanings in Kamba are **-is-** 'get to, happen to, likely to, may', **-ti-** 'be inevitable, must' and **-ail-** 'be fitting, ought' as in (32).

[14] In this case the suffix ***-ık-** refers to the transitive impositive verbal extension and not the homomorphic intransitivizing neuter (cf. SCHADEBERG and BOSTOEN 2019).

(32) Kamba (E55; FARNSWORTH 1957:36)
ni-n-ail-e **ku-neen-a**
FOC-SM1SG-NEC-PFV INF-speak-FV
'I should speak'

Gikuyu (E51), in turn, makes use of the auxiliary verb **-hɔt-** for marking possibility, according to Kihara (2017); see also BARLOW (1961[1951]: 184-185). The translation seems to suggest that **-hɔt-** covers different subcategories of possibility, as also implied (albeit not explicitly stated) by KIHARA (2017: 79-80) in his description of this verb.

(33) Gikuyu (E51; KIHARA 2017: 80)
Maina **a-hɔt-a** **ko-nin-a** **ma-rigo** **ikomi**
Maina 1-POSB-FV INF-finish-FV 6-bananas ten
'Maina may/might finish ten bananas'

To express necessity, Gikuyu uses the verb **-agerer-** 'should/ought' (Kihara 2017), illustrated in (34). It is given as ~ **-agïrïr-** and claimed to have the lexical meaning 'behoove one, be becoming to one' by BARLOW (1951: 185), as in (35) (the "literal" translation here is that of the author). As further seen in the examples below, the authors do not only differ in vowel height but also in which constructional frame the verb occurs in. Whereas KIHARA's (2017) example consists of the active form of the verb + subjunctive, the example provided by BARLOW (1951) seems to consist of the passivized version of the modal verb plus an additional comitative marker.

(34) Gikuyu (E51; KIHARA 2017: 160)
reu **w-a-agere-irɛ** **w-a-nji-ɛ** **go-cari-a** **wera**
now SM2SG-PRS-NEC-PFV SM2SG-PRS-start-FV INF-look-FV 14.work
'Now you should start looking for work'

(35) Gikuyu (E51; BARLOW 1951: 185)
w-a-giriiruo **no** **ku-mu-rih-a**
SM2SG-PRS-NEC ?COM INF-OM1-pay-FV
'You ought to pay him (lit. It behoves you only to pay him)'

Tharaka (E54; LINDBLOM 1914: 28, 47) has a verb **-ut̲-** 'can, be able to', which based on differences in vowel quality does not seem to be related to **-hɔt-** in E51.

3.4 Kilimanjaro-Taita

Moving further south to the Taita Hills and the slopes of Mount Kilimanjaro around the border between Kenya and Tanzania, a salient modal verb attested in several varieties (and also nearby; see §3.5) is the possibility verb **-(i)dim-**,

originally borrowed from Maasai. Section 4.2 offers a further discussion about this etymology. Example (36) comes from Dawida.

(36) Dawida (E74a; SAKAMOTO 2005: 11)
n-a-dim-a **ku-rumaghi-a**
SM1SG-PRS-POSB-FV INF-bear-FV
'I am able to bear (it)'

SAKAMOTO (2005) also notes a (weak?) necessity verb 'ought', expressed by **-fwan-** 'to behove, to concern' (also 'become one'). The verb can be linked to the Proto-Bantu form ***-pú-an-** 'be fitting, resemble each other' (BASTIN et al. 2002). This verb occurs in a construction where the first participant is morpho-syntactically treated as the object and consequently indexed on the verb in the form of an object marker, the subject marker being indexed by an expletive subject marker of class 9.

(37) Dawida (E74a; SAKAMOTO 2005: 11)
y-a-ni-fwan-a **ni-ghal-e**
SM9-PRS-OM1SG-NEC-FV SM1SG-go(home)-SBJV
'I ought to go'

3.5 North-East Coastal Bantu

3.5.1 Swahili(s)

Surprisingly, we were unable to detect any comprehensive description of the modal system and modal auxiliaries in (Standard) Swahili (G42), despite the prominent role it plays socio-culturally in East Africa and within Bantu linguistics. SCHICHO (1995) is the only work we know that specifically tackles modality in Swahili, but his conceptualization of modality and modal verbs in particular (ibid: 140-146) differs quite substantially from the one used in this paper. In lieu of a dedicated study on this topic, what is presented here is a fragmentized and far from complete picture of the Swahili modal system. Often nothing more than the possibility auxiliary **-wez-** is mentioned in grammatical descriptions of the language. NURSE and HINNEBUSCH (1993: 294; probably based on SACLEUX 1939: 1022) trace the etymology of this verb to Arabic **ʿezz** 'power', **ʿazza** 'be powerful'. The auxiliary is used for both pre-core and core modal categories, as seen in the expressions of PE possibility (38), and of deontic (39) and epistemic possibility (40), respectively.

(38) Standard Swahili (G42d; ASHTON 1947: 276)
Mtu **a-ki-ji-pit-i-a** **ovyo,** **a-wez-a** **ku-pond-w-a**
1.person SM1-COND-REFL-cross-APPL-FV carelessly SM1-POSB-FV INF-hit-PASS-FV
'If anyone crosses the road without due care, he is liable to be run over'

(39) Standard Swahili (G42d; KEZILAHABI 1988: 27)
Aksante **sana** **Rosa.** **U-na-wez-a** **kw-end-a**
thanks very Rosa. SM2SG-PRS-POSB-FV INF-go-FV
'Thank you very much, Rosa. You can leave'

(40) Standard Swahili (G42d; MOHAMED 2001: 82)
M-kutano **u-na-wez-a** **ku-maliz-ik-a** **mapema**
3-meeting SM3-PROG-POSB-FV INF-finish-NEUT-FV early
'The meeting can finish early'

ASHTON (1947: 276-277), SACLEUX (1939: 736) and SCHICHO (1995: 143), further present an auxiliary construction with the auxiliary verb **-pat-** 'get' (< ***-pát-** 'hold' (Bastin et al 2002)) plus an infinitive predicate verb. As an "acquisitive modal" (VAN DER AUWERA et al. 2009) verb **-pat-** appears to be confined to expressions of PE and deontic possibility. As indicated in the examples, the construction is used both in independent (41) and dependent clauses (42), the latter with **-pat-** inflected in the subjunctive verb form. Notice furthermore that example (42) shows the (optional) omission of the infinitive marker (i.e., **washinda** instead of **kuwashinda**) formally reflecting a more grammaticalized construction.

(41) Standard Swahili (G42d; SACLEUX 1939: 736)
a-me-pat-a **ku-maliz-a** **kazi**
SM1-PERF-get-FV INF-finish-FV 9.work
'He managed to finish the task'

(42) Standard Swahili (G42d; ASHTON 1947: 276)
ni **vizuri** **m-ji** **w-etu** **u-ku-e,** **adui** **z-etu,**
COP good 3-town 3-POSS1PL SM3-grow-SBJV 10.enemies 10-POSS1PL
wa-ki-tu-shambuli-a, **tu-pat-e** **wa-shind-a**
SM2-COND-OM1PL-attack-APPL-FV SM1PL-POSB-SBJV SM2.PRS-beat-FV(<INF)
'It is good that our town grows, so that if our enemies attack us, we may get the better of them'

The causative derivation of **-pat-**, viz. **-pas-** or **-pash-** (cf. NURSE and HINNEBUSCH 1993: 590) may in turn be used for PE and deontic necessity, either in a passivized form (43) or carrying a dummy subject marker (of class 9) whereas the first participant is object marked on the verb (44).

(43) Standard Swahili (G42d; SACLEUX 1939: 737)
ni-me-pasw-a **kw-end-a**
SM1SG-PERF-NEC.PASS-FV INF-go-FV
'I am obliged to go, i.e., I must go'

(44) Standard Swahili (G42d; MADAN and STEERE 1890: 155)
i-me-ni-pas-a **kw-end-a**
SM9-PERF-OM1SG-NEC-FV INF-go-FV
'It concerns me to go, i.e., I ought to go'

Other necessity verbs include various verbal derivations of the Arabic loan **-lazimu** 'must' (from **lāzim(an)** 'necessary' (SACLEUX 1939: 469; SCHADEBERG 2009), with the first participant again deprived of agentivity, structurally marked either in the form of detransitivizing suffixes such as the neuter **-ik** in (45), or by the shift in indexation from subject to object, as in (46). The verb **-bidi**, yet another Arabic loan (ultimately derived from **budd** 'act of escaping' (SACLEUX 1939: 116) seems to be restricted to constructions with a dummy subject, typically (47) but not necessarily (48) including the first participant as an object marker.[15]

(45) Standard Swahili (G42d; SCHICHO 1995: 143)
Rajabu, **a-li-ye-lazim-ik-a** **ku-zi-ju-a**
Rajabu, SM1-PST-REL1-NEC-STAT-FV INF-OM10-know-FV
desturi **za** **nyumba** **hii ...**
10.habit CONN9 9.house 9.DEM
'Rajabu, who was obliged to know the habits of this house …'

(46) Standard Swahili (G42d; SCHICHO 1995: 141)
i-na-tu-lazimu **tu-fany-e** **m-kataba**
SM9-PRS-OM1PL-NEC SM1PL-make-SBJV 3-agreement
'We have to make an agreement'

(47) Standard Swahili (G42d; SCHICHO 1995: 141)
hadithi **amba=yo** **i-li-ni-bidi** **ni-i-sikiliz-e**
9.story that=REL9 SM9-PST-OM1SG-NEC SM1-OM9-listen-SBJV
'A story that I had to listen to …'

[15] Note that a quick Google search gave several examples with **-bidi** where the first participant is not demoted but is rather indexed in the subject marker slot, for example:

(i) **w-anawake** **wa-na-bidi** **ku-m-shukuru** **m-gunduzi** **wa** **makeup**
2-woman SM2-PRS-NEC INF-OM1-thank 1-discoverer CONN1 makeup
'Women should thank the discoverer of make-up'

This apparent change in argument structure might be due to influence from English but more research is needed in this area.

(48) Standard Swahili (G42d; SACLEUX 1939: 106)
i-ka-bidi **yule** **m-toto** **ku-let-a** **ma-shahidi**
SM9-CONS-NEC 1.DEM 1-child INF-bring-FV 6-proof
'That child needs to bring some proof'

Standard Swahili furthermore makes use of the verb **-faa** as a necessity marker, a reflex of ***-pú-** 'be fitting' (cf. NURSE and HINNEBUSCH (1993: 587), i.e. ultimately the same root as in Dawida (see §3.4) and several other languages discussed further on in this paper.

The set of modal auxiliaries attested in the Congo varieties of Swahili seems to be fairly similar to those used in Standard Swahili, given some general morphological and phonological differences, e.g. **-wez-** > **-ez-**. For Kivu Swahili, however, NASSENSTEIN and BOSE (2016: 56-57) notes an interesting semantic difference in that the modal auxiliary construction **-pashw-** + infinitive which has spread across forces, from marking only necessity (49a), as in Standard Swahili, to also include possibility (49b). Note furthermore that in these examples the infinitive prefix of the main verb has been dropped, a formal indication of the condensation and thus further grammaticalization of this construction.

(49) Kivu Swahili (G40x; NASSENSTEIN and BOSE 2016: 57-58)
a. **mi-na-pashw-a** **ku-on-a**
SM1SG-PRS-NEC-FV OM2SG-see-FV(<INF)
'I have to see you'
b. **Est-ce que** **mi-na-pashw-a** **ku-on-a**
INT SM1SG-PRS-POSB-FV OM2SG-see-FV(<INF)
'Could I (possibly) see you?'

In Kisangani Swahili (another Congo Swahili variety), it is the active form **-pash-** which is employed as a modal auxiliary verb. Again, the verb is used both to denote necessity and possibility. However, these differences in force are formally marked through different constructional embeddings of **-pash-**: A possibility reading arises when **-pash-** stands in the present tense verb form (50a), whereas a necessity reading occur when the verb **-pash-** is further inflected in the future tense verb form (50b). Note furthermore, the apparent lack of the infinitive marker of the predicate verb in (50b) as compared to (50a), which would indicate a further formal specialization of the necessity construction.

(50) Kisangani Swahili (G40x; NASSENSTEIN 2015: 105)
a. **a-na-pash-a** **ku-tu-remonté** **moral**
SM3SG-PRS-POSB-FV INF-OM1PL-restore 9.courage
'S/he can restore our guts'

b. **ni-ta-pash-a** **sema** **i-ko** **partagé**
SM1SG-FUT-NEC-FV say SM9-COP divided
'I have to say it is divided... '

A formally encoded difference with regards the possibility auxiliary **-wez-** is noticed by DEVOS (2008b: 412) for Makwe (G402), where it is marked on the verb form of the predicate verb. Whereas PI possibility is construed with **-wez-** and a predicate verb in the infinitive, as in (51a), PE and deontic possibility is expressed with the second verb inflected in a type of dependent conjugation marking both conditional and consecutive clauses, as in (51b).[16]

(51) Makwe (G402; DEVOS 2008b: 412)
a. **i-wéez-i** **ku-kú-nyákuúl-a**
NEG:SM1SG-POSB-PRS.NEG INF-OM2SG-take-FV
'I cannot carry you'
b. **ń-na-weéz-a** **ni-ka-yaúul-a** **úyóoví**
SM1SG-PRS-POSB-FV SM1SG-CONS-speak thus
'Can I say it like this?'

3.5.2 Ruvu (Central Tanzania)

NYINONDI and LUSEKELO (2020) describes the modal system of Luguru (G35). Here, the modal verb **-dah-** is used for encoding all flavors of possibility, e.g. the deontic possibility as in (52).

(52) Luguru (G35; NYINONDI and LUSEKELO 2020: 33)
ko-dah-a **ku-ghend-a** **ku-dawal-a**
SM1.PRS-POSB-FV INF-go-FV INF-play-FV
'(He doesn't need to stay at home too much) he may go to play (on the street)'

The possibility verb **-dah-** is attested as a modal auxiliary denoting possibility (or at least listed as a verb meaning 'can', 'be able' etc.) in closely affiliated languages like Kagulu (G12; PETZELL 2008: 221), Bondei (G24; WOODWARD 1882: 44) Zigua (G31; WOODWARD 1902: 36), Zaramo (G33; WORMS 1897:306) and Kami (G36; VELTEN 1900: 43; PETZELL and AUNIO 2019: 583-584). We also found attestations further southwards in the Kilombero (G50) language Pogolo (G51) (HENDLE 1907: 137) as well as further north in Rift valley varieties such as Nyaturu (F32), further discussed in §3.6, and Langi (F33; SEIDEL 1898: 432). The etymology of this verb is not clear. However, in Last's grammar of KAGULU (1886: 130), **-dah-** is

[16] Makwe, spoken in Northern Mozambique, is most likely a Swahilized variety of Makonde (see RZEWUSKI 1991, DEVOS 2008b: 2-3 for further discussion and references). As such, it should perhaps have been treated in some other subsection (either §3.7 or 3.9). However, since the modal verb construction is similar to Swahili, we keep Makwe in this section.

provided with the translation 'have power over',[17] a common source meaning for possibility verbs and thus a plausible source meaning. In at least Kami, the Swahili loan **-wez-** appears to occur along with **-dah-** as a possibility auxiliary verb (PETZELL and AUNIO 2019: 584).

The verb **-bam-** 'want' derived with the passive **-igw-** is used for expressions of PE and deontic necessity in Luguru, as illustrated in (53).[18] In these constructions, the predicate verb needs to be inflected in the subjunctive. Again, the semasiological background of the source verb **-bam-** is opaque.

(53) Luguru (G35; NYINONDI and LUSEKELO 2020: 79)
u-ghali **wa-li-bondo** **w-o-bamigw-a** **wa-bik-igw-e**
3-gali CONN3-5-cassava SM3-PROG-NEC-FV SM3-cook-PASS-SBJV
na **tsi-mbogha** **au** **wa-dagala**
with 7-vegetable or 2-sardine
'(and when this dance is played) ugali from cassava must be cooked and served with vegetables or sardine (**dagaa**)'

For Bondei (G24) and Zigua (G31), WOODWARD (1882: 44, 1902: 36) mentions the verb **-agiz-** 'concern', 'behoove' as being used for 'must' and 'ought', viz. PE and deontic necessity meanings. This verb can be set in comparison with the forms ***-lagil-** 'agree, suit', **-agir-** 'be fitting, behoove', and also ***-lag-** 'promise' and ***-lagizy-** 'command' reconstructed for the closely related Sabaki cluster by NURSE and HINNEBUSCH (1993: 595). WOODWARD (1882: 44) furthermore notes, that just as was described for Swahili in the previous section (§3.5.1), Bondei uses **-faa** 'be of use, avail' as a modal verb.

3.5.3 Other North-East Coastal languages

We have also found some brief notes about modal verbs in other North-East Coastal Bantu languages. Thus, PATIN et al. (2019: 609) presents for Ngazidja (G44a) the verb **-ʃind-** 'be able' (= non-core possibility alone?); cp. Swahili **-shind-** 'win, overpower, subject, tame, master, overcome, be stronger, surpass, push-to-end' (Sacleux 1939: 838-839).

NICOLLE (2013: 131) reports the use of **-wez-** as expressing possibility and **-londw-** necessity in the Mijikenda-variety Digo (E73). Whereas the former is the same possibility auxiliary as found in Swahili, the latter stems from a

[17] NYINONDI and LUSEKELO (2020) also refer to this grammar when discussing the semantics of **-dah-**, albeit not this specific lexical meaning.

[18] In the interests of clarity, we have slightly altered the original English translation in this example.

passivized variant of the verb **-lond-** 'want' – ultimately from ***-dònd-** 'search for' and/or 'follow', reconstructable for Eastern Bantu (see BERNANDER 2017: 221 for some elaboration on the doubled meaning of the reconstructed verb; see also footnote 21 in §4.1.2). Interestingly, the same verb occurs with the same function in languages spoken along Lake Nyasa and thus at the very other side of Tanzania (See §3.7).

3.6 West Tanzania and the Rift Valley Bantu languages

Recall from example (11) in §2.3 that Nyamwezi has a possibility verb **-dul-** originating from a lexical source verb meaning 'bore', i.e. a similar etymology to the possibility verb **-tor-/-tur-** in East Nyanza JE40 (§3.2.4) but also **-hotol-** in southern Tanzania (further presented in §3.7). In addition to **-dul-**, JONSSON (1954: 59-60) mentions for Nyamwezi the possibility auxiliary verbs **-duj-**, **-vej-** (copied from Swahili **-wez-** or perhaps even directly from Arabic?) and **-kobol-**. He also describes the use of the verb **-igel-ile** 'be worthy (vara värdig)' for marking necessity (< (petrified) reflexive **-i-** + ***-ged-** 'measure' + perfective **-ile**?). Interestingly, there seems to be a clear semantic difference in modal meaning depending on how the participants are indexed on the modal verb. Where a weak necessity reading 'should' ('bör') is conveyed when the verb is inflected with a dummy morph **shi-** (< class 8?) alone, whereas a strong necessity reading is conveyed when the verb additionally carries the 1st participant indexed as an object marker (54). (When inflected with 1st participant as subject marker, the verb denotes the lexical 'be worthy'). In all cases, the predicate verb is inflected in the subjunctive (but in the infinitive form in negated constructions).

(54) Nyamwezi (F22; JONSSON 1954: 61)
shi-b-igel-ile **boi** **ba-ba-w-ile**
SM?8-OM2-NEC-PFV PERS.PRON2 SM2-OM2-say-PFV
'they must tell them' ('de måste säga till dem')

Like several other languages spoken in the central parts of Tanzania, Nyaturu (Rimi; F32), spoken at the south-western fringes of the Rift Valley, makes use of the auxiliary verb **-dah-** as a possibility auxiliary (OLSON 1964: 213). In addition, Nyaturu has a necessity auxiliary verb **-noney-** 'behoove' < 'deserve, be worthy'. As seen in (55a), the modal necessity meaning is again expressed through the type of construction with an expletive subject marker and the first participant indexed as an object marker. Sentence (55b) is included here to show the "automatic" lexical meaning that arises when the first participant is indexed as a subject (marker).

(55) Nyaturu (F32; OLSON 1964: 214)
a. **i-qu-nónéy-é** **ú-l-á** **mwiimo**
SM9-OM1PL-NEC-PFV INF-eat-FV 3.work
'It behooves us to work'
b. **n-ti-nónéy-é** **u-kuríw-a**
SM1SG-NEG-behoove-PFV INF-praise.PASS-FV
'I don't deserve to be praised'

3.7 Southern Tanzania (The Southern Highlands and Rufiji-Ruvuma)

For the southern parts of Tanzania (G60, N10 + P10-P20), BERNANDER (2016, 2017: ch.10) describes the modal system in Manda (N11) along with comparative notes for several neighboring and closely affiliated languages. In Manda, the possibility verb **-hotol-** has a pre-modal meaning 'overcome, defeat' but also, and originally, 'pierce' (cf. the reconstructed root ***-cotʊd-**; BASTIN et al. 2002). It is used for all subtypes of possibility all the way to marking epistemic meanings, as in (56).

(56) Manda (BERNANDER 2017: 280)
a-hótól-a **kú-y-a** **ku-nyúmba** **ni-wón-a** **baiskéli** **y-áki**
SM1-POSB-FV INF-be-FV 17-9.home SM1SG-see-FV 9.bicycle 9-POSS3SG
'He might be at home; I can see his bicycle'

The modal auxiliary **-hotol-** exists also in several of Manda's closest relatives/neighbors, like Kisi (G67), as illustrated in (57).

(57) Kisi (G67; NGONYANI 2011: 153)
li-boma **lya** **twenga** **lya** **nganda**
5-troupe CONN5 us CONN5 mganda
li-hotol-a **ku-kwel-a** **mw-aka** **ughu**
SM5-POSB-FV INF-win-FV 3-year DEM3
'Our mganda (dance) troupe can win this year'

However, as reported in BERNANDER (2017: 290-291), the expression of possibility seems to be increasingly replaced with the Swahili loan **-wez-** in many of Manda's closest affiliates. In Matengo (N13), for example, the verb **hwes-** (< borrowed and adapted from Swahili **-wez-**) has increased significantly in use (in particular among younger speakers) at the expense of **-hotol-** (YONEDA 2010, BERNANDER 2017: 291).

For the subdomain of necessity, Manda primarily uses the verb **-lond-** 'look for, want, need, desire'. This reflects a similar formally encoded shift of this verb from marking PI to marking PE and deontic flavors, where the two latter forms are expressed with the verb further derived with the medio-passive neuter

suffix **-ɪk-**, thus morphologically coding the semantic shift from a requirement inherent in the first participant, as in (58), to one being externally imposed (59). In the latter example with the neuter **-londek-** (59), the 1st participant is indexed as a subject marker on the modal verb.

(58) Manda (N11; BERNANDER 2017: 296)
ni-lónd-a **kú-nyw-a** **kaháwa**
SM1SG-NEC-FV INF-drink-FV 9.coffee
'I need to drink coffee'

(59) Manda (N11; BERNANDER 2017: 297)
a-lóndék-a **ku-bít-a** **ku-shuli**
SM1-NEC-FV INF-go-FV 17-9.school
'She has to go to school'

Modal **-londek-** may also be inflected with a 'dummy morph' – the subject marker of class 9. In such cases in Manda, however, the 1st participant may not be de-ranked to an object marker on the verb (probably since Manda is quite reluctant to object marking in general). The verb has scope over a whole matrix clause rather than over the single action/event of the predicate verb (cf. BERNANDER 2017: 294-295) and does not really count as an auxiliary (see discussion in §4.1.4). Note also here also the predicate verb occurs in the subjunctive rather than in the infinitive (see BERNANDER 2020 on the special subjunctive form in Manda).

(60) **i-lóndék-a** **a-lút-áyi** **ku-héng-a** **li-héngu**
SM9-NEC-FV SM3SG-go-SBJV INF-work-FV 5-work
'It is necessary (that) he goes to work'

Another necessity verb used by some Manda speakers for PE and deontic flavors is **-yenel-ek-**, which BERNANDER (2017: 298) traces to a verb meaning 'be fitting', again further extended with the neuter suffix. BERNANDER (2017: 299-301) argues that neither **-lond(-ek)-** or **-yenelek-** can be considered core modals as they tend to be more "descriptively" than "performatively" used (NUYTS 2006: 15), and thus seldom with the speaker as the source of obligation, but rather with reference to other external forces (TRAUGOTT and DASHER 2002: 114-117). None of the verbs are used for expressing epistemic meanings.

Just as Manda **-yenelek-**, a verb meaning 'be fitting' used as a necessity marker has also been attested in Mwera (P22; HARRIES 1950)[19]. In this case, however, the

[19] It may be noted that Harries' Mwera data is also sampled and discussed in BYBEE et al.'s (1994: ch.6) cross-linguistic investigation of modal expressions.

verb is not derived with agent-demoting morphology alone as with the neuter in Manda. Instead, the active verb form is used albeit with the first participant indexed as an object (marker) and a dummy morph indexing the subject marker (as, e.g., the Swahili example in (43) above).[20]

(61) Mwera (P22; HARRIES 1950: 109)
i-tu-ŵandicīle **ku-wal-a** **ukoto**
SM9-OM1PL-NEC.PFV INF-dress-FV properly
'We ought to dress properly'

We are not sure how to decompose this verb stem. HARRIES (1950: 67) shows that the neuter + applicative form together is realized as **-icila**, which would make sense semantically and iconically: First, the verb is derived with the neuter extension which de-ranks the 1st participant but also detransitivizes the verb. With the addition of the applicative, the verb construction is made transitive again accommodating the treatment of the 1st participant as an object. Peeling away the applicative and neuter would leave us with a root **-ŵand-**; (from ***-band-** 'flatten'???). Finally, the verb must be a change-of-state verb inflected with the perfect(ive) suffix **-ile**, leading to the present stative reading and to the type of verb-stem modification where the verb base final /*il*/ coalesces with the suffix into forming <**īle**> /*i:le*/ (cf. Harries 1950: 79-81).

For marking possibility, Mwera employs the verb **-many-** originally meaning 'know' (cf. HARRIES 1950: 110), just as in Kamba (§3.3) and Ikoma (§3.2.4) discussed previously. See (§4.1.1) for further discussion on the extension of this verb to marking modal possibility.

Also Bena-Hehe (G63-G62), spoken in the Southern Highlands, use a reflex of **-many-** along with two other verbs originally meaning 'know', namely **-kagul-** and **-dzel-**, for expressions of possibility. This according to PRIEBUSCH (1935: 98-100)[21], who additionally mentions the verb **-vagil-**, derived from **-vag-** 'be enough' with the valency-increasing applicative suffix – thus, 'do to be enough' or similar – used inflected in the perfective **-ile** (with stem modification) for expressions of deontic possibility. For expressing deontic necessity, Bena uses the verb **-nogel-** 'be appealing', also derived with the applicative suffix from the stem **-nog-** 'be/taste good'. The latter verb is

[20] When used with proper subject indexation on the auxiliary verb, the in other terms similar construction has a proximative 'nearly' or 'almost'-meaning, e.g. **wawandiciile kujawula** 'they were almost going' (HARRIES 1950: 109).

[21] Priebusch explicitly refers to his work as a "Bena-Hehe-Grammatik" and claims it to include a mixed set of linguistic data from both Bena (G63) and Hehe (G62).

illustrated in (62). All constructions are formed with the predicate verb in the infinitive. Although mainly associated with possibility and necessity, respectively, Priebusch mentions that both **-vagil-** and **-nogel-** may also be used interchangeably with the opposing force. In (63), for example, -**nogel-** expresses deontic possibility.

(62) Bena (G63; PRIEBUSCH 1935: 98)
u-nogye **ukw-im-a**
SM2SG-NEC.PFV INF-stand-FV
'You have to get up' ('Du musst aufstehen')

(63) Bena (G63; PRIEBUSCH 1935: 100)
a-si-nogye **nda** **uku-l-idz-an-a** **na** **juhwe**
SM2SG-NEG-NEC.PFV NEG INF-eat-CAUS-ASS-FV COM PRON1PL
'He may not eat with us' ('Er darf nicht mit uns mitessen')

See also MORRISON (2011: passim) where **-wes-** (most likely an adapted variant of Swahili **-wez-**) occurs as a possibility auxiliary verb in Bena, indicative of a later introduction of this verb to this language.

3.8 The Lake corridor area + Nyakyusa

For the languages spoken in the "corridor" between Lake Tanganyika and Lake Nyasa (M10-M20) and the Nyakyusa varieties (M30) of southwestern Tanzania and continuing into Malawi and Zambia, we have not found much information on modal auxiliary verbs.

One exception is Lungu (M14), which has the potential prefix **-nga-** for possibility force but uses modal verbs for the expression of necessity. Firstly, by a complex combination of the copula and comitative **na** 'be with' > 'have' plus an infinitive predicate verb, as in (64). Secondly, with auxiliary **-fw-** in the perfective plus an infinitive predicate verb as in (65).

(64) Lungu (M14; KAGAYA 1987; cited in NURSE 2008: appendix)
tú-lí **n-úǃkú-lím-a**
SM1PL-COP COM-INF-cultivate-FV
'We have to farm'

(65) Lungu (M14; KAGAYA 1987; cited in NURSE 2008: appendix)
tú-fw-ííl(e) **úkú-lím-a**
SM1PL-NEC-PFV INF-cultivate-FV
'We ought to farm'

The construction in (64) is reminiscent to one of the necessity constructions in Luganda (cf. §3.2.2), while the one in (65) appears to include a reflex of the same

***-pú-** 'be fitting' (BASTIN et al. 2002), as found in necessity expressions in Dawida (§3.5) and Swahili (§3.5.1) as well as Makhuwa (P31) discussed below in §3.9.

Another exception is Ndali (M301), as described by BOTNE (2008). Whereas a verb-internal prefix, **-nga-**, appears to be the main marker of possibility (see also the note on N20-40 languages in §3.9), modal verbs are used for expressions of necessity. Most prominent is the verb **-bagil-** with the predicate verb in the infinitive, illustrated in (66). Comparing with closely related Nyakyusa (M31), it is clear that **-bagil-** is derived from a lexical source verb (again) meaning 'suit, fit, deserve, be eligible' as indicated in FELBERG's (1996: 9) dictionary.[22]

(66) Ndali (M302; BOTNE 2008: 126)
a-bagiile **u-ku-yi-many'-a** **ii-síla**
SM1-NEC.PFV AUG15-INF-OM9-know-FV 9-way
'S/he must know the way'

In addition, BOTNE (2008) mentions two other verbs being used in complex construction as markers of "obligation" (i.e. PE/deontic necessity), namely the motion verbs + **íse**, from **-is-** 'come', and the (rarer) **íye** from **-y-** 'go', both being inflected in the subjunctive (and with the predicate verb likewise inflected in the subjunctive). Given Botne's examples they would seem to also function as marking subordinate purpose clauses. It is hard to tell whether these erstwhile motion verbs have lost their subject marking or whether they are "raised" constructions with a dummy morph in the subject prefix position (i.e. **í-ye** [SM9-go.SBJV]).

(67) Ndali (M302; BOTNE 2008: 126)
n-aa-tumíw-a **ukuti** **íse** **n-gulísh-e** **ii-ny'ama**
SM1-PRF-be.sent-FV COMPL NEC SM1-sell-SBJV 9-meat
'I have been sent so that I should (come to) sell the meat'

This last construction is reminiscent of a complex verb form expressing possibility in Safwa (M25; VOORHOEVE n.d.: 46). Following VOORHOEVE (ibid: 46-47) it consists of **nza** 'come' followed by the main verb in the present subjunctive. The subject is indexed on the 'auxiliary' only if it consists of a vowel. Accordingly, the examples in (67a) and (68b) show 'double' (SM = **a-**) and "single" agreement (SM = **tú-**).

[22] Peculiarly, FELBERG (ibid) also gives the meaning 'be able, can, be possible' and as an equivalent of Swahili **-wez-**, thus suggesting a reading of possibility rather than necessity in Nyakyusa. We have not been able to investigate this apparent discrepancy further.

(68) Safwa (M25; VOORHOEVE n.d.: 46)
a. **a-nza-á-bal-e**
SM1-POSB-SM1-go-SBJV
'He may go today'
b. **nza-tú-zi-fis-e**
POSB-SM1PL-OM10-hide-SBJV
'We may hide them today'

3.9 Further southwards (N20-40 and P30)

This section closes with a final note on modality in the Eastern Bantu zone N and P languages spoken further southwards in Malawi and Mozambique. Firstly, the N-languages spoken in the south of Tanzania (N20-N40), seem primarily to make use of dedicated verb-internal morphology for the marking of modality, e.g. the use of the post-initial prefix **-nga-** for possibility and **-ka-** and **-zi-** for necessity in Chewa-Nyanja (N31), described by HETHERWICK (1916) and MCHOMBO (2004), as in (69). It is possible that the latter form is related, though, to the 'come'-auxiliary construction described for Ndali and Safwa in the §3.8 above. NGONYANI (2020: 68-69) does report the presence of a possibility verb **-koz-** 'be able' in the Nyanja variety spoken in Tanzania without offering more detail on the matter, however.

(69) Chewa (N31b; MCHOMBO 2004: 30-31)
a. **mu-nga** **ndi-thandiz-e**
SM2PL.POT SM1SG-help-SBJV
'You can/may help me'
b. **m-kango** **u-zi-b-a** **mi-kanda**
3-lion SM3-NEC-steal-FV 4-beads
'The lion must (should) steal beads'

With regard to the Makhuwa group (P30), the most substantial documentation of the modal system is to be found in the aforementioned study by Devos (2008a) on Shangaci (P312). As she shows, different verbs are used to express flavors both within the possibility and necessity domain. The inventory of modal verbs are based on the potential **pode** 'can' borrowed from Portuguese (whose status as an auxiliary verb is a problematic, however, a fact further discussed in §4.2) and the necessity verbs **-na** 'have (to)' for PI flavor, the PE/deontic **-on-** 'see' (used only with an impersonal subject concord, and thus again having a bit of an irresolute status as an auxiliary; see §4.1.4), and **-sal-** 'remain' also marking PE/deontic necessity. The auxiliary construction with **-sal-** illustrated in (70) with both the auxiliary and the predicate verb being inflected in a verb form **(-i-**B**-a)** otherwise only used in subordinate clauses (the 'conditional/subsecutive'). This leads DEVOS (2008a), along with other

concomitant evidence, to suggest that the construction's modal use is the result of insubordination.

(70) Shangaci (P312; DEVOS 2008a: 20)
aphaáno **w-ii-sál-á** **w-ii-khoól-a** **khaázi** **za** **vaánce**
now SM2SG-CON-remain-FV SM2SG-CON-hold-FV 10.work CONN10 16.outside
'Now you have to start working around the house again (after your periods have ended)'

The perhaps most remarkable feature concerning Shangaci modal verbs is what occurs in the epistemic domain, however. Here, the language makes use of a segmentally identical construction consisting of the auxiliary verb **-sakh-** (< 'want') and the predicate verb in the infinitive which is used both for expressing epistemic possibility (71) as well as epistemic necessity (72).

(71) Shangaci (P312; DEVOS 2008a: 15)
leélo **ni-ń-sákh'** **o-khól-á** **n-cúuwa**
today SM5-PRS-want' FV.INF-hold-FV 5-sun
'Today, the sun might shine'

(72) Shangaci (P312; DEVOS 2008a: 23)
násaámbi **ni-ń-sakh'** **o-khol-a** **n-cúuwa**
now SM1SG-PRS-want' FV.INF-hold-FV 5-sun
'Right now the sun will shine (you can already see it appearing from behind the clouds)'

As indicated in these examples, it is only the prosody of the predicate verb that indicates its modal meaning. That is, high tones are missing on the predicate (infinitive) verb in the construction expressing epistemic necessity. DEVOS (2008a) relates this omission of tones (more specifically tone omission of the first high tone and concurrent shifts in the following tone pattern) to "predicative lowering", a strategy, which among other functions is used for marking focus and here, in turn, to strengthen the modal force of the expression.[23] Interestingly this type of shift directly between epistemic possibility and necessity is not accounted for in van der AUWERA and PLUNGIAN's

[23] Notice moreover the word order of these examples with the subject noun in a post-verbal position. With the subject in the canonical pre-verbal position **-sakh-** automatically receives the original source meaning of 'want'. Thus, there is a morpho-syntactic difference between the source construction and the modal constructions, which DEVOS (2008a) suggests is a syntactic reflection of the fact that the epistemic expression has scope over the proposition as a whole, rather than just the complement verb.

(1998) semantic map and thus contradicts their suggested pathways of extension and change.

For other P30-languages the documentation on modal auxiliary verb constructions appears rather slim. For Cuwabo (P34), GUÉROIS (2015: 403) reports an auxiliary verb **-wodh-** 'be able' (< 'defeat'; cf. ibid: 537, 631) used to express possibility, but only dynamic participant-internal possibility it would seem.

(73) Cuwabo (GUÉROIS 2015: 403)
ka-wódh-íle **o-mú-téy-ih-a** **námwáli** **le**
NEG.SM2-POSB-PFV INF-OM1-laugh-CAUS-FV 1a.girl DEM1
'None succeeded in making that girl laugh'

As further discussed along with Shangaci **pode** in §4.1.4 and §4.2, in the necessity domain Cuwabo has also borrowed the Portuguese **ter que** 'have to' for marking obligation, however without morpho-syntactically integrating it as a verb, thus making its position as an auxiliary suspicious.

For Makhuwa (P31) WOODWARD (1926: 303), gives, without exemplifying, **-wory-** 'can' (cognate w. Cuwabo **wodha** above?) as a possibility verb. PIRES PRATA (1990: 410) rather gives **-wer-/-wery-** and adds more translation equivalents suggesting that the auxiliary has an origin in a verb meaning 'win, overcome, be superior' which appears to be a common source for possibility verbs in Eastern Bantu (cf. also § 4.1.1). Next, WOODWARD (1926: 303) gives **-pwanel-** 'must/ought' as a necessity verb. This is again a verb that can be linked up to readings of 'be fitting' (cf. ***pú-an-** 'be fitting', resemble each other'; BASTIN et al 2002). Interestingly, PIRES PRATA (1990: 324) gives the meanings 'be equal, alike, resemble each other' for underived **-pwan-**. The same verb with the applicative suffix **(-pwanel-)** is translated as 'deserve', whereas it is the sequence of the applicative and the detransitivizing stative suffix **(-pwaneley-)** which seems to give the necessity reading. Moreover, PIRES PRATA (ibid: 387) gives yet another necessity verb derived from the verb **-tthun-** 'want, desire' by addition of a stative suffix, i.e., **-tthuney-**, thus confirming the recurrent trait of detransitivizing morpho-syntax in modal constructions expressing necessity (cf. also §4.1.4).

4. East African Bantu modal auxiliary verb constructions – general traits and comparative inferences

After surveying the Eastern African Bantu modal verbs, some comparatively induced patterns emerge that are summarized and discussed in this section.

Starting with the most fundamental traits, it is first and foremost clear that there do exist an extensive amount of modal auxiliary verbs in this Bantu-speaking area, both expressing the concepts of necessity and possibility. However, there is also a bewildering diversity between the various languages with few real cognancies attested beyond very tightly inter-related groups. With that said, there are still striking similarities with regard to the source semantics of the verbs recruited for a specific modal function and their constructional make-up as well as that of the predicate verbs.

The EA Bantu modal auxiliary verb constructions would generally seem to prefer to stick to one and the same force. Thus, although they often vary and may be employed in different categories ("flavors") within a single category of either possibility or necessity, they do not often seem to be intersecting between these overarching categories. Counterexamples do exist, however. Firstly, in Congo Swahili, the auxiliary **-pash-** is a PE/Deontic marker and thus the switch in force can be accounted as adhering to a typical cross-linguistic pathway of expansion; cf. VAN DER AUWERA and PLUNGIAN (1998). A similar case is attested for Bena, although in this case it seems to be both a possibility verb that may be used for necessity and vice-versa a necessity verb used for expressing possibility force. Secondly, for **-esh-** in Ishenyi, the source meaning of the verb would suggest an erstwhile history as necessity verb albeit no language-internal (or comparative) evidence exist to allow us to draw such a conclusion. Thirdly, and most striking, is the case of **-sakh-** in Shangaci where the switch from possibility to necessity occurs within the epistemic subdomain thus breaching the universal pathways of extension postulated by VAN DER AUWERA and PLUNGIAN (1998).

Often, the category of possibility appears to be more cohesive, in the sense that many of the surveyed languages make use of one possibility verb, in one concurring form, dedicated to marking out the full spectrum of possibility flavors. In comparison, the domain of necessity appears as slightly more shattered. This in the sense that different verbs may be employed for marking weak vs strong necessity, but also that what is essentially the same verb is exposed to different morphological and constructional modifications when marking different flavors (as further discussed in §4.1.2). Unlike its equivalents in the possibility domain, auxiliary verbs also appear to be less often extended to be used for expressing epistemic necessity. Additionally, the predicate verb-internal subjunctive suffix, omnipresent in Bantu, may by itself cover the expression of necessity directly on the predicate verb. We could speculate that this feature of the subjunctive leaves less motivation for the developments of auxiliaries within the necessity domain (cf. NURSE and DEVOS 2019 and further references therein). Such circumstances

may, in turn, explain the fact that more focus has been put on the research of possibility verbs in EA Bantu.

4.1 Source verbs and source constructions

In the survey of EA Bantu modal auxiliaries (cf. Tables 3, 4 and 5 in the appendix) it is interesting to note the diversity of lexical source verbs, on the one hand, and, on the other, the similarities in their meaning, the similar constructions they are formed in and the similar ways they tend to be integrated and expanded within the modal domain.

Note that in some cases we have not been able to work out the (exact) source meaning of a modality verb or to trace its lexical etymology further. A case in point is the verb **-dah-** discussed for several languages of our survey.

4.1.1 Possibility verb sources

Starting with possibility verbs in EA Bantu, these may often quite straightforwardly be linked up with the broader non-Bantu typological literature in terms of cross-linguistically common source meanings and further pathways of extension. Most possibility verbs in our survey are derived from verbs roughly meaning 'overcome, defeat, have power (over)'. This adheres to a common cross-linguistic pattern where verbs predicating physical ability extends to cover all types of abilities (cf. BYBEE et al. 1994: 190). Interestingly however, in quite many of these EA Bantu verbs, the lexical etymology can be shown to go even further to a source etymology of 'pierce' or 'bore', namely for **-hotol-** in many languages in Southern Tanzania, **-tur-/-tor-** in the Western Serengeti languages, and **-dul-** in Nyamwezi. BERNANDER (2017: 284) notes with regard to this semasiological background that: "In the same manner as the extended meaning of 'overcome', the more narrow reconstructed original meaning 'pierce' also encompasses the first participant succeeding in forcing his/her way through a physical barrier […]. The recruitment of this verb to denote ability and capacity can thus be seen as a generalization and abstraction from this physical ability to overcome and dominate a concrete referential entity to the mastery of a more abstract activity or event."

There are also some attestations of 'know', mostly represented by cognates of the same verb **-many-** (reconstructable for PB as ***-many-** 'know'). Again, the recruitment of 'know' into marking possibility adheres to a very common cross-linguistic pattern. BYBEE et al. (1994: 190) notes that: "The most commonly documented lexical source for ability is a verb meaning 'to know' or 'to know how to' […] A verb originally restricted to mental ability is extended to apply as

well to physical ability, and thus becomes a general signal of ability". With that said, given the descriptions at hand, the use of **-many-** seems quite marginal in the EA Bantu where it has been attested, often not being the pre-dominant strategy of marking possibility and often not having expanded beyond expressing pre-core (dynamic) concepts of PI and PE possibility.

4.1.2 Necessity verb sources

For the necessity domain, the dominant pattern is the use of a verb expressing 'need' often co-expressing 'want' and other volitional uses, such as **-lond-**, **-tak-/-sakh-**, **-end-**, **-bam-**, **-tun-**.[24]

Used in its original, active, transitive form the verb is used for marking participant-internal necessity. The categories of participant-external and deontic (and occasionally also epistemic) necessity is, in turn, expressed by the additional modification of the verb, either with the dedicated Bantu passive, a reflex of Proto-Bantu ***-(ib)ʊ-**, or less commonly the medio-passive or middle-voice marking neuter, a reflex of Proto-Bantu ***-ɪk-** (cf. SCHADEBERG and BOSTOEN 2019), the latter being the case in Luhya and Manda.[25] Functionally, it signals the detachment of the 1st participant from no longer being co-referential with subject and thus being the agent or the instigator of the event depicted in the predicate (semantic main) verb, but as rather being in the hand of external forces. There exist alternative syntactic strategies for coding this semantic re-perspectivization than to just add the passive and neuter. These are further discussed in the next section (§4.1.4).

Another strand of necessity source constructions are verbs originally carrying denotations of '(make) fit, suit' etc., again with the first participant demoted from a default subject position. To this set of source verbs, we could arguably also accommodate the semantically broader **-téek-** 'put, edict' (+ passive extension) in West Nyanza (JE10 + some JE20). Verbs denoting 'be fitting' is also a common cross-linguistic source for expressing PE/deontic necessity (BYBEE et al. 1994: 182-183).

---

[24] BERNANDER (2017: 295) notes for Manda (N11) that the necessity verb **-lond-** probably at first meant 'follow, look for' as the same meanings have been reconstructed for Proto-Bantu **(*dònd**; GUTHRIE 1967-71. vol3: 178, BASTIN et al. 2002), but that is lexical-semantic range was expanded into additional meanings such as 'want', 'desire', and specifically 'need', based on "the inference that what one is looking for is subsequently something one wants (to have) or desires".

[25] In Manda, these two similar strategies can be said to have collapsed since the neuter has expanded its functional range at the expense of the passive which is rarely used; see BERNANDER (2018).

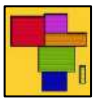
This is also the case with 'have' attested as a necessity source verb in Luganda, Lungu and Shangaci as a source for obligation (see BYBEE et al. 1985, 1994: 183-186 for elaboration on the semantic factors at play behind this extension). These 'have'-constructions are in turn derived, in different ways and in different degrees of lexicalization, from the in Bantu common comitative pattern based on a copula and the comitative conjunction **na** 'with', in which 'be with' > 'have'.

4.1.3 Motion verbs as modal verb sources

There are also attestations in our survey of what we may jointly refer to as motion verbs as source verbs for expressions of both possibility and necessity. For the use of verbs meaning 'arrive, reach' for expressions of (strong) necessity in Luhya varieties in general, and Llogoori in particular, GLUCKMAN and BOWLER (2020) suggest that the extension to a modal verb involves the expansion in referential range from a concretely defined (locational) threshold to an abstract one "of goodness" stipulated by external forces of which the 1st participant reaches or exceeds. In a similar vein, Shangaci's **-sal-** 'remain' could be indicated to implore to stay put in an extended locus of goodness. Other more basic motion verbs used like 'come' and 'go (to)' that we see mostly in the southern languages of our survey, might also fit within this idea. Or, when expressing possibility, with BYBEE et al.'s (1994: 190-191) remark that 'arrive' implies the successful completion of an act which is then easily conventionalized as denoting the capability of the 1st participant to complete the act. More generally, of course, 'come' and 'go'-verbs in Bantu are ubiquitously recruited for expressing all types of grammatical functions, not only associated motion, tense and aspect but also focus, for discourse structuring and in presentational and existential constructions. Apparently, these verbs are quickly undergoing formal truncation, univerbation and morphologization as modal auxiliaries too, something which makes them stand out in relation to the other relatively modestly grammaticalized modal verbs.

4.1.4 Source constructions

Construction-wise it may firstly be noted that almost all auxiliary verb constructions expressing possibility in our survey are formed with the predicate verb in the infinitive. In a few languages (cf. Ishenyi in (26), Makwe in (51b) and Safwa in (68) the auxiliary is followed by a more finite verb form (cf. MIESTAMO 2005: 74 on finiteness being a non-categorical notion) which includes a SM but is still typically found in dependent clauses, i.e. the subjunctive or a conditional/consecutive tense. In at least Ishenyi and Makwe the constructions appear to be reserved for core modal notions. Swahili similarly allows for a



choice between an infinitive, a subjunctive or a consecutive tense after the auxiliary **-wez-** 'be able', but further research is needed to assess what determines the choice. LÉONARD (1980) argues that the consecutive (marked by **-ka-**) implies a lower level of likelihood than suggested by the infinitive (or the subjunctive for that matter), which does point towards an epistemic and therefore core-modal reading. Following LÉONARD (ibid: 218-219), the use of the consecutive (74b) rather than the infinitive (74a) after the auxiliary **-wez-** in (74), taken from the late Tanzanian president Nyerere's tract on socialism, implies that it is less likely for a rich person to have a "socialist soul" than for a poor person to be a capitalist at heart.

(74) Standard Swahili (G42; LÉONARD 1980: 218-219)

a. **Maskini** **wa-na-wez-a** **ku-w-a** **na** **roho** **za** **ki-bepari …**
2a.poor SM2-PRS-POSB-FV INF-be-FV with 10.heart CONN10
'Poor men can have the souls of capitalists …'

b. **Vilevile** **Tajiri** **a-na-wez-a** **a-ka-w-a** **na**
also 1.rich SM1-PRS-POSB-FV SM1-CONS-be-FV with
roho **ya** **u-jamaa**
9.heart CONN9 11-socialism
'Likewise, a rich man can have a socialist spirit'

Another exception is the more complex construction with both an infinitive copula and the predicate verb in the "*conjunctif*" verb form discussed for Rundi epistemic possibility constructions (§3.2.1).

Auxiliary verb constructions devoted to marking necessity vary. Often, the predicate verb is obligatory or in some cases optionally inflected in the subjunctive verb form rather than in the infinitive. This is most straightforwardly connected to the fact that the subjunctive verb form in itself is also often used to form modal expressions of necessity (see §4.1.2), which is a rarer, if at all possible, strategy for expressing possibility cross-Bantu (cf. NURSE and DEVOS 2019).

A significant tendency in the formation of the category of necessity in EA Bantu is the constructional shift in the morphology of the auxiliary verb, between PI and other non-PI flavors. Thus, PI necessity tend to be expressed with an active transitive verb where the subject of the clause also harmonies in semantic terms with the first participant of the SoA, accordingly indexed as the subject marker on the verb. Non-PI flavors, viz PE/deontic and in some rare cases epistemic flavors, are instead expressed with a verb (often the same verb as that employed for PI necessity) somehow associated with the demoting of the first participant of the SoA. As seen, languages vary within themselves and between each other in

the strategies they use for morpho-syntactically coding this type of downgrading of the first participant, and so congenially also marking the fact that the forces – the needs, obligations or necessities, affecting the 1st participant are external to him/her/them (or it). Broadly, three such strategies have been encountered in our survey.

1) The modification of valency of the auxiliary verb by passive or middle voice suffixation, in which case the 1st participant is deprived of its agentivity but is still indexed on the auxiliary as the subject. Examples include **-pas-w-** in Swahili (with the passive suffix) and **-eny-ek-** in Logooli (with the neuter suffix).

2) No shifts in the valency of the verb but instead a shift of the 1st participant from being indexed as a subject to an object on the verb. In these cases, the (obligatory) subject marker slot is filled with an expletive or dummy morph, typically recruited from class 9 as in Swahili (see examples (45) and (47) in §3.5.1) and Mwera (example (61) in §3.7), although examples from other classes also exist, e.g. in Nyamwezi, where it seems that the plural class 8 SM is used, as in example (54) in §3.6.

3) The combination of 1 and 2 insofar that there is de-transitive voice marking on the verb as well as an expletive occupying the subject marker slot. In these cases the 1st participant is thus not indexed anywhere on the verb. Example (60) in §3.7 with SM9**-londeka** in Manda is a case in point. With regard to the choice of subject marker in these constructions, in Logoori, the use of different dummy subject markers from either noun class 9 **e-** or 6 **ga-** even triggers different readings of the proposition including difference in modal force. Thus, (75b), expressed with an expletive **ga-** would convey stronger obligation than (75a) expressed with **e-**. See GLUCKMAN and BOWLER (2016), GLUCKMAN et al. (2017) for further elaboration on this fact.

(75) Llogoori (JE41; GLUCKMAN and BOWLER 2017: 14)

a.	**e-duk-an-a**	**ndee**	**u-zi-ɛ**	**m-skolu**
	SM9-NEC-REC-FV	that	SM2SG-go-SBJV	18-9.school
	'You should go to school'			
b.	**ga-duk-an-a**	**ndee**	**u-zi-ɛ**	**m-skolu**
	SM6-NEC-REC-FV	that	SM2SG-go-SBJV	18-9.school
	'You must go to school'			

To this category we would also assign constructions with no subject marking at all, such as the Portuguese loans discussed in section §4.2, or the constructions in the Luhya cluster expressed with the modal verb in the infinitive (see example (23) in §3.2.4).

The strategy of an expletive subject marker only with the verb modified with some de-transitive voice marking suffix significant for Type 3 constructions exists also for possibility verbs (see SCHADEBERG and BOSTOEN 2019 for more information on these various suffixes, commonly referred to as "extensions", in Bantu). Examples include **-hotol-** derived with the neuter **-ɪk-** to yield **-hotol-ek-** 'be possible (that)' in Manda (see example (76) in the next paragraph for an illustration of its use); Rundi **-shobok-** which is an intransitive variant of **-shobor-**, the final **/ok/** and **/or/** in these verb bases representing the intransitive and transitive pair of the commuting separative suffixes; Swahili **-wezekan-** which is actually a combination of two valency-decreasing suffixes, namely the neuter **-ɪk-** and the associative **-an-**.

Although all these strategies are clearly functionally related, some of them pose a problem in formal terms in relation to our delimitation to auxiliary verb constructions. The structural challenge described here and how it relates to the notion of auxiliary verb constructions definitely merits further investigations. In a simplified manner, however, we would treat these various constructions as on a continuum where we feel that category 1) and, more reluctantly so, category 2), falls within our definitional limits of an auxiliary. Category 3) does not, mainly because being translatable to "it is possible/necessary (that)' it is more reminiscent of a bi-clausal structure where the modal marker does not only operate on the single action/event of a second verb but over a whole matrix clause. This is perhaps most visible in examples like (22), in §3.2.3, and (54), in §3.6, from Llogooli and Nyamwezi, respectively, where an explicit subject noun (i.e. the first participant) interfere between the modal verb and the predicate verb. The Llogori examples above and example (76) from Manda furthermore contain an explicit complementizer.

(76) Manda (N11; BERNANDER *field notes*)

i-hotól-ék-a	**kukɪta**	**va-na**	**va-gón-íli**	**hɪnu**
SM9-POSB-NEUT-FV	COMPL	2-children	SM2-sleep-FV	now

'It is possible that the children are sleeping now'

The regular shift within the necessity paradigm of what is often the same auxiliary verb from a plain form to a passivized or at least intrasitivized variant, has seldom been discussed in the typological literature and might thus be a specific trait significant for (EA) Bantu languages. That passive and middle voice markers are cross-linguistically employed as markers of modality has been pointed out by e.g. NARROG (2012: ch.7). However, he discusses such a function with reference to the morphemes themselves, i.e. in their production with any

verb root.[26] Here, we are instead dealing with the recurrent application of a passive or middle marker as an extra modification of an auxiliary verb already carrying a specialized modal function. The result is the regular shift in the coding of the auxiliary verb accompanying the semantic shift from a participant-internal orientation to what in some way or another is imposed to the 1st participant externally (through circumstances outside the participant's control, whether objectively or subjectively construed).

4.2 On the role of contact

Contact has played a significant part in the formation of EA Bantu modal systems at large, also including the genesis of modal auxiliary verbs. This phenomenon is in line with general cross-linguistic tendencies where the domain of modality is more open to the influx of external influences than to those of tense and aspect (cf. MATRAS 2007; FRIEDMAN 2012). The influx of contact-induced modal verbs could probably in part be explained by the fact that modality is more lexical-like and semantically weighty compared to other markers of the wider TAM-domain (as already mentioned at the outset in §1), lexical items generally being more easily borrowed than grammatical markers. At the same time, the most mature core modal concepts, i.e. deontic and epistemic modality, are also sensitive for contact-induced transfer as they can be taken to belong to MATRAS' (1998, 2020) category of "utterance modifiers", highly (inter-)subjective markers of the pragmatically dominant language to which bilingual speakers tend to switch in order to maintain the assertive authority during conversational interaction. However, in these cases the most prominently transferred markers are not (auxiliary) verbs but rather adverbs and other types of invariable elements. The perhaps most commonly attested one in EA Bantu is the Swahili adverbs/discourse markers expressing core necessity **lazima** 'must' (in turn copied from Arabic into Swahili)[27]. It was attested for Nata already by MEKACHA (1993: 157,161) and is also described as an actively used modal marker in Manda (BERNANDER 2017: 303-304), Nyakyusa (Persohn p.c.), Matuumbi (ODDEN 1996), Tanzanian Nyanja (Ngonyani 2020), Runyambo (Nyinondi and Lusekelo 2020), Logooli (GLUCKMAN and BOWLER 2020) and other Luhya varieties

[26] In which case the modal reading is actually also often different as it more readily seems to target the modal domain of possibility see also the discussion of the "passive-facilitive" from a cross-linguistic perspective in KEMMER (1993), and especially with regard to the neuter in Bantu in DOM et al. (2016, 2018).

[27] In fact, necessity expressions (of stronger force) seem to have been transferred from Arabic to several languages across the globe (see MATRAS 2007).

(GLUCKMAN and BOWLER 2017), and Rufumbira (JD61; NASSENSTEIN 2019: 141).[28] See also GLUCKMAN and BOWLER (2020), also Bernander (forthcoming), on the curious adaptation (replicated from Luo?) of the Swahili loan **mpaka** into a marker of necessity; see also MOUS (2020).

The transfer of the Swahili possibility auxiliary verb **-wez-** 'can' appears as a clear case of Swahili-induced change (or "Swahilization) in several of the surveyed Bantu languages of Eastern Africa. As mentioned in §3.5. the verb in Swahili is in itself borrowed from Arabic. Another minor attendance of this sort, worth mentioning, is the use of **-khitiri** and **-kidiri** for the expression of possibility in Koti (P311; SCHADEBERG and MUCANHEIA 2000: 223) and Mwani (G403; Devos field notes), respectively. The word originates from Arabic **qadara** 'can, be able, appreciate' (SACLEUX 1939: 355). As it does not appear to be regularly used in Standard Swahili, however, it must have been used in the Southern Swahili dialect giving rise to the mixed Koti and Mwani varieties. Alternatively, these varieties borrowed the form directly from Arabic.

Except for Swahilization, modal auxiliary verbs have also been transferred to Bantu languages due to contact with Nilotic at the northern fringes of the Bantu borderland.[29] BERNANDER (forthcoming) compares the Luhya (JE31) modal markers (GLUCKMAN et al. 2017) with those found in Luo (TUCKER 1994: 469-470) and reaches the conclusion that the Luhya language varieties most likely borrowed the possibility auxiliary **-nyal-** 'be able' presented in §3.2 from Luo **nya-lo** with the same meaning (see also BOTNE 2004 who makes this connection for Lusaamia JE34). However, as also indicated in §3.2.3, once integrated into Luhya, its functional range seems to have expanded to also include epistemic possibility unlike in the source language. Another case of a borrowed modality verb is **-(i)dima** 'be able, overcome' which have been transferred to a broad area of languages including the Chaga varieties and various North-East Coast Bantu languages from Eastern Nilotic Maasai (NURSE 1979: 285, 509-511; see also, e.g. KOTZ 1909: 32; HOHENBERGER 1929-1930: 198; RAUM 1909: 274). Nurse (ibid) notes that this transfer must be relatively recent as the Maasai are claimed to have been in the area for no longer than 200-300 years. Although reflexes of **-(i)dima** generally appears to be used for expressions of pre-core dynamic ability, it has at times also been extended in use compared to its source language.

[28] In Rufumbira, **lazima (~razima)** is integrated into a verb-like construction with what appears to be the quotative verb **-ti** (ibid: 141, BERNANDER forthcoming).

[29] To the best of our knowledge there exist no transferred (modal) auxiliary verbs from Cushitic nor the isolate/Khoisan varieties Hadza and Sandawe also in contact with EA Bantu languages.

At least in Kasigau (E74c) where it even appears to have developed post-modal functions of future tense.

(77) Kasigau (E74c; NURSE 1979: 510)
e-dem-a **wa-kab-a**
SM1-FUT-FV ?-hit-FV
'He will hit'

Finally, an influx of markers within the modal domain can also be traced to influences from colonial languages, what appears to most prominently be transfer from Portuguese into Mozambican Bantu varieties. DEVOS (2008a) reports that Shangaci expresses participant-external possibility with **pode** borrowed from Portuguese **poder** 'be able to, be allowed to, may'. Interestingly, DEVOS (2008a) notes that although the Portuguese source originally is a verb it has not been incorporated in the verbal morphotaxis of Shangaci (unlike other Portuguese loan verbs) but acts more like modal particle.

(78) Shangaci (P312; DEVOS 2008a: 10)
pode **o-mw-úrereex-a**
NEG.POSB INF-OM1-eat.with.tea-FV
'You can eat (it) with tea'

The negative equivalent is marked by adding the Portuguese negative particle **não**. A similar case is reported for another Mozambiquan language, Ngoni N12x (Kröger forthcoming), where the nasalized /o/ of **naõ** has been adapted to /m/.

(79) Mozambican Ngoni (KRÖGER forthcoming)
Nampode **ne'** **ku-jenda** **wangali** **li-hona**
NEG.POSB I INF-go without 5-tobacco
'I can't go without tobacco'

In a similar manner of (non-)integration, the Portuguese modal auxiliary verb **ter que** 'have to' has been borrowed and phonologically adapted as a marker of necessity in Cuwabo (P34).

(80) Cuwabo (P34; GUÉROIS 2015: 387)
tinyá **ke-a-ved-úw-e** **áttu**
NEC SM2-look.for-PASS-SBJV 2.people
'They had to look for people'

As was already noted in §4.1.4, the non-integrative character of these Portuguese-originating modal verbs make them resemble modal markers of Category 3, viz. markers originating in verbs but which do not act like auxiliaries.

5. Concluding remarks and directions for further research

This paper has offered a survey of the use of auxiliary verb constructions for expressing modality – here conceptualized as expressions of possibility and necessity – in East African Bantu. That is, in languages spoken from eastern Congo in the north-west to northern Mozambique in the south-east and covering zones JD, JE, E, F, G, M, N and P. For this group of languages, it has been shown that contrary to many traditional standpoints, there does exist a bewildering number of modal verbs.

In many ways, the various language-internal modal systems behave quite similarly. Often, the auxiliary verbs are recruited from lexical sources with analogous meanings, in turn often adhering to cross-linguistically common sources. Interestingly, however, the verbs themselves are seldom cognate which would point towards shared cognitive patterns rather than shared innovations. With that said, some modal verbs derived from cognate lexemes such as reflexes of ***-many-**, ***-dond-** and ***-pú-(an-)** appear across diffuse parts of the surveyed area. Furthermore, some shared cognates do also appear in delimited regions indicative of shared innovation or diffusion, e.g. reflexes of ***-còbʊd-** and ***-báac-** in the western parts of Great Lakes Bantu, **-dah-** in a large part of central Tanzania and **-hotol-** in the very southern parts of Tanzania. Some of the shared modal auxiliaries are in turn externally transferred loans, from Nilotic (**-nyal-** and **-(i)dim-**) or from the highly influential Swahili in the case of **-wez-** (in turn copied into Swahili from Arabic).

The verbs also occur in largely the same constructional frames. The predicate verb, i.e. the verb depicting the main event of the proposition tends to occur in its infinitival form or, occasionally in necessity constructions, in the subjunctive. In some cases, the infinitival prefix may be dropped, indicative of further maturation of the construction along the route to univerbation and morphologization. This factor aside, however, it is striking how the modal verbs in Eastern Bantu seldom, if at all, seem to expose any formal indications of being grammaticalized. This stands in contrast to many European languages like in English where modal verbs often are argued to form a class of their own with characteristic formal features which set them apart from other (auxiliary) verbs (cf. COATES 1983: 4). From a Bantu-internal perspective it also makes the modal verb constructions differ significantly from other auxiliary constructions marking tense and aspect (as well as other functions) which are renown for recurrently and rapidly undergoing formal reductions of univerbation and morphologization (NURSE 2008: 25, ANDERSON 2011, GÜLDEMANN 2003). At the same time, this tendency found in East African Bantu adheres to more general claims in which modal verbs are seen as semantically more weighty, belonging to

a category settled within the lexico-grammatical interphase rather than a clear-cut grammatical domain.

Another typologically striking feature identified in this survey is the regular manipulation of what is essentially the same verb base in order to "rob" the subjecthood and agentivity of the 1st participant, and thus indicating the modal force as participant-external and imposed. Often such constructions also denote a more speaker-oriented (subjective) rather than a participant- and event-oriented (objective) expression. This regular difference in constructional make-up merits a more elaborate investigation and needs to be further situated within the broader cross-linguistic knowledge of modal constructions.

There are several avenues for further research and ways this study could be extended. First and foremost, it could be expanded to include the rest of the Eastern Bantu branch, and, indeed, the whole Bantu language family. It could also be expanded to include other formal types of modal markers than auxiliary verbs, e.g. adverbs, particles or inflectional – affixational – categories like the subjunctive but also potential prefixes like **-oo-** used in some JD-languages or **-nga-** which occurs in several N-languages. This additional category of modal markers would also include erstwhile lexical verbs which have been grammaticalized to a new word class, e.g. the epistemic possibility adverbial **manya(yi)** in N10-languages, and **manyango~manyanga** in Rundi[30], both transparently derived from the lexical verb **-many-** 'know' (i.e. the same source which has also been recruited to an auxiliary verb in some of the languages studied). Elaborating on such items and their diachronic trajectories, would enable a more sophisticated understanding of whether a verb used as a modal marker also qualifies as an auxiliary verb or not.

With that said, it is clear that comparative research on Bantu modality like this one, is still severely impacted by the lack of comprehensive investigation into the modal domain in individual languages. It is our sincere hope that this study in its infancy and rudimentary setting, can provide a helpful overview of the subject matter along with a thorough and up-to-date bibliographic account of work on modality in East African Bantu. To this end, we hope it acts as impetus and inspiration for other scholars to take up or to continue the study of the

[30] See also MBERAMIHIGO et al. (2016) for this curious case of de-grammaticalization of a modal verb construction in Rundi in which a source construction consisting of the modal verb **-meny-** 'know' + quotative **ngo** inflected in sm2sg, viz. 'you know that', was reanalyzed and underwent univerbation into a modal adverb **umenyago~umenyaga** 'probably, apparently'. From there, it further expanded into a lexical verb **-menyag-** 'believe, judge, doubt'.

expression of modality and the use of modal auxiliary verbs in Bantu languages, and indeed more widely.[31]

Abbreviations

Abbreviations follow the Leipzig Glossing Rules with the following additions: 1, 2, 3 = noun classes 1, 2 3 etc.; ASSOC = associative; CJ = conjunctif; CONN = connective; DISJ = disjoint; OM = object marker, NEUT= neuter; SBST = substitutive, SM = subject marker.

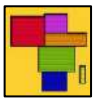
Acknowledgements

We would like to thank the editors and the two anonymous reviewers for helpful comments on an earlier version of the paper. We would also like to thank Heidrun Kröger for sharing her dissertation draft on Mozambique Ngoni and Bastian Persohn for sharing information on Nyakyusa.

The first author would like to thank all Manda and Western Serengeti language consultants. He also wishes to acknowledge his fellow team members of the Mara Project (funded by the Kone Foundation), for help with collecting and analyzing the data on modality in Western Serengeti, as well as the PK Hedborg Foundation and the Salén Foundation whose generous support partly funded this work.

The usual disclaimer applies, and any errors naturally remain our own.

[31] In connection to this plea, we wish to mention Vander KLOK's (2014) methodological contributions to the elicitation of modality. Many of the latter Bantu studies on modality (AUNIO et al. 2022, BERNANDER 2017, KAWALYA et al. 2018 and GLUCKMAN and BOWLER 2017, 2020) draw on this work.



References

ANDERSON, Gregory (2006). *Auxiliary verb constructions*, Oxford: Oxford University Press.

ANDERSON, Gregory (2011). "Auxiliary verb constructions in the languages of Africa", *Studies in African Linguistics* 40(1-2): 1-409.

ASHTON, Ethel O. (1947). *Swahili grammar (including intonation)*, London: Longmans, Green & Co.

AUNIO, Lotta, Antti LAINE, Tim ROTH, Rasmus BERNANDER (2022). *A grammar of Western Serengeti.* Ms.

BARLOW, A. Ruffell. (1951). *Studies in Kikuyu grammar and idiom*, Edinburgh: William Blackwood.

BASTIN, Yvonne. (2003). "The interlacustrine Zone (Zone J)", in Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu Languages*, London & New York: Routledge, 501-528.

BASTIN, Yvonne, André COUPEZ, Évariste MUMBA and Thilo SCHADEBERG. (eds.). (2002). *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3.* Tervuren: Royal Museum for Central Africa, online database.
<http://www.africamuseum.be/collections/browsecollections/humansciences/blr>

BERNANDER Rasmus (Forthcoming). "Language contact and convergence", in Lutz MARTEN, Nancy KULA, Jochen ZELLER and Elisabeth HURST (eds.), *The Oxford Guide to the Bantu Languages*, Oxford University Press.

BERNANDER, Rasmus (2016). "On the modal verb -hotol- in Manda (N.11)", presentation given at the 6th International Conference on Bantu languages, Helsinki 20-23 June.

BERNANDER, Rasmus (2017). Grammar and grammaticalization in Manda: An analysis of the wider TAM domain in a Tanzanian Bantu language, PhD thesis, University of Gothenburg.

BERNANDER, Rasmus (2018). 'The neuter in Manda, with a focus on its reinterpretation as passive', *Southern African Linguistics and Applied Language Studies* 36(3): 175-196.

BERNANDER, Rasmus (2020). "On the "atypical" imperative verb form in Manda", *Studia Orientalia Electronica. Special issue: 'SemVerbMorph1'.* 8(3): 22-42.

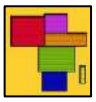
BOSTOEN, Koen, Ferdinand MBERAMIHIGO, and Gilles-Maurice DE SCHRYVER (2012). "Grammaticalization and subjectification in the semantic domain of possibility in Kirundi (Bantu, JD62)", *Africana Linguistica* 18: 5-40.

BOSTOEN, Koen and Rozenn GUÉROIS (Forthcoming). "Reconstructing suffixal phrasemes in Bantu verbal derivation", in Koen BOSTOEN, Gilles-Maurice DE SCHRYVER, Rozenn GUÉROIS, and Sara PACCHIAROTTI (eds.), *Reconstructing Proto-Bantu Grammar*, Berlin: Language Science Press.

BOTNE, Robert. (2004). "Specificity in Lusaamia infinitives", *Studies in Language* 28: 137-164.

BOTNE, Robert (2008). A Grammatical Sketch of Chindali (Malawian Variety). (Lightning Rod Press Series, 2.), Philadelphia: American Philosophical Society.

BOTNE, Robert (2020). "Evidentiality in African Languages", in Chungmin LEE and Jinhho PARK (eds.), *Evidentials and Modals*, Leiden: Brill, 460-501.



BOTNE, Robert Dale, Hannington OCHWADA, and Michael R. MARLO (2006). *A grammatical sketch of the Lusaamia verb*, Köln: Rüdiger Köppe Verlag.

BYBEE, Joan, Revere PERKINS, and William PAGLIUCA (1985). "Cross-linguistic comparison and the development of grammatical meaning", in Jacek FISIAK (ed.), Historical semantics, historical word formation, Berlin: Mouton, 59-83.

BYBEE, Joan, Revere PERKINS, and William PAGLIUCA (1994). *Evolution of grammar: tense, aspect, and modality in the languages of the world*, Chicago: University of Chicago Press.

COATES, Jennifer (1983). *The Semantics of Modal Auxiliaries*, London: Croom Helm.

CRANE, Thera and Bastian PERSOHN (2019). "What's in a Bantu verb? Actionality in Bantu languages", *Linguistic Typology* 23(2): 303–345.

CRANE, Thera Marie, Hilde GUNNINK, Ponsiano KANIJO and Tim ROTH. Forthcoming. "Aspect and Evidentiality in Four Bantu Languages", in Astrid DE WIT, Frank BRISARD, Carol MADDEN-LOMBARDI, Michael MEEUWIS and Adeline PATARD (eds.), *Beyond Time*, Oxford: Oxford University Press.

DAVY, Jim and Derek NURSE. (1982). "Synchronic versions of Dahl's law: The multiple applications of a phonological dissimilation rule", *Journal of African Languages and Linguistics* 4: 157-195.

DEVOS, Maud (2008a). "The expression of modality in Shangaci", *Africana Linguistica* 14: 3-37.

DEVOS, Maud (2008b). *A grammar of Makwe*, Munich: Lincom Europa.

DOM, Sebastian, Leonid KULIKOV and Koen BOSTOEN (2016). "The middle as a voice category in Bantu: Setting the stage for further research", *Lingua Posnaniensis* 58(2): 129-149.

DOM, Sebastian, Leonid KULIKOV and Koen BOSTOEN (2018). "Valency-decreasing derivations and quasi-middles in Bantu: A typological perspective", *Southern African Linguistics and Applied Language Studies* 36(3): 165-173.

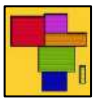
EHRET, Christopher (1998). An African Classical Age: Eastern and Southern Africa in World History, 1000 B.C.~to A.D.~400, Charlottesville: University Press of Virginia.

EHRET, Christopher (1999). "Subclassifying Bantu: The evidence of stem morpheme innovation", in Jean-Marie HOMBERT and Larry HYMAN (eds.) *Bantu Historical Linguistics: Theoretical and Empirical Perspectives*, Stanford: CSLI Publications, 43-147.

FARNSWORTH, Emma (1957). *Kamba Grammar*, Nairobi: Published under the Auspices of the Africa Inland Mission.

FELBERG, Knut (1996). *Nyakyusa-English-Swahili and English-Nyakyusa dictionary*, Dar es Salaam: Mkuki na Nyota Publishing.

FRIEDMAN, Victor (2012). "Tense-Aspect and Language Contact", in Robert BINNICK (ed.), *The Oxford Handbook of Tense and Aspect*, Oxford: Oxford University Press, 398-427.



GLUCKMAN, John and Margit BOWLER (2016). "Expletive agreement, evidentiality, and modality in Logooli", in Mary MORONEY, Carol-Rose LITTLE, Jacob COLLARD and Dan BURGDORF (eds.), *Proceedings of the 26th semantics and linguistic theory conference:* 1063-1082. <https://doi.org/10.3765/salt.v26i0.3935>

GLUCKMAN, John, Margit BOWLER, Maurice SIFUNA, and Michael DIERCKS (2017). *Modality in Luhya: A typological study*, Poster presented at LSA, January 6, 2017.

GLUCKMAN, John and Margit BOWLER (2020). "The expression of modality in Logoori", *Journal of African Languages and Linguistics*, 41(2): 195-238. <https://doi.org/10.1515/jall-2020-2010>

GROLLEMUND, Rebecca, Simon BRANFORD, Koen BOSTOEN, Andrew MEADE, Chris VENDITTI, and Mark PAGEL (2015). "Bantu expansion shows that habitat alters the route and pace of human dispersals", *Proceedings of the National Academy of Sciences of the United States of America* 112 (43): 13296-13301.

GUÉROIS, Rozenn. 2015. *A grammar of Cuwabo (Mozambique, Bantu P34)*, Ph.D. thesis, Lyon: Université Lumière Lyon 2.

GÜLDEMANN, Tom. (2003). *Grammaticalization*, in Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu languages*, London: Routledge. 182-194.

GUTHRIE, Malcolm (1948). *The classification of the Bantu languages*, London: Oxford University Press.

GUTHRIE, Malcolm (1967-1971). Comparative Bantu: an introduction to the comparative linguistics and pre-history of the Bantu languages, London: Gregg.

HAMMARSTRÖM, Harald (2019). "An inventory of Bantu languages", in Mark VAN DE VELDE, Koen BOSTOEN, Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu languages,* London: Routledge, 17-78.

HARJULA, Lotta (2004). *The Ha language of Tanzania*, Köln: Rüdiger Köppe.

HARRIES, Lyndon (1950). *A Grammar of* Mwera, Johannesburg: Witwatersrand University Press.

HEINE, Bernd (1993). *Auxiliaries: Cognitive forces and grammaticalization*, New York: Oxford University Press.

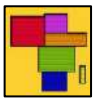
HENDLE, P.J. (1907). Die Sprache der Wapogoro (Deutsch-Ostafrika) nebst einem Deutsch-Chipogoro und Chipogoro-Deutschen Wörterbuch. (Archiv für das Studium der deutschen Kolonialsprachen, 6.), Berlin: Georg Reimer.

HETHERWICK, Alexander (1916). *A practical manual of the Nyanja language*, Glasgow & Blantyre: African Lakes Corp.

HINNEBUSCH, Thomas H. (1981). "Northeast Coastal Bantu", in Thomas H. HINNEBUSCH, Derek NURSE and Martin MOULD (eds.), *Studies in the Classification of Eastern Bantu Languages*, Hamburg: Helmut Buske, 1-126.

HOHENBERGER, J. (1929-1930). "Zur Lautlehre und Grammatik des Asu", *Zeitschrift für Eingeborenensprachen* XX: 175-208.

JONSSON, Erland (1954). *Kinyang'wezi grammatik*, Tabora, Tanganyika: Svenska Pingstmissionen.



KAGAYA, Ryohei (1987). "Tonal patterns of Cilungu predicate verbs", in Yasutoshi YUKAWA (ed.) *Studies in Zambian languages*, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, 379-472.

KAWALYA, Deo, Koen BOSTOEN, and Gilles-Maurice DE SCHRYVER (2014). "Diachronic semantics of the modal verb -sóból- in Luganda", *International Journal of Corpus Linguistics* 19: 60-93.

KAWALYA, Deo, Koen BOSTOEN, and Gilles-Maurice DE SCHRYVER. 2021. "A Diachronic Corpus-Driven Study of the Expression of Possibility in Luganda (Bantu, JE15)", *International Journal of Corpus Linguistics* 26 (3): 336–369.

KAWALYA, Deo, Gilles-Maurice DE SCHRYVER, and Koen BOSTOEN (2018). "Reconstructing the origins of the Luganda (JE15) modal auxiliaries sóból and yînz: A historical-comparative study across the west Nyanza Bantu cluster", *South African Journal of African Languages* 38(1): 13–25.

KAWALYA, Deo, Gilles-Maurice DE SCHRYVER, and Koen BOSTOEN (2019). "A corpus-driven study of the expression of necessity in Luganda (Bantu JE15)", *Southern African Linguistics and Language Studies* 37: 361–381.

KEMMER, Suzanne. 1993. *The middle voice*, Amsterdam: John Benjamins.

KEZILAHABI, Euphrase (1988). *Rosa Mistika*, Dar es Salaam: University Press.

KIHARA, C. Patrick (2017). Aspects of Gĩkũyũ (Kikuyu) Complex Sentences: A Role and Reference Grammar Analysis, PhD thesis, Heinrich-Heine Universität –Düsseldorf.

KO, So Young (2014). *The Vinza language of Tanzania (Bantu, JD67),* MA thesis, Universiteit Gent.

KOTZ, Ernst. (1909). Grammatik des Chasu. (Archiv für das Studium deutscher Kolonialsprachen, 10.) Berlin: Dietrich Reimer.

KRATZER, Angelika (1981). "The notional category of modality", in Hans-Jurgen EIKMEYER and Hannes RIESER (eds.), *Words, worlds, and contexts: New approaches in word semantics*, Berlin: Mouton de Gruyter, 38–74.

KRATZER, Angelika (1991). "Modality", in Armin VON STECHOW and Dieter WUNDERLICH (eds.), *Semantics: An international handbook of contemporary research*, Berlin: Mouton de Gruyter, 639–650.

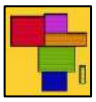
KRUG, Manfred (2011). "Auxiliaries and grammaticalization", in Heiko NARROG and Bernd HEINE (eds.), *The Oxford Handbook of grammaticalization,* Oxford: Oxford University Press, 547-558.

KRÖGER, Heidrun (Forthcoming). *A descriptive Grammar of Mozambican Ngoni*, PhD thesis, Goethe Universität Frankfurt.

LEONARD, Robert A. (1980). "Swahili *e*, *ka* and *nge* as Signals of Meanings", *Studies in African Linguistics* 2-2: 209-226.

LINDBLOM, Gerhard (1926). *Notes on Kamba Grammar. (Archives d'Études Orientales, 10.)*, Uppsala: Appelbergs Boktryckeri.

MADAN, A. C. and E. STEERE (1890). *A Handbook of the Swahili Language as spoken in Zanzibar*, London: Society for the Promotion of Christian Knowledge.



MAHO, Jouni F. (2003). "A Classification of the Bantu Languages: An Update of Guthrie's Referential System", in Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu Languages,* London and New York: Routledge, 639-651.

MAHO, Jouni F. (2009). NUGL Online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages, Gothenburg: Department of Oriental and African Languages.

MATRAS, Yaron (1998). 'Utterance Modifiers and Universals of Grammatical Borrowing', *Linguistics* 36: 281–331.

MATRAS, Yaron (2007). "The borrowability of structural categories", in Yaron MATRAS and Jeanette SAKEL (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*, Berlin, New York: Berlin: Mouton de Gruyter, 31-74.

MATRAS, Yaron (2020). *Language Contact [2nd ed.]*, Cambridge: Cambridge University Press.

MBERAMIHIGO, Ferdinand (2014). *L'expression de la modalité en kirundi. Exploitation d'un corpus électronique*, PhD thesis, Brussels/Gent: Université libre de Bruxelles (ULB), Université de Gand (UGent).

MBERAMIHIGO, Ferdinand, Gilles-Maurice DE SCHRYVER, and Koen BOSTOEN (2016). "Entre verbe et adverbe : Grammaticalisation et dégrammaticalisation du marqueur épistémique umeengo/umeenga en kirundi (bantou, JD62)", *Journal of African Languages and Linguistics*, 37(2): 247-286.

MCHOMBO, Sam. (2004). *The Syntax of Chichewa*, Cambridge: Cambridge University Press.

MEKACHA, Rugatiri D.K. (1993). The sociolinguistic impact of Kiswahili on ethnic community languages in Tanzania: a case study of Ekinata, Bayreuth: Bayreuth University.

MIESTAMO, Matti (2005). Standard negation: The negation of declarative verbal main clauses in a typological perspective, Berlin: Mouton de Gruyter.

MOHAMED, Abdulla Mohamed (2001). *Modern Swahili Grammar*, Nairobi/Kampala/Dar es Salaam: East African Educational Publishers.

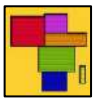
MORRISON, Michelle Elizabeth (2011). *A reference grammar of Bena*, PhD thesis, Houston: Rice University.

MOUS, Maarten (2019). "Language Contact", in Mark VAN DE VELDE, Koen BOSTOEN, Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu Languages*, London and New York: Routledge, 355-380.

MOUS, Marten (2020). "Transfer of Swahili 'until' in contact with East African languages", in Norval SMITH, Tonjes VEENSTRA and Enoch Oladé ABOH (eds), *Advances in Contact Linguistics: In honour of Pieter Muysken*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 218-233.

MUNRO, Pamela (2006). "Modal expression in Valley Zapotec", in William FRAWLEY (ed.), *The expression of modality*, Berlin: Mouton de Gruyter, 177-209.

NAMYALO, Saudah, Alena WITZLACK-MAKAREVICH, Anatole KIRIGGWAJJO, Amos ATUHAIRWE, Zarina MOLOCHIEVA, Ruth MUKAMA, and Margaret ZELLERS (2021). *A dictionary and grammatical sketch of Ruruuli-Lunyala*, Berlin: Language Science Press.



NARROG, Heiko (2005). "Modality, mood, and change of modal meanings: A new perspective", *Cognitive Linguistics* 16(4): 677-731.

NARROG, Heiko (2012). *Modality, subjectivity, and semantic change*, Oxford: Oxford University Press.

NASSENSTEIN, Nico (2015). Kisangani Swahili: Choices and Variation in a Multilingual Urban Space, München: LINCOM.

NASSENSTEIN, Nico (2019). *Speaking with a Difference: Border Thinking in Rufumbira*. PhD thesis, Universität zu Köln.

NASSENSTEIN, Nico and Paulin Baraka BOSE (2016). Kivu Swahili Texts and Grammar Notes. (Languages of the World/Text Collections, 37.), München: LINCOM.

NAUZE, Fabrice (2008). *Modality in typological perspective*, Institute for Logic, Language and Computation, Amsterdam: Universiteit van Amsterdam.

NGONYANI, Deo (2011). *A grammar of Kikisi*. Unpublished manuscript.

NGONYANI, Deo (2020). Chinyanja of Tanzania: A Grammatical Sketch. (Languages of the world / Materials, 514.), München: LINCOM.

NICOLLE, Steve (2013). *A grammar of Digo: a Bantu language of Kenya and Tanzania*, Dallas, Texas: SIL International.

NURSE, Derek (1979). Classification of the Chaga dialects: language and history on Kilimanjaro, the Taita Hills, and the Pare Mountains, Hamburg: Buske.

NURSE, Derek (1999). "Towards a historical classification of East African Bantu languages", in Jean-Marie HOMBERT and Larry HYMAN (eds.) *Bantu historical linguistics: theoretical and empirical perspectives*, Stanford: CSLI, 1-41.

NURSE, Derek (2008). *Tense and aspect in Bantu*, Oxford: Oxford University Press.

NURSE, Derek and Maud Devos (2019). "Aspect, Tense and Mood", in Mark VAN DE VELDE and Koen BOSTOEN and Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu Languages*, London: Routledge, 204-236.

NURSE, Derek and Thomas J. HINNEBUSCH (1993). *Swahili and Sabaki: a linguistic history*, Berkeley: University of California Press.

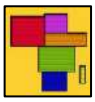
NURSE, Derek and Franz ROTTLAND (1991/92). "Sonjo: Description, classification, history", *Sprache und Geschichte in Afrika* 12/13: 171-289.

NUYTS, Jan (2006). "Modality: Overview and linguistic issues", in William FRAWLEY (ed.), *The Expression of Modality,* Berlin: Mouton de Gruyter, 1-26.

NUYTS, Jan (2015). "Subjectivity: Between discourse and conceptualization", *Journal of Pragmatics* 86: 106-110.

NUYTS, Jan (2016). "Analyses of the modal meanings", in Jan Nuyts and Johan VAN DER AUWERA (eds.), *The Oxford handbook of modality and mood,* Oxford: Oxford University Press, 31-49.

NYINONDI, Onesmo and Amani LUSEKELO (2020). "Lexical verbs expressing modality in Bantu languages: the case of Runyambo and Luguru", *Journal of West African Languages* 47(2): 63-84.



ODDEN, David. (1996). *The Phonology and Morphology of Kimatuumbi*, Oxford: Clarendon Press.

OLSON, Howard S. (1964). *The phonology and morphology of Rimi*, PhD thesis, Hartford: Hartford Seminary Foundation.

PATIN, Cédric, Kassim Mohamed-Soyir and Charles Kisseberth (2019). "Ngazidja G44a", in Mark VAN DE VELDE and Koen BOSTOEN and Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu Languages*, London & New York: Routledge, 591-614.

PETZELL, Malin (2008). The Kagulu Language of Tanzania: Grammar, Texts and Vocabulary, Köln: Rüdiger Köppe.

PETZELL, Malin and Lotta Aunio (2019). "Kami G36", in Mark VAN DE VELDE and Koen BOSTOEN and Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu Languages*, London and New York: Routledge, 563-590.

PIRES PRATA, António (1990). *Dicionário Macua-Português*, Lisboa: Ministério do Planeamento e da Administração do Território, Secretaria de Estado da Ciência e Tecnologia, Instituto de Investigação Científica Tropical

POLAK-BYNON, Louise (1975). A Shi Grammar: Surface Structures and Generative Phonology of a Bantu Language. (Annalen van het Koninglijk Museum van Belgisch-Congo: Reeks in 8, 86.), Belgique: Tervuren.

PRIEBUSCH, Martin (1935). *Bena-Hehe-Grammatik*, Berlin: Kommissionsverlag der Buchhandlung der Berliner Miss.-Ges.

RAUM, J. (1909). *Versuch einer Grammatik der Dschaggasprache (Moschi-Dialekt)*. (Archiv für das Studium deutscher Kolonialsprachen, 11), Berlin: Georg Reimer.

RZEWUSKI, Eugeniusz (1991). "Mother tongue / father tongue convergence: on swahilization and deswahilization in Mozambique", in James R. DOW and Thomas STOLZ, *Akten des 7. Essener Kolloquiums über "Minoritätensprachen/ Sprachminoritäten"*, Bochum: Brockmeyer, 267-305.

SACLEUX, Charles (1939). Dictionnaire swahili-français. (Travaux et mémoires de l'Inst. d'Ethnologie, 36-37.), Univ. de Paris.

SAKAMOTO, Kunihiko (2005). "An Introduction to Kidawida: The Language of the Taita in south-east Kenya", *Bulletin of policy and management, Shobi-Gakuen University* I, II, III. 11-21, 1-11, 1-16.

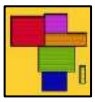
SCHADEBERG, Thilo C. (1992). *A sketch of Swahili morphology*. Cologne: Rüdiger Köppe.

SCHADEBERG, Thilo. C. (2009). "Loan words in Swahili", in Martin HASPELMATH, and Uri TADMOR (eds.), *Loanwords in the world's languages: a comparative handbook*, Berlin: De Gruyter Mouton, 76-102.

SCHADEBERG, Thilo and BOSTOEN, Koen (2019). "Word Formation", in Mark Van de Velde, Koen BOSTOEN, Derek NURSE and Gérard PHILIPPSON (eds.), *The Bantu* Languages, New York: Routledge, 172-203.

SCHADEBERG, Thilo C. and Francisco Ussene MUCANHEIA (2000). *Ekoti. The Maka or Swahili language of Angoche,* Cologne: Rüdiger Köppe.

SCHICHO, Walter (1995). "Modalität und Sprecherintention", in Gudrun MIEHE, and Wilhelm J. G. MÖHLIG (eds.), *Swahili-Handbuch*, Köln: Rüdiger Köppe, 125-166.



SCHOENBRUN, David L. (1990). Early history in Eastern Africa's Great Lakes region: Linguistic, ecological and archaeological approaches, ca. 500 B.C. to ca. 1000 A.D., PhD thesis, Los Angeles: UCLA.

SCHOENBRUN David L. (1994). "Great Lakes Bantu: Classification and settlement chronology", *Sprache und Geschichte in Afrika* 16: 91-152.

SCHOENBRUN, David L. (1997). The historical reconstruction of Great Lakes Bantu cultural vocabulary: Etymologies and distributions, Köln: Rüdiger Köppe.

SEIDEL, August (1898). "Grammatik der Sprache von Irangi nebst Wörterverzeichnis", in C. Waldemar WERTHER (ed.), *Die mittleren Hochländer des nördlichen Deutsch-Ost-Afrika*, Berlin: Hermann Paetel, 387-434.

SHETLER, Jan B. (2003). Telling our own stories: Local histories from South Mara, Tanzania, Leiden: Brill.

TRAUGOTT, Elizabeth Closs and Richard DASHER (2002). *Regularity in semantic change*, Cambridge: Cambridge University Press.

TUCKER, Archibald N. (1994). *A Grammar of Kenya Luo (Dholuo)*, Köln: Rüdiger Köppe.

VAN DER AUWERA, Johan, Petra KEHAYOV and Alice VITTRANT (2009). "Acquisitive modals", in Lotte HOGEWEG, Helen DE HOOP, and Andrej MALCHUKOV (eds.), *Cross-linguistic semantics of tense, aspect, and modality,* Amsterdam/Philadelphia: John Benjamins, 271–302.

VAN DER AUWERA, Johan and Vladimir PLUNGIAN. (1998). "Modality's Semantic Map", *Linguistic Typology* 2: 79-124.

VAN OTTERLOO, Roger (2011). The Kifuliiru Language. Volume 2. A Descriptive Grammar, Dallas: SIL International.

VANDER KLOK, Jozina (2014). Questionnaire on modality for cross-linguistic use, retrieved from:

<https://www.eva.mpg.de/lingua/tools-at-lingboard/pdf/Modal_Questionnaire_CrossLing_JVK.pdf>

VELTEN, Carl (1900). "Die Sprache der Wakami in Deutsch-Ostafrika", *Mittheilungen des Seminars für Orientalische Sprachen* III: 1-56.

VERSTRAETE, Jean-Christophe (2005). "Scalar quantity implicatures and the interpretation of modality. Problems in the deontic domain", *Journal of Pragmatics* 37: 1401-1418.

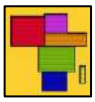
WHITELEY, Wilfred Howell and Muli, M. G. (1962). *Practical Introduction to Kamba*. London: Oxford University Press.

VOORHOEVE, Jan (n.d.). *A grammar of Safwa*. Ms.

WOODWARD, Herbert Willoughby (1882). *Collections for a handbook of the Boondéi language*, London: Society for Promoting Christian Knowledge (SPCK).

WOODWARD, Herbert Willoughby (1902). *Collections for a handbook of the Zigula language,* Msalabani: Universities' Mission, Magila Deutsch-Ostafrika.

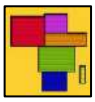
WOODWARD, Herbert W. (1926). "An outline of Makua grammar", *Bantu Studies* 2: 269-325.



WORMS, A. (1897). "Grundzüge der Grammatik des Ki-Zaramo in Deutsch-Ostafrika", *Zeitschrift für afrikanische und oceanische Sprachen* III: 298-310.

YONEDA, Nobuko (2010). "Swahilization" of ethnic languages in Tanzania: the case of Matengo", *African Study Monographs* 31(3): 139-148.

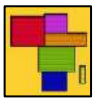
ZIEGELER, Debra (2011). "The Grammaticalization of Modality", in Heiko NARROG and Bernd HEINE (eds.), *The Oxford Handbook of Grammaticalization*, Oxford: Oxford University Press, 595-604.



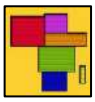
Appendix

The following three tables list the modal auxiliaries discussed in this paper. Table 3 and 5 concern modal auxiliaries used for the expression of possibility and necessity, respectively. Table 4 gives 'hybrid' auxiliaries that can be employed in both domains in particular languages. Within each table, the verbs are grouped following similarity in source meanings (cf. the second column). The starred forms are reconstructions to Proto-Bantu or lower nodes within the Bantu tree (cf. the main text). Concerning the construction types, we make a distinction between two main types of constructions according to the formal characteristics of the main verb. Constructions with infinitival main verbs are referred to as 'compound', whereas constructions with an inflected main verb are referred to as 'complex' (cf. SCHADEBERG 1992 for a discussion of compound and complex verb constructions in Standard Swahili). Next, we also indicate whether the construction involves agent-demoting morphology (typically passive or neuter derivational suffixes) or syntax (typically an expletive subject marker with or without demotion of the first participant to the object marker slot).

*reconstruction/(*reflex*)/<source language	non-modal (source) meaning(s)	languages	modal meanings	construction types
***cotʊd**	'pierce' > 'overcome, defeat'	Manda N11, Matengo N13, Kisi G67	PI, PE, deontic epistemic	compound
***túd**	'pierce'	Ngoreme JE401, Ikoma JE45, Ishenyi JE45, Nata JE45	PI, PE, deontic, epistemic	compound
(dul)	'bore'	Nyamwezi F22	?	compound?
***còbʊd**	'be able'	Rundi JD62, Fuliiru JD63, Ha JD66, Vinza JD67, Luganda Runyoro-Rutooro JE11, Luganda JE15, Lusoga JE16, Lugwere JE17 Haya JE22	PI, PE, deontic epistemic	compound complex (AUX + INF(be) + DEPENDENT)
(yînz)	'be powerful, overcome, manage, control'	Luganda JE15, Lusoga JE16, Lugwere JE17	PI, PE, deontic, epistemic	compound
(nyal) < Luo	'manage'	Lubukusu	PI, PE,	compound



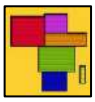
		JE31c, Luwanga JE32a, Lunyore JE33, Lusaamia JE34, Llogoori JE41, Lutiriki JE413	deontic, epistemic	
(dah)	'have power over'	Kagulu G12, Bondei G24, Zigua G31, Zaramo G33, Kami G36, Kilombero G50, Pogolo G51, Nyaturu F32, Langi F33, Nyaturu F32	PI, PE, deontic, epistemic	compound
(nagy)	'have strength, power' (+ causative morphology)	Ngoreme JE402	PI	compound
(wez) < Arabic	'be powerful'	Standard Swahili G42, Kivu Swahili G40x, Makwe G402, Kami G36, Digo E73, Nyamwezi F22, Matengo N13, Bena G63	PI, PE, deontic, epistemic	compound/complex
(wodh)	'defeat, be able'	Cuwabo P34, Makhuwa P31	PI	compound
((i)dim) < Maasai	'be able, overcome'	Dawida E74, Kasigau E74c	PI	compound
(zìgir)	'be able'	Fuliiru JD63	PI	compound
(tony)	'be able'	Kamba E55	PI	compound
(ut̪)	'be able'	Tharaka E54	PI	compound?
(ʃind)	'be able'	Ngazidja G44a	PI?	?
(kithiri, kidiri) < Arabic	'be able, appreciate'	Koti P311, Mwani G403	PI?	compound
(pode) < Portuguese	'be able'	Shangaci P312	PE, deontic epistemic	compound
***màny**	'know'	Ikoma JE45, Kamba E55, Mwera P22, Hehe G62, Bena G63	PI, PE	compound



(kagul)	'know'	Hehe G62, Bena G63	PI, PE	compound
(dzel)	'know'	Hehe G62, Bena G63	PI, PE	compound
(is)	'get to, happen to'?	Kamba E55	epistemic	compound
(long)	'get'	Fuliiru JD63	PI	compound
***pát**	'hold'	Standard Swahili G42	PI	compound
***báac**	'be active'	Runyakore-Rukiga JE13/14, Nyambo JE21, Haya JE22, Rundi JD62, Fuliiru JD63	PI, PE, deontic	compound
(esh)	'become fitting'	Ishenyi JE45	deontic, epistemic	complex
(vag)	'be enough' (+ applicative morphology)	Hehe G62, Bena G63	deontic	compound
(ezy)	?	Lugwere JE17	PI	?
(hɔt)	?	Gikuyu E51	PI, PE, deontic, epistemic	compound
(kuhích)	?	Kerewe JE24, Sinza JE23	?	compound

Table 3 – Possibility auxiliaries

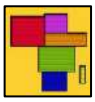
***pát-i**	'hold' (+ causative morphology)	Kivu Swahili G40x	**possibility** deontic **necessity** deontic	compound + participant-demoting morphology compound + participant-demoting morphology
***pát-i**	'hold' (+ causative morphology)	Kisangani Swahili G40x	**possibility** deontic **necessity** deontic	compound (AUX in PRS) compound (AUX in FUT)



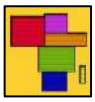
(bagil)	'suit, fit deserve, be eligible'	Ndali M301	**necessity** deontic? (also **possibility?**)	compound
***càk**	'want'	Shangaci P312	**necessity** epistemic **possibility** epistemic	compound compound
***jìj**	'come'	Ndali M302 Safwa M25	**necessity** PE, deontic **possibility** ?	complex (+ participant-demoting syntax) complex (+ participant-demoting syntax)

Table 4 – 'Hybrid' auxiliaries

***pu(-an)**	'be fitting, resemble each other'	Dawida E74, Standard Swahili G42, Bondei G24, Lungu M14 Makhuwa P31	(weak) deontic	complex/compound ? + participant-demoting morphology
***lagil**	'agree, suit, be fitting' (+ causative morphology)	Bondei G24, Zigua G31	PE, deontic	?
(agerer)	'behoove one, be becoming to one'	Gikuyu E51	deontic	complex/compound + participant-demoting morphology
(ail)	'be fitting'	Kamba E55	deontic	compound
(igel)	'be worthy'	Nyamwezi F22	weak strong	complex + participant-demoting (no OM) complex + participant-demoting (with OM)
(noney)	'behoove, deserve, be worthy'	Nyaturu F32	deontic	compound + participant-demoting syntax
(yenel)	'be fitting'	Manda N11		compound + participant-demoting morphology
(wandicil)	'be fitting'	Mwera P22		compound + participant-demoting morphology&syntax



***dònd**	'search for, follow'	Manda N11, Digo	PI PE, deontic	compound compound + participant-demoting morphology/syntax
(dakh) < Swahili ?	'want'	Lusaamia JE34	weak	?
(eny)	'want'	Lubukusu JE31c, Luwanga JE32a, Lunyore JE33, Lusaamia JE34, Llogoori JE41, Lutiriki JE413	PI (weak) PE, deontic, (epistemic)	compound complex + participant-demoting morphology
(end)	'look for, want desire'	Ikoma JE45, Ishenyi JE45, Nata JE45	PI (weak) PE, deontic	compound compound + participant-demoting morphology
(bam)	'want'	Luguru G35	PE, deontic	complex + participant-demoting morphology
(tun)	'need'	Ngoreme JE401	PI (weak) PE, deontic	compound compound + participant-demoting morphology
(tthun)	'want, desire'	Makhuwa P31	deontic?	? + participant-demoting morphology
***pát-i**	'hold' (+ causative morphology)	Standard Swahili G42	**necessity** deontic	compound + participant-demoting morphology/syntax
***pát-ık**	'wedge in'	Ikoma JE45, Nata JE45	(strong) deontic	compound +participant-demoting morphology/syntax
***tég**	'set (trap)'	Rundi JD62	PI, PE, deontic, epistemic	compound (participant-demoting morphology)
(téek)	'make a law, bind by law, edict'	Runyakore-Rukiga JE13/14, Luganda JE15	PE, deontic, epistemic	compound (participant-demoting morphology)
***(dı) nà**	'have'	Luganda JE15, Lungu M14, Shangaci P312	deontic	compound
(duk, tukh, khoy, khoyer, ol, l)	'arrive, reach'	Llogoori JE41, Lutiriki JE413, Lubukusu JE31c, Lusaamia JE34, Lunyore JE33, Luwanga JE32a	PI, PE, deontic, epistemic	compound/complex + participant-demoting syntax



***gì**	'go'	Ndali M302	PE, deontic	complex (+ participant-demoting syntax)
(ti)	'be inevitable'?	Kamba E55	epistemic	compound
(lazimu) < Arabic	'be necessary'	Standard Swahili	PE, deontic	compound/complex + participant-demoting morphology/syntax
(bidi) < Arabic	'be necessary'	Standard Kiswahili G42	PE, deontic	compound/complex + participant-demoting syntax
(nog)	'be/taste good' (+ applicative morphology)	Hehe G62, Bena G63	deontic	compound
***bón**	'see'	Shangaci P312	PE, deontic	compound (+participant-demoting morphology)
***cíad**	'remain'	Shangaci P312	PE, deontic	complex

Table 5 – Necessity auxiliaries

JOURNAL OF AFRICAN LANGUAGES AND LITERATURES
3/2022, 86-105

Microvariation in nominal plurality in Northern Masa

GUILLAUME GUITANG [1] / OUSMANOU [2] / PIERRE DAVOUNOUMBI [3]

[1] Université libre de Bruxelles
guillaume.guitang@ulb.be

[2] University of Yaoundé I
ousmafed@yahoo.fr

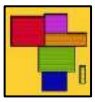
[3] University of Maroua
davounoumbip@yahoo.com

ABSTRACT

This paper describes nominal plurality and examines microvariation in the marking of plural on nouns across three closely related Northern Masa languages, namely in Gizey, Masana, and Musey. These three languages use the same set of nominal plural exponents: **-Vj**, **-ii/-ij**, and **-Vgi** which are reflexes of Proto-Chadic ***-ai**, ***-i** and ***-aki**, respectively (NEWMAN 1990). In addition to suffixal formation, the three languages have restricted sets of suppletive plural nouns in which two further formatives can be identified, namely **-n** and **-u**. Finally, traces of an erstwhile vowel internal ablaut can be observed in a few vestigial plurals. Although these languages constitute a more or less homogeneous lectal continuum, they have developed different plural assignment systems ranging from fully morphological (e.g., Masana) to fully phonological (e.g., Gizey). The three languages also differ in their potential for number-marking (numerality). Generally, nominal plurality seems to be on the decline.

KEY WORDS: microvariation, plural assignment systems, nominal plurality, Masa, numerality

DOI https://doi.org/10.6093/jalalit.v3i3.9145
ISSN 2723-9764



1. Introduction[1, 2]

We know from available descriptions of Northern Masa languages (e.g., CAÏTUCOLI 1983, MELIS 1999, OUSMANOU 2007, DE DOMINICIS 2008, DAVOUNOUMBI 2017) that reflexes of reconstructed Proto-Chadic plural suffixal formatives ***-ai**, ***-i** and ***-aki** (NEWMAN 1990) are well attested in that subbranch. What we lack, however, is an understanding of the productivity of nominal plural formation in general, and of the nature of variation pertaining to this inflectional process in this subbranch. The sources just mentioned do not provide any quantitative indication of the productivity of nominal plural marking. MELIS (1999: 94) only signals the existence of a 'limited number' ('un nombre limité', our translation) of nouns which use a plural suffix, without any quantitative backing. The question of the productivity of plural marking is nonetheless important, especially given the pervasive trend away from morphological marking of grammatical categories within Chadic. For example, some Chadic languages (e.g., the ones under study) have lost morphological marking of gender although one can confidently assume that they had some means of indicating gender morphologically, as part of their Proto-Chadic and Afroasiatic inheritance (NEWMAN 2006). Also, while many contemporary Chadic languages still maintain rich plural marking systems, the erosion of this inflectional process in some branches (e.g., in West Chadic) has been reported (see BLENCH 2021). In West Chadic, the decrease of plural marking correlates to an increase of feminization i.e., the assignment of feminine gender to erstwhile masculine nouns (BALDI and LEGER 2011). The facts we present strongly suggest that the Northern Masa languages examined experience an erosion of number marking in nominals. However, there is no evidence to suggest that this erosion correlates to the feminization of nouns.

Generally, comparative studies highlighting microvariation are rare within Masa. One major exception is MELIS (2019) whose aim, however, has been to provide evidence that Gizey is a distinct language within Masa. Also, MELIS (2019) chiefly focusses on Masana and Gizey phonology.

The present paper addresses these research gaps by fully describing and comparing the systems underlying nominal plurality in three major languages of the subbranch, namely in Gizey, Masana, and Musey. We use four parameters to compare nominal plurality in these languages: a) numerality, b) inventory of markers, c) assignment systems, and d) phonological processes fed by plural

[1] We thank Abbie Hantgan-Sonko, Izabela Jordanoska, and Philippine Geelhand de Merxem for reading our manuscript and making extremely useful suggestions.

[2] The research of Guillaume Guitang is funded by the Université libre de Bruxelles (ULB) under the MINI ARC framework.

marking. These parameters are discussed under §2, §3, §4, and §5, respectively. By 'numerality' we mean, the ability for nouns to mark number morphologically. It appears from examining this parameter that plural marking is on the decline, especially in Masana where only a few nouns mark plural morphologically.

Gizey, Masana, and Musey are part of a language continuum spanning Cameroon and Chad. MELIS (1999, 2006, 2019) classifies the three languages under the Northern subgroup of Masa (Chadic), which is to be distinguished from the Southern subgroup composed of Zimé, Lame, Ngedé, and Mesmé. The Northern subgroup also involves Ham, the Marba-Lew-Monogoy lectal continuum, and Zumaya. MELIS (2006b) has argued for classifying the highly endangered Ham as a distinct language within Masa. Zumaya is now an extinct language (SEIGNOBOS and TOURNEUX 2002, MELIS 2019).

The status of Gizey as a separate language, distinct from Masana, has also been argued for by MELIS (2019). NEWMAN (2013) also lists Gizey as a separate language within Masa. On the contrary, sources like BARRETEAU and DIEU (2005) describe Gizey as a western dialect of Masana. BARRETEAU and DIEU (2005) also include three Central (Baygana, Gagana, and Kayamna) and one Eastern (Gumayna) Masana dialects. OUSMANOU (2007) lists Gizey as a dialect of Masana along with the Yagwa, Muzuk, Walya, Buguɗum, Domo, and Wina varieties. MELIS (1999, 2006a) counts yet another Masana dialect, Harra, spoken in Chad, south of Bongor.

While the nominal plurality facts examined here do not permit to ascertain the place of Gizey within Masa, one can observe some micro-level morphological variation which may provide additional support to MELIS' (2019) claim that Gizey and Masana constitute different languages.

Musey has two main dialect groups: **vùn kúr vòò=na** 'mouth interior house=ART' – 'interior language'; and **vùn ngòò=na** 'mouth bush=ART' – 'bush/exterior language' (BERTONI 2018). These dialect groups are spoken in Chad. In Cameroon, Musey is represented by the Pee variety, which is reportedly influenced by Masana (BERTONI 2018).

Gizey, Masana, and Musey have very little inflectional morphology. Gender is not marked morphologically. However, with the exception of the ones denoting mass, nouns have gender; and nouns distribute according to whether they have variable or inherent gender. Nouns with variable gender denote animates and count inanimates. Nouns that mark plural generally have variable gender. Nouns with inherent (fixed) gender, for their part, generally denote inanimates. These nouns may have fixed feminine gender if they derive from verbs (state/action

nominalizations) or if they denote time divisions, body parts, meteorological phenomena, and some celestial bodies. Nouns denoting mass have plural morphosyntax i.e., their agreement targets show plural marking. Generally, in Chadic, "[p]lurals represent a third category that is impervious to gender" (NEWMAN 2006: 194).

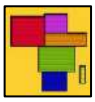
As for number, only plural is marked; and morphological marking of plural only concerns some nouns. Thus, the three languages contain a considerable number of transnumeral nouns i.e., nouns maintaining the same form irrespective of whether they denote a singular or plural referent. There are also nouns which do not take part in the number system. At this stage, we cannot accurately predict number marking and transnumeral nouns, however, there is a trend concerning which nouns are excluded from the number system. These semantic trends are discussed under §2. It seems there is a connection between gender and exclusion from the number system.

For nouns that mark plural morphologically, plurality is expressed via suffixal formation. A restricted set of nouns have dedicated suppletive plurals, some of which may further admit a plural suffix. When this happens, there are two additional formatives which show up: **-u**, and **-n** (Proto-Chadic ***-a**w and ***-n-?**). The **-n** formative behaves like a morphological separator occurring between a plural stem and an additional plural suffix in cases of double plural marking. Both formatives are discussed under §3.

The primary data used for this paper were collected as part of on-going descriptions of the languages under study, by the authors. Secondary data were extracted from existing literature: MELIS (1999, 2006) and OUSMANOU (2007) for Masana; AJELLO and MELIS (2008) for Gizey; and DAVOUNOUMBI (2017) and SHRYOCK (n.d.) for Musey. Material extracted from these sources is included with appropriate references.

2. Numerality

We use the term 'numerality' in this paper to refer to the ability for a category to express number values morphologically. For the purpose of this paper, we distinguish between nouns that are involved in the number system and nouns that are not. This distinction is relevant because, on the one hand, there are nouns involved in the number system that have not retained a morphological means for expressing plural, and, on the other hand, there are nouns which do not mark plural morphologically only from being excluded from the number



system. While both types do not have overt morphological marking, they exhibit different morphosyntactic behaviour. The main criterion for distinguishing between both types is agreement (see CORBETT 2000) i.e., the ability for given nouns to impose the expression of plural on modifiers or other interrelated clausal constituents. Nouns involved in the number system while being transnumeral can trigger singular or plural (two agreement values) on targets. Nouns excluded from the number system, by contrast, can trigger only one value i.e., either only singular or only plural. Nouns excluded from the number system are of three kinds: a) nouns denoting objects considered as being unique (e.g., celestial bodies, time divisions), b) gerundial nouns i.e., action/state nominalisations, and c) nouns conceived as being inherently plural.

The first two sets i.e., nouns denoting objects considered as being unique (1) and gerundial nouns (2) constantly trigger singular agreement. All such nouns have inherent feminine gender.[3]

(1) Masana (MELIS 2006a)

a.	**bàlák**	'Orion'
b.	**fàt**	'Sun'
c.	**tìl**	'Moon'
d.	**fàlèj**	'day'
e.	**vìl(ì)dí**	'evening'

(2) Gizey

a.	**mìt=tà**	'dying'
b.	**kùl=dà**	'stealing'
c.	**ɓɔ̀=dà**	'growing'
d.	**mùt=tà**	'eating'
e.	**hàt=tà**	'teaching'
f.	**ŋɔ̀k=kà**	'becoming thin'

Nouns denoting body parts also tend to be excluded from the number system. Interestingly, these nouns also have inherent feminine gender. However, there seems to be variation amongst the languages as to whether nouns denoting body parts are included in the number system or not. For example, while the Masana and Musey body part names in Table 1 are excluded from their number systems, in Gizey, they mark plural morphologically (e.g., **gàj** > **gìj-ɛ́j** 'foot' > 'feet').

[3] However, related word forms with the opposite gender may occur to capture a related reality. For example, the word **tìl** (**tìl=dà** 'moon=ART.SF) is inherently feminine when it refers to the Moon and it is inherently masculine (**tìl=là** 'moon=ART.SM) when it refers to the lunar month.

Masana	Musey	Gloss
ɛ̀r	**íí**	'eye(s)'
hʊ̀m	**húm**	'ear(s)
sɛ̀j	**síí**	'tooth/teeth'
ɓām	**kō**	'hand(s)
gâj	**gáj**	'foot/feet'
tīrèk	**ʧívìɗ**	'nail(s)'
hūʔ	**húɗ**	'ball(s)
mîr	**pō**	'breast(s)

Table 1 – Eight Masana and Musey nouns denoting body parts

The last set of nouns excluded from the number system are nouns denoting mass (e.g., Gizey: **wú** 'millet', **sùm** 'beer'). These nouns are inherently plural as they constantly impose plural on agreement target. For example, in (3) below, the plural suppletive verb **ʧúk** 'to throw.PL' is required for the structure to be grammatical. With a singular object NP, the verb would have been **gì** 'to throw.SG'.

(3) **ʧùg=ùn** **wú=n** **ʧūk**
throw.PL.N[4]=1S millet=ART.PL throw.PFV
'Throw me some millet'

There is also some variation amongst the languages under study as to which nouns have inherent plural number or not. For example, the Masana nouns in (4) all have inherent plural number. Such nouns generally denote objects with a high degree of plasticity, and which, as a result, create an impression of plurality.

(4) a. **gùj** 'snake (s)'
b. **zèw** 'rope (s)'
c. **lúwán** 'fishing net(s)'
d. **ɬígár** 'cloth(es)'
e. **bìrìm** 'bag(s)'

Evidence that these nouns have plural interpretation comes from the fact that agreement targets generally have to occur in their plural form. Observe in the Masana example in (5) that only a plural demonstrative is allowed after **gùj** 'snake'.

[4] N= Neutral aspect used in the imperative.

(5) **gùj-n** **lāknʲí/*wānní**
snake-ART DEM.PL/*S
'This is a snake'

However, in Gizey and Musey, most of these nouns are involved in the number system, and some even express number morphologically.

	GIZEY	MUSEY
'snake (s)'	**gùwj-íj**	**bòndɔ̀r-íí**
'rope (s)'	**zìgɛ̀w-ɛ́j**	**zéw-íí**
'fishing net(s)'	**bìj-ɛ́j**	**báj**
'cloth(es)'	**sàkr-ɛ́j**	**bāráw-íí**
'bag(s)'	**bìrìm-íj**	**bìrìm-íí**

Table 2 – Gizey and Musey plural marking nouns which are excluded from the number system in Masana

We compared the productivity of plural number marking in Gizey, Masana, and Musey with a list of 100 simple nouns. Chart 1[5] below shows the proportion of nouns which do not mark plural and that of number marking nouns. The red slots in each column represent the proportion of non-marking nouns, and the blue slots the proportion of number marking nouns. The prediction made by this chart is that there are more nouns in Gizey and Musey which mark plural than there are in Masana.

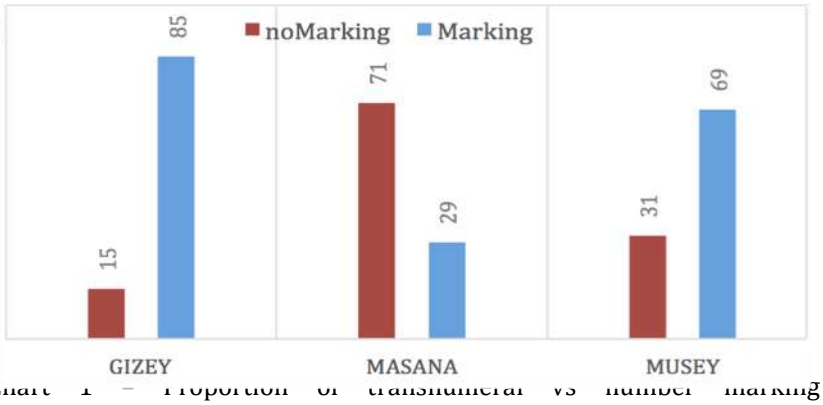


Chart 1 – Proportion of transnumeral vs number marking (morphological) nouns in Gizey, Masana, and Musey, 100 nouns compared

[5] We used convenience sampling to draw comparable data from wordlists used by the authors. When any language used a compound noun for a concept, (with potential partial marking on one of the constituents), that concept was removed from the count. We also maintained word forms known to be excluded from number marking in order to have a general picture.

Table 3 below provides some concrete data illustrating variation in number marking amongst nouns included in the number system. In this table, Musey and Gizey have the most forms with plural formatives (**-éj**, **-íj**, **-íí**, and **-gí**). Most non-marking nouns come from Masana. As can be seen, non-marking nouns in Masana may express plural morphologically in Gizey and Musey as with **kỳw-éj/kìw-éj** (Gizey) **- kēw** (Masana) **-ŋgèw-íí** (Musey).

	GIZEY	MASANA	MUSEY
'knives'	**kỳw-éj/kìw-éj**	**kēw**	**ŋgèw-íí**
'hippos'	**gárìjàm-éj**	**gàrjàm**	**gàrjàm-íi**
'okras'	**zùlù**	**zùlò**	**dlòònò-gí**
'feathers'	**ɮìmd-íj**	**ɮìmìt**	**ɬímíɗ**
'smiths'	**ʧáf**	**ʧāf**	**ʧàf-íí**

Table 3 – five words illustrating variation in plural marking in Gizey, Masana, and Musey

Based on the figures computed for Chart 1, it seems neither the absence of marking, nor overt marking can generalise as the rule or the exception for Masana and Musey. In line with the *Tolerance Principle* (SCHULER et al. 2021), a rule *R* generalises if the number of exceptions to *R* does not exceed the quotient of *N*/ln*N*, where *N* stands for the number of words in a given category and ln*N*, the natural log of *N*. For the 100 nouns computed, the number of exceptions should not exceed 21.7 items (100/ln101). In Gizey, if one considers number marking to be the rule, then the number of exceptions does not surpass 21.7. This implies that learners of Gizey will straightforwardly learn only a few exceptions while forming plurals productively via suffixation. This does not seem to work for Masana and Musey where, whether one considers absence of marking or overt marking as the rule, the number of exceptions exceeds 21.7. However, the tendency in Musey is clearly towards number marking, while it is the opposite in Masana.

The relatively important number of transnumeral nouns in the dataset, especially in the Masana data may be due to a general decline of overt morphological number marking on nouns, which may have started at an early diachronic stage (Proto-Masa). The presence in Masa of reflexes of reconstructed plural suffixes is clear evidence that Masa inherited at least one plural marking strategy (suffixation) from Proto-Chadic. However, the fact that languages like Gizey, Masana and Musey contain a high amount of transnumeral nouns, while other languages like Pévé simply lack morphological means for expressing singular/plural distinction (SHAY 2019) may be an indication that an erosion of number marking occurred within Masa. However, reports of a similar

morphological loss have been made for West Chadic (BLENCH 2021), which suggests that the decline started at an even earlier stage, namely in Chadic; unless one thinks of it as having emerged independently in different branches and geographical areas.

In the Masa languages under study, a number of changes affecting the morphology of number marking seem to have occurred only recently. In Masana for example, as Antonio Melis has pointed out (personal communication), there are plural inflections which are now obsolete. For example, the suppletive form **fɔ́k** 'caprinae' is now used as plural for **ɲù** although it had the more complex form **fúk-ú-n-ɛ́j** as noted in the 70s in grammar notes by Jean Goulard. The complex form **fúk-ú-n-ɛ́j**, still sporadically attested, combines several markings, namely, the suppletive stem **fúk**, the formatives **-u** and **-n**, and finally, the "regular" suffixal marking **-ɛ́j**. Additional examples with eroded plural marking in Masana, i.e., present-day transnumeral nouns include **hùrùm < *hùrùm-áj** 'crocodile(s)', **bàlàk < *bàlàk-áj** 'shed(s)', **tùlùm < *tùlùm-áj** 'ficus sycomorus' (Antonio Melis, personal communication). Other changes in Masana involve the progressive abandonment of plural forms with **-gáj**, which productively suffixed on stems ending with an open syllable; and the loss of loan plural formation with **-ga-n-ɛj**, in favour of the reduced form with allomorphs of **–aj** (6).

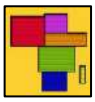
(6) Masana (Melis, personal communication)

a.	**per**	'father'	**per-ga-n-ej**	<	**per-ije**	'fathers'
b.	**ser**	'nun'	**ser-ga-n-ej**	<	**ser-ije**	'nuns'

3. Inventory of markers

Gizey, Masana, and Musey do not use consonant gemination and reduplication as methods for forming nominal plurals, although these are well attested across Chadic. This somehow confirms NEWMAN's (2006) hypothesis that gemination and reduplication constitute recent developments within Chadic. One other strategy found pervasively across Chadic, but which is absent from the languages under study is 'a-infixation'. However, the ablaut mechanisms generally associated with 'a-infixation' also manifest in some vestigial forms. Thus, the primary means for forming nominal plurals in Gizey, Masana, and Musey is suffixation. Vowel mutation appears only scarcely.

As concerns suffixation, number marking nouns in these languages generally distribute in two classes, except for the Yagwa variety (OUSMANOU 2007) of Masana which has only one class.



	MUSEY	GIZEY	MASANA (GUMAY & HARRA)	MASANA (YAGWA)
Class 1	**-ii/-i**	**-ij**	**-ij**	**-aj**
Class 2	**-Vgi**	**-V**[-high]**j**	**-aj**	

Table 4 – Plural markers used in Musey, Gizey, and Masana

These markers have phonologically conditioned and dialectal variants.

In Gizey, the plural marker **-V**[-high]**j** has two surface forms, **-ɛj** and **-ɔj**, resulting from harmony in roundness (§5).

In Masana, the marker **-aj** has surface forms **-aj**, **-ej**, and **-oj** resulting from harmony. These surface forms also alternate with **-ja**, **-je**, and **-jo**, respectively. The latter forms result from the application of metathesis in the vicinity of sonorants **n, l**, and **r** (§5). Finally, there is dialectal variation between **-aj**, **-ej**, and **-ɛj**.

It is not entirely clear whether the short **-i** marker occurring in Musey plurals found in SHRYOCK (n.d.) represents an orthographic choice by the author or an attested variant. The Pee variety spoken in Cameroon uses a long **-ii**. The marker **-Vgi**, for its part, alternates with **-gi**. The alternate form **-gi** results from the resolution of hiatus when **-Vgi** follows a vowel.

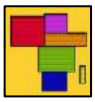
The markers identified for these languages correspond to markers which are attested across Chadic and which have been reconstructed for Proto-Chadic (NEWMAN 1990). Table 5 below provides the corresponding Proto-Chadic forms and their Northern Masa reflexes.

FORM	PROTO-CHADIC (NEWMAN 1990)
-ii/-i; -ij	***-i**
-V[-high]**j**, **-aj**	***-ai/*aj**
-Vgi	***-aki**

Table 5 – Northern Masa plural markers and reconstructed Proto-Chadic forms

There is an additional formative, **-n**, which appears in suppletive plurals admitting double marking (suppletion + suffixation). This formative can be seen in forms for 'dog' and 'goat' under Table 6 which provides examples of suppletive plurals for the three languages.[6]

[6] Generally, suppletive plurals are identical across the three languages with a few exceptions coming from Musey. Observe in Table 6 that while Gizey and Masana use suppletive plurals for



	GIZEY		MASANA		MUSEY	
	SG	PL	SG	PL	SG	PL
'person'	**sì**	**sūmū**	**sà**	**sūmū**	**sā**	**súú**
'woman'	**ʧì**	**ʔúrbɔ́j**	**ʧà**	**bɔ́j**	**ʧā**	**bójó-gíí**
'thing'	**vù**	**ɬúk-ɔ́j**	**và**	**ɬɛ́**	**vā**	**ɬē**
'cow'	**pūt**	**lúw-ɛ́j**	**pút**	**lúw-ɛ́j**	**mbūɬ**	**mbūɬ-íí**
'dog'	**dìj**	**dúːr-ɛ́j/dúːr-*n*-ɛ́j**	**dīj**	**dùr-*n*-ɛ́j/dìr-*n*-ɛ́j**	**dì**	**dík-*n*-íí**
'goat'	**ɦù**	**fúk-ɔ́j/fɔ́k-ɔ́j/fɔ́k-*n*-ɛ́j**	**ɦù**	**fɔ́k/fùg-ù-*n*-ɛ́j**	**ɦù**	**ɦù-*n*-íí**
'fowl'	**ɬēk**	**ɬàk-ŋ-ɛ́j**	**ɬèk**	**ɬùk-ù-n-ɛ́j/ɬìgì-n-ɛ́j**	**ɬék**	**ɬék-n-íí**

Table 6 – Some suppletive plurals for Gizey, Masana, and Musey

We analyse **-n**, not as a plural marker, but rather, as a kind of morphological separator occurring between two methods for expressing plurality, namely between suppletion and suffixation. Note that, except for Musey, **-n** appears only in such forms with double marking. However, there are forms like those for 'cows' and 'things' which do not show **-n** while admitting double marking. It is worth also noting that NEWMAN (1990) describes an identical form for Proto-Chadic which has reflexes in the West, East, and Biu-Mandara branches, but not in Masa. It is not clear how related the two forms are.

There is a further formative, **-u**, which shows up also only in a few suppletive plurals. In Table 6, it can be seen in the forms **fùg-*ù*-n-ɛ́j** '*caprinae*' and **ɬùk-*ù*-n-ɛ́j** 'fowls'. Other examples include:

(7)	a.	Masana	**dò**	**dù-g[7]-*ù*-n-ɛ́y[8]**	'multiparous'	(MELIS 2006a)
	b.	Gizey	**gùnɛ̀j**	**gòny-*ù*-g-ɛ̀j**	'zeroparous'	(AJELLO and MELIS 2008)

It is not clear whether the Gizey, Masana, and Musey word forms for 'people' (8)can also be analysed as containing that formative.

(8)	a.	Gizey	**sì**	**sūm-ū**
	b.	Masana	**sà**	**sūm-ū**
	c.	Musey	**sā**	**sū-ū**

Like **-n**, the **-u** formative is never attested as the sole plurality index; it always occurs as part of a complex plural marking which includes suppletion and

'goats', 'fowls', and 'cows', Musey rather uses regular plural marking as **ɦù-n-íí**, **ɬék-n-íí**, and **mbuɬ-íí**, respectively.

[7] The sound **-g** occurring in these examples generally occurs when suffixal plural markers attach to stems ending with an open syllable.

[8] **dù** and **dòknogéj** in Gizey .

suffixal formation. **-u** clearly appears as a vestigial formative, given its extreme rarity. SOUAG (2021) also describes a plural suffix **-u** in Mubic languages (East Chadic) which shows up in a number of plural relics. In Mubi, the phonological effects of **-u** suffixation is quite evocative of an erstwhile morphological labial prosody which got fully realised as a final vowel.[9] The term 'prosody' is used in Chadic linguistics to refer to a suprasegmental unit (phonological or morphemic) assumed to cause the labialization (rounding) or palatalisation (fronting) of segments within a word (ROBERTS 2001, 2007; GRAVINA 2014). Observe in the following Mubi plurals that the suffixation of **-u** triggers the labialisation of unround vowels as can be expected from a labial prosody. Of course, one need not resort specifically to prosodies to account for these Mubi data which could simply be analysed as 'regular' metaphony, i.e. vowel quality changes imparted by a neighbouring segment.

SG	PL	GLOSS
ƒùbáàg-ò	**ƒùbòog-ú**	'blind'
mársíy-ò	**mòrs-ù**	'lazy'
sùwàng-ót	**sùwòong-ú**	'Arab Shuwa'
sìŋàar-ó	**sìŋoor-ú**	'cat'
gàayìm-ó	**gòoyùm-ú**	'wild cat'
sògòryàk	**sògòryùg-ú**	'squirrel'

Table 7 – Plural formation with **-u** suffix in Mubi (data from SOUAG 2021: 254)

Finally, there are words which seem to form their plural via vowel mutation, i.e. unconditioned vowel quality change. The effect of this mutation process is raising, as can be seen from the following examples.

(9) a. **ŋól ŋúl** 'elder/s'
b. **gòr gùr(u)** 'child/ren'

Note that this raising process is to be distinguished from the one triggered by the suffixation of markers containing **-j**, which is discussed in detail under §5. Here too, this formation is rare and is probably a Proto-Chadic retention.

The word for 'child' in (9) above, **gɔ̀r/gòr**, can actually be pluralised as **gùr(ù)** (Gizey)/**gùrò** (Masana) or **gùr-ɛ́j/-ɛ́j** 'children'. However, both plural forms have distinct interpretations and morphosyntactic behaviour. The form using suffixal formation (**gùr-ɛ́j/-ɛ́j** 'children') patterns with collective plurals. Collective plurals denote groups consisting of at least two members. These nouns have a split agreement system whereby they select singular definite (enclitic) articles

[9] Diachronic segmental realisations of prosodies are not uncommon (see Gravina 2014).

while also requiring plural agreement on other targets like pronouns. Observe in (10) that the collective nouns (from Masana) select a singular definite article instead of the expected plural **=na**. However, in (11), note that only a coreferential plural pronoun is admitted in that construction. As for the form using vowel mutation, it has full plural morphosyntax and semantics i.e., it imposes plural on all agreement targets and does not refer to a group taken as forming a single unit as is the case with collectives.

(10) Masana
a. **kém=bā/*mā** 'group of children'
b. **ŋòn-áj=dā/*nā** 'twins'
ŋùlò=dā/*nā 'group of great/important people'

(11) Masana

gùr-éj=tā	***nígíj***	**ká**	**ŋàà**	**sū**
children=ART.SF	2PL	EXIST	good	Q

'Children, are you okay?'

4. Assignment systems

The languages under study have different assignment systems for the set of suffixal formatives discussed previously.

In Musey, the criterion for assigning plural markers is formal; it refers to the nature of the final syllable of the noun; and specifically, to whether it is open or closed. The marker **-Vgi** is used with open syllables and **-ii/-i** with closed syllables. The V position of **-Vgi** is filled by the last stem vowel. Illustrations for each of these markers are provided below.

(12)	a.	**mūl**	**mūl-íí**	'chief'
	b.	**ŋgèw**	**ŋgèw-íí**	'knife'
	c.	**bāráw**	**bāráw-íí**	'cloth'
	d.	**tōgōlōm**	**tōgōlōm-íí**	'flute'
	e	**fūl**	**fūl-íí**	'spirit'
	f.	**ɦèdèw**	**ɦèdèw-íí**	'necklace'
(13)	a.	**ʤìdʒì**	**ʤìdʒì-gí**	'African locust beans'
	b.	**óóhóó**	**óóhóó-gí**	'boy'
	c.	**màtàwì**	**màtàwì-gí**	'girl'
	d.	**gàmlà**	**gàmlà-gí**	'ram'
	e	**dlòònò**	**dlòònò-gí**	'okra'
	f.	**mbòjmò**	**mbòjmò-gí**	'bastard'

The **-Vgi** formative feeds hiatus, which is resolved in Musey by deleting its V_1. Thus, a form like **màtàwì-gì** 'girls' has a previous underlying shape **màtàwì-ìgì** in

which the final vowel of the noun spreads unto the empty vowel slot of the plural marker. Subsequently, this vowel is deleted to produce the surface form in (13).

Evidence that this plural marker has an underlying empty V-slot comes from words like those below which form their plural in the two Musey classes. It can be seen from the Plural 2 column that an agreeing vowel intervenes between the noun and the plural marker. The parentheses in the Plural 2 column represent the fact that some varieties of Musey (especially in Cameroon) delete that vowel.

SG	PL1	PL2	GLOSS
ŋgèw	**ŋgèw-íí**	**ŋgèw-(è)gí**	'knife'
mbìrwìn	**mbìrwìn-íí**	**mbìrwìn-(ì)gí**	'whirlwind'
ŋgùs	**ŋgùs-íí**	**ŋgùs-(ù)gí**	'tree'
sēēsēlēw	**sēēsēlēw-íí**	**sēēsēlēw-(ē)gí**	'swing'

Table 8 – Musey nouns admitting two plural markers

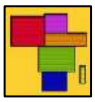
We also analyse the unexpected formation in **-Vgi** of the nouns in Table 8 to be a case of extension i.e., the plural marker, which is originally only used with nouns ending with an open syllable, is now being extended to a few nouns not meeting that requirement. This partly explains why nouns having two plural forms are limited in number. The triggering mechanism of this extension is unknown at this stage.

In Gizey, the assignment system is formal, however, it is different from what obtains in Musey. As indicated previously, Gizey has two plural formatives: **-ij** and **-V**$^{[-high]}$**j**. The criterion retained in Gizey is noun root vowel (or V_1, specifically) aperture. **-ij** is selected by nouns with [+high] V_1 (14) and **-V**$^{[-high]}$**j** is selected by nouns with [-high] V_1 (15).

(14) a. **mùl** **mùl-íj** 'chief'
b. **fùl** **fùl-íj** 'spirit'
c. **dìf** **dìf-íj** 'necklace'
d. **sìnìl** **sìnìl-íj** 'tongue'
e. **nìrwìn** **nìrwìn-íj** 'whirlwind'

(15) a. **dèl** **dìl-éj** 'vagina'
b. **bàk** **bɨk-éj** 'skin/leather'
c. **màt** **màt-íj/mɨt-éj** 'evil spirit'
d. **ŋɔ̄t** **ŋùt-ɔ́j** 'calabash'
e. **ɲù** **fɔ́k-ɔ́j/fúk-ɔ́j** 'goat'

The surface high vowels in the plural forms in (15) result from an active metaphony rule in Gizey which raises root V_1 when roots are combined with suffixes containing the palatal /-j/.



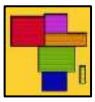
In the Gumay and Harra varieties of Masana described by MELIS (1999), morphological marking of plural has semantic underpinnings. The marker **-ij** is generally used with nouns denoting kin (16), while **-aj** is used elsewhere as can be seen under examples (17)-(18) culled from MELIS (1999). As already pointed out, **-aj** may surface as **-aj** or **-ej**.

(16)	a.	**sùgòl**	**sùgòl-íj**	'ally'
	b.	**nàsú**	**nòs-íj**	'uncle'
	c.	**tànà**	**tàn-íj**	'brother-in-law'
	d.	**ɓùsú**	**ɓùs-íj**	'brother/sister'
	e	**gòrbú**	**gùrób-íj**	'half-brother'
	f.	**kùnò**	**kùn-íj**	'in law'
(17)	a.	**gàwlàŋ**	**gàwlàŋ-éj**	'prostitute'
	b.	**ʧèɗ**	**ʧèɗ-éj**	'axe'
	c.	**ʧùt**	**ʧùt-éj**	'*acacia albida*'
	d.	**dùt**	**dùt-éj**	'calabash'
	e	**lùm**	**lùm-éj/lùm-áj**	'canoe'
	f.	**ŋàl**	**ŋàl-éj**	'sauce pottery'
(18)	a.	**gòlòŋ**	**gòlòŋ-áj**	'side-stream'
	b.	**ʤùf**	**ʤùf-áj**	'husband'/'male'
	c.	**hùrùm**	**hùrùm-áj/Ø**	'crocodile'
	d.	**bàlàk**	**bàlàk-áj/Ø**	'shed'
	e	**tùlùm**	**tùlùm-áj**	'ficus sp'
	f.	**gúm**	**gúm-áj**	'hoe'

In the Yagwa variety of Masana, there is just one marker with different surface forms that will be discussed in detail in the next section.

(19)	a.	**gàwlàŋ**	**gàwlàŋ-áj**	'prostitute'
	b.	**dùt**	**dùd-áj**	'calabash'
	c.	**gùk**	**gùg-áj**	'dove'
	d.	**ɬōk**	**ɬōg-ój**	'elephant'
	e	**gòlòŋ**	**gòlòŋ-ój**	'river'
	f.	**vèt**	**vēt-éj**	'hare'

The general parameters used in all three languages are summarised in Table 9 below.



	MUSEY	GIZEY	MASANA (G/ɦ)	MASANA (Y)
Class 1	**-ii/-i** elsewhere	**-ij** [+high] root V or V1	**-ij** kin	**-aj**
Class 2	**-Vgi** open syllable	**-V**[-high]**j** [-high] root V or V1	**-aj** elsewhere	
Parameter	syllable structure	vowel aperture	kin/non-kin	

Table 9 – Plural assignment parameters for Musey, Gizey, and Masana (Gumay, Harra, and Yagwa varieties)

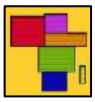
5. Phonological processes fed by plural formation

Nominal pluralization triggers different phonological processes in the languages under study. These phonological processes apply with different degrees of systematicity. The most notable processes include: metaphony, vowel harmony, and metathesis.

We use the term *metaphony* to refer to the effects of a trigger segment onto a target vowel, resulting in quality change on the part of the target. In the languages under study, the suffixation of plural markers may trigger the rising of non-high first vowels (V_1) of the nominal root. In Gizey, this process is systematic; in Masana, it occurs only occasionally (on specific nouns e.g., **súgól** > **súgúl-ój** 'parents'). Metaphony is not attested in the Musey dataset. Some Gizey data is given under (20). The word for 'husband' is analysed as having the underlying form **ʤɔ̀f**, although it mostly occurs with a high vowel as **ʤùf**. This analysis in necessary to account for the fact that it selects a [-high] plural marker. It is not uncommon to come across the form **ʤɔ̀f** in spontaneous data.

(20)	a.	**dɛ̀l**	**dìl-ɛ́j**	'vagina'
	b.	**bàk**	**bɨ̀k-ɛ́j**	'skin/leather'
	c.	**màt**	**màt-íj/mɨ̀t-ɛ́j**	'evil spirit'
	d.	**ŋɔ̄t**	**ŋùt-ɔ́j**	'calabash'
	e	**ɦù**	**fɔ́k-ɔ́j/fúk-ɔ́j**	'goat'
	f.	***ʤɔ̀f**	**ʤùv-ɔ́j**	'husband'

Some vowel harmony phenomena occur in Gizey and Masana. In the Musey data, no such phenomena have been observed. In Gizey there is harmony in roundness: if the root V is [– round], the non-high marker surfaces as **[-ɛj]**, if it is round, then the marker surfaces as **[-ɔj]**. This roundness harmony is formalised in (21) and exemplified with the data under (20) repeated as (22) below. Examples (22a-c) show agreement between the plural marker and [– round] root vowels; and (22d-f) show agreement in the [+ round] feature.



(21) V [αround, – high] > V [αround, – high] / V [αround, – high] ___

(22)	a.	**dɛ̀l**	**dìl-ɛ́j**	‘vagina’
	b.	**bàk**	**bɨ̀k-ɛ́j**	‘skin/leather’
	c.	**màt**	**màt-íj/mɨ̀t-ɛ́j**	‘evil spirit’
	d.	**ŋɔ̄t**	**ŋùt-ɔ́j**	‘calabash’
	e	**ɦù**	**fɔ́k-ɔ́j/fúk-ɔ́j**	‘goat’
	f.	***ʤɔ̀f**	**ʤùv-ɔ́j**	‘husband’

In Masana, there is parasitic harmony: if the root V or V_1 is [– high] like the plural marker, then there is total harmony (23c-e). If V_1 is [+ high], then the default **-aj** is used (23a-b).

(23)	a.	**dùt**	**dùd-áj**	‘calabash’
	b.	**gùk**	**gùg-áj**	‘dove’
	c.	**gòlòŋ**	**gòlòŋ-ój**	‘river’
	d.	**vèt**	**vēt-éj**	‘hare’
	e	**gàwlàŋ**	**gàwlàŋ-áj**	‘prostitute’

The last process observed in our nominal pluralization datasets is metathesis and this occurs only in Masana. The segments /a/ and /j/ of the plural marker invert their positions when this marker is preceded by /l/, /r/, and /n/. The harmony properties of the vocalic segment /a/ are preserved i.e., it may surface as [a], [e], or [o] depending on the nature of the noun V_1. Resyllabification may require i-insertion, such that words like **mūljá** ‘chiefs’ surface as **mūlījá**.

(24)	a.	**mūl + -aj**	**mūl-(ī)já**	‘chief’
	b.	**mār + -aj**	**mār-(ī)já**	‘old’
	c.	**múr + -aj**	**múr-(í)já**	‘wild animal’
	d.	**fér + -aj**	**fér-(í)jé**	‘attic’
	e	**géń + -aj**	**gén-(í)je**	‘pitcher’
	f.	**rāj + -aj**	**ŋòr-(ì)jó**	‘sorcerer’

Variation in the application of these processes is summarised under Table 10 below.

	GIZEY	MASANA	MUSEY
Metaphony	systematic	occasional	not applicable
Metathesis	not applicable	systematic: **-aj** > **-(i)ja**	not applicable
vowel harmony	applicable: αround harmony	applicable: total parasitic harmony	not applicable

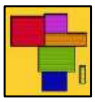
Table 10 – Summary of variation in terms of phonological processes fed by nominal plurality

6. Conclusion

Gizey, Masana and Musey differ in their proportion of nouns marking number morphologically. Our datasets indicate there are more nouns marking plural morphologically in Gizey and Musey (in this order) than there are in Masana. Plural marking thus seems to be more eroded in Masana. Gizey, Masana and Musey have two underlying suffixal plural markers, except for the Yagwa variety of Masana which has only one. These markers are reflexes of Proto-Chadic markers attested across present-day Chadic languages (Newman 1990). Musey and Gizey use a formal system to assign nouns into different classes: Musey uses syllable structure (closed vs open final syllable) and Gizey vowel aperture (high vs non-high). The Gumay and Harra varieties of Masana use a semantic criterion (kin vs non-kin). In the non-kin class, there are two surface forms **-aj** and **-ej** relating to dialectal variation. In the Yagwa variety there is only one class and surface realizations of the plural marker depend on aperture and syllable structure. The languages also differ in the kinds of phonological processes fed by pluralization. Gizey pluralization systematically feeds metaphony, while this is only seen occasionally in Masana data. Both languages also exhibit vowel harmony processes, which however, produce different effects. Masana uses a metathesis rule which is seen neither in Gizey nor in Musey. Generally, the pluralization of nouns does not trigger important segmental changes in Musey.

Abbreviations

ART: article; DEM: demonstrative; EXIST: existential; F: feminine, G: Gumay variety of Masana; ɦ: Harra variety of Masana; IPFV: imperfective, N: neutral aspect, PFV: perfective; PL: plural; Q: question operator; S: singular; Y: Yagwa variety of Masana.



References

AJELLO, Roberto and Antonino MELIS (2008). *Dictionnaire gizey-français, suivi d'une liste lexicale français-gizey*, Pisa: Edizioni ETS.

BALDI, Sergio and Rudolf LEGER (2011). "Some diachronic observations on gender and number in Bole-Tangale languages", in Doris LÖHR and Ari AWAGANA (eds.), *Topics in Chadic linguistics VI. Comparative and descriptive studies*, Köln: Rüdiger Köppe, 23-31.

BARRETEAU, Daniel and Michel DIEU (2005). "Linguistique", in *Atlas de la province Extrême-Nord Cameroun*, Marseille: IRD Editions, 1-32.

BERTONI, Marco (2018). *La culture musey (Tchad). Naissance, mariage et funérailles*, Paris: L'Harmattan.

BLENCH, Roger (2021). "The erosion of number marking in West Chadic nominal", paper presented at the *10th World Congress of African Linguistics* (7 July 2021).

CAÏTUCOLI, Claude (1983). *Lexique masa (Tchad et Cameroun)*, Paris and Yaounde: ACCT-CERDOTOLA.

CORBETT, Greville G. (2000). *Number*, Cambridge: Cambridge University Press.

DAVOUNOUMBI, Pierre (2017). *Phonologie et morphologie du museyna parlé à Gobo (Extrême-Nord Cameroun)*, Maroua: Université de Maroua Mémoire de master 2.

DE DOMINICIS, Amedeo (2008). "Phonological Sketch of Gizey", *Studi Linguistici e Filologici Online* 6: 1-78.

GRAVINA, Richard (2014). *The phonology of Proto-Central Chadic : the reconstruction of the phonology and lexicon of Proto-Central Chadic, and the linguistic history of the Central Chadic languages*, Utrecht: Lot.

MELIS, Antonino (1999). *Description du Masa (Tchad): Phonologie, Syntaxe et Dictionnaire Encyclopédique*, PhD thesis, Université de Tours.

MELIS, Antonino (2006a). *Dictionnaire Masa-français: dialectes gumay et ɦaara (Tchad)*, Pisa: Edizioni Plus.

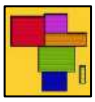
MELIS, Antonino (2006b). "Ham: une langue et une culture en danger de disparition (Tchad)", in Amedeo DE DOMINICIS (ed.), *Undescribed and Endangered Languages the Preservation of Linguistic Diversity*, Newcastle: Cambridge Scholars Press, 44-59.

MELIS, Antonino (2019). "Storia e identità linguistica", in Luigi GAFFURI, Antonino MELIS and Valerio PETRARCA (eds.), *Tessiture dell'identità. Lingua, cultura e territorio dei Gizey tra Camerun e Ciad*, Napoli: Liguori Editore, 43-104.

NEWMAN, Paul (1990). *Nominal and verbal plurality in Chadic*, Dordrecht: Foris.

NEWMAN, Paul (2006). "Comparative Chadic revisited", in Paul NEWMAN and Larry HYMAN (eds.), *West African linguistics: Papers in honor of Russell G. Schuh (Studies in African Linguistics Supplement 11)*, Columbus: Ohio State University, 188-202.

NEWMAN, Paul (2013). 2013. *The Chadic Language Family: Classification and Name Index*, Mega-Chad Research Network / Réseau Méga-Tchad, accessed on 21/8/2019, <http://www.megatchad.net/publications/Newman-2013-Chadic-Classification-and-Index.pdf>



OUSMANOU. 2007. *Phonologie lexicale du masa: dialecte yagua parlé au Cameroun*, Cameroun: Yaoundé I Mémoire de DEA.

ROBERTS, S. James. 2001. *Phonological features of Central chadic languages*, in Ngessimo MUTAKA and Beban Chumbow SAMY (eds.), *Research mate in African linguistics: focus on Cameroon*, 93-118. Köln: Rüdiger Köppe.

ROBERTS, S. James. 2007. "Palatalization and Labialization in Mawa (Eastern Chadic)", in Eva ROTHMALER (ed.), *Topics in Chadic Linguistics V* (Papers from the 4th Biennial International Colloquium on the Chadic Languages, Bayreuth, October 30–31, 2007), Köln: Rüdiger Köppe, 129-140.

SCHULER, Kathryn, Charles YANG and Elissa NEWPORT (2021). "Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient", *PsyArXiv* (March 30), doi:10.31234/osf.io/utgds.

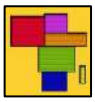
SEIGNOBOS, Christian and Henry TOURNEUX (2002). *Le Nord-Cameroun à travers ses mots: dictionaire des termes anciens et modernes*, Paris: IRD Editions & Karthala.

SHAY, Erin (2019). *A Grammar of Pévé*, Leiden: Brill.

SHRYOCK, Aaron. n.d. *A Dictionary of Musey*, accessed on 11/11/2021, <https://silo.tips/download/a-dictionary-of-musey>

SOUAG, Lameen (2021). "Pattern borrowing and hybridization in Mubi (East Chadic): The importance of congruence", *Word Structure* 14 (2): 246-270.

JOURNAL OF AFRICAN LANGUAGES AND LITERATURES
3/2022, 106-119



Attenuation by reduplication in Hausa and beyond

GEORG ZIEGELMEYER

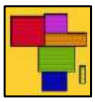
University of Vienna
georg.ziegelmeyer@univie.ac.at

ABSTRACT

A first examination of patterns expressing attenuated qualities in Hausa and some other mainly West Chadic languages reveals that reduplication is the most wide-spread morphological process to reduce the degree of a quality. For instance, in Hausa full reduplication and shortening of the final vowel on each component derives denominal "*X-like*"-expressions, or "*X-y*" and "*X-ish*" adjectives, i.e. concrete nouns may form adjectives with the meaning "*characterized by*", sometimes retaining nominal status with an attenuated meaning, e.g. **gishirī** 'salt' > **gishiri-gishiri** 'salty', but **barcī** 'sleep(ing)' > **barci-barci** 'a nap'. In a similar manner abstract nouns of sensory quality (ANSQs) allow reduplicated forms, with a detensified "*X-ish*" meaning, e.g. **zāƙī** 'sweetness' > **zāƙi-zāƙi** 'sweetishness' (**lɛ̀mō mài zāƙi-zāƙi** 'a sweetish soft drink'). According to SCHUH, GOGE and DOLE (n.d.) Ngamo employs partial reduplication to expresses the concept "*NOUN-like*", "*NOUN-ish*". Apart from this, some other Chadic languages allow reduplication of simple or derived adjectives, usually denoting colours or physical attributes, in order to get a detensified/attenuated meaning of a quality, e.g. Hausa: **bàbba** 'big' > **bàbba-bàbba** 'biggish' (cf. JAGGAR 2001); Bade: **ɓuwâ** 'red' > **ɓuwâ-ɓuwâ** 'reddish' (cf. ZIEGELMEYER 2015); Bole: **dài** 'red' > **dài-dài** 'reddish' (cf. GIMBA and SCHUH Ms.); Malgwa: **dzáyye** 'white' > **dzáy-dzáyye** 'whitish' (cf. LÖHR 2002). Our survey of something like 45 Chadic languages from all branches reveals that if we find information on attenuation processes, then either full or partial reduplication is involved.

KEY WORDS: attenuation, Chadic, Hausa, reduplication, semantic weakening

DOI https://doi.org/10.6093/jalalit.v3i3.9143
ISSN 2723-9764



GEORG ZIEGELMEYER

Attenuation by reduplication in Hausa and beyond

1. Introduction

In this paper we illustrate morphologic attenuation (reduced degree of a quality) in Hausa and some other mainly West Chadic languages. It has been shown in typological studies of derivational morphology that attenuation is among the most frequent adjectival derivational categories in the languages world-wide (see e.g. BAUER 2002: 42). While some languages have specific suffixes which derive attenuated adjectives (e.g. English: **-ish** as in **green-ish**; French: **-âtre** as in **blanch-âtre**; Kanuri: **-àràm** as in **cìmê-àràm** 'reddish'), Hausa and at least some related West Chadic languages use full or partial reduplication for the expression of attenuated qualities.

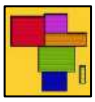
A certain fly in the ointment is that attenuation processes are scarcely described in many grammars on Chadic languages. In our survey of something like 45 Chadic languages from all branches only for eight languages, i.e. Hausa, (Gashua) Bade, Bole, Ngamo, Zaar, Kwami, Malgwa, and Margi, some information on attenuation processes was found. While in Hausa, which is probably one of the best described languages of the African continent, morphologic attenuation is well documented, in most other Chadic languages information is rather scanty. Our scrutiny, nevertheless, shows a certain tendency, i.e. if processes that produce semantic attenuation are documented for a Chadic language then morphological reduplication is involved.

In paragraph 2 we will briefly outline how qualities are expressed in Chadic languages before we draw our attention to Hausa in paragraph 3. Paragraph 4 illustrates attenuation processes in some other Chadic languages, followed by our conclusions in paragraph 5.

2. A note on expressing qualities in Chadic

Generally Chadic languages show a great variety in their possibilities to express a quality, e.g. some languages have a class of adjectives defined by morphosyntactic criteria (e.g. Hausa), others are said to lack a class of adjectives (e.g. Sakun (Sukur)). Some Chadic languages basically employ verbs to express qualities (e.g. Goemai), and others also have a robust class of quality nouns often called abstract nouns of sensory quality (e.g. Hausa). According to FRAJZYNGIER and SHAY (2012: 270)

> "not all Chadic grammars state explicitly how the lexical class of adjectives differs from that of verbs, nouns, or other lexical items. Some Chadic languages are reported to have many adjectives and others are reported to have only a few."



Often a language also has various means to express quality concepts, e.g. Hausa whose most common constructions are illustrated in Table 1.

adjectives	noun + adjective	**zāɓì nagàri**[1]	'good choice'
	adjective-linker[2] + noun	**sābuwâ-r̃ mōtā**	'new car'
quality nouns	"being with" + quality noun	**tanà dà kyâu**	'she is beautiful'
	mài + quality noun	**mōtā mài tsàdā**	'an expensive car'

Table 1 – Common constructions expressing qualities in Hausa

Basically Hausa employs either adjectives, or nouns (often called abstract nouns of sensory quality) in specific constructions to express qualification. While in Hausa adjectives morphologically behave like nominals, according to NEWMAN (2000: 22) they

> "can, nevertheless, be distinguished from nouns. First, there are some derivations whose sole function is to create adjectives, not nouns. Second, adjectives have syntactic properties that set them apart from nouns […] Functionally, adjectives serve as noun modifiers rather than head words […] Moreover, gender and number in adjectives are agreement features determined by the category of the head noun rather than being intrinsic properties, as is the case with nouns."

3. Expressing attenuated qualities in Hausa

In Hausa full reduplication accompanied by shortening of the final vowel in each component is a morphological process which derives attenuated qualities. Simple and derived adjectives, certain nouns, as well as some adverbs undergo semantic attenuation by means of this process.

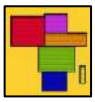
3.1 Attenuated adjectives in Hausa

According to JAGGAR (2001: 144)

> "Some adjectives, both simple and derived, and typically denoting colours or physical attributes, allow fully-reduplicated forms with a short final vowel on each component, including the inflected feminine and plural forms. These fully reduplicated adjectives have a detensified/attenuated meaning, equivalent to English "X-ish"".

[1] Transcription: **ā**, **ī**, etc. = long vowel; **a**, **i**, etc. = short vowel; **ə** = high central vowel; **à** = low tone; **á** = high tone; **â** = falling tone; **ǎ** rising tone; **ɓ**, **ƈ** = laryngeal implosives; **'y** = glottalized palatal glide; **ɬ** = voiceless lateral fricative, **r̃** = apical tap/roll, **c** and **j** = palato-alveolar affricates.

[2] Singular masculine adjectives, as well as plurals take the linker **-n**, while singular feminine adjectives are linked with **-r̃**.



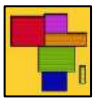
Following NEWMAN (2000: 27) this process "produces forms that are semantically "attenuated", i.e., have the quality of the simple adjective but to a lesser extent, e.g. **tsanwā** 'green', **tsanwa-tsanwa** 'light green, greenish'". The example **yalò-yalò** 'yellowish' from **yalò** 'yellow', which is a recent loanword from English, shows that the process is still active in the Hausa language. Further examples[3] are given in Table 2 where we can see that many attenuated adjectives form also feminine and plural forms. The examples also show that the processes of feminine and plural inflection precede the attenuation derivation.

MASCULINE	FEMININE	PLURAL	GLOSS
bàbba-bàbba **bàbba**	**bàbba-bàbba** **bàbba**	**mânya-mânya** **mânyā**	'biggish' 'big'
baƙi-baƙi **baƙī**	**baƙa-baƙa** **baƙā**	**baƙàƙe-baƙàƙe** **baƙàƙē**	'blackish, grey' 'black'
dōgo-dōgo **dōgō**	**dōguwa-dōguwa** **dōguwā**	**dōgwàye-dōgwàye** **dōgwàyē**	'tallish' 'tall'
ɗanye-ɗanye **ɗanyē**	**ɗanya-ɗanya** **ɗanyā**	**ɗànyu-ɗànyu** **ɗànyū**	'rawish' 'raw'
fari-fari **farī**	**fara-fara** **farā**	**faràre-faràre** **farārē**	'whitish, off-white' 'white'
gàjēre-gàjēre **gàjērē**	**gàjēra-gàjēra** **gàjērā**	**gàjèru-gàjèru** **gàjèrū**	'shortish' 'short'
jāja-jāja **jā**	**jāja-jāja** **jā**	**jājàye-jājàye** **jājàyē**	'reddish' 'red'
shūdî-shūdî **shūdǐ**	**shūdîya-shūdîya** **shūdîyā**	**shûɗɗa-shûɗɗa** **shûɗɗā**	'light blue, bluish' '(dark) blue'

Table 2 – Attenuated adjectives derived from primary adjectives

While the above outlined attenuation process most often applies to primary adjectives, it is sometimes also encountered with various derived adjectives, e.g. adjectival past participles (adj.pp), derived adjectives of sensory quality (DASQs), ethnonymic adjectives (ethn.adj), and agential adjectives (agent.adj), cf. NEWMAN (2000), and JAGGAR (2001). In Table 3 some examples are given. As with primary adjectives this process applies also to feminine and plural forms of derived or secondary adjectives.

[3] All Hausa examples in this paper come from NEWMAN (2000) and JAGGAR (2001).



dàfaffē m.	'cooked'	>	**dàfaffe-dàfaffe**	'somewhat cooked'	adj.pp
sǒyayyiyā f.	'roasted'	>	**sǒyayyiya-sǒyayyiya**	'slightly roasted'	adj.pp
tsàttsāmā sg.	'very sour'	>	**tsàttsāma-tsàttsāma**	'somewhat sour'	DASQ
kàkkaurā sg.	'very strong, hefty'	>	**kàkkaura-kàkkaura**	'somewhat stout'	DASQ
mahàukàtā pl.	'mad'	>	**mahàukàta-mahàukàta**	'a bit mad'	agent.adj
malàlàtā pl.	'lazy'	>	**malàlàta-malàlàta**	'lazyish'	agent.adj
bàhagǒ m.	'left-handed'	>	**bàhagò-bàhagò**	'semi-left-handed'	ethn.adj
bàgidājìyā f.	'naive'	>	**bàgidājìya-bàgidājìya**	'a bit naive'	ethn.adj

Table 3 – Attenuated adjectives derived from secondary adjectives

Syntactically attenuated adjectives, like other heavy reduplicated adjectives, generally follow the noun they modify (see examples 1-5), and according to JAGGAR (2001: 145) "speakers will often insert an additional diminutive modifier (m./f./pl.) **ɗan/'yař/'yan** in position before the attenuated adjective" (see examples 6-8). Some attenuated, phonologically short and masculine adjectives, however, may also modify a noun in pre-head position with the linker **-n**, cf. NEWMAN (2000), (see example 9). This is not possible with heavy reduplicated adjectives (see example 10).

(1) **mōtà** **jāja-jāja**
car reddish
'reddish car'

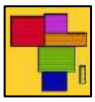
(2) **wani** **dōkì** **baƙi-baƙi**
SID horse blackish
'a blackish horse'

(3) **nāmà** **ɗanye-ɗanye**
meat rawish
'rawish meat'

(4) **wata** **màcè** **mahaukacìya-mahaukacìya**
SID woman slightly crazy
'a slightly crazy woman'

(5) **wata** **màcè** **bàgidājìya-bàgidājìya**
SID woman somewhat naive
'a somewhat naïve, dim-witted girl'

(6) **wasu** **rīgunà** **('yan)** **jājàye-jājàye**
SID gowns DIM reddish
'some reddish gowns'



(7) **sōjà (ɗan) dōgo-dōgo**
soldier DIM tallish
'medium-height soldier'

(8) **wani mùtûm (ɗan) mahàukàci-mahàukàci**
SID man DIM somewhat.crazy
'a somewhat crazy man'

(9) **wani fari-fari-n kèkè = wani kèkè fari-fari**
SID whitish-LNK bicycle SID bicycle whitish
'some off-white bicycle'

(10) ***mahaukacìya-mahaukacìya-ɾ̃ màcè**
slightly.carzy-LNK woman
'a slightly mad woman'

3.2 Attenuated abstract nouns of sensory quality in Hausa

In addition to the above illustrated attenuated adjectives, Hausa has also a set of abstract nouns of sensory quality (ANSQs) which, probably on semantic grounds, also may undergo the same attenuation process, i.e. full reduplication accompanied by shortening of the final vowel on each component, cf. (NEWMAN 2000), and JAGGAR (2001). According to NEWMAN (2000: 28) "this is also true in the case of a few other semantically appropriate words that are not ANSQs in the strict morphological sense". In this group we find a few semantically attenuated colour terms derived from source nominals, e.g. **ƙasa-ƙasa** 'brownish' < **ƙasā** 'earth, land', **tòka-tòka** 'greyish' < **tòkā** 'ashes', cf. JAGGAR (2001). Note also that compound colour adjectives whit the structure **ruwan** 'lit. water of' + colour simply copy the second member of the compound to create the corresponding attenuated form, e.g. **ruwan hōdà** 'pink' > **ruwan hōdà-hōdà** 'pinkish'. Table 4 illustrates some attenuated ANSQs.

ɗācī	'bitterness'	>	**ɗāci-ɗāci**	'somewhat bitter'
sanyī	'cold'	>	**sanyi-sanyi**	'coldish'
kaurī	'thickness'	>	**kauri-kauri**	'slightly thick'
zāfī	'heat'	>	**zāfi-zāfi**	'warmish'
tsāmī	'sourness'	>	**tsami-tsami**	'sourish'
taurī	'toughness'	>	**tauri-tauri**	'kind of tough'
zāƙī	'sweetness'	>	**zāƙi-zāƙi**	'sweetish'
ɗùmī	'warmth'	>	**ɗùmi-ɗùmi**	'lukewarm'

Table 4 – Attenuated abstract nouns of sensory quality (ANSQs)

It is important to note that the outputs are grammatically still nouns, i.e. syntactically they behave like nouns, and not like adjectives. For instance, when modifying a noun attenuated abstract nouns of sensory quality do not follow the

noun immediately but insert the particle **mài** sg./**màsu** pl. 'being characterized by'. When used in predicative functions attenuated abstract nouns of sensory quality take HAVE sentences.

(11) **àbinci** **yā** **yi** **sanyi-sanyi**
food 3M.COMP do coldishness
'the food has gone a little bit cold'

(12) **dà** **zāfi-zāfi** **yâu**
with warmishness today
'it's warmish today'

(13) **māgànī** **mài** **ɗāci-ɗāci**
medicine characterized.by bitterness
'kind of bitter medicine'

(14) **kàtīfà** **tanà** **dà** **laushi-laushi**
mattress 3F.CONT with softness
'the mattress is somewhat soft'

Consider the syntactic use of attenuated adjectives in examples 15, and 16.

(15) **kàtīfà-ř** **tsàgaggiya-tsàgaggiya** **cè**
mattress-PRM somewhat.ripped SATB.F
'the mattress is somewhat soft'

(16) **bàbûř** **shūdì-shūdì**
motorcycle bluish
'a light blue motorcycle'

Last but not least, we can observe that the same morphological process, i.e. reduplication and final vowel shortening, also applies to mainly singular concrete nouns to derive adjectives with the meaning "X-like, X-y, X-ish" (where X represents the source noun), cf. NEWMAN (2000), and JAGGAR (2001), e.g. **gishirī** 'salt' > **gishiri-gishiri** 'salty', **ruwā** 'water' > **ruwa-ruwa** 'watery', **yàshī** 'sand' > **yàshi-yàshi** 'sandy'. Like other reduplicated adjectives they are generally employed in post-head position, i.e. after the noun they modify, e.g. **madařā ruwa-ruwa** 'watery/skimmed milk'. However, some typically eventive-dynamic nouns also retain their nominal status after application of this morphological process, e.g. **barcī** 'sleep(ing)' > **barci-barci** 'a nap', **màganà** 'discussion' > **màganà-màganà** 'a brief discussion'. According to JAGGAR (2001: 144) "semantically, this subset could be viewed as similar to either reduplicated ("type of") nouns …, or compared to the attenuated "X-ish" adjectives …", or as NEWMAN (2022: 133) puts it:

> "what really seems to tie the two manifestations together [i.e. reduplication of adjectives or abstract nouns of sensory quality and reduplication of concrete nouns G.Z.] is the notion of '-ish' or '-like'. Given that adjectives in Hausa are very noun-like, the essence of the process appears to be converting independent nominals into adjective-like or adverb-like modifiers, a use of reduplication that typologically fits nicely into the Chadic mould. Beyond that, there is not much more that we can say at the moment. There is clearly an interesting historical story or stories waiting to be told regarding these phenomena, but this remains a task for future investigation".

3.3 Attenuated adverbs in Hausa

Finally we can observe that in Hausa reduplication of adverbs basically has two functions. While reduplicated simple adverbs of time, place, and manner typically intensify the meaning, e.g. **dàbam** 'differently' > **dàbam-dàbam** 'very differently', some denominal adverbs (typically locative) undergo semantic weakening when fully reduplicated. According to NEWMAN (2000: 42) "this is the same attenuation process found with adjectives". As denominal adverbs already have a short final vowel the shortening rule observed with adjectives and ANSQs does not apply, and "in some cases, the nonreduplicated adverb does not exist, i.e., one has only the noun and the reduplicated adverb [...]" (NEWMAN 2000: 42). According to JAGGAR (2001: 657) "these detensified reduplicates can combine with the downscaling adverb **kàɗan** 'a little, slightly', e.g. **gàba-gàba kàɗan** 'just slightly in front' (cf. **gàba** 'in front' < **gàbā** 'front of body')". Some more examples are given in Table 5.

bāyā	'back'	>	**bāya**	'behind'	>	**bāya-bāya**	'slightly behind'
ƙasā	'earth, ground'	>	**ƙasà**	'below'	>	**ƙasà-ƙasà**	'a bit lower'
samà	'sky, heavens'	>	**samà**	'above'	>	**samà-samà**	'a bit higher'
			nēsà	'far away'	>	**nēsà-nēsà**	'a bit far away'
tsakiyà	'middle, centre'	>				**tsakiyà-tsakiyà**	'around the centre'
dāmā	'chance'	>	**dà dāmā**	'quite a lot'	>	**dà dāma-dāma**	'moderately'

Table 5 – Attenuated (locative) adverbs in Hausa

4. Expressing attenuated qualities in other Chadic languages

Our scrutiny reveals that some Chadic languages other than Hausa express attenuated qualities by reduplication. As already stated above information on attenuation processes in many languages is scanty, nevertheless, examples from Gashua Bade, Bole, Ngamo, Zaar, Kwami, Malgwa, and Margi show that full, or in the case of Ngamo, partial reduplication is at hand.

4.1 Bade

In Gashua Bade the morphological process of reduplication may derive attenuated qualities amongst others (cf. ZIEGELMEYER 2015). Examples recorded in my data comprise colour terms which are semantically attenuated by reduplication shown in Table 6.

ɓuwâ-ɓuwâ	'reddish'	<	**ɓuwâ**	'red'
peetà-peetà	'whitish'	<	**peetà**	'white'
pəlkâ-pəlkâ	'blackish'	<	**pəlkâ**	'black'
tlərtà-tlərtà	'greenish'	<	**tlərtà**	'green'

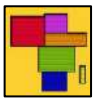
Table 6 – Attenuated colour terms in Gashua Bade

Note that Bade varieties are diverse to the extent that one is tempted to speak of different languages (cf. SCHUH 2007). Quite surprisingly, in Western Bade reduplication is not attested as a process which derives attenuated qualities, instead we find a few examples where reduplication of colour terms intensify their meaning, e.g. **hēta** 'white' > **hethētán** 'very white', **pəlka** 'black' > **pələk-pələkà** 'very black'.

4.2 Bole

Bole adjectives are discussed in a draft chapter of a reference grammar by Russell G. Schuh, which unfortunately could not appear.[4] Again we find that reduplication of adjectives may give them an attenuated meaning, like English "-ish", although not all adjectives allow reduplication, e.g. ***ɗòle-ɗòle** "smallish", ***gòŋ"-gòŋ"** "niceish" seem unacceptable. Bole usually employs full reduplication in this attenuation process, but there are also some adjectives which undergo partial, rather than full reduplication, e.g. **a'algàji** 'greenish' < **algàji** 'green'. The adjective **pètìlà** 'white' allows full reduplication (see below), but the partially reduplicated form **pèpètìlà** 'whitish' is also attested, and according to SCHUH (2005/Adjectives: 8) "the adjective root **ɓûl** 'yellow', which is usually used in the reduplicated form **ɓulɓùl** in the base meaning 'yellow', has a reduplicated form **ɓuɓulɓùl**". In Bole all formally plural adjectives also allow reduplication, however, unlike what we have seen in Hausa, reduplicated plural adjectives in Bole "distributes the base quality over the members of a group, with the additional implication of there being many such referents. Thus, **dāndên ɗolle-ɗollè** means 'a large number of children, each of whom is small, not 'children who are smallish'" (SCHUH 2005/Adjectives: 8).

[4] Glottolog 4.5 dates the draft chapters with 2005 (cf. HAMMARSTRÖM et al. 2021).



dài	'red'	>	**dài-dài**	'reddish'
bù'ùm	'black'	>	**bù'ùm-bù'ùm**	'blackish'
zòi	'pleasant, nice, tasty'	>	**zòi-zòi**	'rather pleasant'
ɓènèm	'sour, fermented'	>	**ɓènèm-ɓènèm**	'sourish'
pètìlà	'white'	>	**pètìlà-pètìlà**	'whitish'

Table 7 – Attenuated adjectives in Bole

4.3 Ngamo

SCHUH with Umaru MAMU Goge and Jibir Audu Janga DOLE (n.d.)[5] devote a whole manuscript to "-ish" reduplication in the Gudi variety of Ngamo. According to them (n.d.: 1) "the Gudi dialect of Ngamo has a productive nominal reduplication process that expresses the concept 'like a NOUN', 'NOUN-like', 'NOUNish'". Note that the same process also exists in the Yaya dialect of Ngamo, however, it is not described systematically in the manuscript because of insufficient data. In their manuscript the authors provide tables with reduplicated "-ish" forms without translation into English, i.e. the focus is not on the semantics of the reduplicated forms but rather on syllable weight principles which apply. Interestingly, in Gudi Ngamo it is reduplication of the first syllable the word, not full reduplication as observed in other languages discussed here. According to SCHUH, GOGE and DOLE (n.d.: 2)

> "the most striking feature of the "-ish" reduplicants is the consistent application of SYLLABLE WEIGHT POLARITY: if the first syllable is light, the reduplicated syllable copies the first CV but lengthens the vowel; if the first syllable is heavy (either CVC or CVV), the reduplicated syllable copies the first CV with the short counterpart of the vowel".

Table 8 repeats some of the examples given in SCHUH, GOGE and DOLE (n.d.: 2).

	BASE FORM	"-ISH" FORM	GLOSS
light syllable	**dàhù**	**dàdahù**	'cornstalk door panel'
	mèrì	**mèmerì**	'thatching needle'
	rùɗù	**rùruɗù**	'charcoal'
	kùrì	**kùkurì**	'red clay'
	ìdò	**ì'idò**	'eye'
heavy syllable	**rèndî**	**rèrendî**	'spear'
	kèrwò	**kèkerwò**	'fish'
	lànjà	**làlanjà**	'yellow clay'
	hetrè	**hèhetrè**	'white'
	ūnù	**ʔu'ūnù**	'black'

Table 8 – "-ish" reduplication in Gudi Ngamo

[5] Glottolog 4.5 dates the manuscript with 2009 (cf. HAMMARSTRÖM et al. 2021).

4.4 Zaar

According to CARON (2005: 227) in Zaar four adjectives are attested which "have a reduplicated form with a diminutive meaning, which can be translated into English by a [sic!] «-ish» suffix". The examples given in CARON (2005) are repeated in Table 9 and suggest that rather semantic attenuation than diminution is at place.

jì:	'black'	>	**jì:-jì:**	'blackish'
fyà:li	'white'	>	**fyà:li-fyà:li**	'whitish'
zhèlì	'red'	>	**zhèlì-zhèlì**	'reddish'
mbóci	'variegated'	>	**mbóci-mbóci**	'somewhat variegated'

Table 9 – Attenuated colour terms in Zaar

4.5 Kwami

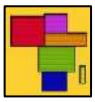
Following LEGER (1994) adjectives in Kwami, which are formed by reduplication from nouns, semantically usually get a "X-like, X-y, X-ish" meaning, e.g. **ʔàm(ì)** 'water' > **ʔàm(ì)-ʔàm(ì)** 'watery', **mór(í)** 'oil, fat, grease' > **mór(í)-mór(í)** 'greasy, fatty'. Furthermore, colour terms in Kwami, if fully reduplicated, have a detensified/attenuated meaning. In Table 10 examples from LEGER (1994: 132) are replicated.

tín	'black'	>	**tín-tín**	'blackish'
púr(í)	'white'	>	**púr(í)-púr(í)**	'whitish'
táy	'red'	>	**táy-táy**	'reddish'
shúdí	'blue'	>	**shúdí-shúdí**	'bluish'

Table 10 – Attenuated colour terms in Kwami

4.6 Malgwa

According to LÖHR (2002) Malgwa has a class of quality nouns which nevertheless may be distinguished syntactically from nouns and adverbs. Full reduplication accompanied by apocope of the final vowel in the first component derives attenuated meanings. LÖHR (2002) gives only examples of colour terms recapitulated in Table 11, and states that loanwords as well as "describing expressions" [beschreibende Ausdrücke] do not undergo reduplication. While it is not totally clear what is meant with "describing expression", the fact that loanwords are excluded might be a hint that the process is not productive anymore.



dzáyye	'white'	>	**dzáy-dzáyye**	'whitish'
ɬárəkə	'green'	>	**ɬárək-ɬárəkə**	'greenish'
kyáŋe	'red'	>	**kyáŋ-kyáŋe**	'reddish'
dóŋwe	'black'	>	**dóŋ-dóŋwe**	'blackish'

Table 11 – Attenuated colour terms in Malgwa

Like in Hausa reduplication in Malgwa applies to mainly singular concrete nouns to derive adjectives with the meaning "X-like, X-y, X-ish" (where X represents the source noun), e.g. **yáwe** 'water' > **yáwyáwe** 'watery', **ʔíízɜe** 'salt' > **ʔíízíízɜe** 'salty' (LÖHR 2002: 101).

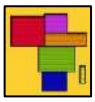
4.7 Margi

According to HOFFMANN (1963: 67) "in Margi the adjectives are fairly numerous and must be considered a special part of speech, which is often characterized by the possibility of forming a plural by reduplication". While he does not provide examples of attenuated adjectives, we nevertheless observe again that adjectives may be derived from nouns typically meaning "X-like, X-y, X-ish", e.g. **màl** 'oil' > **màlmàl** 'fat, oily', **ə́ntəbṵ̀** 'rubber' > **ə́ntəbə́ntəbṵ̀** 'rubber-like, sticky', **màlà** 'woman' > **màlàmàlà** 'of the kind of a woman' (HOFFMANN 1963: 68).

5. Conclusions

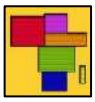
In this paper we illustrated attenuation by reduplication in Hausa and some other Chadic languages. We have seen that the process of full reduplication often accompanied by shortening or apocope of the final vowel produce forms that are semantically attenuated, i.e. they have the quality of the simple adjective but to a lesser extent. Attenuation by full reduplication is found basically in West Chadic languages like Hausa, Gashua Bade, Bole, Zaar, and in Margi, while West Chadic Ngamo shows reduplication of the first syllable of a word to the left. Attenuation by full reduplication with apocope of the final vowel in the first component is also attested in the Central Chadic language Malgwa.

As Hausa is probably one of the best documented languages of the African continent, it is not surprising that also attenuation processes are well described in various grammars (e.g. NEWMAN (2000), and JAGGAR (2001)). Apart from that our survey of something like 45 Chadic languages from all branches yields that information on processes that produce semantic attenuation is very scarce. Needless to say that attenuation probably is not the first and most important phenomenon one looks at when documenting a language. However, what becomes apparent from our scrutiny is that if we find information on attenuation



in a Chadic language, then it is full reduplication which is employed, or in the case of Gudi Ngamo reduplication of the first syllable.

Full and partial reduplication are very common processes in Chadic languages, see e.g. AL-HASSAN (1998). While various reduplication processes most often are associated with pluractional verbs, nominal plurality, and functions like intensification, distributiveness, etc., it is still puzzling how semantic attenuation may be linked to these processes.



References

AL-HASSAN, Bello S.Y. (1998). *Reduplication in the Chadic Languages – A Study of form and function*, Frankfurt am Main, Berlin, Bern, New York, Paris & Wien: Peter Lang.

BAUER, Laurie (2002). "What you can do with derivational morphology", in Sabrina BENDJABALLAH, Wolfgang U. DRESSLER, Oskar E. PFEIFFER and Maria D. VOEIKOVA (eds.). *Morphology 2000. Selected Papers from the 9th Morphology Meeting, Vienna, 24-28 February 2000*, Amsterdam & Philadelphia: Benjamins, 37-48.

CARON, Bernard (2005). *Za:r – Dictionary, Grammar, Texts*, Ibadan: IFRA.

FRAJZYNGIER, Zygmunt and Erin SHAY (2012). "Chadic", in Zygmunt FRAJZYNGIER and Erin SHAY (eds.), *The Afroasiatic Languages*, Cambridge: Cambridge University Press, 236-341.

GIMBA, Alhaji Maina and Russell G. SCHUH (n.d.). Ms. "Adjectives", in Draft chapters of a Bole reference grammar.

HAMMARSTRÖM, Harald, Robert FORKEL, Martin HASPELMATH and Sebastian BANK (2021). *Glottolog 4.5*, Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://doi.org/10.5281/zenodo.5772642>

HOFFMANN, Carl (1963). *A Grammar of the Margi Language*, Oxford: Oxford University Press.

JAGGAR, Philip J. (2001). *Hausa*, Amsterdam & Philadelphia: John Benjamins.

LEGER, Rudolf (1994). *Eine Grammatik der Kwami-Sprache (Nordostnigeria)*, Köln: Rüdiger Köppe.

LÖHR, Doris (2002). *Die Sprache der Malgwa (Nárá Málgwa)*, Frankfurt am Main: Peter Lang.

NEWMAN, Paul (1990). *Nominal and Verbal Plurality in Chadic*, Dordrecht & Providence: Foris.

NEWMAN, Paul (2000). *The Hausa Language – An Encyclopedic Reference Grammar*, New Haven & Lodon: Yale University Press.

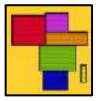
NEWMAN, Paul (2022). *A History of the Hausa Language – Reconstruction and Pathways to the Present*, Cambridge: Cambridge University Press.

SCHUH, Russell G. with Umaru Mamu Goge and Jibir Audu Janga Dole (n.d.). *"-ish" Reduplication in Gudi Ngamo* (Ms).

SCHUH, Russell G. (2005). Draft Forthcoming Bole Reference Grammar (Ms).

ZIEGELMEYER, Georg (2015). "On the adjective class in Gashua Bade", *Afrikanistik-Aegyptologie online* 2015.

JOURNAL OF AFRICAN LANGUAGES AND LITERATURES
3/2022, 120-132



Soldiers, coffee and markets: On Baldi's *Dictionary of Arabic Loanwords in the Languages of Central and East Africa*

GIAN CLAUDIO BATIC

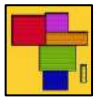
University of Naples L'Orientale
gcbatic@unior.it

The Dictionary of Arabic Loanwords in the Languages of Central and East Africa (hereafter DoAL) (2021) is Sergio BALDI's second magnum opus on Arabic loanwords in the languages of Africa. Preceded by Dictionnaire des emprunts arabes dans les langues de l'Afrique de l'Ouest et en Swahili (hereafter DEAr) (2008), the DoAL shifts the focus to the eastern and central quadrant of the African continent, also including and extending the data on Arabic loanwords in Swahili already present in the DEAr. From a general perspective and ideally, the DEAr and the DoAL should be considered as two parts of a single work in progress aimed at compiling all existing Arabic loanwords in the languages of the African continent. Of course, an ultimate goal of this kind clashes with the type of data available, so that any compilation work is inevitably limited by the nature of the sources. In the next section we will devote some space to the discussion of the actual representativeness of the DoAL in relation to the target languages and to their reality beyond the lexical descriptions we can provide of them.

The DEAr and the DoAL adopt the same indexing system for Arabic terms and present an isomorphic structure, so that the user can easily move from one volume to the other using the same numbering system and without having to go through new consultation guidelines. The end user will appreciate the structural consistency of the two dictionaries, in addition to the value given by the scientific compilation work that makes the DoAL – as it was for its predecessor – a reference work in the field of African lexicography.

This article has a twofold purpose: on the one hand it aims to provide critical observations of a theoretical and practical nature, and on the other hand it intends

DOI https://doi.org/10.6093/jalalit.v3i3.9168
ISSN 2723-9764



to outline possible uses and lines of research centred on the lexico-graphic data provided by the DoAL.

Before starting the discussion on some aspects of the DoAL, it seems important to frame the work in the perspective of the Author, i.e. to state his proposed objectives. The DoAL

> "[…] aims to provide a tool for researchers of African languages, to offer them a lexicon of these languages if they want to indicate Arabic loanwords in their work" (p. 1)

BALDI also states that his

> "[…] objective was to collect all loanwords, such that this work is as complete as possible, though some words may not be commonly used by speakers of a given language that seem to be pure Arabic xenisms […]" (p. 5)

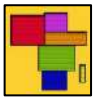
And it is precisely as a tool that we will look at the DoAL, trying to understand how it was designed and what uses it lends itself to.

1. Observations

Three observations can be made about the DoAL. They concern, in order, a) the languages examined, b) the problem of representativeness, and c) the nature of the sources.

1.1 Target languages

In the DoAL the Author considers 57 languages (plus Arabic). The list of languages analysed (pp. 2-4) provides the names used by the authors of the individual lexical or lexico-grammatical descriptions. Since many of the sources consulted are – with varying degrees of value – pioneering or dated, the naming of languages is sometimes obsolete and misleading and differs from the standard denominations used in modern language catalogues such as, for example, the ISO e Glottolog databases and the SIL's *Ethnologue*. In some cases the Author indicates groups of languages (e.g. Batéké, i.e. Teke languages; Gbéa, i.e. Gbaya languages), and in other cases the name of the most widespread variant whereas the source deals with a specific dialect or sub-variant. In Table 1, for each target language, the name of the language used by the Author, the standard name used in scientific literature (if different from that provided by the Author) or the name of the language actually covered in the source, the Glottolog code (HAMMARSTRÖM et al. 2021), and the ISO 639-3 code are provided. The languages are sorted alphabetically according to their genetic affiliation (the top-level family is indicated).



GIAN CLAUDIO BATIC

Soldiers, coffee and markets

TOP-LEVEL FAMILY	LANGUAGE	LANGUAGE NAME	GLOTTOCODE	639-3
Afro-Asiatic	Rendille	–	rend1243	rel
Atlantic-Congo	Ankole Lunyankole, Runyankore [1]	Nyankole	nyan1307	nyn
Atlantic-Congo	Batéké	Teke languages	kasa1251	–
Atlantic-Congo	Bemba	–	bemb1257	bem
Atlantic-Congo	Bende	–	bend1258	bdp
Atlantic-Congo	Ciluba	Luba-Lulua	luba1249	lua
Atlantic-Congo	Digo[2]	–	Digo1243	dig
Atlantic-Congo	Gbéa	Gbaya languages	gbay1279	–
Atlantic-Congo	Gikuyu	Kikuyu	kiku1240	kik
Atlantic-Congo	Gmbwaga	Ngbaka Ma'bo	ngba1284	nbm
Atlantic-Congo	Ngh'wele	Kwere	gwer1238	gwr
Atlantic-Congo	Haya	–	haya1250	hay
Atlantic-Congo	Ila	–	ilaa1246	ilb
Atlantic-Congo	Jita	–	jita1239	jit
Atlantic-Congo	Kamba	–	kamb1297	kam
Atlantic-Congo	Kikongo[3]	–	(core1256)[4] yomb1244	kon[5]
Atlantic-Congo	Kiluba	Luba-Katanga	luba1250	lub
Atlantic-Congo	Kinyarwanda	–	kiny1244	kin
Atlantic-Congo	Kiw'oso	Bosho	(mach1266)	(jmc)
Atlantic-Congo	Kuria	–	kuri1259	kuj
Atlantic-Congo	Lega	(Beya-Musange dialect)	lega1253	–
Atlantic-Congo	Lhukonzo	Konzo	konz1239	koo
Atlantic-Congo	Lingala	Kinshasa Lingala	ling1263	lin
Atlantic-Congo	Lomongo	Mongo	mong1338	lol
Atlantic-Congo	Lotuxo	Otuho	otuh1238	lot
Atlantic-Congo	Luena	Luvale	luva1239	lue
Atlantic-Congo	Luganda	Ganda	gand1255	lug
Atlantic-Congo	Lunyoro	Nyoro	nyor1246	nyo

[1] The Author lists the Nyankole language under three different names: Ankole, Lunyankole and Runyankore.

[2] Although found in the DoAL (e.g. entry 1862, **'askāri** 'military, army'), the language is not present in the initial list (pp. 2-4).

[3] The Author is considering the Kikongo (i.e. Kongo) cluster. At least one of BENTLEY's two works (1887, 1895) on the Kikongo (i.e. Kongo) variant spoken in M'banza-Kongo (San Salvador) has been consulted (the author is listed as one of the sources for Kikongo, p. 3), but the reference is missing.

[4] Out-dated.

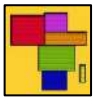
[5] Inclusive code (kon > kng, ldi, kwy, yom).

Atlantic-Congo	Macua	Makhuwa	makh1264	vmw
Atlantic-Congo	Matengo	–	mate1258	mgv
Atlantic-Congo	Ndogo	–	ndog1248	ndz
Atlantic-Congo	Ndonga	–	ndon1254	ndo
Atlantic-Congo	Ngh'wele	Kwere	gwer1238	gwr
Atlantic-Congo	Ngombe	–	ngom1268	ngc
Atlantic-Congo	Nyakyusa	Nyakyusa-Ngonde	nyak1260	nyy
Atlantic-Congo	Pokomo	–	poko1261	pkb
Atlantic-Congo	Sango	–	sang1238	sag
Atlantic-Congo	Shona	–	shon1251	sna
Atlantic-Congo	Sukuma	–	suku1261	suk
Atlantic-Congo	Swahili	–	swah1253	swh
Atlantic-Congo	Xhosa	–	xhos1239	xho
Atlantic-Congo	Zande	–	zand1248	zne
Atlantic-Congo	Zulu	–	zulu1248	zul
Austronesian	Malagasy	Kibosy Kiantalaotsy-Majunga	bush1250	buc
Central Sudanic	Madi	Ma'di	madi1260	mhi
Khoe-Kwadi	Khoi-Khoin	Nama	nama1264	naq
Kuliak (Nilo-Saharan)	Ik	–	ikkk1242	ikx
Nilotic	Acooli	Acoli	acol1236	ach
Nilotic	Alur	–	alur1250	alz
Nilotic	Anywa	Anuak	anua1242	anu
Nilotic	Ateso	Teso	teso1249	teo
Nilotic	Bari	–	bari1284	bfa
Nilotic	Dholuo	Luo	luok1236	luo
Nilotic	Dinka	Dinka languages	dink1262	din
Nilotic	Pokot	Pökoot	poko1263	pko
Nilotic	Shilluk	–	shil1265	shk

Table 1 – Languages analysed in the DoAL

1.2 Representativeness

The second issue to be addressed emerges – almost immediately – from the title and concerns the very object of the lexicographic compilation and its place in the real world. Since dictionaries are inevitably partial – albeit in many cases accurate and valuable – representations of the lexical inventory of a language, the DoAL could actually be labelled as a 'dictionary of Arabic loanwords *in the dictionaries* of the languages of Central and East Africa': that is, it is a refined product of partial and cultural products that cannot convey to us the ultimate truth, that is, the adherence of the data to the lexical reality as it manifests itself in the community of speakers. The use of compilations as intermediaries between the data and the



final product (i.e. the DoAL) is in itself unavoidable when operating on a large scale, but it is worth remembering that the existence of the loanwords dealt with in the DoAL is anchored not so much to the languages as to their available representations (lexical descriptions, dictionaries, glossaries). Certainly, it can be argued that any compilation or survey work of this kind is a product of products, or indeed that research is often a high-refined product built on sources (data collected and somehow processed) and not on raw data.

1.3 Nature of sources

In addition to the structural limitation of lexical descriptions, the inhomogeneity of the available sources must be pointed out. This inhomogeneity acts as much on a synchronic as on a diachronic level. It happens in fact that relatively recent lexical descriptions set themselves different goals of completeness, thus describing a current language but with a different degree of depth, or that descriptions far back in time are used alongside recent or less recent descriptions. The qualitative difference of the sources makes them difficult to compare: does HULSTAERT's (1952) impressive *Dictionnaire lɔmɔ́ngo-français* have the same value as an article devoted to loans, such as KABUTA's (1998) for the Ciluba (i.e. Luba-Lulua) language? Of course, the issue has not escaped the Author, even with regard to publications that fall into the same category:

> "The reliability of the works I consulted varies: the Swahili dictionary of Johnson is less reliable than that of father Sacleux with regard to quoting loanwords in general and Arabic in particular" (p. 5)

The Author worked with the sources he had at his disposal and was able to consult, being forced in many cases to rely on a single lexical description to find Arabic loans in a given language. This limitation, in the landscape of publications devoted to the lexical description of minority languages, is physiological. We should not be surprised if for Swahili, the Author lists ten sources, while for languages such as Lotuxo (i.e. Otuho) and Sukuma only one. To what extent is the abundance of sources for some languages and the scarcity of sources for others reflected in the final compilation? If for Swahili (and, in the DEAr, for Hausa) we can work with various sources that present a satisfactory degree of exhaustiveness, for other languages (most of them) we are somehow forced to rely on the quality of the only available source, resetting to zero all the parameters that usually guide us in the selection of sources (in no particular order: period of publication, type of publication, prestige of the publishing house, reputation of the author, number of entries, etc.).

2. Loanword mapping

The DoAL is an extremely useful tool for mapping the distribution of loans and their diffusion centres. The distribution of the languages analysed in the DoAL is shown in Figure 1 (the coordinates used to build the maps are taken from HAMMARSTRÖM et al. 2021):

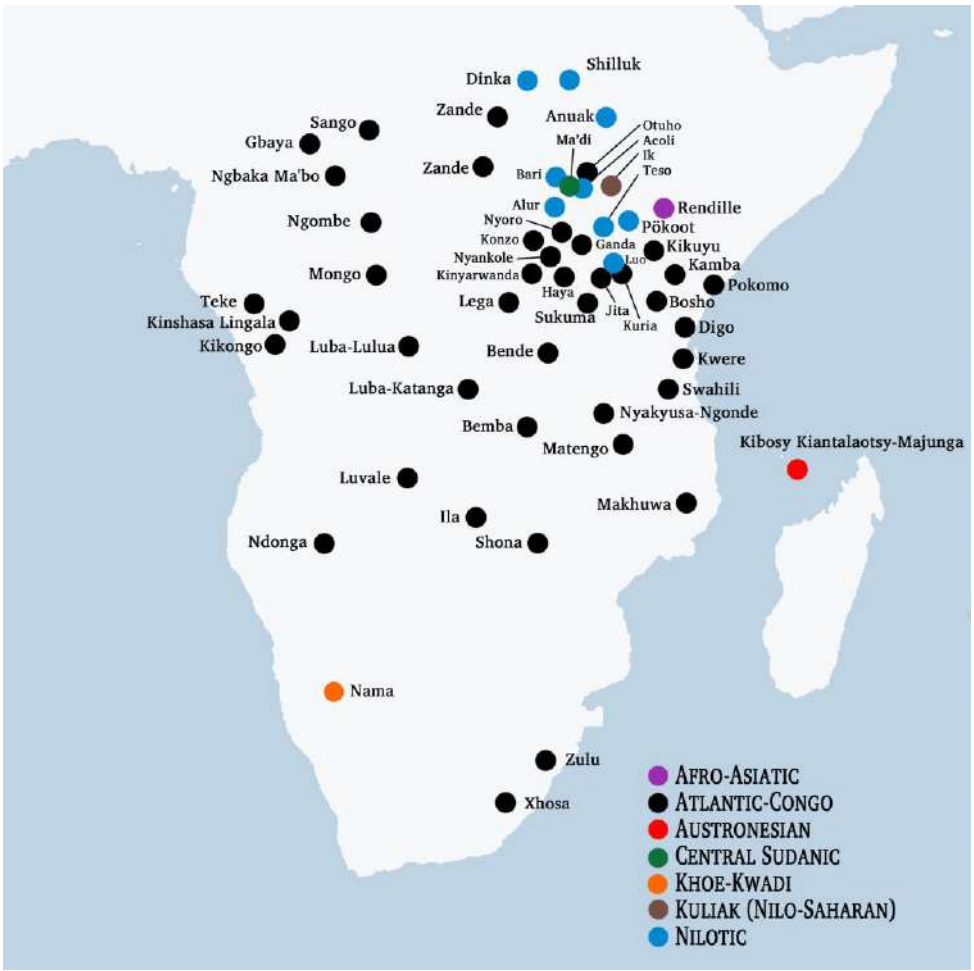
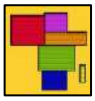


Figure 1 – Distribution and genetic affiliation of the languages examined in the DoAL

In order to analyse the diffusion paths, or at least the major centres of diffusion, it will be necessary to consider 'successful' Arabic loans, i.e. loanwords integrated into a significant number of languages. Figures 2-7 (see pages 127-129) show the distributions of the following Arabic loanwords: **sāʿa** 'while, hour, clock' (#1398),



sūq 'market' (#1402), **ʻaskāri** 'military, army' (#1862), **qalam** 'pen' (#2335), **qahwa** 'coffee' (#2350), and **māl** 'wealth, money' (#2684).[6]

Where the Arabic term has been integrated through another language (as indicated in the DoAL), a line connects the intermediate donor language to the language(s) that borrowed the term. In this case, the ISO codes of the intermediate donor language and the recipient languages are indicated.

As can be seen, Swahili plays a prominent role as a second donor language. Two observations can be made in this regard:

(a) the influence of Arabic acts primarily on the east coast. Arabic terms are integrated into Swahili and later, through Swahili, into other languages in contact with the Swahilophone world;

(b) the spread of Arabic loanwords through Swahili follows a north-west direction. The range of Swahili as a donor language finds its limit in the belt that includes north-central Kenya and the area bordering the northern and western shores of Lake Victoria. A first analysis based exclusively on the mapping of recipient languages allows us to formulate the hypothesis of the existence of two centres of diffusion: a southern centre corresponding to the Swahili-speaking area and exerting its influence towards the north and north-west, and a northern centre corresponding to the Nilotic and Afro-Asiatic area (although the latter is hardly represented in the DoAL) acting towards the south. This observation is quite evident if we consider the close contiguity with the Arabic language and the Arab-Islamic world of both the Swahilophone coastal area and the region corresponding to present-day Sudan and South Sudan. It is also important to stress the role played by the contact of the Nilotic and Afro-Asiatic languages with Sudanese Arabic (apd; suda1236) in Sudan, South Sudan and Eritrea.

The intermediate donor language is not always Swahili. Figures 2 and 6 indicate the case of the Kinshasa Lingala and Sango languages: the Arabic loanwords **sāʻa** 'while, hour, clock' and **qahwa** 'coffee' are supposed to have entered the Sango language via Kinshasa Lingala (in the case of **qahwa**, the Lingala term is not indicated in the DoAL).

[6] The six loanwords were selected to exemplify an informative mapping and provide elements for further analysis. Other terms that would have proved equally significant are, among others, **bunduqīya / bundūg / bundug** 'rifle, gun' (#317), **ẖaima** 'tent' (#869), and **sukkar** 'sugar' (#1314).

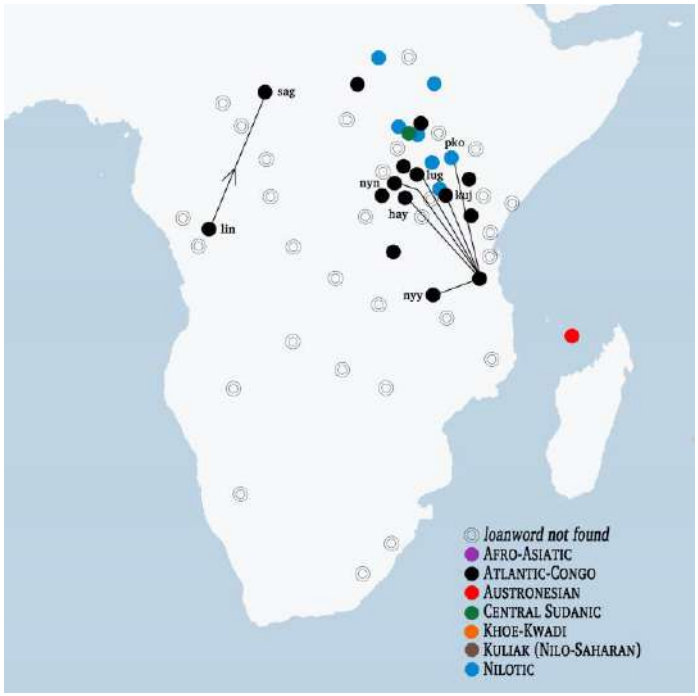


Figure 2 – **sā'a** 'while, hour, clock' (#1398)

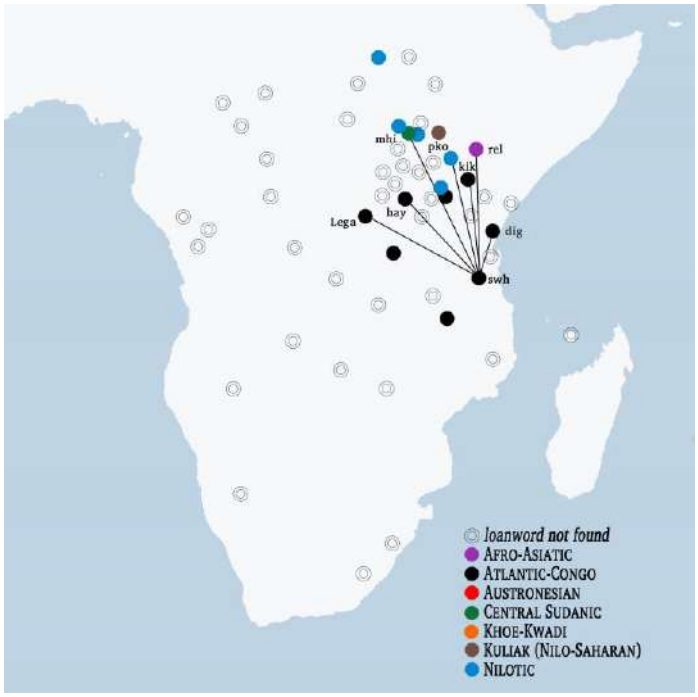


Figure 3 – **sūq** 'market' (#1402)

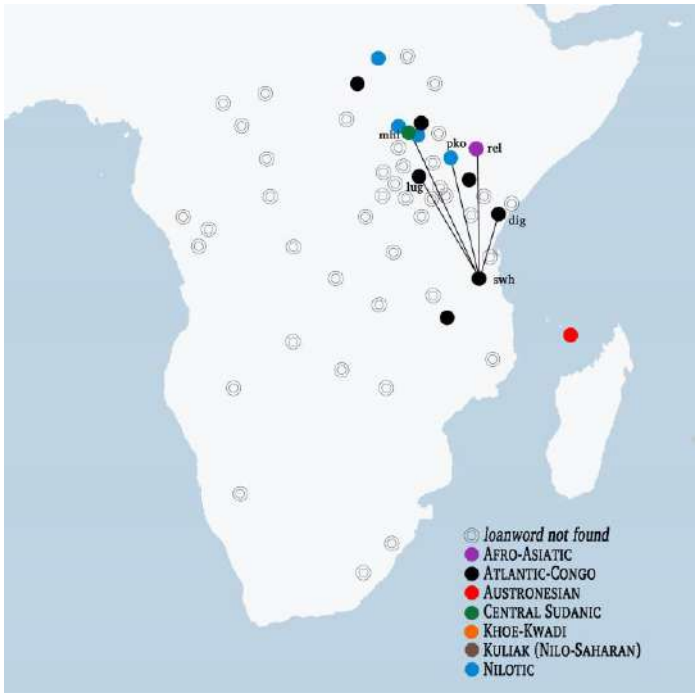


Figure 4 – **ʻaskāri** ‘military, army’ (#1862)

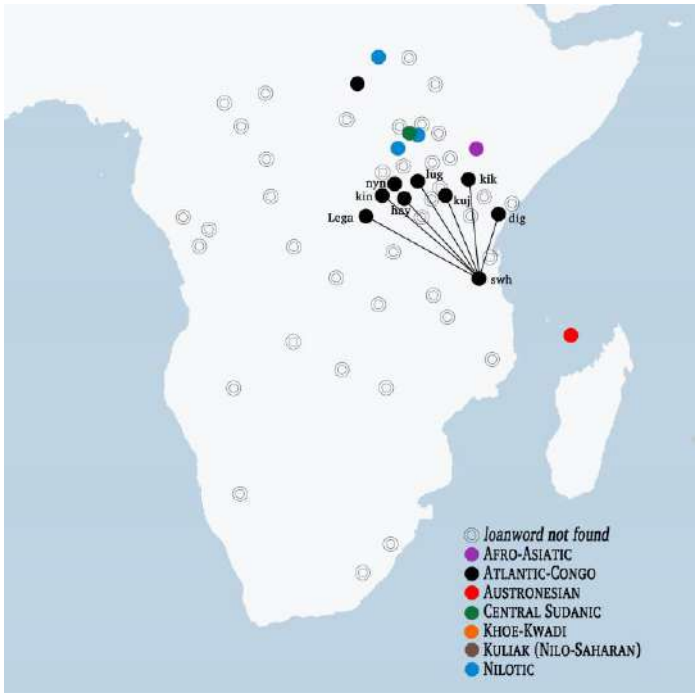


Figure 5 – **qalam** ‘pen’ (#2335)

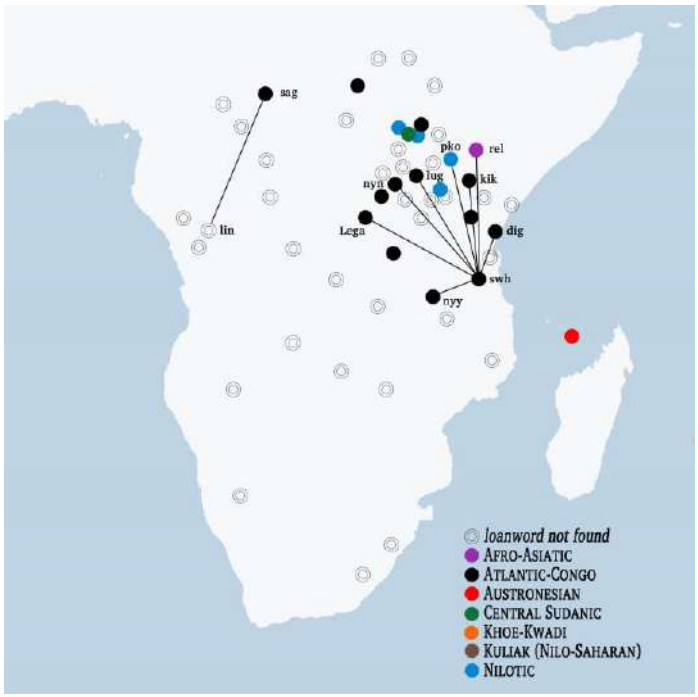


Figure 6 – **qahwa** 'coffee' (#2350)

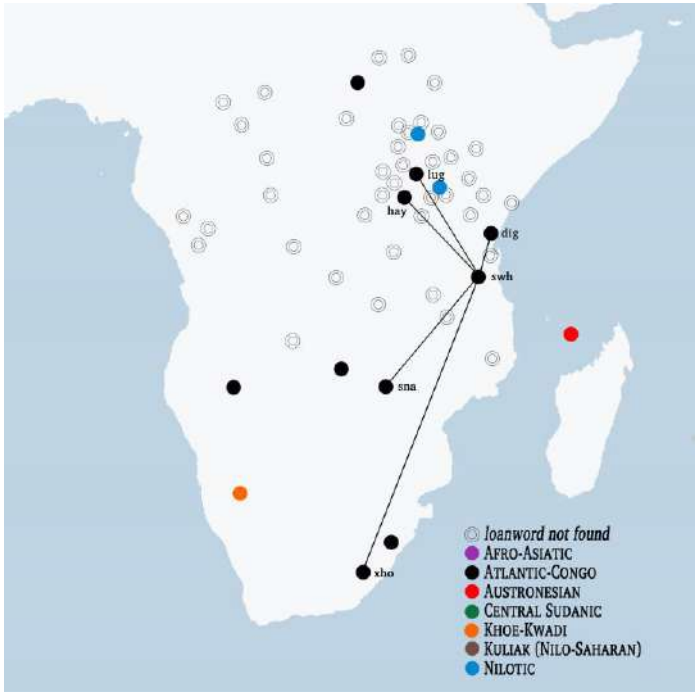
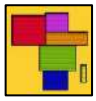


Figure 7 – **māl** 'wealth, money' (#2684)



The influence of a language as a vehicle of loanwords or the extension of material and immaterial objects resulting from direct or indirect contact with Arab-Islamic culture can be measured not so much by mapping individual loanwords, but rather by mapping entire lexical fields. The mapping of loanwords related to money (spending money, changing money, obtaining money as a result of a business transaction, etc.) and currency (coin(s), silver, dinar, drachma, piaster, etc.), for instance, would give us the situation illustrated in Figure 8:

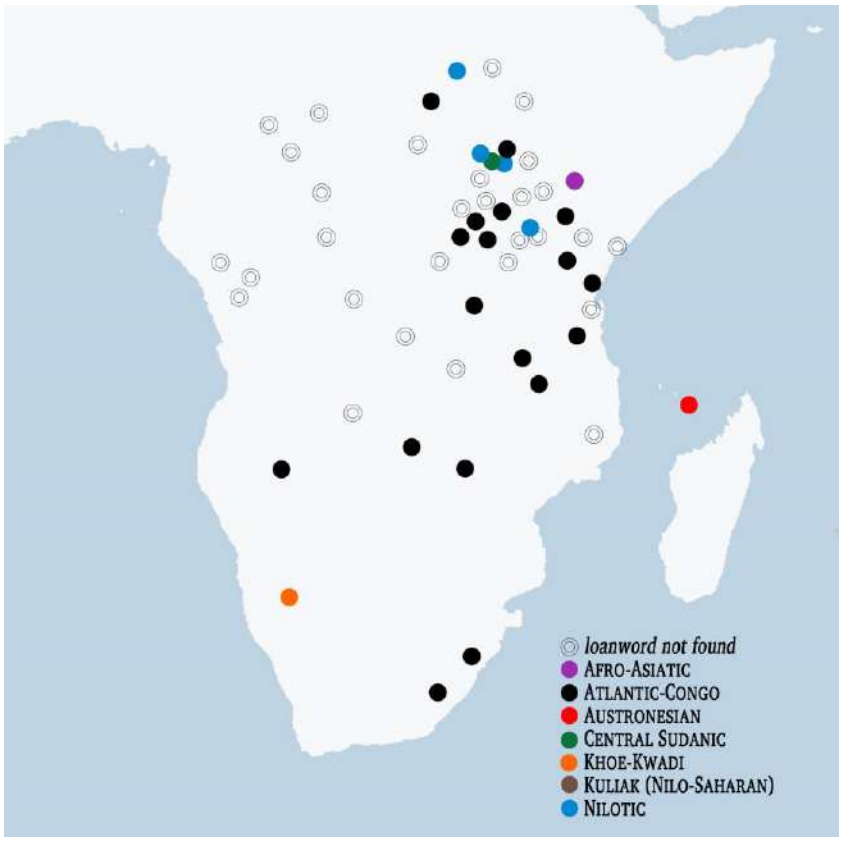


Figure 8 – Distribution of the money-currency lexical field

The lexical field considered consists of 19 Arabic terms distributed across 26 languages. The most receptive language is, predictably, Swahili (16 loans), followed by Kibosy Kiantalaotsy-Majunga (4), and by the Acoli, Ma'di and Diga languages (3 loans each). All other languages do not have more than two loanwords.

The DoAL can be easily used to conduct this kind of analysis by resorting to the Arabic, English and French indices at the end of the volume. Of particular interest is the comparison of purely lexical data (such as those obtained from the DoAL)

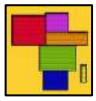
with historical data concerning the diffusion of material and immaterial culture elements of Arab and Arab-Islamic origin pertaining, for example, to trade, administrative practices, and religion.

We know that the Author has "consulted thousands of dictionaries" (p. 1). Unfortunately, we do not know to what extent the non-reporting of loanwords in certain languages or areas is due to the absence of adequate sources or to the absence of loans. This limits the ability to carry out a complete mapping of loanwords: it is not possible to determine, relying solely on the DoAL, which languages (apart from the problem of source reliability discussed above) have demonstrated a degree of impermeability to the penetration of Arabic loanwords (and indirectly to Arab and Arab-Islamic culture).

3. A conclusive remark

The main aim of this brief discussion was to formulate critical observations and explore some potentialities offered by the DoAL.

The DoAL is the result of years of research in the field of Arabic loans in African languages. A solid work of bibliographic research as well as a rigorous compilation, the volume promises no more than what it gives – and what it gives is a lot. Indeed, the DoAL presents itself as a solid basis for exploring the dynamics and patterns of penetration and integration of Arabic terms in African societies, which makes it a valuable tool for anyone wishing to approach the topic of loanwords from a lexical, historical, anthropological, semantic, or semantic-cognitive perspective.



References

BALDI, Sergio (2008). *Dictionnaire des emprunts arabes dans les langues de l'Afrique de l'Ouest et en Swahili*, Paris: Karthala.

BALDI, Sergio (2021). *Dictionary of Arabic loanwords in the languages of Central and East Africa*, Leiden & Boston: Brill.

BENTLEY, W. Holman (1887). *Dictionary and grammar of the Kongo language, as spoken at San Salvador, the ancient capital of the old Congo empire, West Africa*. London: Baptist Missionary Society.

BENTLEY, W. Holman (1895). *Appendix to the dictionary and grammar of the Kongo language, as spoken at San Salvador, the ancient capital of the old Congo empir*e, West Africa, London: Baptist Missionary Society.

HAMMARSTRÖM, Harald and Robert FORKEL, Martin HASPELMATH, Martin and Sebastian BANK (2021). *Glottolog 4.5*, Leipzig: Max Planck Institute for Evolutionary Anthropology, <https://doi.org/10.5281/zenodo.5772642>, last access on 1/3/2022, <http://glottolog.org>.

HULSTAERT, Gustaaf (1952). *Dictionnaire lɔmɔ́ngo-français (lonkundo)*, Tervuren: Annales du Musée Royal du Congo Belge.

KABUTA, Ngo Semzara (1998). "Loanwords in Cilubà", *Lexikos* 8: 37-64.