ACAL in SoCAL

Selected papers from the 53rd Annual Conference on African Linguistics

Edited by

Yaqian Huang

Nina Hagen Kaldhol

Jun Jie Lim

Sharon Rose

Anthony Struthers-Young

Contemporary African Linguistics 11

language
science
press

Contemporary African Linguistics

Editors: Michael R. Marlo, Laura J. Downing, Kristina Riedel

In this series:

ACAL in SoCAL

Selected papers from the 53rd Annual Conference on African Linguistics

Edited by

Yaqian Huang

Nina Hagen Kaldhol

Jun Jie Lim

Sharon Rose

Anthony Struthers-Young

language
science
press

Freie Universität Berlin

Contents

Contents

Preface

The 21 papers in this volume were presented at the 53rd Annual Conference on African Linguistics, hosted by the University of California San Diego from April 7–9, 2022. The conference was held online due to the ongoing effects of the COVID-19 pandemic. In recognition of UC San Diego's location in Southern California, colloquially known as 'SoCal', we chose to title the volume *ACAL in SoCAL*. We regret that our conference participants were not able to visit SoCal in person and appreciate all it has to offer.

ACAL53 featured 85 talks, and four invited talks by plenary speakers. The invited speakers were:

- Pius Akumbu (LLACAN - INALCO)

- James Essegbey (University of Florida)

- Fridah Kanana Erastus (University of Nairobi)

- Gertrud Schneider-Blum (Cologne University)

We take this opportunity to thank them once again for their contributions, which showcased the varied facets of research on African languages and linguistics, and stimulated lively discussion. Pius Akumbu's thoughtful and reflective article challenging current community-based language documentation practices is featured in this volume.

Acknowledgments

Of the 28 papers submitted for inclusion in the volume, 21 were accepted for publication, three were rejected and four were withdrawn by the authors following review. Publication of the *Selected Papers from the Annual Conference on African Linguistics* relies on the expertise and willingness of reviewers to assess submitted papers and ensure high quality. We are extremely grateful to fifty-three reviewers who submitted thoughtful and helpful reviews. They are listed below.

Acknowledgments

Chapter 1

A community approach to language documentation in Africa

Pius W. Akumbu

LLACAN (CNRS - INaLCO - EPHE)

Although still not common, some recent language documentation projects in Africa have adopted the community-based approach (Good 2012, Bischoff & Jany 2018) which allows for the participation of community members at various stages of the documentation process. This approach is highly favored over the traditional method of linguistic inquiry that was based on the speech of one or a few individuals, i.e. the "ancestral code" (Woodbury 2011). However, it appears that even within this current framework, the linguist reaps the greatest academic and financial benefits while community members end up with linguistic outputs that do not meet their current livelihood needs. I argue in this paper that instead of focusing on the documentation of African languages while neglecting current survival needs of community members, and thereby legitimizing and accompanying language death, linguists and funding agencies should conceive linguistic projects as community development projects. If linguistic projects assist in community development and maintenance, languages and cultures are more likely to be preserved, making linguistic work meaningful and useful to African communities.

1 Background: My personal story

For the purpose of this paper, I would like to provide some background information about myself. I am that village boy who was born to parents who practiced subsistence farming in Babanki Tungo in North-West Cameroon. My mom and my sisters went out early to the fields each working day and tilled the soil all day with hoes, usually under the scorching sun. In the evening they returned with food and cooked, using firewood, for us to eat before going to bed. My father and

the rest of the male folk were responsible, among other things, for clearing the fields and gathering firewood. This circle continued daily until the country Sunday, which is the resting day of the week (we have an eight-day week, one of the days is set for rest, another for business at the local market, and the remaining six for farming). That was and still is the typical lifestyle of a Babanki family. When I turned six and my right hand could go over my head and touch my left ear, I was ready for school where I met and was forced to use English although I knew and spoke only Babanki. I trekked barefooted for approximately 3.5 miles each way. After school, I went another mile to fetch water and eventually firewood, which will be used for cooking. By the time food was ready and I ate, I would be so exhausted that all I could do was listen to a few stories before sleeping off on my bare bamboo bed without any mattress, just to start all over the following morning. In the end it worked out favorably for me because I somehow was able to make it through primary school and move on to secondary school. Only four of my 72 classmates did the same, and I was the only one who went on to the graduate level. The secondary school was 15 miles away from home. Since I could not trek that distance on a daily basis, I had to go live there and walk home every Friday to gather food which I carried back on my head on Sunday. I eventually went to the University of Yaoundé 1, in a distant French-speaking part of Cameroon some 350 miles away from Babanki. There I studied linguistics. Interestingly, two decades later, I returned to Babanki for fieldwork and what I observed was the same cycle for those Babanki children who were still in the community. The situation has remained the same several decades down the line. The lands have become barren, there is no electricity, no potable water, children still walk several miles to school, women still till the fields with hoes, and so on. On the other hand, there are tons of journal articles and books on the Babanki language. Many linguists, including myself, have become successful and famous, with various promotions because of data from Babanki and other language communities whose lots have not changed. The issue I struggle with is how linguists can contribute to improving the lives of people who produce the languages they study. In this paper, I try to make a case for linguistic work as a part of community development, something different from the data mining that continues to happen in various forms in parts of Africa.

2 The state of African languages and cultures

While the dynamics of language use and evolution in multilingual Africa provide some optimism about the future of languages of the continent (see, for example,

Di Carlo & Good 2020, Lüpke 2019, Mufwene 2004, 2016), the endangerment of the languages and cultures of Africa continues to accelerate. From personal experience, I can attest to the decline in use of the Babanki language and the loss of many cultural practices. I speak as a Babanki community member who spent most of his first 20 years in Babanki before moving to continue school and eventually work in various parts of Cameroon. While growing up forty-fifty years ago we all mostly spoke Babanki in the community, a scenario which has changed drastically due to the influence of modernism which has ushered in Cameroon Pidgin English, English, and French.[1] Babanki is not favored since it is neither a language of education, work, nor business. This scenario is true of other rural communities across Cameroon and elsewhere in Africa. The persistence of language endangerment across the African continent can be linked to several factors, including, but not limited to, economic pressures, increased mobility,[2] spread of mass/social media, and pressure from dominant languages.[3] It can also be seen as one of the consequences of colonialism since the colonial project for Africa mostly portrayed whatever was African (language, culture, religion, etc.) as negative. As elsewhere in the world, e.g. among Native American Indians (see Holmes 2018), the colonizer set out to erase and make invisible what belonged to the colonized. The French assimilation policy, for instance, was meant to kill the African in us, to make us hate ourselves and believe that we have to become like our colonizers to survive (Zambakari 2021, Eko 2003). After more than a century of enforcing and reinforcing that, aspects of traditional African systems and practices such as education and religion have been transformed radically. Africans have made so much effort and progress to be civilized and modernized; to transition from "primitive", "pagans", "savages", and "uneducated" to fit colonial standards. Some Africans go as far as hating themselves, hating their skin color, knowledge, innovation, skills, and ultimately lack belief

---

[1]Cameroon Pidgin English, the lingua franca of Anglophone Cameroon, is common among youth and students who also speak some English and to a lesser extent French. English and French are the two official languages of Cameroon. English is the language of education and administration in Anglophone Cameroon (Northwest and Southwest Regions) while French is used in the other eight Regions of the country (Francophone Cameroon). The Babanki people who have been to school and learned English (since it is the language of instruction) may speak it with each other whereas French is occasionally used by those who have been exposed to it by living in Francophone Cameroon or learning it in school as a foreign language.

[2]A related factor is the modern capitalist economy (connected to colonialism) where lingua francas became important. People learned Pidgin on the plantations, for example, and brought it back to the villages.

[3]The argument that the evolution of a language is dependent on its ecology has been developed by Mufwene (2001, 2005, 2008).

in themselves. Others, including parents, siblings, peers, and teachers shame and punish children for speaking their languages (Bwesigye 2014). People are shamed for remembering and practicing their cultural and traditional customs and rites. Consequently, many Africans have remained in their shells, hiding their true selves and pretending to be who they are not and can never become. Ngugi (1986: 9) argues that "language was the means of the spiritual subjugation" such that the moment Africans lost their languages was also the moment they lost their bodies, gold, diamonds, copper, coffee, tea, and many other natural resources. In brief, the domination of African languages by the languages of the colonizers was the means to dominate the African mental universe and all that goes with it – culture, art, dances, religions, history, geography, education, and so on.

This is to say that colonialism contributed its share to language and culture endangerment and death in Africa. As mentioned above, Africans have been made, directly or indirectly to believe that all that is African, including our languages and cultures is worthless, and that we need to be transformed, that we need to develop not just socio-economically but mentally. In such a context it is difficult to get people to regain interest and rebuild a sense of self-worth for themselves, their languages and cultures. Africans are now rushing for formal education and all the modern technological advances, and it does not seem reasonable for us to be convinced that our erstwhile informal education, passed orally from generation to generation through storytelling around a fireplace, as well as our traditional technology, including our means of communication such as the talking drums, are worthy and should be preserved and promoted. Some people used all the means at their disposal at some point to successfully make Africans know and believe that their practices were primitive and ugly and should be abandoned. We are now working so hard to get better at abandoning and taking up what is "best", as we have been made to understand.

On the other hand, current linguistic research (description and documentation) kind of validates, stands by, watches, accompanies, and even accelerates the language endangerment process. Keeping aside SIL linguists who mostly do linguistics and language development over extended periods for religious purposes, linguists are rushing to document African languages in order to have something to look back to when the languages would have died completely. Many Western researchers obtain funding, go to Africa, gather some community members, do community-based documentation, prepare and archive recordings, conduct analyses, attend and present at conferences, publish, advance linguistic theory, gain academic promotion and fame, while the community continues to lose its language and culture. It is not about the community; all that matters is the language for current and future exploitation by those with the skills and resources.

Newman (1998: 15) insists that "the justification for doing research on an endangered language has to be the scientific value of providing that documentation and in preserving aspects of that language and culture for posterity. The purpose cannot be to make the few remaining speakers feel good." As Harrington (1941) put it, "if you can grab these dying languages before the old timers completely die off, you will be doing one of the few things valuable to the people of the remote future. You know that. The time will come and soon when there won't be an Indian language left in California, all the languages developed for thousands of years will be ashes, the house is afire, it is burning. That's why I said to go through the blinding rain, roads or no roads...". While this is laudable as a call to keep records of languages before they die, it can be seen as a way of condoning and accompanying death much more than working to prevent it. By documenting languages and watching them die instead of seeking to conserve and preserve, the colonial practice is reinforced. The linguistic world has evidence and is so convinced that many of the world's languages are in danger of disappearing (Krauss 1992, Crystal 2000). Language endangerment is used to justify language documentation – it ensures that researchers (especially those trained in European and American universities) obtain language documentation funding.[4] As required by their donors, funding agencies are able to fund only the documentation of highly or critically endangered languages (the Endangered Languages Documentation Programme (ELDP), for example, would not even consider funding revitalization projects).[5] This is quite unusual. Humanity does not have the habit of standing by and watching death happen. Usually, efforts are made to

---

[4]The general lack of influence of native speaker linguists in the field of linguistics has been pointed out previously (Crowley 2007, Tsikewa 2021) and this is quite obvious in language documentation in Africa. For example, I found that of the 108 ELDP grants in Africa between 2003 and 2019, 75 percent had been granted to those trained in Europe or America and only 25 percent to those trained in Africa (https://www.eldp.net/en/our+projects/projects+list, accessed on December 3, 2021). Overall, those trained in Africa manage to get the Small Grants while those trained elsewhere obtain the Major Grants.

[5]Similarly, NSF projects may involve one or more of the following three emphasis areas: (1) Language Description: To conduct fieldwork to record in digital audio and video format one or more endangered languages; to carry out the early stages of language documentation including transcription and annotation; to carry out later stages of documentation including the preparation of lexicons, grammars, text samples, and databases; to conduct initial analysis of findings in the light of current linguistic theory; (2) Infrastructure: To digitize and otherwise preserve and provide wider access to such documentary materials, including previously collected materials and those concerned with languages which have recently lost all fluent speakers and are related to currently endangered languages; to create other infrastructures, including conferences to make the problem of endangered languages more widely understood and more effectively addressed; (3) Computational Methods: To further develop standards and databases to make this documentation of a certain language or languages widely available in

find a cure. The world has not stood by to watch COVID-19 gradually destroy humanity. All kinds of efforts are being made, particularly between 2020 and 2022, to find a way out of the virus. We cannot afford to and should not continue to accompany endangered (sick) languages to death.

This is the context in which language documentation activities and linguistic fieldwork in Africa have been taking place over the last 60 years or so. In some sense the colonial experience has been carried over to linguistic work in the African continent. While African languages have contributed largely to the advancement of linguistic knowledge (see, for example, Hyman 2003a,b, Lionnet & Hyman 2018), Africa itself has probably benefitted only in very minor ways from the study of its languages. Tsikewa (2021) has demonstrated that until now, linguistic fieldwork training has largely neglected the needs of communities such that colonial approaches to linguistics continue to prevail. Most concerns identified and addressed in fieldwork courses and training revolve around the well-being of the fieldworker, not that of the community. She further points out that critiques of the kind of training provided to those who do linguistic fieldwork are many (Macaulay 2004, Ahlers 2009, Grenoble 2009, Newman 2009, Brickell 2018), reflecting the reality that community members' needs and wellbeing are not equally prioritized. At best, what is said about community needs is limited to giving back products of linguistic work such as dictionaries, grammars, reading materials, as well as providing training to community members, gifts that are not immediately useful to the community, as I will further discuss later. Linguists focus on language and its documentation ignoring the interests and needs of communities. They argue for the usefulness of documentation for future generations, ignoring the current users and their needs. It is not entirely helpful to document a language to keep it in archives while allowing the community to disappear. In this sense, SIL linguists generally seek to promote literacy, education, socio-economic opportunities, even if the overall goal is Bible translation and spiritual growth.[6] I submit in this paper that if linguists should have the

---

consistent, archivable, interoperable, and Web-based formats; to develop computational tools (taggers, parsers, speech recognizers, grammar inducers, etc.) for endangered languages, which present a particular challenge for those using statistical and machine learning, especially deep learning methods, since such languages do not have the large corpora for training and testing the models used to develop those tools; and to develop new approaches to building computational tools for endangered languages, which make use of deeper knowledge of linguistics, including language typology and families, and which require collaboration among theoretical and field linguists, and computational linguists, and computer scientists and engineers (https://www.nsf.gov/pubs/2020/nsf20603/nsf20603.htm, accessed on March 3, 2022).

[6]Linguistic work by missionary linguists has come under criticism (e.g. Dobrin & Good 2009, Grenoble & Whaley 2005, Handman 2009, Keane 2002, Pennycook & Makoni 2005, Rehg 2004) but my concern here is on linguists' contribution to community development in general.

interest of communities in mind and consequently, think of ways to preserve the communities where they work, then language endangerment will be slowed down or averted to some extent. As a member of a community that has been the object of linguistic research for decades, I propose ways in which linguistic work can be transformed from being purely linguistic projects to community development projects in which language documentation occupies some space in the scale of preference. Projects that encourage community members to desire to continue living in their community or attract outsiders into the community should be highly encouraged. Linguists should do all they can to accompany these languages and practices along with their speakers into modernity.

A number of questions come to mind, including how linguistic work in Africa can be done in a way as to create a lasting positive impact, not only on communities, but in the global effort to attain equality of some sort for all and promote cultural diversity. I certainly do not have the perfect answer to the question but will propose a community approach which takes into account the development needs of communities. Inspiration for this comes from the fact that in rural Africa, the most linguistically vibrant areas are those with potable water, health centers, electricity, roads, and schools. Therefore, linguistic projects should seek to provide some of these facilities in order to help maintain communities and encourage the preservation and use of their languages. Before presenting the community approach to language documentation, I will first discuss why the popular community-based approach on which I build needs to be revisited.

3 The community-based approach to language documentation

Community-based research has become a valued model in linguistic research in recent years, particularly in the areas of language documentation and revitalization. According to Rice (2018: 15), "community-based research begins with a research topic of practical relevance to the community and is carried out in community settings. Second, community members and researchers equitably share control of the research agenda through active and reciprocal involvement in the design, implementation, and dissemination. Finally, the process and results can transform and mobilize diverse ideas, resources, and experiences to generate positive action for communities."

This research model that emphasizes collaboration between linguists and language communities encourages research on a language, conducted for, with, and by the language-speaking community within which the research takes place and

which it affects. This kind of research involves a collaborative relationship between researchers and members of the community. It is community-based because a researcher joins efforts with community members to carry out activities in that community for the benefit of both parties (Bischoff & Jany 2018, Cameron et al. 1992, Czaykowska-Higgins 2009). According to Ochocka & Janzen (2014), community-based research is community-driven (i.e. it begins with a research topic of practical relevance to the community and promotes community self-determination), participatory (i.e. community members and researchers equitably share control of the research agenda through active and reciprocal involvement in the research design, implementation and dissemination), and action-oriented. The goals of the researcher and community members must be clearly defined in order to establish a productive long-term collaboration in which both parties benefit from the interaction. Leonard & Haynes (2010) stress the importance of collaborative consultation in defining research roles and goals. Ameka (2006: 70) insists that "unless the records of the languages being documented are the product of collaboration between trained native speaker and non-native speaker (anthropological) linguists, they will not be real, or optimal descriptions representing the realities of the languages."

A closer examination of this current approach to language documentation reveals a series of pitfalls that beg urgent reconsideration. It appears that in community-based projects in the African context, the researcher brings the expertise, equipment, tools, finances while the community actually just provides the language – the mining field (Kadanya 2006, Akumbu 2020). In some cases, to be allowed to supply the language data, i.e. participate in a project, community members are required to sign various papers or provide their finger prints – sometimes an extremely traumatizing experience to those who have never been to school or who dropped out and may not even be able to write their names. Holmes (2018: 153) highlights some of the shortcomings of the community-based approach raised by McDonald (2003: 84) who argues that "first, community-based research is located in communities. So what? Almost by definition ethnographic research is located in communities... The issue is that community-based research needs to be about something more than location". Continuing from McDonald's theorizing, Evans (2004: 60) notes that "...the term 'community based' says nothing about the role of the community in the research process." Community members in rural parts of Africa may, at best, give an opinion on some aspects of the project but rarely participate in crucial decision making since they do not have the academic ability or financial power to do so. I examine some of the issues in detail in order to demonstrate that community members in the African context do not benefit much from current linguistic work.

Concerning the identification of community needs, the expectation is usually that the linguist establishes contact with the community prior to submitting a grant proposal. This means that the linguist has the opportunity to find out what the community needs and to aim at meeting such needs. In general, the focus is on linguistic needs since they fall within the scope of the researcher's focus. Hardly ever are social and community development needs taken into account (Ngué Um 2019: 377-383), presumably because linguists are not social workers or because such "participation exacts a great deal from the fieldworker" (Samarin 1967: 14). It might also be that the linguist simply wishes to focus on pure fundamental research and "resist the ever-present pressure to justify our work on grounds of immediate social relevance" as recommended by Newman (1998: 15). Therefore, the linguist, sometimes in collaboration with some community members, identifies some linguistic product such as a dictionary that will be given back to the community as a means of meeting the community's desire to have its language in written form. This might have been done by asking people yes-no questions like, "Would you like a dictionary for your language"? I find this misleading because in most parts of Africa, linguistic needs do not feature as priority among the livelihood and survival needs of communities. In February 2022, using WhatsApp, I asked Cameroonians living both in cities and rural areas what they thought were the needs of their communities. Of the 126 responses I got, the top five needs were potable water (126), roads (125), electricity (124), health facilities (117), schools (115). A few others mentioned food, security, and peace, as shown in Figure 1.

Interestingly, no respondent mentioned anything related to language and culture. Of course, if people have survival needs and are given a choice they will certainly point to pressing current needs; they won't ask for language development and preservation which cannot contribute to their livelihood at the moment. The results of this survey point to the fact that community needs identified and focused on in linguistic projects do not actually reflect the real needs of communities in most parts of Africa. The linguistic needs that communities are made to identify or accept are constrained by the options linguists present to those who represent communities in the projects. When the researcher eventually meets the need by "giving back" some output of the project, e.g. dictionary, storybook, etc. to the community, it doesn't do much good to them because more than 90 percent of community members are probably illiterate in the language of the publication and cannot consume the product(s). The limits of such intellectual materials have been pointed out by several researchers, e.g. Mufwene (2016, 2017), and Akumbu (2020: 84) who observes that "sometimes, a linguist can think of giving back to the community and some copies of the grammar may be made available – but of
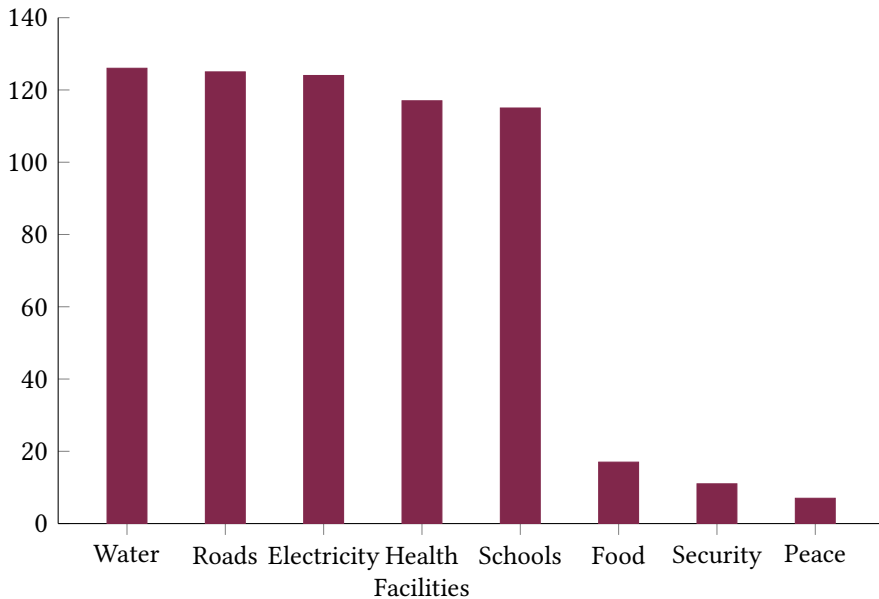
Figure 1: Community needs identified by Cameroonians

what use is this to people who in most cases are illiterate in the foreign language used to write the grammar, and are also unable to read and write in their own language?"

Tsikewa (2021: 309) notes that viewing "giving back" via pedagogical resources as equivalent to collaboration is a widely perpetuated misconception. She adds that according to Sapién (2018: 208) "giving back" via pedagogical resources is not representative of "true collaboration [that] seeks to 'work together' to set goals and undertake projects that are of balanced mutual benefit and depend on contributions from all stakeholders." Without doubt, community members in the African context usually welcome and celebrate linguistic work and its products since that is, at least, better than nothing. It is often so when someone does not have a choice and typically does not have the opportunity or privilege to decide, as I will demonstrate shortly. Obviously, not many people will choose what may only be useful for future generations since they also have and want to meet current needs. If linguistic products can only preserve the language or culture for future generations when the language would have died, people will welcome them only because they cannot have what will serve their immediate needs. In fact, as suggested by Ngué Um (2019), language preservation efforts should be combined with social work. He proposes that "...in critical language endangerment situations of West Africa where language survival and economic

welfare are intertwined, both actions should inform each other, and be carried out collaboratively, so that the people whose language is endangered may be 'saved' along with their language, as opposed to being left to perish while the language is preserved" (Ngué Um 2019: 391).

This is to say that linguists (and funders) need to give community members a chance to identify their real needs and then integrate them in their language development projects. Akumbu (2020: 91) mentions the Pig for Pikin intiative of the KPAAM-CAM project in Lower Fungom[7] and the water supply initiative implemented by the Beezen Language Documentation Project[8] which resulted from researchers fulfilling the wishes of two remote communities in North-West Cameroon to provide basic needs revealed by community members themselves (see Good 2012). Linguistic products would have been of no immediate use to these communities and the researchers, being outsiders, could not know exactly what the communities needed most. Therefore, giving community members a chance to identify what is useful for the entire community at that point in time is the best way to understand and integrate community needs in a language documentation project. People could be asked open ended questions such as "what do you need?" Responses to such a question will most likely point to urgent community needs. This leads to the second issue, that is, who benefits from community-based language documentation projects?

The aim of a language documentation project is to provide a comprehensive record of the linguistic practices characteristic of a given speech community (Himmelmann 1998). The materials collected are archived so that they can be accessible to other interested parties. Funders of linguistic work desire to sponsor projects that meet the goal of documenting and archiving. One of ELDP's objectives is to "create a repository of resources for linguistics, the social sciences, and the language communities themselves".[9] When funding is obtained, the researcher travels to the community and collaborates with a handful of community members to create recordings of the language in use. The researcher goes ahead to process the collected materials, and eventually archives them out of the community in high standard digital formats.[10] The researcher most likely produces

---

[7]https://ubwp.buffalo.edu/kpaamcam/research-communities

[8]https://www.aai.uni-hamburg.de/afrika/medien/beldop.pdf

[9]https://www.eldp.net, accessed March 4, 2021.

[10]Community members who provided the data may have little knowledge of where the materials are archived, nor what can be done with them. Even if the researcher had clearly informed them and, of course, obtained their verbal consent that the materials will be kept at ELAR in the UK or CERDOTOLA in Cameroon, that really does not have any implications for the community at that moment. The open access materials are available to the researcher and other interested persons, but are not as accessible for the community members.

publications of various sorts, and gains recognition and academic advancement. The successful completion of one project increases the researcher's chances of obtaining more grants. This is to say that the researcher clearly obtains both financial and academic benefits, some in the short term and others in the future. On the other hand, the funders use the success stories of researchers to justify expenses and secure further funding from their donors.

An interesting activity that deserves attention is archiving. The current archiving paradigm is where the neo-colonial aspects of language documentation are most visible since the archive is where the "extracted" language data is stored, mostly for use by researchers, use which can be construed as exploitation. While it is possible that archived materials may someday be useful to future generations, current community members have little access or control over what is archived. As mentioned earlier, community members are most likely to have given their consent for making their materials accessible, but they may have little or no formal education or ability to understand fully what is involved before giving "informed" consent. Instead they simply agree to the request of the linguists who have financial power and are visiting the community.[11] Meissner (2018: 273-274) discusses the problem of archiving and access to community members and it emerges clearly that archives do not primarily serve current community members' needs.

Regarding financial benefits of community-based linguistic projects, it is obvious that communities benefit the least. Most budgets cover costs related to personnel (e.g. collaborating researchers, (post)doctoral candidates) salaries or allowances, equipment, travel, lodging, feeding, and payment of consultants. It is often the case that the greater part of the budget will cover the researcher's expenses and other project costs while consultant payment will be in the neighborhood of 10 percent of the budget. Figure 2 roughly shows how a language documentation budget of 100,000 USD is most likely distributed in some community-based projects.

Arguments for, as well as reasons for not letting consultants have more money are many and varied. While offering practical advice on the payment of consultants, Bowern (2008: 162-163) notes that "it might not be appropriate in all cultures to pay people in money; that may be considered insulting... it's also useless if there's nowhere to spend the money... your consultants may wish to be paid in alcohol, or in cigarettes. Paying with cigarettes can be more valuable to the recipient than paying in cash, because cigarettes can be traded or used to 'buy' favours."

---

[11]Part of African hospitality involves being kind to guests to the extent that someone may give consent just because it is requested by the outsider.
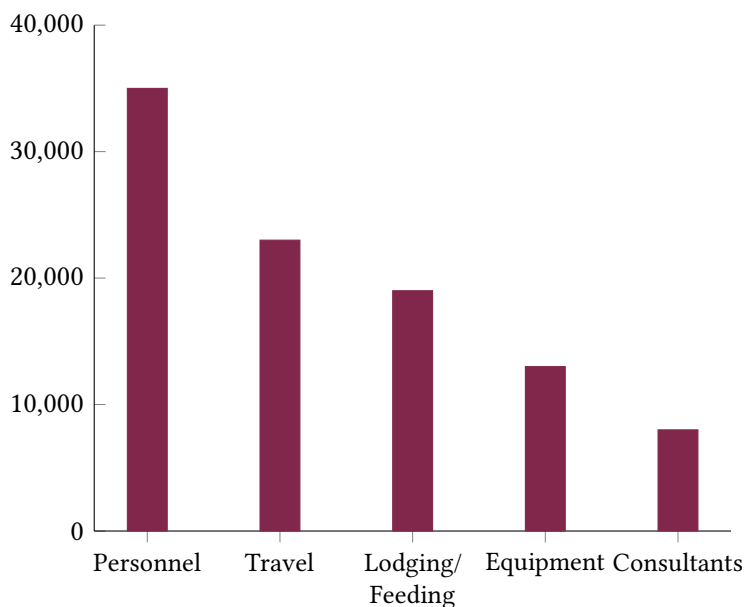
Figure 2: Allocation of 100,000 USD language documentation budget

For these and several well justified reasons, community members who serve as consultants may end up with a tiny part of the budget (1 or 2 USD an hour) which in some cases even creates conflict with those community members who do not benefit directly. As Ngué Um (2017: 13) points out, "bargaining with consultants implies negotiating with only a few of them, usually less than a dozen in the course of a project. This may be a source of felt discrimination and frustration by non-involved community members." What is obvious is that the payments are for individuals and not for the entire community. Therefore, the community as a whole does not directly benefit from linguistic projects. Some researchers do not spend substantially in the community as they avoid buying local food items. For obvious reasons related to health, some researchers go to communities with potable water and food items such as energy bars and may not buy or consume local products. In such communities, linguistic projects leave very little economic and financial impact and, clearly, do not contribute much to community development. Therefore, even if the project is community-based in some sense, it only enables the exploitation or mining of the community's linguistic resources and does not leave the community with any considerable benefits. This explains why communities in Africa where linguistic projects and linguistic fieldwork in general have been carried out for several decades have not been transformed as a

result of the linguistic work. At the same time, the researchers working on the language(s) have become highly successful and built solid careers based on the data from the communities. Apart from the Pig for Pikin intiative of the KPAAM-CAM project in Lower Fungom and the water supply initiative of the Beezen Language Documentation Project which I mentioned earlier, I am not aware of community development projects that have been initiated by outside linguists working in the Cameroonian Grassfields.[12] At the same time, and since the establishment of the Grassfields Bantu Working Group in the early 1970s (Elias et al. 1984), Grassfields Bantu languages have contributed significant amounts of data that have led to advances in the field of linguistics and produced world class linguists. This is to say that the community benefits the least (keeping aside the argument that in the long run when the language would have died future generations will have products of language documentation to turn to). Even in contexts where the researcher "gives back" some output of the project, e.g. dictionary, storybook, etc. to the community, these do not serve any immediate purpose because, as pointed out earlier, it is probably the case that the majority of the community members are illiterate in the language and cannot read.

It is necessary to reflect on how linguistic work can be done in the African context in a way that will be more beneficial to the community of speakers whose languages are documented and exploited. In the next section I propose that linguistic projects should be conceived and implemented as part of community development projects.

4 Community approach to language documentation in Africa

The idea of doing linguistic work for the benefit of the target language communities is not new. Various scholars have sought to suggest ways in which community needs and interests should be taken into consideration by linguists (e.g. Dobrin 2008, Henderson et al. 2014, Ngué Um 2017, 2019). One of the suggestions has been to move from community-based to community-centered research (McDonald 2003, Evans 2004, Holmes 2018). In this approach, the research is

---

[12]The dynamics and level of involvement of SIL linguists in communities where they work are quite complex. In particular, many SIL linguists spend years or decades living and working in Africa and, therefore, generally integrating and contributing to community development in various ways that I do not consider in this study. It also appears that as missionaries, most SIL linguists are funded by their home churches and related organizations rather than by language documentation funding agencies.

"both located at the community, and one that centralizes community concerns and participation" (Evans 2004: 60). In addition, community members have the possibility to "interpret and take ownership of research in which they can see themselves reflected and named" (Holmes 2018: 157). Holmes adds that when she followed an "indigenous community-centered process, they [community members] were in control of this research process, and however they felt comfortable participating was up to them, during the process itself and after it was over, at any time." (Holmes 2018: 157). While the community-centered approach will ensure that linguistic goals are met in ethically acceptable ways, it falls short of meeting the survival needs of communities in the African context (see also Ngué Um 2017). I believe we need to do more humanistic linguistic work that takes into account and strives to meet the needs of each community that is being studied. In other words, linguists should seek sustainable ways of doing linguistic work while also contributing to community development. Linguists could, at least, desire and find ways to engage in interdisciplinary (Hill & Ameka 2022) or cross-disciplinary work (Mufwene 2022). Work of this nature involves collaboration across disciplines (see, for example, the activities of Wuqu' Kawoq|Maya Health Alliance, a healthcare NGO in Guatemala (Henderson et al. 2014), and experiences on collaboration between linguists and communities in North-West Cameroon (Good 2012)). If linguistic projects do not cater for community needs, languages will be documented and described but they will eventually die out. It is preferable to want languages to be alive and to evolve based on their ecology (Mufwene 2001, 2005, 2008). Once a community is vital, its language will also be. As Ngué Um (2017: 10) puts it, "in Cameroon, there appears to be a correlation between language vitality and the community's wellbeing ... the less economically empowered a community, the less the members are inclined to asserting and performing the group's identity through language use, and the more exposed and endangered their cultural heritage. African communities whose languages are most endangered also almost happen to be the most economically and politically marginalized: e.g. the Bakola, the Bati, the Bezen, etc."

I suggest, therefore, that the researcher should accompany the community in its quest for survival by seeking funds for community development beyond what is used to pay a few consultants. Donors and funding agencies should not continue to think that they are only responsible for linguistic work since they consider linguists to not be in charge of community development. They should finance community development projects in order to slow down or reverse language endangerment and death while supporting linguistic projects. To consider this point seriously is to confront and desire to decolonize linguistic work in general and language documentation in particular. It requires going against "the

colonial, linguist-focused model that has been so widely critiqued" (Tsikewa 2021: 306). Therefore, there is a need to come to terms with the fact that challenges with language preservation in Africa are, in part, attributable to the impact of colonialism on the continent. The reality is that where African languages find themselves today, and therefore need reclamation, is not the sole responsibility of community members. The colonial history that is partly responsible for the present state of affairs is well known and demands honesty about it. Decolonizing linguistics can be challenging considering where the expert knowledge and funding come from. Nevertheless, the desire to maintain cultural diversity and ensure the respect of linguistic rights globally makes it possible at this point in time to seek ways to do linguistic work in Africa for the good of target community members.

As I mentioned earlier, meaningful linguistic work in Africa should be embedded in community development work. Linguistic work should be conducted alongside activities that seek to retain community members in their original settings. Without potable water, roads, electricity, schools, healthcare facilities, internet connection, jobs, etc., rural exodus is inevitable, and once someone leaves a community they are unlikely to return and live there anymore, especially if they succeed in a city.[13] One way to counter rural exodus is to contribute to the provision of the facilities people go looking for. A linguistic project with community development in mind will identify those community needs that lead to exodus and seek ways to provide them to ensure the maintenance and survival of languages and cultures. Situations where the availability of schools helped in the retention of young community members and promoted the continued use of languages are found in parts of Africa. In North-West Cameroon, for example, communities whose languages have continued to be more vibrant are those where secondary schools were established between the 1950s and 1970s, i.e during the independence era. Languages such as Lamnso' (ISO 639-3 [lns]), Kom (ISO 639-3 [bkm]), Bafut (ISO 639-3 [bfd]), Limbum (ISO 639-3 [lmp]), and Aghem (ISO 639-3 [agq]) are known to be spoken much more than other languages in the region. The presence of schools in these communities accelerated the provision of other amenities such as water, electricity, roads, and health facilities enabling the retention of several children in the communities until they became about 20 years old

---

[13]Jeff Good (pc) has pointed out the fascinating and impressive ability of Cameroonian Grassfielders to maintain their connection to their villages while in the urban diaspora. This is done, for instance, by sending children to live with relatives in the villages for some time and through country meetings (regular monthly meetings of members of the same village living in a specific urban area). There are also several people who choose to retire to their village (or near their village). This relationship between community members in the diaspora and their language and culture certainly helps and should be strengthened.

before leaving to pursue tertiary (university) education. As young people below the age of 20 stayed in their communities they continued to experience their culture for an extended period of time and to use their languages more extensively. They studied in English while living in their community, speaking their language and practicing community customs and traditions such as hunting, participating in traditional dances, doing farming, etc. The children continued to receive informal education from their parents, family and the community at large, thereby getting rooted in their culture but also getting exposure to the world through formal education. In urban centers across Cameroon it is very common to hear people from the communities listed above speaking their languages freely in public spaces. Even in university milieus students from some of these communities are often heard discussing other subjects such as chemistry or literature in their mother tongues. Communities in North-West Cameroon where schools were established early enough have greater rates of language and culture transmission than those that lacked schools and had to send their teenage children elsewhere for school.

Communities that lacked schools until recently sent their teenagers aged approximately 12–15 to other places where they could attend secondary school. Of course, the children left their languages and cultural practices behind or, at best, practiced them only minimally whenever it became possible. Most of those who were successful in school only returned occasionally for a few days during holidays and eventually settled in urban centers where they work and have built their own families away from their original communities. The consequences are obvious: the language is not used frequently and, therefore, not transmitted to younger generations, making it endangered and requiring urgent documentation before extinction. What I say here is true of a majority of communities in North-West Cameroon including Babanki where I come from. I left at the age of 18 and almost four decades later I have not had many opportunities to live in the community for more than one month at a given moment since I had to continue studying and working hundreds of kilometers away from the community. The consequence is lack of transmission of my Babanki language to my children who were all born and raised in urban areas far from other Babanki speakers. While the non-transmission of my mother tongue to my children may be a surprise to people who know me and my engagement in the promotion of mother tongue based multilingual education it represents the reality of many indigenous community parents who are sometimes blamed for not speaking their languages. As Phyak (2022) puts it, "blaming indigenous communities and parents for not speaking their languages is unfounded because they would like to, but structures, systems and ideologies do not allow spaces for indigenous languages."

Concretely, therefore, I propose that linguistic grant applications should include at least 20 percent of the budget for community development projects. The linguist should truly allow community members the liberty to select an urgent community development need that should be funded. From the results of the survey I presented in §2 above, it is very unlikely that an African community will desire language or cultural preservation which Ngué Um (2017: 5) describes as "more or less, often peripheral". They are more likely to go for crucial needs such as those he thinks allow for "coping with daily survival (very strong!), ensuring a better future for the kids (very strong!), socio-economic empowerment and security of the group (strong!)" (Ngué Um 2017: 5). A linguist with an interest in community development will find ways to justify the request of 20 percent of the budget for this non-linguistic work, which they might describe as "community compensation" (Anna Belew, pc). Funding agencies should become sensitive to community needs and willfully approve funds for community compensation rather than simply allocating money for documenting and archiving languages while ignoring the owners and producers of the knowledge. I do not understand why a funding agency can give, say 100,000 USD to a linguist to document a language but cannot give 20,000 USD to assist the speakers of that language. If some funders make provision for overheads of up to 40 percent to host institutions I believe they can do more for host communities, e.g. by setting aside funds for community overheads. This has nothing to do with the one or two dollars given to the few select consultants who work directly with the researcher. If a linguist is interested in a language, as well as in the speakers of that language, 20 percent should be a good minimum request for the interest of that community. Once the funding is obtained, the responsibility of executing the chosen project should be given to the village or community development association so that the linguist can play only a supervisory role (allowing more time for research). Development associations are found in many African communities and carry out development work of all sorts, e.g. construction of schools, health centers, roads, etc. If this is done, at the end of the project, the linguist will fulfil their agenda of collecting data for scientific inquiry, safeguarding the language heritage, assisting the community in language development efforts, obtaining academic benefits and building capacity in scholarship, while the target language community will have either potable water supply, a school, health facility, road, electricity, or whatever they needed most. The success of the project will also be used by the funding agency to convince its donors to continue supporting linguistic work. It appears to be a win-win situation for all involved.

Another thing that needs to be considered is the payment of consultants. In general, consultants are paid according to local rates and standards measured

in diverse ways. Bowern (2008: 162) suggests to "pay consultants in scale with the local economy, and tie the rate to the closest equivalent job (e.g., a teacher)." While there are many complications involved (see, for example, McLaughlin & Seidou Sall 2001), I believe consultants can be paid reasonably well, if not at international rates but high enough to compensate for their knowledge, which is indispensable for linguistic analyses. I do not think it is terribly bad to pay a consultant more than a teacher, especially because the consultant's job is for a limited period of time. Paying a consultant about 20 USD a day in a rural African setting is very likely to empower them in unimaginable ways. I have heard of a consultant in a locality in North-West Cameroon who was able to replace the grass roof of his house with zinc sheets and prevent water from dripping through each time it rained. Another was able to pay for his education after working as a consultant over an extended period.[14] If linguists consider the amount they themselves earn per hour while working with those consultants to whom they pay one or two dollars an hour, it will become obvious that more needs to be done. Consultants can be made to feel that by speaking their language(s) they are doing important work and can earn reasonably well.

5 Conclusion

In this paper, I have argued that instead of focusing on the documentation of African languages while neglecting current survival needs of community members, and thereby legitimizing and accompanying language death, linguists and funding agencies should conceive linguistic projects as community development projects. If linguistic projects assist in community development and maintenance, languages and cultures are more likely to be preserved, making the discipline meaningful and useful to African communities. Identifying community projects for funding should be the community's responsibility and, as demonstrated, community members are most likely to choose basic survival needs such as potable water, electricity, roads, schools, and health facilities instead of language development. If any of these survival needs are provided while doing linguistic work, the linguist can consider that they have given back something useful to the current generation. Such an accomplishment will most likely help retain community members and ensure continued use and transmission of the target language(s).

---

[14]As I mentioned earlier, community members who are not involved in a project and, therefore, are not paid, may be disgruntled. It is possible that such adverse effects will be minimized in situations where social work that can benefit the entire community is also implemented.

While the linguist will continue to advance their academic career, target community members will also gain substantially from having better living conditions. I have proposed that funding agencies should become sensitive to community needs and approve 20 percent of budgets for community development rather than simply allocating money for documenting and archiving languages while ignoring the owners and producers of the knowledge. I also suggest that consultants should be paid reasonably well, if not at international rates but high enough to compensate for their knowledge which is indispensable for linguistic work.

Abbreviations

CERDOTOLA	Centre International de Recherche et de Documentation sur les Traditions et les Langues Africaines
ELAR	Endangered Languages Archive
ELDP	Endangered Languages Documentation Programme
KPAAM-CAM	Key Pluridisciplinary Advances on African Multilingualism – CAMeroon
NSF	National Science Foundation
SIL	Summer Institute of Linguistics

Acknowledgements

References

Ahlers, Joycelyn. 2009. The many meanings of collaboration: Fieldwork with the Elem Porno. *Language & Communication* 29(3). 230–243. DOI: 10.1016/j.langcom.2009.02.005.

Akumbu, Pius W. 2020. Reflections on a community-based approach to writing grammars of endangered African languages. *Studies in African Languages and Cultures* 54. 71–96.

Ameka, Felix K. 2006. Real descriptions: Reflections on native speaker and non-native speaker descriptions of a language. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching grammar: The standing challenge of grammar writing*, 69–112. Berlin, New York: Mouton de Gruyter.

Bischoff, Shannon T. & Carmen Jany (eds.). 2018. *Insights from practices in community-based research: From theory to practice around the globe.* Berlin: Mouton de Gruyter.

Bowern, Claire. 2008. *Linguistic fieldwork: A practical guide.* 1st edn. New York: Palgrave MacMillan.

Brickell, Timothy C. 2018. Linguistic fieldwork: Perception, preparation, and practice. *Language Documentation and Description* 15. 179–207. http://www.elpublishing.org/PID/165.

Bwesigye, bwa Mwesigire. 2014. *Why are schools punishing children for speaking African languages?* https://thisisafrica.me/african-identities/schools-punishing-children-speaking-african-languages/ (28 March, 2022).

Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M. B.H. Rampton & Kay Richardson (eds.). 1992. *Researching language: Issues of power and method.* London: Taylor & Francis.

Crowley, Terry. 2007. *Field linguistics: A beginner's guide.* New York: Oxford University Press.

Crystal, David. 2000. *Language death.* Cambridge: Cambridge University Press.

Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian indigenous communities. *Language Documentation & Conservation* 3. 15–50. http://hdl.handle.net/10125/4423.

Di Carlo, Pierpaolo & Jeff Good (eds.). 2020. *African multilingualisms: Rural linguistic and cultural diversity.* Lanham: Lexington Books.

Dobrin, Lise M. 2008. From linguistic elicitation to eliciting the linguist: Lessons in community empowerment from Melanesia. *Language* 84(2). 300–324. http://www.jstor.org/stable/40071035.

Dobrin, Lise M. & Jeff Good. 2009. Practical language development: Whose mission? *Language* 85(3). 619–629.

Eko, Lyombe. 2003. Freedom of the press in Africa. *Encyclopedia of International Media and Communications.* 95–116. DOI: 10.1016/B0-12-387670-2/00101-1.

Elias, Philip, Jacqueline Leroy & Jan Voorhoeve. 1984. Mbam-Nkam or Eastern Grassfields. *Afrika & Übersee* 48. 31–107.

Evans, Mike. 2004. Ethics, anonymity, and authorship in community centred research or anonymity and the island cache. *Pimatisiwin: A Journal of Aboriginal and Indigenous Community Health* 2(1). 59–75.

Good, Jeff. 2012. Community collaboration in Africa: Experiences from Northwest Cameroon. In Stuart Gill & Peter K. Austin (eds.), *Language documentation and description*, vol. 11, 28–58. London: SOAS.

Grenoble, Lenore. 2009. Linguistic cages and the limits of linguists. In Jon Reyhner & Louise Lockard (eds.), *Indigenous language revitalization: Encouragement, guidance, and lessons learned*, 53–61. Flagstaff: Northern Arizona University.

Grenoble, Lenore A. & Lindsay J. Whaley. 2005. Review of language endangerment and language maintenance, and language death and language maintenance. *Language* 81. 965–974.

Handman, Courtney. 2009. Language ideology and Christianization. *Language* 85(3). 635–639.

Harrington, John P. 1941. A field comparison of Northwestern with Southwestern Indians. *Explorations and field work of the Smithsonian Institution in 1940*. 91–94.

Henderson, Brent, Peter Rohloff & Robert Henderson. 2014. More than words: Towards a development-based approach to language revitalization. *Language Documentation & Conservation* 8. 75–91. http://hdl.handle.net/10125/4611.

Hill, Deborah & Felix K. Ameka. 2022. *Languages, linguistics and development practices*. London: Palgrave Macmillan.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36. 161–195. DOI: 10.1515/ling.1998.36.1.161.

Holmes, Amanda L. 2018. *Geographies of home, memory, and heart: Mohawk Elder praxis, land, language, and knowledge woven in place*. The University of Arizona. (Doctoral dissertation).

Hyman, Larry M. 2003a. African languages and phonological theory. *GLOT International* 7(6). 153–163.

Hyman, Larry M. 2003b. *Why describe African languages? Keynote address at the 4th World Congress of African Linguistics and the 34th Annual Conference on African Linguistics*. New Brunswick: Rutgers University.

Kadanya, James L. 2006. Writing grammars for the community. *Studies in Language* 30(2). 253–257.

Keane, Webb. 2002. Sincerity, "modernity", and the protestants. *Cultural Anthropology* 17. 65–92.

Krauss, Michael E. 1992. The world's languages in crisis. *Language* 68(1). 4–10.

Leonard, Wesley Y. & Erin Haynes. 2010. Making "collaboration" collaborative: An examination of perspectives that frame linguistic field research. *Language Documentation & Conservation* 4. 268–293. http://hdl.handle.net/10125/4482.

Lionnet, Florian & Larry M. Hyman. 2018. Current issues in African phonology. In Tom Güldemann (ed.), *The languages and linguistics of Africa*, 602–708. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110421668-006.

Lüpke, Friederike. 2019. Language endangerment and language documentation in Africa. In Ekkehard Wolff (ed.), *The Cambridge handbook of African linguistics*, 468–490. Cambridge: Cambridge University Press. DOI: 10.1017/9781108283991.015.

Macaulay, Monica. 2004. Training linguistics students for the realities of fieldwork. *Anthropological Linguistics* 46. 194–209. https://www.jstor.org/stable/30029028.

McDonald, James A. 2003. *People of the Robin: The Tsimshian of Kitsumkalum: A resource book for the Kitsumkalum Education Committee and the Coast Mountain School District 82 (Terrace).* Canadian Circumpolar Institute, Alberta ACADRE Network, & School District No. 82 (Coast Mountains, B.C.). First Nations Education Centre. Edmonton: CCI Press.

McLaughlin, Fiona & Thierono Seidou Sall. 2001. The give and take of fieldwork: Noun classes and other concerns in Fatick, Senegal. In Paul Newman & Martha Ratliff (eds.), *Linguistic fieldwork*, 189–210. Cambridge: Cambridge University Press.

Meissner, Shelbi Nahwilet. 2018. The moral fabric of linguicide: Un-weaving trauma narratives and dependency relationships in indigenous language reclamation. *Journal of Global Ethics* 14(2). 266–276. DOI: 10.1080/17449626.2018.1516691.

Mufwene, Salikoko S. 2001. *The ecology of language evolution.* Cambridge: Cambridge University Press.

Mufwene, Salikoko S. 2004. Language birth and death. *Annual Review of Anthropology* 33. 201–222.

Mufwene, Salikoko S. 2005. *Créoles, écologie sociale, évolution linguistique.* Paris: L'Harmattan.

Mufwene, Salikoko S. 2008. *Language evolution: Contact, competition and change.* London: Continuum Press.

Mufwene, Salikoko S. 2016. A cost-and-benefit approach to language loss. In Luna Filipović & Martin Pütz (eds.), *Endangered languages and languages in danger: Issues of documentation, policy, and language rights*, 115–143. Armsterdam: John Benjamins.

Mufwene, Salikoko S. 2017. Language vitality: The weak theoretical underpinnings of what can be an exciting research area. *Language* 93(4). 202–223. DOI: 10.1353/lan.2017.0065.

Mufwene, Salikoko S. 2022. The linguist, language, and economic development: A commentary on languages, linguistics and development practices. In Deborah Hill & Felix K. Ameka (eds.), *Languages, linguistics and development practices*, 267–289. London: Palgrave Macmillan.

Newman, Paul. 1998. We has seen the enemy and it is us: The endangered languages issue as a hopeless cause. *Studies in the Linguistics Sciences* 28(2). 11–20.

Newman, Paul. 2009. Fieldwork and field methods in linguistics. *Language Documentation & Conservation* 3(1). 113–125. http://hdl.handle.net/10125/4428.

Ngué Um, Emmanuel. 2019. Achieving sustainable language preservation through economic empowerment in endangered language settings in West Africa. In Pius W. Akumbu & Esther P. Chie (eds.), *Engagement with Africa: Linguistic essays in honor of Ngessimo M. Mutaka*, 375–395. Cologne: Köppe.

Ngué Um, Emmanuel. 2017. Beyond expanding the functions of threatened languages; expanding the life opportunities of communities who speak them: Economic empowerment and language revitalization in critical socio-economic conditions in Africa. http://hdl.handle.net/10125/42068.

Ngugi, wa Thiong'o. 1986. *Decolonising the mind: The politics of language in African literature.* London: James Curray.

Ochocka, Joanna & Rich Janzen. 2014. Breathing life into theory: Illustrations of community-based research hallmarks, functions, and phases. *Gateways: International Journal of Community Research and Engagement* 7. 18–33.

Pennycook, Alastair & Sinfree Makoni. 2005. The modern mission: The language effects of Christianity. *Journal of Language, Identity, and Education* 4. 137–155.

Phyak, Prem. 2022. Language policy in Nepal. ELP Festival of indigenous languages.

Rehg, Kenneth L. 2004. Linguists, literacy, and the law of unintended consequences. *Oceanic Linguistics* 43. 498–518.

Rice, Keren. 2018. Collaborative research: Visions and realities. In Shannon T. Bischoff & Carmen Jany (eds.), *Insights from practices in community-based research: From theory to practice around the globe*, 13–37. Berlin: Mouton de Gruyter.

Samarin, William J. 1967. *Field linguistics: A guide to linguistic field work.* New York: Holt, Rinehart & Winston.

Sapién, Racquel-María. 2018. Design and implementation of collaborative language documentation projects. In Kenneth L. Rehg & Lyle Campbell (eds.), *The Oxford handbook of endangered languages*, 203–224. New York: Oxford University Press. DOI: 10.1093/oxfordhb/9780190610029.013.12.

Tsikewa, Adrienne. 2021. Reimagining the currrent praxis of field linguistics training: Decolonial considerations. *Language* 97(4). 293–319. DOI: 10.1353/ lan.2021.0072.

Woodbury, Anthony C. 2011. Language documentation. In Peter K. Austin & Julia Sallabank (eds.), *The Cambridge Handbook of Endangered Languages*, 158–186. Cambridge: Cambridge University Press.

Zambakari, Christopher. 2021. Mission to civilise: The French West African federation. https://www.accord.org.za/conflict-trends/mission-to-civilise-the-french-west-african-federation/ (14 December, 2022).

Chapter 2

Anglicization of personal names: The case of Akan

Kwasi Adomako

University of Education, Winneba

John Odoom

University of Education, Winneba

Akan personal names are losing their cultural or ethnopragmatic values and meanings through anglicization. Anglicization affects the form, structure, pronunciation, and original meanings of some indigenous Akan personal names. Personal names are not merely arbitrary labels but sociocultural tags that have sociocultural functions and meanings and provide linguistic information about the Akan culture. This study seeks to examine orthographic changes identified in some anglicized Akan personal names and their implication for the Akan language and culture. This is a descriptive qualitative study. The data for this study were collected from university records of written Akan personal names. The paper shows that when Akan personal names are anglicized, they undergo orthographic modifications such as vowel replacement and consonant replacement. Two prosodic processes such as resyllabification and tone alterations were also observed in the anglicized names. The paper concludes that these processes have affected the meanings of the names by eroding their cultural or ethnopragmatic values.

1 Introduction

This paper provides a linguistic analysis of anglicization in Akan personal names (APNs). According to Mensah (2022), anglicization is a cultural and linguistic assimilation process that entails the diffusion of English words (or names) through

borrowing and adapting to other languages. In this study, we look at anglicization from the perspective of the adaptation of personal names to English spelling, syllable structure, and pronunciation.

Anglicized written names impact pronunciation, producing changes in segments, syllable structure and tone. Anglicization has largely impacted the writing system of Akan, particularly names, therefore affecting the sociocultural meaning of the name (being anthroponymy, toponymy, or ethnonymy). For instance, Akan (Fante dialect) personal names whose original, non-Anglicized, spelling and pronunciations are *Esuon* [èsóń] 'seventh born' and *Gyasi(e)* [dʑèsí] 'father's keepings' have been anglicized as *Eshun* [eʃən] and *Gaisie* [ɡeizi] respectively, and have lost the original cultural value. This writing change leads to a different pronunciation and, affects the meaning of the name. This adulteration of personal names is a product of several centuries of contact between the Akan people and Europeans (Owu-Ewie 2017).

Personal names among the Akan are not merely arbitrary labels, but sociocultural tags that have sociocultural functions and meanings and provide linguistic and typological information (Essien 1986, Agyekum 2006). The cultural and social contexts identify the bearer or convey a wide range of invaluable information about the bearer. According to Kachru (1994), a name is a valuable source of information, which can indicate gender, birthplace, nationality, ethnicity, religion and position within a family and the society at large. The concept of naming within the African context is highly valued because names echo the bearers' cultural values and identities. This has aroused scholars' interest to research Akan anthroponomy over the years from different perspectives (Boadi 1984; Obeng 1997, 1998, 2001; Ansu-Kyeremeh 2000; Agyekum 2006, 2010; Adomako 2015, 2017, 2019; Ofori 2019; among others). For instance, Agyekum (2006) discusses the typology of Akan personal names from a sociolinguistic perspective. Obeng (1998) describes the morpho-syntactic processes of death prevention names in Akan, while Adomako (2015, 2019) examines the truncation of some Akan personal names and some morphological and phonological processes embedded in Akan day-name formation respectively. Mireku-Gyimah & Mensah (2015) analyze the anglicization of some town names in Tarkwa, a prominent mining community in Ghana.

This paper discusses orthographic changes identified in some anglicized Akan personal names that have led to changes in the pronunciations of such names. The study shows that when personal names in Akan are anglicized, it affects the orthography and pronunciation of the names, and, by implication, it affects the meanings and sociocultural values of such names.

2 The Akan people and language

Akan refers to an ethnic and linguistic group of people in Ghana. Geographically, this group of people predominantly occupies the southern and middle parts of Ghana and Côte d'Ivoire (Dolphyne 2006). Nine out of the sixteen regions in Ghana are predominantly occupied by the Akan peoples. These regions are Asante, Bono, Bono East, Central, Eastern, Oti, Western, Western North, and the Ahafo regions. However, due to trading and resultant migration, Akan speakers can be found across the length and breadth of the country.

Akan is a Central Tano language. About 80% of Ghana's population can speak Akan, and about 45.7% of Ghanaians are native speakers (Ghana Statistical Service 2021). This, therefore, arguably makes Akan the most widely spoken language in Ghana. The various subdivisions of Akan share a similar culture and there is mutual intelligibility between the dialects. According to Agyekum (2006), the language consists of the related dialects Asante, Akuapem, Fante, Agona, Assin, Ahanta, Akyem, Bono, and Wassa, among others. Out of these dialects, it is only the three major dialects, Akuapem, Asante, and Fante, that have achieved literary status.

3 The Akan naming system

Every Akan child is named after the day he/she was born, except in a few cases (Agyekum 2006: 213). For example, when a male is born on Sunday, he can be called Kwesi, Kwasi, Akweesi, or Akwasi, and if the child is female, she can be called Esi, or Akosua depending on the ethnic subgroup she belongs to (see Obeng 2001; Agyekum 2006; Odoom 2013; Adomako 2015, 2019, among others). In addition, a child receives a given name from the father, and at times a religious name (Agyekum 2006; Odoom 2013). The given name is given on the seventh day of his or her birth. The religious name, which is often a first name, depends on the religion of the child's parents. If the parents of the child are Christians, they will give him or her a Biblical name and if they are Muslims, they will give him or her an Islamic name. Aside from the given name, there are other categories of Akan names. Some of these categories include circumstantial names, proverbial names, birth order/positional or ordinal position names, deity names, death prevention names, etc. (see Ansu-Kyeremeh 2000; Agyekum 2006; Odoom 2013). These kinds of names express the sociocultural values of the Akan people through their structure and meaning. However, through anglicization, such sociocultural values have been eroded. Orthographic modifications have led to

tonal alteration and different pronunciations of such names. In this study, we discuss some of these orthographic changes embedded in the anglicized Akan personal names, specifically both day-names and family-names. Again, our use of the term "anglicized" simply refers to the variants usually produced by the Akan (especially Fante) educated elites. Therefore, the examples we will provide under "anglicized names" do not actually reflect a native English speaker's pronunciations, instead, those of the educated elites. One of the characteristics of the variants by the educated elites is that sometimes an expected diphthong in the native English speaker's pronunciation is simplified to a monophthong. Before we begin the discussions, let us consider the sound system of Akan.

4 Akan segmental inventory

Akan has nine contrastive oral vowels /i, ɪ, e, ɛ, u, ʊ, o, ɔ, a/, and one allophonic oral vowel [æ], as well as five contrastive nasal vowels /ĩ, ɪ̃, ũ, ʊ̃, ã/ (Schachter & Fromkin 1968; Dolphyne 2006; Odoom & Adomako 2021, among others) and one phonetic nasal vowel [æ̃] in Fante only (Abakah 2013). The allophonic [æ] is in complementary distribution with [a] (Dolphyne 2006; Odoom & Adomako 2021). The vowel chart below illustrates the ten oral vowels and their nasal counterparts. It can be seen from Table 1 that there are no nasal mid vowels in Akan.

Table 1: Akan vowels

	Oral			Nasal		
	Front	Central	Back	Front	Central	Back
+ATR	i		u	ĩ		ũ
−ATR	ɪ		ʊ	ɪ̃		ʊ̃
+ATR	e		o			
−ATR	ɛ		ɔ			
+ATR	[æ]			[æ̃]		
−ATR		a			ã	

At the consonantal level, Akan has fourteen consonant phonemes (Table 2) , namely /p, b, t, d, k, g, m, n, r, f, s, h, j, w/ (Abakah 2005, 2012). According to Adomako (2018), there are additional consonants such as [t͡ɕɥ, d͡ʑɥ, ɕɥ, ɲɥ, d͡z, t͡s, t͡ɕ, d͡ʑ, ɕ, ɲ] that might have gone or are undergoing phonemicization in Akan.

According to de Jong & Obeng (2000), there is a contrast between plain, labialized, palatalized and labio-palatalized sounds, even though there are restrictions before certain vowels. Moreover, there are phonetic consonants such as [ŋ], [ɱ] and [ɥ] which are allophones of the phonemes /n/, /m/ and /w/, respectively. /l/ is only used in Akan in borrowed words, except in the Gomoa variant of Fante where it is phonemic. The consonants [d], [r] and [l] are used as free variants (Abakah 2005; Odoom 2013). Table 2 summarises all the phonemic and phonetic consonants of Akan.

Table 2: Akan Consonants

	Bilabial	Labio-dental	Alveolar	Alveo-palatal	Palatal	Labial-palatal	Velar	Labial-velar	Glottal
Stop	p b		t d				k g		
Affricates			t͡s d͡z	tɕ dʑ		t͡ɕɥ d͡ʑɥ			
Fricatives		f	s	ɕ		ɕɥ			h
Nasals	m	ɱ	n		ɲ	ɲɥ	ŋ		
Approximants					j	ɥ		w	
Trill			r						
Lateral			l						

5 Akan Orthographic System

The current orthographic system of Akan uses seven vowels, ⟨i e ɛ a o ɔ u⟩. Orthographic symbols will be shown between ⟨⟩. Vowel ⟨e⟩ is realized as either [e] or [ɪ], ⟨o⟩ is realized as [o] or [ʊ] and ⟨a⟩ is realized as [a] and [æ] and at times [e] in Fante. Akan orthography was developed around the 17th and 18th centuries mostly by German and British missionaries. It was not until the 1850s that a maiden standard alphabet (orthography) was developed for Akan (Christaller 1933: viii-ix). There was an important change made to the orthography in the elimination of diacritics on ⟨e⟩, ⟨o⟩, ⟨u⟩, and ⟨a⟩ for [ɪ], [ʊ], [u], and [a], respectively from the Gold Coast orthographic system to the Unified Akan orthography. The last attempt to have a unified Akan orthography was in the late 1970s but it has not been successful (Dolphyne 2006).

It should be noted that the [æ] which mostly occurs before [+ATR] vowels in Asante and Akuapem dialects (Twi) is raised to [e] in the Fante dialect (Abakah 2012, Adomako 2015). This phenomenon occurs in the nasal vowels too. The nasal vowels [ĩ] and [ʊ̃] are written as ⟨e⟩ and ⟨o⟩, respectively. The nasality of vowels is not encoded in the writing system. Akan indicates vowel length by doubling the letter, as in ⟨kɔ⟩ 'go' and ⟨kɔɔ⟩ 'reddish'. Akan vowels occur in word-initial, medial and final positions. However, not all vowels occur word-initially in all the dialects. The vowels [i], [ɪ], [u] and [ʊ] do not occur word-initially in Twi, while [u] and [ʊ] fail to occur word-initially in Fante.

At the consonantal level, Akan uses twenty-six consonants consisting of 15 single glyphs, ⟨b, d, f, g, h, k, (l), m, n, p, r, s, t, w, y⟩ and 11 digraphs, ⟨ky, gy, kw, dw, dz, tw, ts, hy, hw, ny, nw⟩ (Abakah 2005, Dolphyne 2006). The diagraphs ⟨ts⟩ and ⟨dz⟩ are used by the Fante dialect of Akan. Moreover, ⟨l⟩ is used in Akan for borrowed words, such as ⟨lɛtɛ⟩ [lɛtɛ] 'letter', ⟨bɔɔl(o)⟩ [bɔɔl(ʊ)] 'ball'.

All these consonants occur word-initially and medially. The consonant ⟨r⟩ occurs word-initially only in progressive verb stems. The only consonants that occur word-finally in Akan are [m, n, w, r, ŋ, l]. Among these consonants, only [m, n, w, r, l] occur word-finally in Fante, [m, w, ŋ] in Akuapem and [m] in the Asante dialect. When ⟨n⟩ occurs word-finally, it is realized as [ŋ] in the Akuapem and Agona dialects of Akan, nasalized high vowels [ĩ, ĩ, ũ, ʊ̃] in Asante and [n] in Fante (Dolphyne 2006; Abakah 2005, 2012). Akan does not have consonant clusters. Thus, any [+nasal] consonant followed by a [−nasal] consonant constitutes a syllable (Dolphyne 2006; Marfo & Yankson 2008; Marfo 2013).

The consonant sounds [tɕ], [dʑ], [d͡ʑɥ], [t͡ɕɥ], [ɕ], [ɕ͡ɥ], [ɲ], and [ɲ͡ɥ] are written as ⟨ky⟩, ⟨gy⟩, ⟨dw⟩, ⟨tw⟩, ⟨hy⟩, ⟨hw⟩, ⟨ny⟩, ⟨nw⟩, respectively in the orthography. The Akan [tɕ] is matched to [tʃ] in English. [dʑ] and [ɕ] are also matched to [dʒ] and [ʃ] in English, respectively, and this will affect the way names are anglicized.

6 Methodology

This study adopts a qualitative approach to data collection and analysis. Data for the study were collected from primary and secondary sources. For the primary source, we sampled 60 respondents comprising 40 males and 20 females with their ages ranging from 30–60 years. The 60 respondents were made up of three groups. The first two groups were purposively selected based on the criteria of using Akan as their first language and also bearing Akan personal names that have undergone anglicization. The third group comprised foreign Americans in Ghana. We provide explanations for methods of selecting the respondents in the following paragraph.

Out of the 60 respondents, 20 had no formal education. They were all from typical farming and fishing communities in the Ashanti and Central Regions of Ghana. 8 out of the number were subsistent farmers, 7 were fishermen and 5 were petty traders. The purpose of selecting this group of respondents was based on two objectives; first to find if their pronunciations of the target Akan family names would be influenced by the anglicization as is often done by the educated elite or not. Second, we sought to find out their knowledge of the meaning of those target names. We went with a list of the sampled anglicized names and asked about their knowledge of the names based on the two objectives. We mentioned the names in their frequently used anglicized tone after which we sought their opinions. They then provided the indigenous pronunciations and the meanings of those names. Their responses were recorded on a voice recorder and field notes. To solicit views of respondents with formal education, we sampled 30 respondents from the class lists of Akan students of the Department of Akan-Nzema, University of Education, Winneba, Ghana where Akan (Twi and Fante) is taught. We sampled 30 respondents (14 females, 16 males) who themselves bear the anglicized names. The objectives of their selection were to determine whether having been exposed to formal education up to the tertiary level, they still possess knowledge of the indigenous Akan pronunciations of the anglicized names or not, and also to elicit their knowledge of the meaning of their anglicized names. This group of respondents have both their first names and family names Akan. For example, full orthographic names such as *Kojo Essel*, *Kwamina Daadzie*, etc. The third group of the respondents were ten Americans; five students on exchange programs in Ghana, and five missionaries of the Church of Jesus Christ of Latter-day Saints at Agona Swedru in the Central Region of Ghana. Even though it is safe to assume that if indigenous Akan names have been anglicized, then it should be in British English due to the long-standing contact between the Akan people and the British during colonialism, the purpose for sampling these respondents was to find out whether orthography might have influenced the anglicization of those names (even by the Akan educated elites). For this group, our interest was solely to elicit the pronunciations of the anglicized names.

For the secondary source, in addition to the class name lists, a book on Fante personal names, namely Crayner (1988) was used to add to the data. In all, 300 personal anglicized names were collected for the present study comprising class lists of Akan students of the Department of Akan-Nzema, University of Education, Winneba and some names from Crayner (1988) (see the Appendix for the list of the names).

Out of the list of 300 anglicized names, about 60 names were sampled, recorded, categorized and coded according to themes and subthemes as are presented in §7 of the present paper. These sampled names were checked for accuracy, transcribed according to respondents' pronunciations and glossed. It is these sampled names that served as the primary source of data for the present study. We employed our native speakers' intuition (Asante and Fante dialects) in our analysis of the sampled names. The analysis of the data is descriptive. The anglicization of Akan personal names in this study is categorized and discussed under three main threads: orthographic or segment modifications, resyllabification, and tonal alteration. Each of these processes is analyzed in detail in what follows.

7 Orthographic or segment modifications

This section focuses mainly on the anglicization of Akan personal names through English orthographic representations. Akan names can be written with Akan orthography but there are anglicized versions which are characterized orthographically based on the English alphabet. This technique was introduced by early missionaries due to a "misconception" of the original names (Ukpong 2007: 227).

The following spelling modification strategies demonstrate how APNs have been anglicized with English spellings. In this section, tone will not be indicated in the phonetic transcriptions. Tone will be discussed in §7.4.

7.1 Vowel replacement

Vowel replacement occurs when a phonemic vowel in Akan is replaced with a different vowel in the anglicized version. Unlike the Akan orthographic system, the English orthographic system does not have the means to represent [ɔ] and [ɛ] with separate symbols. In Akan names that have such sounds, they are replaced with ⟨o⟩ and ⟨e⟩ respectively and pronounced as given in the transcription (Abakah 2013), as shown in (1a-d) below.[1] In addition, the vowels [ɪ] and [ʊ] are written as ⟨i⟩ and ⟨u⟩ respectively as shown in (1e-g). Finally, in (1h), the sound [e] is written ⟨ai⟩ and pronounced as a diphthong [ai].[2]

---

[1]*Ɔsɛɛ* is derived from *ɔ+sa+yɛ*. The *ɔ-* is a nominal prefix or nominalizer. The *sa* means 'war' or 'battle' and the *yɛ* means 'good'. It has a historical interpretation that we will not go into detail about here. These morphological components show that the name has gone through phonological processes such as assimilation due to vowel hiatus.

[2]The raising of the /a/ to [e] in Fante is a result of a well-attested phonological process of vowel raising/replacement before high [+ATR] vowels.

(1)   **Original name**   **Anglicized name**

a.	Ɔsɛe	[ɔsɛɪ]	'battle is good'	Osei	[osei]
b.	Dɔnkɔ	[dɔŋkɔ]	'a slave'	Donkor	[doŋko]
c.	Abɛbrɛsɛ	[abɛbrɛsɛ]	'struggler'	Abebresse	[abebɹese]
d.	Akyerɛ	[atɕɪrɛ]	'delayed in womb'	Achere	[atʃiɹe]
e.	Dadze	[dad͡zɪ]	'iron/metal'	Daadzie/Dadzi	[daadzi]
f.	Apea	[apɪja]	'spear'	Appiah	[apia]
g.	Afor	[aful]	'tender'	Afful	[aful]
h.	Badu	[bedu]	'tenth born'		

These replacement processes can be summarized as in (2).

(2)
a.	⟨ɔ⟩	[ɔ]	→ ⟨o⟩	[o]
b.	⟨ɛ⟩	[ɛ]	→ ⟨e⟩	[e]
c.	⟨o⟩	[ʊ]	→ ⟨u⟩	[u]
d.	⟨e⟩	[ɪ]	→ ⟨ie⟩ or ⟨i⟩	[i]

It is worth pointing out that despite these orthographic modifications, native speakers of Akan pronounce the anglicized names the same as the original form. However, non-native speakers and elites pronounce them based on the orthographic forms they see, as shown above.

7.1.1 Replacing a single vowel ⟨u⟩ with double vowels ⟨oo⟩

This vowel modification process involves replacing a single vowel symbol with double vowels as shown in (3).

(3)   **Original name**   **Anglicized name**

a.	Otu	[otu]	'gun'	Otoo	[otu]
b.	Owu	[owu]	'death'	Owoo	[owu]
c.	Badu	[bedu]	'tenth born'	Baidoo	[beidu]
d.	Adu	[edu]	'tenth born'	Aidoo	[eidu]
e.	Ewur	[ewur]	'key/padlock'	Ewool	[əwuol]
f.	Odum	[odum]	'teak tree'	Odoom	[odum]
g.	Nkum	[ŋkum]	'do not kill me'	Incoom	[ɪŋkum]

The original vowel ⟨u⟩ in Akan is replaced with ⟨oo⟩ in the English orthography, though this orthographic change does not necessarily affect the pronunciations of the vowels. This replacement strategy is seen as a "writing style" and it mostly affects disyllabic and trisyllabic names.

7.1.2 Replacing Asante ⟨ɔ⟩ with ⟨our⟩ or ⟨uor⟩

In Akan, rounding harmony occurs only in Asante and Fante (see Dolphyne 2006; Abakah 2012). Dolphyne (2006) explains that in addition to having two sets of verbal prefixes depending on whether the vowel of the verb stem is an advanced or unadvanced tongue root, Fante verbs have an additional feature of round vowel harmony. In Asante, the nominalizing suffix /-E/ alternates for both ATR and rounding.

(4)　　　**Asante round harmony**
a.	owu-o	'death'
b.	nsu-o	'water'
c.	ɛbʊ-ɔ	'stone'
d.	ɔfʊ-ɔ	'buffalo'
e.	adɪ-ɛ	'thing'
f.	asɪ-ɛ	'the base of'
g.	esi-e	'anthill'

This phenomenon is also applicable to personal names in Asante Twi. The nominal suffix -ɔ when it follows ʊ is written as ⟨our⟩ or ⟨uor⟩ in the English version and pronounced as [uɔ:]. Compare (5) to (6).

(5)　　　**Original name**
a.	Ɔboɔ	[ɔbʊɔ]	'stone'
b.	Sereboɔ	[srɪbʊɔ]	'sharpen stone'
c.	Kufoɔ	[kufʊɔ]	'warrior/fighter'
d.	Ɔkofoɔ	[ɔkʊfʊɔ]	'warrior'
e.	Adubɔfoɔ	[ædubɔfʊɔ]	'great hunter'
f.	Fɔdwoɔ	[fɔdʑɥʊɔ]	'deity name'

(6)　　　**Anglicized name**
a.	Obour/Obuor	[obuɔ:]
b.	Srebour/Srebuor	[sribuɔ:]
c.	Kuffour/Kuffuor	[kəfuɔ:]
d.	Okoffour/Okoffuor	[okəfuɔ:]
e.	Aduboffour/Aduboffuor	[ədubɔfuɔ:]
f.	Fordjour/Fordjuor	[fɔdʒuɔ:]

The added ⟨r⟩ with the final ⟨our⟩ in English is in parallel with words like ⟨your⟩ [jɔ:]. So, this could be the source of this added ⟨r⟩ and the spelling with

⟨uo⟩ or ⟨ou⟩ . The ⟨r⟩ in the orthography would ensure a long vowel or a diphthong pronunciation if one is speaking with a non-rhotic British dialect. These processes do not affect pronunciations by uneducated native speakers of Akan but rather non-native speakers and educated elites.

7.1.3 Other replacements

There are some other cases in which a vowel in Akan is realized differently at the phonetic level of the anglicized version, as illustrated in (7).

(7) **Original name** **Anglicized name**

a.	Badu	[bedu]	'tenth born'	Baidoo	[beidu]
b.	Adu	[edu]	'tenth born'	Aidoo	[eidu]
c.	Gyasi	[dʑesi]	'father of all'	Gaisie	[geizi]
d.	Mensa	[mensã]	'third male born'	Mensah	[mənsa]
e.	Esar	[esar]	'warrior'	Essel	[esəl]
f.	Adɛɛ	[adɛɛ]	'cloud'	Adae	[adaɪ]

We can see from (7) that the phonetic vowel [e] in Akan is written ⟨ai⟩ and pronounced either [ei] (7a-c) or [ə] (7d). The vowel [a] is written ⟨e⟩ and pronounced [ə] (7e), and the vowel [ɛɛ] is written ⟨ae⟩ and pronounced [aɪ]. Vowel replacement is one of the common segmental processes of anglicizing Akan personal names.

7.2 Consonant modifications

There are two consonant modification processes discussed in this section: consonant replacement and consonant addition. Under consonant replacement, a consonant in Akan has been replaced with a different orthographic consonant in the anglicized version. With consonant addition, a different orthographic consonant is added to the already anglicized Akan personal names. This subsection discusses the consonant replacement strategies in the anglicized Akan personal names.

7.2.1 Replacement of ⟨kw⟩ with ⟨qu⟩

There are many Akan consonant sounds, which are modified or replaced with different orthographic sounds in the anglicized version. In the following data, ⟨kw⟩ is replaced with a ⟨qu⟩ in the anglicized version. This orthographic strategy is because English uses ⟨qu⟩ to write the [kw] sound sequence; compare (8) and (8).

(8)   **Original name**
    a.   Kwansa   [kʷãnsã]   'there is no way'
    b.   Kwame   [kʷaamĩ]   'Saturday male born'
    c.   Kwam   [kʷãm]   'someone born in a road'
    d.   Kwao Ata   [kʷaata]   'twin born on Thursday'
    e.   Kwaku   [kʷeku]   'Wednesday male born'
    f.   Kwasi   [kʷesi]   'Sunday male born'
    g.   Kwao Otu   [kʷeetu]   'Thursday male born with a gun'
    h.   Kwao Anu   [kʷeenũ]   'second male born on Thursday'

(9)   **Anglicized name**
    a.   Quansah   [kwãnsa]
    b.   Quarmi   [kwamĩ]
    c.   Quarm   [kwãːm]
    d.   Quarttah   [kwata]
    e.   Quaicoo   [kwakuː]
    f.   Quaisi   [kwasi]
    g.   Quaitoo   [kwatuː]
    h.   Quainoo   [kwanũː]

The spellings ⟨ar⟩ and ⟨ai⟩ that follow the ⟨qu⟩ reflect the ATR quality of the vowel. The ⟨ar⟩ represents the [-ATR] vowel [a], written ⟨a⟩ in the original name as in (8a-d), but pronounced with no [r]. Conversely, the ⟨ai⟩ represents the [+ATR] harmonized vowel [e] (Fante variant) and [æ] (Asante variant) in the original name, but is pronounced as [a] in the Anglicized form. Moreover, data (8d), (8g) and (8h) have been assimilated to form a long vowel in the original name and this bears a direct influence on the anglicized version, which is written as a single word.

7.2.2 Replacement of palatalized and labialized consonants

There are other consonants in Akan that undergo replacement processes in the anglicized version. The palatalized and labialized consonants in the Akan consonant system ⟨ky⟩ [tɕ], ⟨gy⟩ [dʑ], ⟨tw⟩ [tɕɥ], ⟨dw⟩ [dʑɥ], ⟨kw⟩ [kʷ], ⟨sw⟩ [sʷ] are written in the anglicized version as ⟨ch⟩ [tʃ], ⟨dj⟩ or ⟨dg⟩ [dʒ], ⟨tch⟩ [tʃw], ⟨dj⟩ [dʒ], ⟨cq⟩ or ⟨k⟩ [k], ⟨sh⟩ [ʃ] respectively. This modification process deviates from the orthographic and phonotactic systems of Akan as shown in (10).

(10)

		Original name			Anglicized name	
	a.	Ɔkyere	[ɔtɕɪrɪ]	'delayed male born'	Ochere	[otʃʟɪɪ]
	b.	Takyi	[tetɕi]	'a male twin'	Techie	[tətʃi]
	c.	Otwe	[otɕɥe]	'eighth born'	Otchie/Otchey	[otʃwie]
	d.	Botwe	[botɕɥe]	'variant of eighth born'	Botchey	[botʃwie]
	e.	Agyapɔn	[edʑapɔ̃ɤ̃]	'great father'	Agyapong	[ədʑapɔŋ]
	f.	Fɔdwoɔ	[fɔdʑɥɔ]	'an Akan deity name'	Fordjour	[fɔdʑuɔ:]
	g.	Esuon	[esʷuon]	'seventh born'	Eshun	[eʃən]

Albeit the differences in the orthographic representation, phonetically, the replaced sounds are similar. The native speakers of Akan pronounce the names correctly and people can easily know to whom they are referring (Agyekum 2006). This is because the replaced orthographic symbols have similar phonetic realizations. The exception is (10g) wherein the original sound is realized as a labialized alveolar fricative [sʷ] while the anglicized version is realized as a post-alveolar fricative [ʃ].

7.2.3 Replacing a single consonant with double consonants

Some APNs are written with double consonants in the anglicized version as demonstrated in (11) below.

(11)

		Original name			Anglicized name	
	a.	Ɔpɔn	[ɔpɔ̃n]	'great one'	Oppong	[ɔpɔ̃ŋ]
	b.	Apea	[apɪja]	'spear'	Appiah	[apia]
	c.	Aban	[abãn]	'a fortress'	Abban	[abãn]
	d.	Asan	[asãn]	'a warrior'	Assan	[asãn]
	e.	Anan	[anãn]	'4th male born after three males'	Annang	[anãŋ]
	f.	Akufo	[ekufʊ]	'a warrior'	Akuffo	[akufu]
	g.	Afor	[afʊr]	'tender'	Afful	[aful]

This data shows the intervocalic consonants ⟨p⟩, ⟨b⟩, ⟨s⟩, ⟨n⟩, and ⟨f⟩ being replaced with double consonants ⟨pp⟩, ⟨bb⟩, ⟨ss⟩, ⟨nn⟩ and ⟨ff⟩ in the anglicized version respectively. This writing convention has no bearing on the pronunciation of the anglicized variant.

7.2.4 Addition of /g/ after /n/

Akan personal names in which the final consonant is an alveolar nasal /n/ have an additional ⟨g⟩ in the English orthography. In Akan, the alveolar nasal /n/ is realized as [n] in Fante, [ŋ] in Akuapem and Agona Mfantse and nasalized vowels [ĩ, ĩ, ũ, ʊ̃] in Asante in word-final position. Thus, this may be the source of the perceived [ŋ] in the English variant. This sound would be written ⟨ng⟩ in English orthography; compare (12) and (13).[3]

(12)

	Original name	Mfante	Akuapem	Asante	
a.	Ɔpɔn	[ɔpɔ̃n]	[ɔpɔ̃ŋ]	[ɔpɔ̃ʊ̃]	'great one'
b.	Ɔten	[ɔtsĩn]	[ɔtĩŋ]	[ɔtĩĩ]	'righteous'
c.	Ɔben	[ɔbĭn]	[ɔbĭŋ]	[ɔbĩĩ]	'scholar'
d.	Akon	[akʊ̃n]	[akʊ̃ŋ]	[akʊ̃ʊ̃]	'ninth born'
e.	Enyan	[eɲãn]	[æɲãŋ]	[æɲãĩ]	'resurrect'
f.	Agyapɔn	[ædʑapɔ̃n]	[ædʑapɔ̃ŋ]	[ædʑapɔ̃ʊ̃]	'great father'
g.	Akyampɔn	[atɕampɔ̃n]	[ætɕampɔ̃ŋ]	[ætɕampɔ̃ʊ̃]	'great shield'

(13)

	Anglicized name		
a.	Oppong	[opɔŋ]	'great one'
b.	Oteng	[oteŋ]	'righteous'
c.	Obeng	[obeŋ]	'scholar'
d.	Akong	[akuŋ]	'ninth born'
e.	Anyang	[əɲãŋ]	'resurrect'
f.	Adjepong	[ədʑapɔŋ]	'great father'
g.	Acheampong	[ətʃiampɔŋ]	'great shield'

Akan orthographic principles do not allow ⟨g⟩ at the word-final position. However, English adds a ⟨g⟩ because that is how the final [ŋ] is written in English orthography.

7.2.5 Addition of /r/ after the first vowel

There are instances where an /r/ is inserted right after the first vowel of the name as shown below. This, we assume, is influenced by the orthographic ⟨r⟩ after the initial vowels.

---

[3]The anglicized variants of the names in (13) were the forms produced by the non-native (American) respondents.

(14)		Original name			Anglicized name	
	a.	Ahen	[ahĩn]	'a king'	Arhin	[aɹhĩn]
	b.	Ebo	[ebo]	'Tuesday male born'	Erboh	[əɹbo]
	c.	Ato	[atʊ]	'Saturday male born'	Artoh	[aɹtu]
	d.	Ako	[akʊ]	'a warrior'	Arkoh	[aɹku]

The alveolar approximant sound [ɹ] is inserted to function contextually as a lengthening segment. In (14b), the close-mid front unrounded vowel [e] in the original form is swapped with the close-mid central unrounded vowel [ə] in the anglicized form. Moreover, in (14c-d), the close back -ATR rounded vowel [ʊ] in the original form is replaced with the close-back tense rounded vowel [u] in the anglicized version.

7.2.6 Addition of ⟨h⟩ at the end of the original name

Here, Akan personal names have an extra letter ⟨h⟩ at the end of the original name. The ⟨h⟩ is redundant and does not represent any sound in the name. Orthographically, English has final ⟨h⟩ only in 2 or 3-letter interjections like "ah", "eh", "oh" and "duh". Thus, the use of ⟨h⟩ to end the names looks like a strategy to represent a final [a] (Williamson 1984) as shown in (15).

(15)		Original name			Anglicized name	
	a.	Baa	[bãã]	'defense/guard /fence/wall'	Baah	[baa]
	b.	Apea	[apɪja]	'spear'	Appiah	[apɪa]
	c.	Ansa	[ansã]	'it does not finish'	Ansah	[ansa]
	d.	Ammoa	[ammõa]	'it does not help'	Amoah	[amua]
	e.	Mɛnsa	[mẽnsã]	'third male born after two males'	Mensah	[mənsa]
	f.	Brako	[brakõ]	'life is war'	Brakoh	[braku]
	g.	Damoa	[damõa]	'grave'	Damuah	[damua]

The use of this sound does not have any systematic explanation other than being analyzed as a stylistic convention. If there were no final ⟨h⟩ in these words an English speaker might pronounce the final [a] as unstressed [ə] in bisyllabic words, for example, [ansə] for Ansa. To avoid this, the ⟨h⟩ is added, ensuring [a] production. Though this paper does not focus on the morphological patterns of APNs, in this data, particularly (15c-d), there are nominal and negative morphemes that need to be pointed out. The stems *sa* 'war/battle' and *boa* 'help' have a nominal prefix morpheme or nominalizer *a-* and adjacent nasal consonant [n]

in (15c) and [m] in (15d) function as negative morphemes. This shows that some Akan personal names are derived through affixation.

7.2.7 Replacement of an approximant with a lateral ⟨l⟩

A name-final approximant consonant /r/ in the original version has been replaced with a lateral /l/ consonant in the anglicized version. This occurs exclusively with the Fante dialects of Akan as illustrated in (16) below.

(16)		**Original name**			**Anglicized name**	
	a.	Ewur	[ewur]	'key/padlock'	Ewool	[əwuol]
	b.	Esar	[esar]	'warrior'	Essel	[əsəl]
	c.	Afor	[afʊr]	'tender'	Afful	[aful]
	d.	Mbir	[mbir]	'not black'	Mbil	[mbɪl]
	e.	Otabir	[otebir]	'identical twins'	Otabil	[otæbɪl]
	f.	Pantsir	[pentsir]	'stock of knowledge	Paintsil	[pantsɪl]
	g.	Enkyir	[eɲtɕir]	'do not hate it'	Enchil	[əɲtʃɪl]
	h.	Ntsifor	[ntsifʊr]	'new mind/idea'	Ntsifful	[ntsɪful]

The alveolar trill /r/ in the original form is replaced with an alveolar lateral /l/ in the anglicized form. The [i] in Akan is written as ⟨i⟩ in the Anglicized form, but pronounced as [ɪ] before /l/, reflecting an English pronunciation, as in ⟨pill⟩ [pɪl].

7.3 Resyllabification

The anglicization of Akan personal names does not affect only the orthography and segments in Akan, but syllable structure, too. Akan has three basic syllables: V, CV, and C. The C is a syllabic consonant that can be a nasal, trill, lateral, or labial-velar glide as in /asɛmʊ/ [a.sɛ.m] 'issue', /asɪwʊ/ [a.sɪ.w] 'parent-in-law', /atarɪ/ [a.ta.r], /ɔhɪnɪ/ [ɔ.hɪ.n] 'chief' correspondingly. Akan does not have CC or VC syllables (Marfo 2013; Marfo & Yankson 2008; Dolphyne 2006), i.e. Akan does not permit coda consonants. The Wassa dialect differs from the Fante dialect on this point, as illustrated in (17).

(17)		**Wassa**	**Fante**	
	a.	ɔ.ta.nĩ	ɔ.tã.n	'enmity'
	b.	a.ta.rɪ	a.ta.r	'cloth'
	c.	pa.mʊ̃	pã.m	'send him/her away'
	d.	sa.wʊ	sa.w	'bed'

The Fante form of the data has gone through a phonological process of final vowel deletion. When the vowel deletes, the adjacent non-nucleus segment becomes syllabic. Like in Wassa, the Asante dialect maintains the CV syllable structure. Dolphyne (2006) proposes that every vowel in Akan constitutes a syllable and each vowel in a vowel sequence belongs to a different syllable, whether the vowels are pronounced on different pitches or not, and whether the vowels are of the same quality or not as shown in (18).

(18) **Syllable Structure**

a.	[ti.e]/[tsi.e]	CV.V	'listen'
b.	[da.a]	CV.V	'everyday, daily'
c.	[a.ba.a]	V.CV.V	'stick'

Anglicization sometimes modifies the syllable structure of the original indigenous Akan personal names; compare (19) and (20).

(19) **Original name**   **Syllable Structure**

a.	Badu	[be.du]	CV.CV
b.	Gyasi	[dʑe.si]	CV.CV
c.	Adu	[e.du]	V.CV
d.	Kwentsir	[kwe.n.tsi.r]	CV.C.CV.C
e.	Dadze	[da.d͡zɪ]	CV.CV

(20) **Anglicized name**   **Syllable Structure**

a.	Baidoo	[be.i.du]	CV.V.CV
b.	Gaisie	[ge.i.zi]	CV.V.CV
c.	Aidoo	[e.i.du]	V.V.CV
d.	Crenstil	[kɹən.stil]	CCVC.CCVC
e.	Daadzie	[dad.zi]	CVC.CV

The syllable structure of the original form is not the same as the syllable structure of the anglicized version. In examples (19a-c), the syllable structure of the original form is (C)V.CV, however, in the anglicized form, the syllabification component is (C)V.**V**.CV, because the vowel is diphthongized. In examples (20c-e), the syllable structure of the anglicized form violates the syllable structure of Akan. We have explained that Akan does not permit CC or CVC syllable structure. Akan has a voiced alveolar affricate [d͡z] sound in the Fante dialect but does not have /z/. But, this sound is treated as a cluster in the anglicized version. Akan does not have a [st] or [ts] consonant cluster but rather a voiceless alveolar affricate [t͡s], which can occur in the onset slot. This is realized as [st] in the anglicized version of Kwentsir.

7.4 Tonal Alteration as an Anglicization Strategy

Tonal alteration or shift is one of the common strategies of anglicizing APNs. So far, we have not been indicating tone in the forms. We will now consider it. The tones born by the indigenous personal names are carriers of the socio-cultural meaning of the names. When the speaker's tone changes from the original lexical tone, the socio-cultural meaning of the names diminishes. The following data illustrate the tone used to pronounce the original and the anglicized variant of the Akan personal names.

(21)		**Original name**			**Anglicized name**	
	a.	Dadze	[dàd͡zí]	'iron/metal'	Dadzie	[dáádzì]
	b.	Owu	[òwú]	'death'	Owoo	[ówù]
	c.	Otu	[òtú]	'gun'	Otoo	[ótù]
	d.	Yɛboa	[jɛ́bʊ̀á]	'we help'	Yeboah	[jɛ́búà]
	e.	Ofori	[òfòrí]	'convener'	Ofori	[òfóɹì]
	f.	Ammoa	[àmʊ̀á]	'does not help'	Amoah	[ámúà]
	g.	Asare	[àsàrì]	'warrior'	Asare	[àsáɹì]
	h.	Nkum	[ŋkǔḿ]	'do not kill me'	Incoom	[ŋ́kǔm̀]
	i.	Mɛnsa	[mɛ̀ǹsá]	'third male born after two males'	Mensah	[mɛ̌nsà]
	j.	Odum	[òdúḿ]	'oak tree'	Odoom	[ódùm̀]
	k.	Ankra	[àŋkɹ̀á]	'unnoticed while coming'	Ancrah	[áŋ́kɹà]
	l.	Asante	[àsàǹtí]	'name among Asante ethnic group'	Asante	[àsáńtì]

It can be seen from (21) that the tone for the original names has one of six tone melodies: LH, LLH, LLL, LHH, HLH, and LLLH, namely a L tone followed by a H tone somewhere in the word, or all L tone. However, the anglicized version has HL, HHL, LHL, HLL, LHHL tone melodies, that is, always a H tone followed by a L tone somewhere in the word. Now the issue is why the tones change under anglicization. The answer is that English stress matches high tone in tone languages (Leben 1996, Owino 2003, Mugabe 2006, among others). So, if one considers where English would place stress naturally in words like this, it is typically the place where a high tone appears in some of the words, while unstressed syllables receive low tone.

These tonal variations cause a change in the meaning of APNs' names. While the names with tones carry meaning to the bearer, the anglicized versions do not

convey any socio-cultural meaning to the bearer. When we asked our respondents to tell us the meaning of their names, out of the 50 Akan respondents, only 6 people (constituting 12%) were able to tell the socio-cultural meaning of their names. These people have backgrounds from rural areas and are teachers. This shows that people bear names but they do not understand the original meaning of the names they bear.

The alteration of the tone of the various names carries social tags. Our respondents revealed that an anglicized pronunciation indicates a high social class whereas original pronunciation indicates a low or inferior class. The anglicized names show that the person who is pronouncing the name is educated or has gone to school. Anglicized names can also confer social power (Mensah 2022).

8 Implication of Anglicization on Akan language and culture

As has been noted already, the three main dialects of Akan that have achieved writing status, namely Akuapem, Asante and Fante are studied at all levels of education in Ghana, from the basic up to the tertiary level. Even at the lower primary level, Akan is used as a medium of instruction (Asare et al. 2012; Adika 2012; Owu-Ewie 2017). The switch from not using Akan orthography to write personal names leads to cultural endangerment. Culture is the way of life of a group of people, and people portray this life through their names, language, food, dress, songs, buildings, etc. The most common mean of knowing someone's background is to know the name he or she bears and the language he or she speaks. This is because names show the country and tribe one is coming from. When people change or modify the orthography of their names to that of the English language, it does not affect the meaning alone, but their identity as an African (Dzameshie 1998; Adebanwi 2012).

Another implication of anglicization for the Akan language is language endangerment. Language endangerment refers to a language on the verge of extinction. Nowadays, English has become a lingua franca among Ghanaian youth, including some young Akan speakers. Though the Ghanaian language is used as a medium of instruction in lower primary, few students with elite backgrounds living in cosmopolitan areas proudly use the language (Owu-Ewie 2017, Fairclough 2004). It is only adult speakers and children from relatively smaller communities who use Akan frequently and fluently in public discourses. This is because the English language is equated to literacy and civilization; hence, the preference of the young Akan speakers to anglicize their names in order to sound and feel

more modernized. As a result, the original meanings that the indigenous Akan names bear are being eroded. Young Akan speakers cannot tell the meanings of the names they bear themselves or pass on to their children.

From the study, it was inferred that the respondents recognized the fact that anglicized Akan personal names are less meaningful compared to their non-anglicized counterparts. Although one must recognize that these names are not English names, they are Akan names that have been altered, yet it was observed that the young respondents especially still prefer being addressed by these anglicized personal names. It further came out that the majority of the respondents do not seem to appreciate the impact of anglicization on their linguistic and cultural identities. The implication for this is that although the Akan language in general is widely spoken, the Akan culture, through the indigenous Akan personal names, faces a threat of losing some value if the status quo perpetuates.

9 Conclusion

In this study, we have examined orthographic changes underlying some anglicized personal names in Akan and their implication for the Akan language and culture. We have shown that the anglicization of Akan personal names affects Akan orthography and pronunciation of indigenous names. This eventually results in the alterations of the names and in some cases, the meanings of the names are no longer apparent. Some of the processes that Anglicization goes through include orthographical modifications such as vowel replacement, vowel addition, and consonant replacement. For modifications beyond the segments, resyllabification and tone alterations were observed to occur in the anglicization process. The latter is the influence of anglicization on the sociocultural values and meanings of the affected person's names. Though Akan speakers are aware of the implications of the changes to the Akan language and culture, especially those of higher social class prefer the anglicized names as an expression of social identity. We conclude that this attitude by the bearers of these names' has a dire consequence for the future of indigenous Akan personal (family) names in particular and the culture in general.

Acknowledgements

got during our paper presentation. Again, we are indebted to the invaluable review comments by the anonymous reviewers and especially by Prof. Sharon Rose, the co-editor of the proceedings, in the preparation of this chapter. However, all shortcomings that may be found in the chapter are solely ours.

Appendix: List of Akan personal names

1. Ɔsɛe	20. Baa	39. Benya	58. Kɔba
2. Dɔnkɔ	21. Ansa	40. Gyeduwa	59. Bɛɛko
3. Badu	22. Amoa	41. Aban	60. Amoonu
4. Abɛbrɛsɛ	23. Mɛnsa	42. Sapɔn	61. Duodu
5. Akyerɛ	24. Damoa	43. Ako	62. Bɛntsil
6. Dadze	25. Brako	44. Edua	63. Tombo
7. Apea	26. Kwansa	45. Egyaa	64. Tumbil
8. Afor	27. Ahoma	46. Akɔndɔ	65. Danfor
9. Otu	28. Enin	47. Ayedan	66. Etuafor
10. Owu	29. Enamil	48. Sakyi	67. Ogow
11. Adu	30. Krampa	49. Abotar	68. Etua
12. Ewur	31. Nyan	50. Boadi	69. Abaka
13. Odum	32. Bɛntsil	51. Amisa	70. Ɔko
14. Nkum	33. Sɛsa	52. Asafua	71. Agre
15. Ɔbo	34. Efum	53. Takyi	72. Asan
16. Sereboɔ	35. Esuman	54. Afedzi	73. Kwaw
17. Fɔdwoɔ	36. Safo	55. Efe	74. Atoapem
18. Ɔkofoɔ	37. Anaman	56. Asamoa	75. Dankwa
19. Kufoɔ	38. Kɔbena	57. Gaban	76. Asɛnso

77. Ampɔnsa
78. Yiadom
79. Amakye
80. Ababio
81. Akomanin
82. Ado
83. Aka
84. Awua
85. Amankwaa
86. Amokwandɔ
87. Banyin
88. Esenyi
89. Ɔnoma
90. Aborkwa
91. Esandɔ
92. Ayɛtse
93. Okwan
94. Agyakwa
95. Dɛtse
96. Kyɛm
97. Abagya
98. Nyim
99. Nimako
100. Abonyi

101. Kwao Otu
102. Kwao Anu
103. Ewur
104. Esar
105. Mbir
106. Otabir
107. Pantsir
108. Ntsifor
109. Enkyir
110. Ɔkyere
111. Takyi
112. Botwe
113. Otwe
114. Agyapɔn
115. Akyampɔn
116. Esuon
117. Ɔpɔn
118. Aban
119. Asan
120. Ɔten
121. Ɔben
122. Ato
123. Ebo
124. Ahen

125. Ako
126. Akon
127. Adɛɛ
128. Esar
129. Gyasi
130. Hama
131. Koma
132. Nyameatse
133. Nyamesɛm
134. Kwentsil
135. Esiaw
136. Gɔɔman
137. Aduamoa
138. Ano
139. Dwuma
140. Bagyina
141. Kweku
142. Nkroma
143. Yanki
144. Panki
145. Yanka
146. Amoako
147. Ndome
148. Ndo

149. Aboagye
150. Akron
151. Ampa
152. Andɔ
153. Obosu
154. Dantsil
155. Daako
156. Manko
157. Dɛw
158. Krampa
159. Ewudzi
160. Bampo
161. Beeden
162. Abaasa
163. Benyin
164. Benya
165. Danso
166. Bekyɛm
167. Ɔkosa
168. Otsiwa
169. Apem
170. Brebo
171. Bɔtse
172. Akorabo
173. Etudur

174. Amoanin
175. Otoba
176. Eduam
177. Gyatɛ
178. Apaadu
179. Eyisa
180. Ɔsompa
181. Ɔsomnya
182. Saakwa
183. Aboa
184. Akaekyir
185. Ayɛyie
186. Anamoa
187. Anamɔn
188. Apɛnkwa
189. Konya
190. Akyerɛfi
191. Yɛboa
192. Kontɔ
193. Amoadu
194. Fɔ
195. Awer
196. Ananse
197. Ogua

198. Nimo
199. Ɔmano
200. Akɛsɛ
201. Adu
202. Ankoma
203. Kwentsir
204. Crentsil
205. Ofori
206. Asare
207. Ankra
208. Asante
209. Kwakye
210. Abor
211. Ɔtoboa
212. Boaben
213. Andɔ
214. Esiaw
215. Ata
216. Abam
217. Sakyi
218. Nyamekyɛ
219. Asuako
220. Etsiwa
221. Adom

222. Atakora
223. Amoesi
224. Aggrey
225. Akonu
226. Ano
227. Boadu
228. Boafo
229. Asaam
230. Kuma
231. Kyeasi
232. Okran
233. Ɔfɛɛ
234. Sago
235. Nuunu
236. Patu
237. Gyau
238. Kwaano
239. Kɛse
240. Kwabo
241. Kwadu
242. Ɔbo
243. Ahor
244. Ahuren
245. Ahun

246. Esuman
247. Taatu
248. Amoasi
249. Adabo
250. Beesi
251. Baafi
252. Ato
253. Maafo
254. Kwaafo
255. Anto
256. Ahɔe
257. Yaaba
258. Werɛko
259. Tsipa
260. Tsikwa
261. Amanin
262. Kuntu
263. Nkansa
264. Ekuhyia
265. Ɔfen
266. Pra
267. Ayensu
268. Adoko
269. Boakye
270. Anɔkye

271. Ɔmane	279. Gura	287. Edonu	295. Amoakwa
272. Esua	280. Anɔkwa	288. Akonya	296. Amosa
273. Afoa	281. Sãã	289. Brenu	297. Bosompem
274. Pɔw	282. Ampɔnsɛm	290. Boako	298. Ɔkoampa
275. Onkyir	283. Prɛko	291. Brako	299. Akono
276. Wɔntwe	284. Esiedu	292. Apeasa	300. Boaful
277. Dɔntwe	285. Ewusi	293. Asampɔn	
278. Dɔntɔ	286. Edumadze	294. Amampɔn	

References

Abakah, Emmanuel N. 2005. Phonological analysis of word-final consonants in Akan. *Africa & Asia* 5. 47–65.

Abakah, Emmanuel N. 2012. Some assimilatory processes in Akan. *Journal of West African Languages* 39(2). 47–82.

Abakah, Emmanuel N. 2013. Vowel replacement pattern in the Mfantse dialect of Akan. *Journal of Universal Languages* 14(2). 7–51.

Adebanwi, Wale. 2012. Glocal naming and shaming: Toponymic (inter) national relations on Lagos and New York's streets. *African Affairs* 111(445). 640–661.

Adika, Gordon Senanu Kwame. 2012. English in Ghana: Growth, tensions, and trends. *International Journal of Language, Translation and Intercultural Communication* 1. 151–166.

Adomako, Kwasi. 2015. Some phonological processes in an Akan linguistic game. *Linguistik Online* 72(3). 3–22. DOI: 10.13092/lo.72.1970.

Adomako, Kwasi. 2017. Morphophonological analysis of Akan female family-name formation. *Ghana Journal of Linguistics* 6(3). 1–32.

Adomako, Kwasi. 2018. *The phonology of Akan loanwords in Ga and Dangme*. University of Ghana. (Doctoral dissertation).

Adomako, Kwasi. 2019. Some morphological and phonological processes in Akan day-names formation. *Journal of the Linguistic Association of Nigeria* 22(1). 116–136.

Agyekum, Kofi. 2006. The sociolinguistic of Akan personal names. *Nordic Journal of African Studies* 15(2). 206–235.

Agyekum, Kofi. 2010. The ethnopragmatics of Akan compliments. *Legon Journal of Humanities* 21. 13–38.

Ansu-Kyeremeh, Kwasi. 2000. Communicating nomination: Some social aspects of Bono personal names. *Research Review New Series* 16(2). 19–33.

Asare, Emmanuel O., Damian K. Mereku, Jophus Anamuah-Mensah & George K. T. Oduro. 2012. *In-service teacher education study in Sub-Saharan Africa: The case of Ghana* (Report of a study funded by UNESCO Commonwealth Secretariat, Association for the Development of Education in Africa (ADEA)). Accra: Teacher Education Division, Ghana Education Service.

Boadi, Lawrence Addai. 1984. How to derive Araba and Abenaa from a common underlying representation: Some comments on historical methodology. *Anthropological Linguistics* 26(4). 435–444.

Christaller, Johann Gottlieb. 1933. *A dictionary of the Asante and Fante language called Tshi (Chwee, Twi).* 2nd edn. Basel: Basel Evangelical Missionary Society.

Crayner, John Brandford. 1988. *Yeehyiahyia oo!* Tema: Ghana Publishing Corporation.

de Jong, Kenneth & Samuel G. Obeng. 2000. Labio-palatalisation in Twi: Contrastive, quantal, and organizational factors producing an uncommon sound. *Language* 76(3). 682–703.

Dolphyne, Florence Abena. 2006. *The Akan (Twi-Fante) language: Its sound systems and tonal structure.* Accra: Woeli Publishing Services.

Dzameshie, Alex K. 1998. *Sociocultural and sociolinguistic dimensions of Ewe personal names.* Paper presented at the GILLBT Seminar Week in Tamale; 16th-20th February.

Essien, Okon E. 1986. *Ibibio names: Their structure and meaning.* Ibadan: Daystar Press.

Fairclough, Norman. 2004. *Analyzing discourse: Textual analysis for social research.* London: Routledge Taylor & Francis Group.

Ghana Statistical Service. 2021. *2021 Population and Housing Census preliminary report.* Accra. http://www.census2021.statsghana.gov.gh/.

Kachru, Braj B. 1994. Englishisation and contact linguistics. *World Englishes* 13(2). 134–154.

Leben, William R. 1996. Tonal feet and the adaptation of English borrowings into Hausa. *Studies in African Linguistics* 25. 139–154.

Marfo, Charles. 2013. Optimizing structure: The case of the 'CrV' syllable of Akan. *Ghana Journal of Linguistics* 2(1). 63–78.

Marfo, Charles & Solace Yankson. 2008. The structure of the CCV syllable of the Akan. *Concentric studies in Linguistics* 34(2). 85–100.

Mensah, Eyo O. 2022. The Englishisation of personal names in Nigeria: What Englishisation of Efik and Ibibio personal names suggests about English in Nigeria. *English Today* 38(3). 152–164.

Mireku-Gyimah, Patricia Beatrice & Anthony Armstrong Mensah. 2015. Anglicized places names in Ghana: The case of towns in the Tarkwa mining areas. *The Extractive Industries and Society* 2(1). 19–23.

Mugabe, John (ed.). 2006. *Tone loans: The adaptation of English loanwords into Yoruba*. Somerville, MA: Cascadilla Proceedings Project.

Obeng, Samuel G. 1997. From morhophonology to sociolinguistics: The case of Akan hypocoristic day-names. *Multilingua* 16. 39–56.

Obeng, Samuel G. 1998. Akan death-prevention names: A pragmatic and structural analysis. *Names* 46(3). 163–187.

Obeng, Samuel G. 2001. *African Anthroponymy: An ethnopragmatic and morphophonological study of names in Akan and some African societies*. Munich: Lincom Europa.

Odoom, John K. 2013. *Nkodzen: Mfantse Anokasadwin na Amambra for SHS*. Agona Swedru: Klinsmart Publications.

Odoom, John K. & Kwasi Adomako. 2021. Progressive vowel harmony in Gomoa. *Ghana Journal of Linguistics* 10(2). 58–83.

Ofori, Seth Antwi. 2019. Rules and rule relations in Akan day names. *Journal of West African Languages* 46(1). 47–73.

Owino, Daniel. 2003. *Phonological nativization of Dholuo loanwords*. University of Pretoria. (Doctoral dissertation).

Owu-Ewie, Charles. 2017. Language, education and linguistic human rights in Ghana. *Legon Journal of the Humanities* 28(2). 151–172.

Schachter, Paul & Victoria Fromkin. 1968. *A phonology of Akan: Akuapem, Asante and Fante* (UCLA Working Papers in Phonetics 9). University of California, Los Angeles.

Ukpong, Edet. 2007. *An inquiry into Ibibio culture: Ibibio names*. Uyo: Dorand Publishers.

Williamson, Kay. 1984. *Practical orthography in Nigeria*. Ibadan: Heinemann.

Chapter 3

Plural marking in Nigerian Pidgin English: A sociolinguistic study of diaspora speakers

Precious Affia

York University

This study examines plural marking in Nigerian Pidgin English (NPE) in light of an ongoing change in the language. Many older studies hold that *dem* and zero markings are the dominant plurals used in NPE (Mafeni 1971, Faraclas 1989) while more recent studies hold that *-s* is the dominant plural in NPE (Deuber 2005, Ogunmodimu 2014). This apparent time study assesses the choice of plurals in the spontaneous speech derived from sociolinguistic interviews with 20 older and younger NPE speakers living in Winnipeg to bring to light the patterning of this change. The interaction between age and gender is found to be statistically significant as *-s* is the dominant plural used among the educated middle-class speakers of NPE in Winnipeg, with younger speakers and female speakers leading the change while the older male speakers are slower to adopt this change.

1 Background

1.1 Nigerian Pidgin English (NPE)

Nigerian Pidgin English (NPE) is part of a continuum of West African Pidgin Englishes that stretches from Sierra Leone in West Africa to Gabon in Central Africa (Barbag-Stoll 1983). In 2016, Nigeria was estimated to have between three and five million people who use primarily NPE in their day-to-day interactions, and up to seventy-five million people (about 35% of the population) who use NPE as their second language (BBC 2016). However, in a more recent study, Faraclas (2021) claims that there are about 110 million speakers of NPE in Nigeria,

making it not just the most widely spoken language in Nigeria but one of the most widely spoken languages in the world. NPE has come to dominate urban spaces in Nigeria and has widespread acceptance. It has even become the first language of many Nigerians, especially in Edo state, Warri, Port Harcourt and Sapele (Elugbe & Omamor 1991, Igboanusi 2008). Although English is the official language in Nigeria, NPE has almost taken over the role of lingua franca in informal domains (Ogunmodimu 2015). NPE and its lexifier language English are "separate varieties". Evidence of this comes from Deuber (2006) who empirically investigates the possible existence of a continuum in Nigeria by looking at NPE spoken by educated speakers in Lagos, Nigeria in the areas of copula constructions, tense-aspect marking and verbal negation. She finds NPE and English to be two separate varieties with no evidence for intermediate varieties.

Despite NPE being an informal lingua franca, which transcends regional, ethnic and linguistic boundaries, it has no official status or recognition in Nigeria (Faraclas 2021). NPE is not acquired through formal education and has no standardized orthography. Until rather recently, it was accorded a low prestige in Nigeria and relegated to the less educated. It is now increasingly used also by and among educated people and in tertiary institutions (Agbo & Plag 2020). It cuts across religious affiliations, and even Nigerians in the diaspora use NPE as a language that reflects their "Nigerianness", a marker of identity and solidarity (Akande & Salami 2010, Affia 2022). Affia (2022) conducted a preliminary attitudinal study with ten Nigerians in Toronto (five males and five females) to assess the role of NPE in identity formation in the diaspora. The participants in the focus group expressed in different ways how NPE is tied to their identity as Nigerians when identifying with other Nigerians in Toronto. In light of the widening use of NPE, the present study will examine how the language is used in the diaspora with a focus on one salient feature of the language, plural marking.

1.2 The variable: Plural marking in NPE

Plural in NPE is marked in a number of ways. Faraclas (1989) argues that most nouns in NPE are assumed to be singular, except when indicated by morphosyntactic or pragmatic contextual cues, and bare nouns are given a generic reading. The (morpho-)syntactic means of signaling plural in NPE as identified in Tagliamonte et al. (1997) are the use of inflectional plural *-s*, the postnominal plural *dem*, and prenominal determiners which include numerals and indefinite quantifiers. Reduplication of adjectives/nominals is another plural marking strategy in NPE, and NPs can also occur without any overt plural marker (zero marking) where

the plural reading is derived from context based on prior information (Ogunmodimu 2014).

2 Literature review

2.1 Plural marking in NPE

Both descriptive and variationist studies have examined the use of plural markers in NPE. Faraclas (1989) analyzes the grammar of NPE based on transcribed interviews of Nigerians who were living in Port Harcourt between 1985–1986. He explains that in NPE, most nouns are assumed to be singular except when indicated by pragmatic or morphosyntactic cues. Numerals, quantifiers, the postnominal plural *dem* and null marking (zero marking: where the plural noun is bare and has no suffix or plural modifications) were reported as the plural markers in his study. He found that *dem* and null marking are the most common plurals in NPE (Faraclas 1989: 352), and that *-s* can be used as a plural indicator but referred to it as 'borrowing'. Mafeni (1971: 110) also identifies *dem* as the most common if not the only plural marker in NPE. Deuber (2005) is a more recent study that investigated the grammar of NPE speakers in Lagos, a metropolitan city in Nigeria. According to Deuber (2005: 119), although people in Lagos use *dem*, it is very limited, has disappeared from the speech of the average educated Nigerian and is restricted to formal radio broadcasts as a conscious pidginization strategy. One possible reason for the difference in the frequency and use of *dem* between the studies by Faraclas and Mafeni on the one hand, and the study by Deuber on the other hand, is that the former studies were conducted in areas where NPE is creolized and used as the first language of many, while the latter was an L2 variety. It may also be a result of an ongoing change in the language.

In the variationist study of Tagliamonte et al. (1997), the authors look at linguistic factors that may influence the choice of plural markers. The data collected for analysis was from 12 NPE speakers (11 adults and one child) living in Ontario, Canada. They were all born and raised in NPE-predominant areas in Nigeria, (Bendel, Rivers, Lagos) and immigrated to Canada between 1991 and 1996. The age range for the adults was from 28 to 52 years and they were all middle-class educated Nigerians. In the study, the variable context was individual nouns with plural reference, accompanied by overt plural marker or not. In their data, the plural marker *dem* accounted for less than 1% while *-s* was the dominant marker used. Animate nouns favored overt *-s* plural marking, and generic-referenced NPs favored zero marking. "Definiteness" was not a distinguishing factor in their

data and linguistic factors such as the preceding or following phonological segment were not significant factors in the choice of plural the speakers used. Tagliamonte et al. (1997) suggest that knowledge of other languages may play a role in the choice of plural used and the speakers of Igbo in their study were more likely to use zero marking as a plural strategy as an influence of a dominant Igbo grammar. The researchers suggested that the more exposed to English a speaker is, measured by their acquired level of education, the more likely it is that they will use more inflectional plural -*s*. Social factors such as age and gender, which are the main focus of the present study, were not investigated, and an apparent time analysis was not the goal of their study.

Ogunmodimu (2014) is another variationist study of plural marking in NPE. In contrast with Tagliamonte et al. (1997), Ogunmodimu's data collection was carried out in Nigeria. The variants investigated were inflectional plural -*s*, zero marking, numerals/quantifiers, *dem* and double-marking which is when more than one plural marking strategy is used to mark plural on a noun phrase. The participants were native speakers of Hausa, Yoruba and Igbo, and minor languages Urhobo and Edo. He analyzed the social factors such as ethnicity, education, and region of residence, and a linguistic factor of animacy of antecedent. His results showed that -*s* was the dominant plural used by the speakers while *dem* was rare (Ogunmodimu 2014: 10). Animate referent NPs strongly favored -*s* plural over inanimate referents (Ogunmodimu 2014: 13). Speakers with higher education levels strongly preferred the -*s* plural marker over others. The use of *dem* was highly disfavored by those with higher levels of education (Ogunmodimu 2014: 13). Speakers of Igbo, Urhobo and Edo, all have lexical plural markers in their native languages that function somewhat like *dem* and these speakers used *dem* more than speakers of Hausa, which does not have such a plural marker. Yoruba speakers, on the other hand, produced no *dem* even though they use the third person plural pronoun to mark plural in Yoruba. For double marking, speakers with university education used double marking more than those who only had high school education (Ogunmodimu 2014: 14). The author noted that education is the major factor determining the use of -*s*, suggesting it is as a result of contact with English (Ogunmodimu 2014: 15). There was an absence of significant differences correlating with speaker's gender, L1 or region of residence, which, according to the researcher, suggests a broad consistency of NPE. Language change using an apparent time framework was not investigated by Ogunmodimu (2014).

2.2 Social variables - age, gender and dominant language

The social variables investigated in this study are gender, age and dominant language. The gender of a speaker has been found to be one of the most important social factors predicting sociolinguistic differences. Studies investigating speaker gender go back to foundational studies (Labov 1966, Trudgill 1972), both showing that gender is a predictor of linguistic differences. Trudgill (1972) explains that women produce linguistic forms which closely approach that of the standard language or have higher prestige than those produced by men (Trudgill 1972: 180), which could be the case in NPE. His sociolinguistic study of British English has a striking feature in common with the Labov (1966) study of American English, with both investigating the variable production of *-ing* as either /iŋ/ or /in/ in English words like *singing*. Trudgill's results showed that in Norwich, males used a higher percentage of non-standard [in] forms than females in British English. The same results were found for American English in Labov (1966). His results suggested that working-class non-standard speech is more highly valued and prestigious for male speakers, which he calls covert prestige. This was reinforced not only by usage but by the fact that male participants under-reported their use of the standard or prestige form while the women over-reported theirs. He therefore argued that women and men respond to different norms; men to covert vernacular prestige norms and women to overt, standard-language, prestige norms. In a more recent study by Kristian (2018), she investigated the same variable *-ing* among Newfoundland youths. Her results showed that gender was significant, with male speakers favoring the non-standard [in] compared to their female counterparts. Analyzing the gender effect in the present study is important because it has not been well investigated in previous sociolinguistic studies of plural marking in NPE. By looking at gender, I intend to see whether women will use more of the variant (*-s*) that has the prestige of the superstrate, English, than men.

Age is another important factor that is analyzed in this study. A gradually increasing or decreasing frequency in the use of linguistic features according to speaker age can be interpreted as a change in progress, where the apparent time hypothesis is used to investigate this change (Tagliamonte 2012). Apparent time is the study of language change by examining the distribution of variation by age groups like the older and younger generation of speakers at one point in time. For example, Labov (1963) carried out sociolinguistic interviews with the residents of Martha's Vineyard in 1961 to study (*ay*) and (*aw*) diphthongs. Younger and older people (69 in total) were interviewed assuming that the younger group would show how the dialect had changed. The age ranges analyzed were 14 to

30, 31 to 45, 46 to 60, 61 to 75 and over 75 (Labov 1963: 291). In his analysis, Labov calculated an index value for the height of the vowel nucleus of each speaker for each diphthong. The results showed that the nuclei of both diphthongs were progressively higher with each younger age cohort. The interpretation was that the older speakers learned the language at a time the speech community as a whole had lower values for the height of the diphthongs. So, the older speakers' speech reflected the state of the language at that earlier date. The regular increase across the generations investigated represents a generational change in progress with younger speakers leading the change (Labov 1963: 291). In the case of plural marking in NPE, older studies give *dem* and zero marking as the primary plurals (Faraclas 1989, Mafeni 1971) while newer studies show *-s* as the primary plural used in NPE (Ogunmodimu 2014, Tagliamonte et al. 1997). One way to investigate this change is through an apparent time study, where the speech of different generations is examined. This study seeks to infer from comparisons of two generations that any differences may be attributed to linguistic change in progress.

Finally, since Nigerians are mostly multilingual, the languages they speak could have an impact on their use of NPE. If the dominant language of an NPE speaker is Yoruba, for instance, the speaker may tend to use features in NPE that are similar to that of Yoruba when speaking NPE. This makes the variable 'dominant language' an important one when investigating variation in NPE. We saw that previous researchers suggest the influence of some of these languages as factors in their studies (Ogunmodimu 2014, Tagliamonte et al. 1997). In my study, apart from English and NPE, Yoruba and Hausa are the two Nigerian languages reported as dominant languages by some of the participants. Most of the participants are Yoruba speakers, which is a West African language different from Igbo, the native language of the majority of the participants in Tagliamonte et al. (1997). Since the authors suggested that substratum[1] influence from Igbo may have been the reason for the high rate of zero marking, because zero marking is a strategy

---

[1]I chose to include "dominant language" as a factor instead of "substratum language" because, although all the participants reported that they spoke at least one substratum Nigerian language, some of them, especially participants in the younger group, expressed that they did not speak their native language fluently and frequently, using it only to exchange pleasantries and to converse with their grandparents. Hence, I decided to analyze their reported dominant language – the one they speak fluently and frequently with other Nigerians – as a factor that can influence their use of NPE, instead of a substratum language that some may not speak at all. So, the participants expressed greater preference, fluency and use of a particular language in their linguistic repertoires. Eight out of ten older speakers preferred Yoruba and spoke it in social settings, and in their homes every day, which could affect their use of NPE. It contrasted with the younger speakers where more variability of dominant language is found.

in Igbo, it will be interesting to see if the plural marking of Yoruba speakers and participants in my study will suggest substratum influence for those who speak the substratum language fluently and frequently. The plural marking patterns of NPE, Yoruba and Hausa are presented below in section §3.2.

2.3 Linguistic variables

Animacy, type of nominal reference, and determiner type are the three linguistic variables investigated in this study. In the data, animacy was divided into animate NPs (human and animate nouns, ex. *boy*, *dog*) and inanimate NPs (*table*). For the factor nominal reference, I divided NPs into definite NPs (*the boy*) indefinite NPs (*a boy*) and generic referent NPs (*boys*). Determiner type is divided into possessive (*our house*), no determiner (*house*), demonstrative (*that house*), numeric (*two houses*) non-numeric or quantifier (*many houses*) and definite determiner (*the house*). The primary reason for investigating these linguistic variables is that both determiner type and animacy were found to be significant in Tagliamonte et al. (1997)'s study. I adopt these linguistic factors here to see if the patterns found in their study would be mirrored in my data.

3 Methodology

3.1 Data collection and subjects

Twenty participants took part in the study, ten males (five younger and five older) and ten females (five younger and five older). Two age categories were targeted: those in the older category were 50 years old and above (born in 1963 and before) while the younger category included individuals between the ages of 18–35 years (born between 1985–2001). They were all NPE speakers living in Winnipeg, Manitoba at the time of the interviews in the fall of 2019. They were recruited and selected through the snowballing method, targeting those who lived in Nigeria long enough to acquire NPE and speak it fluently. I went to a Nigerian church in Winnipeg and discussed the study with the pastor, who then recommended church members who fit into the older group age category. The individuals who spoke NPE signed up to participate in the study and recommended their friends to participate as well. For the younger group, I approached Nigerian students at the University of Manitoba where I was obtaining my master's degree in linguistics at the time. The people who fit the criteria for the study participated and recommended others. Sociolinguistic interviews were conducted with the participants in NPE, and the recordings were the source of data for this study. As a

young female speaker of NPE, I conducted all the interviews mostly presenting a cheerful countenance while the conversations in the interviews progressed. The interviews with the younger speakers, who were mostly University of Manitoba undergraduate students, were conducted in the sociolinguistics lab of the University of Manitoba. The older speakers were interviewed in their own homes or in quiet places of their choice, like libraries and churches in Winnipeg. Topics discussed in the sociolinguistic interviews were informal and centered on issues such as language use, school activities, hobbies, sports, TV shows, friends, family, etc. The interviews were semi-structured to allow the participants to be flexible with their responses, and each interview lasted for about an hour. The recording device used for the interviews was the Zoom H4N Handy recorder. Sanken COS-11D Miniature Omnidirectional Lavalier Microphones were used as external microphones. All the speakers were multilingual with fluency in English, one major or minor Nigerian language, and NPE. The major languages reported were Hausa, Yoruba and Igbo, and the minor language was Ẹ̀dó/Bini. The dominant languages of the speakers which were self-reported in the interview were Yoruba (10), NPE (5), English (4) and Hausa (1). More information about each individual participant is provided in Appendix 2 and this is presented in Table 1 below.

Table 1: Distribution of speakers by dominant language

Dominant Language	older	younger
English	1F	1M 2F
Yoruba	5M 3F	1M 1F
NPE	1F	2M 2F
Hausa	-	1M

3.2 Data

In this section, I present examples from my data for the following types of plural: inflectional *-s*, *dem*, reduplication, zero marking, and numerals/quantifiers. Although these are oral recordings, the transcription is in NPE orthography, not IPA. I also provide data from the literature on plural marking in Hausa and Yoruba.

The inflectional plural affix *-s* is similar to the [s] and [z] allomorphic forms of the plural marker in English in voicing assimilation which is suffixed to the noun in the NP. However, from the recordings and also from my knowledge as

a native speaker of NPE, [əz] the third allomorphic form of the English plural marker, cannot appear after stridents in words like *church, judge*, etc. As seen in (1) below,[2] *-s* is attached to the noun *Nigerian* to give it a plural reading.

(1) Hia dey espensiv fo os Nigerian-s (06m2001)
here is   expensive for us Nigerian-PL
'We Nigerians find it expensive to live here'.

3.2.1 *Dem*

In (2), the *dem* plural marker which immediately follows the noun *fren* signifies or provides a plural reading to the noun in the sentence.

(2) Yu no   ol of mai fren   dem. (14f1970)
you know all of my friend PL
'You know all of my friends'.

3.2.2 Reduplication

In NPE, nominals or adjectives can be reduplicated to mark plurals. The adjectives usually precede the noun. In (3), the reduplication of the noun *ashawo* shows that the word signifies more than one prostitute, providing a plural reading to the NP. In (4) below, the adjective *fain* is reduplicated to indicate plurality of the noun *haus*.

(3) Ashawo-ashawo   na im   yu no.   (20f1985)
prostitute-prostitute is   what you know
'Prostitutes are the people you know'.

(4) Im   go go biud fain-fain haus. (20f1985)
him go go build fine-fine house
'He will go and build fine houses'.

3.2.3 Zero marking

Plural reading here is derived from context due to the absence of overt plural marking and it relies on prior information to show the number of the noun in question. An example is seen below in (5) where the sentence, which is definite, does not have any overt plural marking on the noun *lab*, but the plural reading is derived from the context of the sentence.

---

[2]Speaker 06 male born in 2001.

(5)   An mai skul   get ol di   lab. (19f2000)
      and my school get all the lab
      'And my school has all the laboratories'.

3.2.4 Numerals and quantifiers

Numerals are used in NPE to show plural. The sentence in (6) below shows the use of a numeral in the NP *sevun yes* 'seven years' and the inflectional plural suffix *-s* to mark plural. This could be viewed as a type of *double marking*.[3]

(6)   Ai wok fo dem fo sevun ye-s   bifo   ai kom Canada. (11m1974)
      I   work for them for seven year-PL before I   come Canada
      'I worked for them for seven years before I came to Canada'.

Quantifiers are also used to indicate plural in NPE. The sentence in (7) shows the use of a quantifier and zero marking on the noun in question, such that the plural reading is derived from the word *meni* 'many' preceding the noun in the NP.

(7)   So meni choch no bi laik hie. (18f1969)
      so many church no be like here
      'So many churches are not like the ones here'.

3.3 Plural marking in Yoruba and Hausa

To adequately investigate possible dominant language influence, I present the plural marking strategies found in Yoruba and in Hausa. In Yoruba, there is a 3rd person plural pronoun that is used as a plural marker, but unlike the one found in NPE, it precedes the noun (8). Yoruba also has reduplication of adjectives (9) and zero marking (10) for plural marking (Ogunmodimu 2014).

(8)   (Yoruba, Ogunmodimu 2014)
      Awon   oluko yin   oga ile   iwe.
      3PL.PRO teacher praise head house book
      'The teachers praised the principal of the school'.

---

[3] For double marking, the plan was to look at them separately to determine their patterning and possible statistical significance. However, only six cases occurred in the data, five of which were a combination of *reduplication of adjectives + -s*. It was unclear in the context of use whether the reduplication of adjectives recorded for double marking were intensifying the adjectives or pluralizing the nouns. Therefore, given the small number of occurrences of this plural marking strategy in the data, and uncertainty, this testing was not done.

(9) (Yoruba, Ogunmodimu 2014: 10)
Ile nla nla po ni ilu eko.
house big big many in town Lagos
'There are big houses in Lagos'.

(10) Kole ri aja ninu igbo.
Kole see dog inside bush
'Kole saw a dog inside the bush', or
'Kole saw dogs inside the bush'. (Yoruba, Ogunmodimu 2014: 10)

In Hausa, the most widely spoken language in northern Nigeria, plural is marked by nominal inflection and nominal reduplication (Table 2).

Table 2: Hausa (Ogunmodimu 2014: 10)

	singular	plural	
a.	zabo	zabi	'guinea fowl'
b.	bàkā	bakunkunâ	'bow'

3.4 Analysis

Each interview recording was coded in ELAN transcribing software, a transcription tool that can be used for extracting and coding tokens of linguistic variables for quantitative analysis (Nagy & Meyerhoff 2015: 5). After coding the variables, I exported transcription annotations and token codes into an Excel sheet used in the statistical analysis program, RStudio. A total of 811 tokens of plural nouns were retained for quantitative analysis excluding tokens from unintelligible recordings, situations where the number is ambiguous, where there is ambiguity over whether the pluralized NP (*-s*) is English or NPE in the context of the sentence (for instance, the phrase *'I hate books'* can be English or NPE). In such cases where it is not clear from the context if the speaker codeswitched to English or not, the tokens were excluded. In R, I ran the generalized linear mixed model, or glmer. Within the model, I used random and fixed effects to determine the choice of plural marker used by the speakers from all the variables investigated. *Speaker* and *noun* as random effects were fit to the dataset in the R model. Fixed effects in the model were a two-way interaction between age category and gender, all social factors (age, gender and dominant language) and the linguistic factors (animacy, type of nominal reference and type of determiner).

4 Results

As seen in Table 3, *-s* is the dominant plural marker used in the data at 80%, while the postnominal plural *dem* is used about 1% of the time. Zero marking made up about 17% and reduplication occurred in 1.5% of the tokens counted.

4.1 Age, gender and dominant language

Although *-s* is the most frequently used variant, each variant's rates of usage differ by age. As seen in Figure 1, the younger speakers use the inflectional plural *-s* 90% of the time, while the older speakers use it 60% of the time. Zero marking is more common among the older speakers (35%) than among the younger speakers (7%). Although *dem* is not used much overall, it is especially worth noting that it is absent in the younger speakers' speech.

Figure 2 breaks down the distribution of the plural marking pattern across male and female speakers. The female speakers use somewhat more *-s* (83%) than the male speakers (79%). The male speakers use zero marking (20%) more than the female speakers (13%). The data suggests that there is slightly less *-s* usage among men and they make up for it with zero marking.

The dominant language of the speakers were Yoruba (10), NPE (5), English (4) and Hausa (1). Participants that have English as their dominant language, as seen in Figure 3, use more *-s* (91%) than speakers of the other groups. The Yoruba speakers use *-s* 85% of the time while zero marking is used 14% of the time. For the NPE speakers, *-s* is the dominant plural, used 76%, and zero marking is used 20% of the time. The Hausa speaker used the inflectional plural *-s* 55% of the time and zero marking 43% in his speech. Although this is a marked difference, nothing much can be said concerning his distribution as it could be his idiolect. Overall, the findings show that English language dominant speakers use *-s* more than the speakers in the other groups. Furthermore, Yoruba, NPE and Hausa speakers use zero marking more than the English speakers. The statistical analysis will show whether any of these differences are statistically significant.

Table 3: The overall distribution of variants

overall distribution	%	N
-s	80.6	655
dem	0.9	7
zero	17.0	137
redup	1.5	12
Total N		811
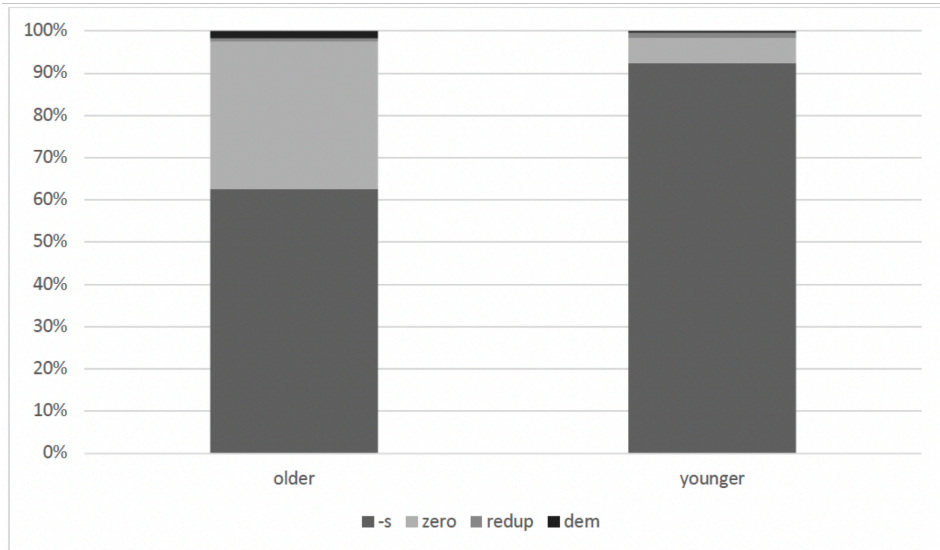
Figure 1: Use of plural variant by age group
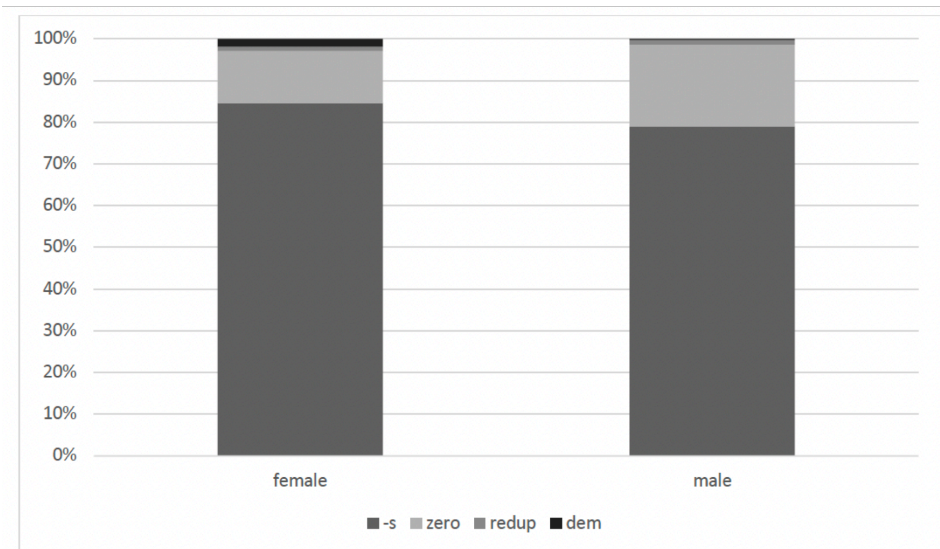
Figure 2: Use of plural variant by gender
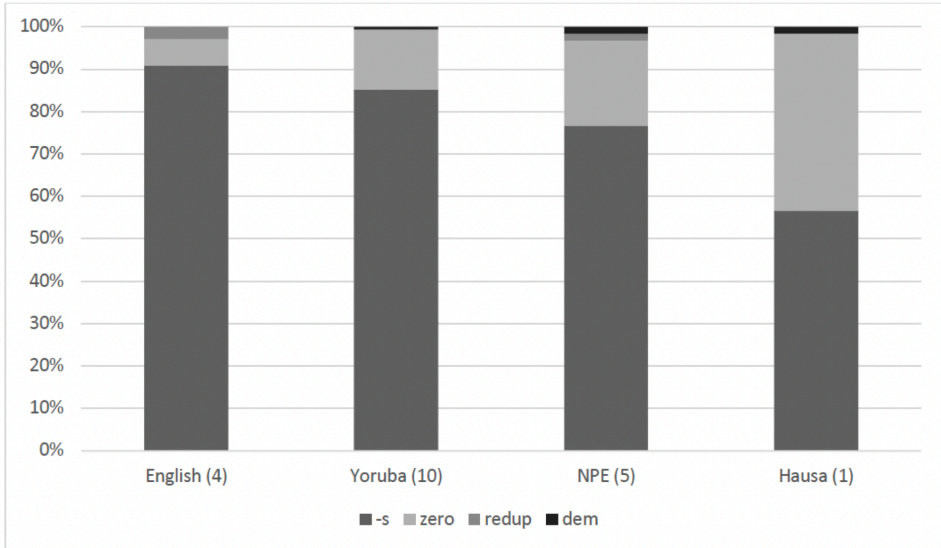
Figure 3: Use of plural variant by dominant language
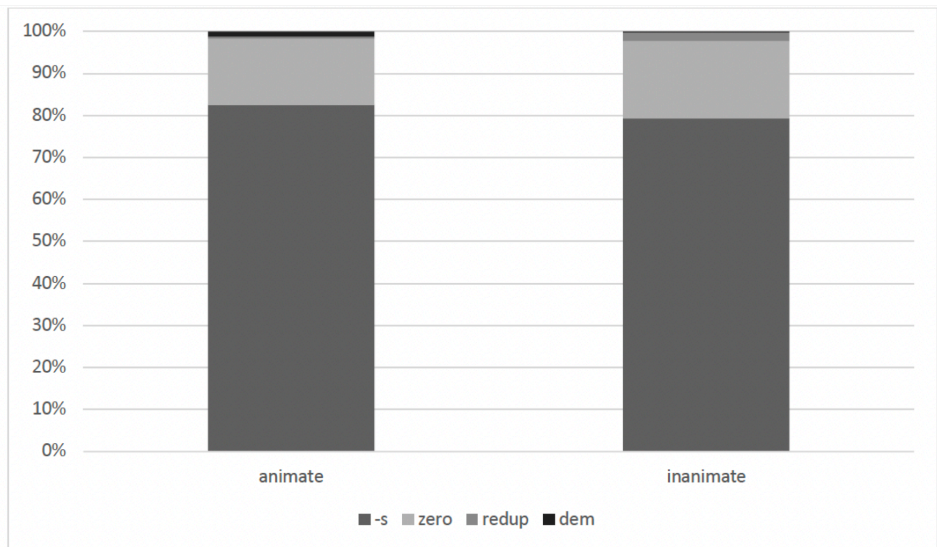
Figure 4: Use of plural variant by animacy

4.2 Linguistics variables

The linguistic variables investigated are animacy (animate and inanimate), type of nominal reference (definite NPs, indefinite NPs, generic NPs), type of determiner (definite, possessive, demonstrative, non-numeric/quantifier, numeric, no determiner). I did not find any interesting distributions or patterning with reference to these linguistic variables in the data, which is evident in the statistical result where no linguistic variable is significant. Figures 4, 5, and 6 show the distribution of plural NPs by the linguistic variables analyzed in this study.

4.3 Statistical analysis

For the statistical analysis, the variants examined were limited to two: the inflectional plural *-s* and zero marking as they made up 95% of the selected tokens. The R code for my model is:

```
glmer(plural~gender * age_category
          + animacy
          + type_of_nominal_ref
          + type_of_determiner
          + dominant_language
```
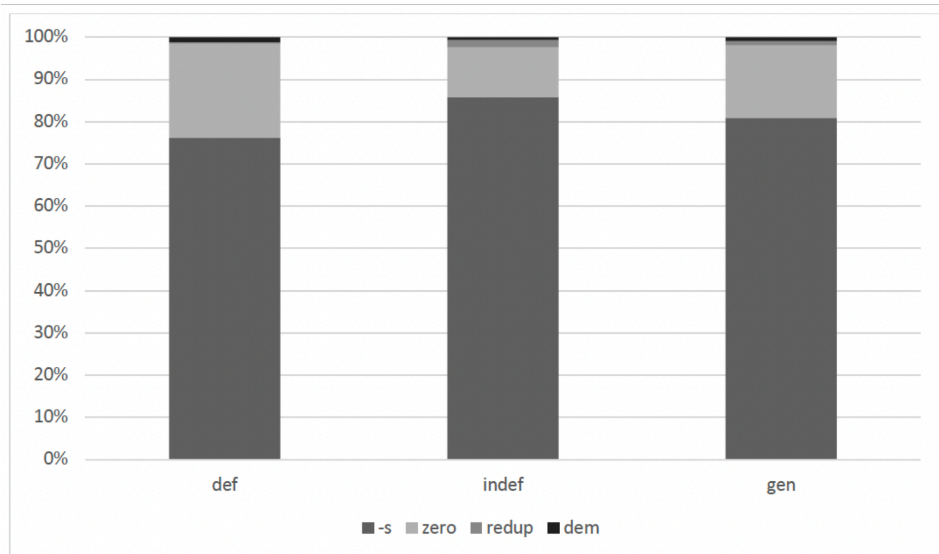

Figure 5: Use of plural variant by type of nominal reference
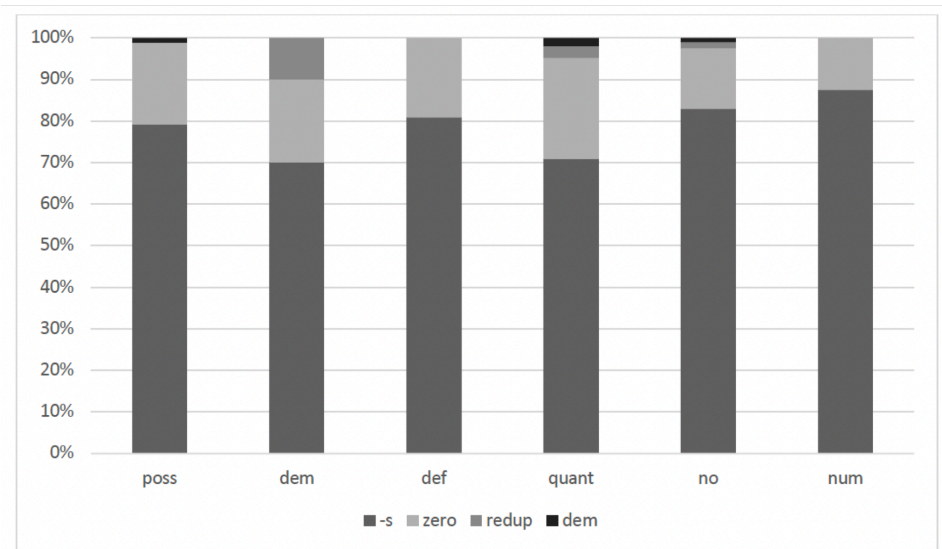
Figure 6: Use of plural variant by type of determiner

```
        + (1|speaker)
        + (1|noun),
  family = binomial,
  data = NPEFInal,
  control = glmerControl(optimizer = "bobyqa")
)
```

Table 4: Output of the best fit model

	Estimate	SE	z	$\Pr(> |z|)$	
intercept	−4.98	1.35	−3.67	0.000234	***
Ageyounger	−0.95	0.75	−1.26	0.20	
Gendermale	4.45	0.93	4.77	1.83×10^{-6}	***
Gendermale:Ageyounger	−5.57	1.26	−4.39	1.11×10^{-5}	***
dominant langHausa	6.32	1.08	5.85	4.77×10^{-5}	***
dominant langNPE	4.31	0.89	4.81	1.49×10^{-6}	***
dominant langYoruba	2.77	0.84	3.26	0.001083	**

The statistical results in Table 4 show the variables that remained after non-significant variables were removed from the model. The asterisks (*) show the levels of significance for each variant. The variables remaining are gender and dominant language, indicated by the asterisks under the Pr>|z| column. None of the linguistic predictors – animacy, nominal reference and type of determiner– had a significant effect on the choice of plural forms used by the speakers, even though predictors like animacy and type of determiner were statistically significant in the study by Tagliamonte et al. (1997).

Although age on its own is not significant, the interaction of age and gender is. Specifically, while men are statistically less likely to use the plural form *-s*, this is especially true among older males. As shown in Figures 7 and 8, the distribution of the two plural forms used by older males is approximately equal, whereas, for all other speakers, *-s* is the clear preferred variant. The tokens (zero marking and *-s*) produced by individual speakers are seen below in Figure 7. These are organized by gender and age, as indicated by (M/F) and birth year on each individual's code. This means that there could be a change in progress going on from the use of zero marking to *-s*, especially among the older group, where the female speakers are ahead in this change.

Dominant language is also a significant factor in the choice of plural used by speakers of NPE. The statistical analysis confirms the significance of the descriptive results, where zero marking is used more by speakers who have Hausa, NPE and Yoruba as dominant languages than those whose dominant language is English. The statistical analysis suggests that the interaction of age and gender is the most relevant potential indicator of a change in progress, and that the dominant language of a speaker may play a role in the choice of plural used.

5 Discussion

5.1 Age, gender and dominant language

As the statistical analysis shows, the interaction between age and gender is significant in determining the use of plural markers in NPE. This means that the interaction of both variables is predictive in the choice of plural marker used, not simply the individual variable age. The older male participants tend to use zero marking more than any other speaker group, as is apparent in Figure 8. The fact that older speakers (specifically the males) favor zero marking more than the younger group and the women, may be evidence that zero marking is the older variant. This is because from the literature regarding change in progress, there is often an interaction where women lead a change in progress, while older
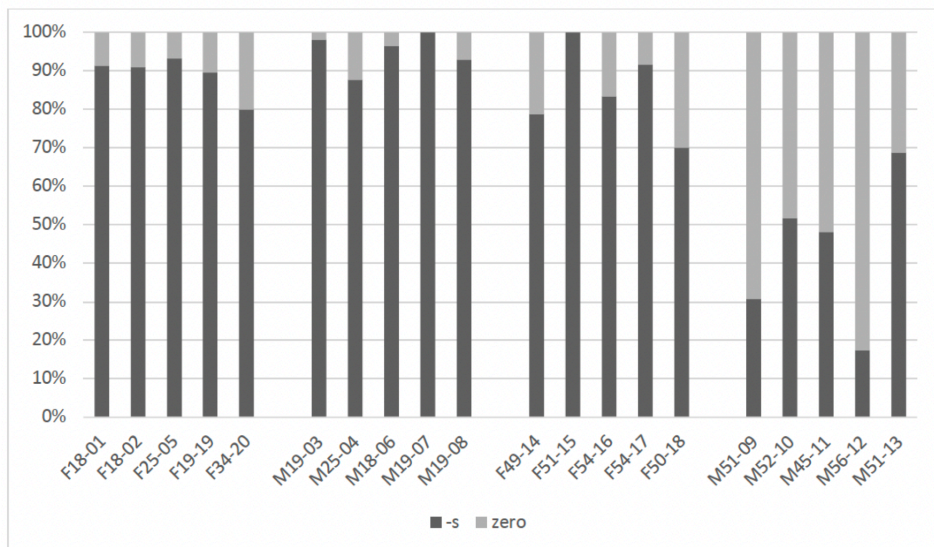
Figure 7: Distribution of *-s* and zero marking by individual speakers
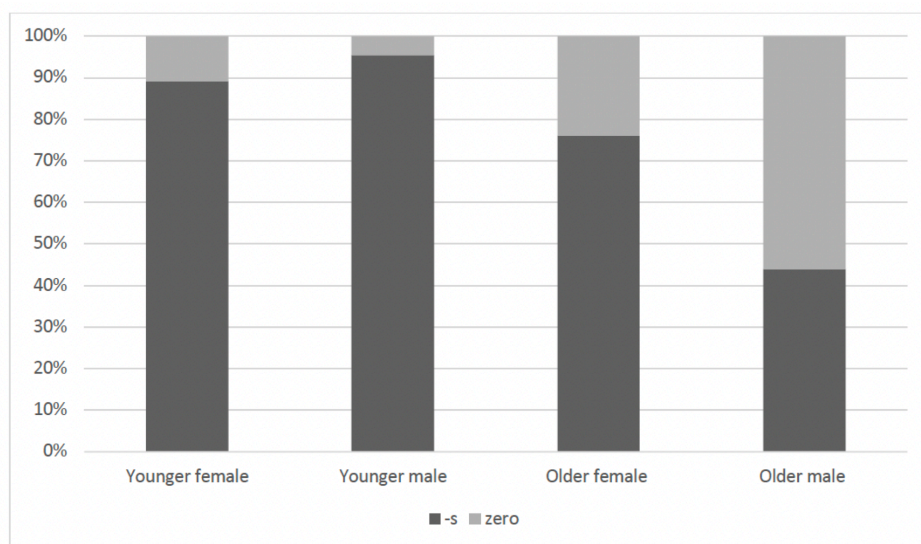
Figure 8: Age and gender distribution of *-s* and zero

men lag behind (Labov 1963, Tagliamonte et al. 1997). This interpretation may be strengthened by the contrast between older studies which show zero as the main plural in NPE (Faraclas 1989, Mafeni 1971) and more recent studies with *-s* as the main plural in NPE (Tagliamonte et al. 1997, Ogunmodimu 2014, Deuber 2005). Recall that the apparent time analysis shows language change through a comparison of the frequency of use of a language feature among different generations of speakers. The difference points to language change. Therefore, this apparent time study suggests a change in progress in reference to plural marking in NPE from zero marking to *-s. -s* is more dominant in the grammar of the younger generation than that of the older generation, with very little use of the older variants (zero marking).

Women used *-s* more than men, who in turn used zero marking more. As has been shown in this study, *-s* is a plural marker that is available in the English language. English has a higher prestige than NPE because English is an official language in Nigeria and in Canada. The common everyday language in Winnipeg, Manitoba, where the participants reside, is English. It is understood that women produce linguistic forms that more closely approach the standard language than men (Trudgill 1972, Kristian 2018). Based on the results of my study, and since *-s* has higher prestige as it is affiliated with the English language, it could be why women, particularly in the older generation, use *-s* more than men.

Apart from speaker age and gender, it is important to discuss the factor dominant language, since it was found to be significant in the statistical analysis. As seen in Figure 3, Yoruba, NPE and Hausa-dominant language speakers used zero marking more than the English-dominant speakers. This may be because, in Yoruba and NPE, zero marking can be used as a plural marking strategy and as such, it may be a result of substratum influence, as suggested in previous research (Tagliamonte et al. 1997, Ogunmodimu 2014). Not much can be said concerning the impact of Hausa as a dominant language, since only one participant spoke Hausa as a dominant language, and the pattern observed in his speech could be a result of his individual idiolect. Since the Yoruba, NPE and Hausa dominant language speakers use zero marking more than the English speakers, dominant language could be an important factor in the choice of plural used in NPE.

All the participants of this study are fluent English speakers. It may very likely be that the more NPE speakers are exposed to the formal English language, the more the grammatical features of English like *-s* are borrowed into NPE (Ogunmodimu 2014). This is especially so because English is the lexifier language for NPE and most of its vocabulary is English. Transfers between both languages are therefore common. One reason why *-s* is more predominant among the younger speakers may be because in Nigeria today, most children first learn the English

language in school and are later exposed to NPE as adults by association. This was expressed by most of the younger speakers in the discussion of language use in the interviews. As such, it could be common for them to use some rules of the grammar of English in NPE, like in plural marking. On the other hand, the older speakers were first exposed to their native language and NPE and were introduced to English as a class subject/course when they enrolled in school. This could explain why the older speakers used more zero marking than *-s*. It is important to note that eight out of the ten younger speakers were concluding their first year in the university and had lived in Canada for about a year (as shown in appendix 2), making their length of time and exposure to English in Canada shorter than the older speakers who had lived longer in Canada. This further strengthens the result of this apparent time study where the older speakers' speech, particularly the male speakers, may be reflecting the state of the language at a much earlier time during their acquisition of the language.

5.2 Linguistic factors

Contrary to Tagliamonte et al. (1997), which found determiner and animacy to be significant in the choice of plural marker used, I did not find any linguistic factor that predicts the use of a particular plural variant in a statistically significant way. As seen in Tagliamonte et al. (1997), animate NPs appear to be somewhat more likely to take *-s*, but animacy was not significant in the present study. Also, the received wisdom about plural marking in creoles is that overt marking should be favored in contexts where plurality has not been otherwise disambiguated in the NP headed by the noun in question (Tagliamonte et al. 1997: 114). Nouns with demonstrative determiners and non-numeric determiners or quantifiers which are not supposed to favor overt marking (in this case *-s*) are at least seen as slightly favoring zero marking more than the others in my data (see Figure 6 above). The lack of statistical significance may be due to limited amount of data. It is possible that with more data, a different result may be obtained.

5.3 The plural marker *dem*

This study arose from the question of the disappearance of *dem* as a plural marker in NPE. Since NPE is spoken in different geographical areas of Nigeria and has been nativized in some areas like the Niger Delta part of the country (Faraclas 2021), it could be the case that the variety of Pidgin spoken and the education of the speaker, play major roles in the use of plural marking in the language. According to Ogunmodimu (2014), *dem* is still commonly used as a plural marker in the

variety of NPE spoken in regions where NPE is nativized, like in the Niger Delta and the southeastern part of the country. Members of this community, seem to rate speakers of NPE highly with regard to the level of education attained, modernity and general sophistication (Ihemere 2006: 205). In a society like this where NPE is highly regarded, it is possible for features of the grammar such as *dem* to still be preserved over time. Faraclas (1989) was based on transcribed interviews of Nigerians living in Port Harcourt between 1985 and 1986 and he found *dem* to be the main plural used in NPE. The *dem* plural may still be a strategy used often by people in this area who are educated and are native speakers, and may be more frequent among uneducated speakers who may not have a command of the formal English language. This is supported in the Goldvarb result for *dem* in Ogunmodimu (2014: 13) that shows that the speaker's education level is a significant factor, with lower levels of education favoring *dem*.

6 Conclusion

This quantitative study investigates the use of plural markers in NPE. The fact that older literature (Faraclas 1989, Mafeni 1971, Agheyisi 1971) recognized *dem* and zero marking as the dominant plural markers, while more recent studies (Tagliamonte et al. 1997, Deuber 2005, Ogunmodimu 2014) find *-s* to be the dominant plural marker with a rare occurrence of *dem* in the data, indicates that there is a change going on in the language. This apparent time study has shown that the younger generation of speakers and the female speakers are the leaders of this change, and the older male speakers are slower to adopt this change, which conforms with results from previous sociolinguistic studies (Labov 2001, Tagliamonte 2012). This study is therefore a good indication of patterning for change in progress in the plural marking in NPE, at least for middle-class educated individuals. The change could also be a result of familiarity with English or due to the prestige of English which is the official language in both Nigeria and in Canada, which are the countries where the participants used to reside and currently reside respectively. Further studies could rigorously investigate education and geographical area (specifically Port Harcourt, Nigeria) as factors that could affect the use of plural markers in NPE.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

DEM	demonstrative determiner	NO	no determiner
		NPE	Nigerian Pidgin English
GENR	generic	NP	noun phrase
NON-NUM/QUANT	non-numeric determiner/quantifier	REDUP	reduplication

Acknowledgements

Appendix A Sociolinguistics interview questions (NPE and English)

A.1 Nigerian Pidgin English

- Abeg tel mi somtin abaut ur bakgroun?

  1. Hu bi ur papa an mama, granpapa an granmama?

  2. Wia dem (bin) from?

  3. Wia dem de naw an wetin dem de du?

- Abeg tel mi abaut haw yu take gro.

  1. Wia yu bin de, gro op?

  2. Haw e bin de wen yu de di ples? Yu get memories wey sti de fresh fo yur main yu fo laik sher?

  3. Wia u bin de go skul? tel mi abaut di skul, di students, di tichas.

  4. Haw yu bin de tek go di skul? Haw di skul bin de? Which kain subject dem bin de tich dia?

  5. Haw meni klas wey yu finish dia, or yu bin cheng skul?

  6. Tel mi abaut di frend dem wey yu bin get dat taim? How dem bin de?

  7. Wetin yu bin de like du fo fun wit yur frend dem?, laik d kain ple una bin de laik?

8. Wich religion una de praktis? Wich choch or mosk una de go, if eni de?

- Wich language yu bin de spik wen yu de grow op?

  1. Wich language dem wey pipu araund yu bin de spik?

  2. Wich language dem yu de spik naw?,

- (For people who indicate they speak more than one language)

  1. If yu de spik pas wan language, wich wan yu sabi speak pas?

  2. Yu de use difrent language fo diferent tins?

  3. You de mix di language wen you de tok somtaims?

  4. Giv ezampul of haw yu de tek mix d language dem?

  5. Yu de eva mix am wit persn wey sabi spik ol of dem?

- Mek we tok abaut yur experience fo Winnipeg

  1. Haw e tek difrent from haus (Nigeria)

  2. Tel mi abaut yur fest winta espiriens

  3. Haw yu si skul hia an skul fo haus? Wich wan yu laik pas? Wai? (for students)

  4. Wetin yu mis pas fo Nigeria

  5. Wetin yu fo laik go bak haus go du?

  6. Wat of yur frends fo haus, yu de mis dem?

  7. Yur papa, mama an yur family membas dem, wetin yu mis abaut dem?

- You feel se awa traditions dem de impotant?

- Wetin mek yu tink laik dat?

  1. Yu fit tel me abaut di wans yu sabi?

- (for people with children)

  1. Wia yur chudren (grandchudren) de/bin go skul?

  2. Yu fit tel me tins abaut di skuls?

  3. Wetin dem tich dem fo di skul?

  4. Yu fil se dem de teach dem wel?

5. Wich language dem de spik or len for di skul?

- (if the children are still very young)

    1. wetin yu hope se dem go bicom?

    2. Wetin bi d difrens bitwin di laif yu bin get nd di wans wey yur chudren get naw?

    3. Wich kain values and sense wey yu de tich yur chudren?

    4. E get eni oda tin yu fo laik gist me?

A.2 English
- Please tell us something about your background.

    1. Who are your parents and grandparents?

    2. Where did they come from?

    3. What brought them to Winnipeg?

- Please tell us something about how you grew up.

    1. Where did you grow up?

    2. What was it like there? Do you have any memories?

    3. Where did you go to school? Tell us something about the school, the students, the teachers.

    4. How did you get to school? What did the schoolday look like? What subjects were you taught?

    5. Which grades did you complete?

    6. Who were your friends? What were they like?

    7. What kinds of things did you do for entertainment, such as hobbies or sports?

    8. What church or religious group did you and your family attend, if any?

- Which languages did you grow up with?

    1. Which languages did the people around you speak?

    2. Which languages do you speak now?

- (For people who indicate they speak more than one language)

1. If you speak more than one language, which one do you feel you speak best?
2. Do you use different languages in different circumstances or with different people?
3. Do you ever mix up your languages?
4. Do you ever accidentally use a word from the wrong language?
5. Do you ever use both languages in a conversation with someone who is fluent in both?

- Let us talk about your experience in Winnipeg.

    1. How does it differ from home (Nigeria)?
    2. Tell me about your first winter experience
    3. How do find schooling here compared to schooling in Nigeria? (for students)
    4. What do you miss most about/in Nigeria?
    5. What would you love to go back to Nigeria to do if you have the opportunity?
    6. Tell me about your friends back home. Do you miss them?
    7. What about people in your family (if they are still in Nigeria) do you miss them? If yes, what do you miss most about being with them?

- Do you feel that tradition is important? Why or why not?

    1. Can you tell me about the ones you know?

- (for people with children)

    1. Where do/did your children go to school?
    2. Can you tell us something about the school(s)?
    3. What do/did they teach them?
    4. Do they get the programming they need?
    5. Which languages are/were they learning?

- (if the children are still not adults) What do you hope for your children's future?

1. What is different between your children's lives and your life as a child?

2. What values and knowledge do you teach your children?

3. Is there anything else that you would like to tell us?

Appendix B Information about the individual participants

Table 5: More information about individual participants. "age$_M$": Age at migration.

Speaker	State of origin	Born and raised	age	age$_M$	gender	languages spoken	
						other	dominant
01f2001	Imo	Port Harcourt	18	17	F	Igbo	English
02f2001	Osun	Lagos	18	17	F	Yoruba	English
19f2000	Edo	Binin	19	18	F	Bini	NPE
05f1994	Oyo	Oyo	25	21	F	Yoruba	Yoruba
20f1985	Abia	Edo	34	32	F	Igbo	NPE
06m2001	Edo	Warri	18	17	M	Urhobo	NPE
03m2000	Anambra	Lagos	19	18	M	Igbo, Yoruba	English
07m2000	Gombe	Gombe	19	18	M	Hausa	Hausa
08m2000	Edo	Binin	19	18	M	Bini	NPE
04m1999	Ogun	Ogun	20	19	M	Yoruba	Yoruba
14f1971	Lagos	Lagos	48	40	F	Yoruba	Yoruba
18f1969	Lagos	Lagos	50	36	F	Yoruba	Yoruba
15f1968	Ondo	Ogun	51	41	F	Yoruba	Yoruba
16f1965	Edo	Benin	54	50	F	Yoruba	NPE
17f1965	Ogun	Lagos	54	48	F	Yoruba, Bini	English
11m1974	Lagos	Lagos	45	32	M	Yoruba	Yoruba
09m1968	Ogun	Abeokuta	51	39	M	Yoruba	Yoruba
13m1968	Osun	Lagos	51	45	M	Yoruba	Yoruba
10m1967	Ogun	Abeokuta	52	40	M	Yoruba	Yoruba
12m1963	Ekiti	Lagos	56	49	M	Yoruba	Yoruba

References

Affia, Precious. 2022. Nigerian Pidgin English: The identity of a Nigerian away from home. Presentation at the 2nd Conference for Graduate Students in Language, Communication, and Culture, University of Alberta.

Agbo, Ogechi Florence & Ingo Plag. 2020. The relationship of Nigerian English and Nigerian Pidgin in Nigeria: Evidence from copula constructions in ICE-Nigeria. *Journal of Language Contact* 13(2). 351–388.

Agheyisi, Rebecca N. 1971. *West African Pidgin English: Simplification and simplicity*. Stanford University. (Doctoral dissertation).

Akande, Akinmade T. & Lasisi Oladipo Salami. 2010. Use and attitudes towards Nigerian Pidgin English among Nigerian university students. In Robert McColl Millar (ed.), *Marginal dialects: Scotland, Ireland and Beyond*, 70–89. Aberdeen: Forum for Research on the Languages of Scotland & Ireland.

Barbag-Stoll, Anna. 1983. *Social and linguistic history of Nigerian Pidgin*. Tübingen: Stauffenberg Verlag.

BBC. 2016. Pidgin—West African Lingua Franca. *BBC*. https://www.bbc.com/news/world-africa-38000387 (16 November, 2016).

Deuber, Dagmar. 2005. *Nigerian Pidgin in Lagos: Language contact, variation and change in an African urban setting*. Colombo: Battlebridge Publications.

Deuber, Dagmar. 2006. Aspects of variation in educated Nigerian Pidgin. In Ana Deumert & Stephanie Durrleman (eds.), *Structure and variation in language contact* (Creole Language Library 29), 243–261. Amsterdam: John Benjamins.

Elugbe, Ben Ohiọmamhẹ & Augusta Phil Omamor. 1991. *Nigerian Pidgins: Background and prospect*. Ibadan: Heinemann Educational Books Nigeria PLC.

Faraclas, Nicholas G. 1989. *A Grammar of Nigerian Pidgin*. Berkeley: University of California.

Faraclas, Nicholas G. 2021. Naija: A language of the future. In Akande T. Akinmade & Oladipo Salami (eds.), *Current trends in Nigerian Pidgin English: A sociolinguistic perspective*, 9–38. Boston/Berlin: De Gruyter.

Igboanusi, Herbert. 2008. Empowering Nigerian Pidgin: A challenge for status planning? *World Englishes* 27. 68–82.

Ihemere, Kelechukwu Uchechukwu. 2006. An integrated approach to the study of language attitudes and change in Nigeria: The case of the Ikwerre of Port Harcourt City. In Olaoba F. Arasanyin & Michael A. Pemberton (eds.), *Proceedings of the 36th Annual conference on African linguistics*, 194–207. Somerville, MA: Cascadilla Press.

Kristian, Sarah. 2018. Social aspiration and traditional speech features among rural Newfoundland youth. In Elizabeth Kelley Seale & Christine Mallinson (eds.), *Rural voices: Language, identity, and social change across place*, 143–164. Lanham: Lexington.

Labov, William. 1963. The social motivation of a sound change. *Word* 19(3). 273–309.

Labov, William. 1966. *The social stratification of English in New York City*. Cambridge: Cambridge University Press.

Labov, William. 2001. *Principles of linguistic change: Social factors*, vol. 2. Malden, MA: Wiley-Blackwell.

Mafeni, Bernard. 1971. Nigerian Pidgin. In John Spencer (ed.), *The English language in West Africa*, 95–112. London: Longman.

Nagy, Naomi & Miriam Meyerhoff. 2015. Extending ELAN into variationist sociolinguistics. *Linguistics Vanguard* 1(1). 271–281.

Ogunmodimu, Morakinyo. 2014. Plural marking strategies in Nigerian Pidgin English (Naija). *Fleur de Ling: Tulane University Working Papers* 1(1). 143–160.

Ogunmodimu, Morakinyo. 2015. Language policy in Nigeria: Problems, prospects and perspectives. *Journal of Humanities and Social Science* 5(9). 154–160.

Tagliamonte, Sali. 2012. *Variationist sociolinguistics: Change, observation, interpretation*. Malden/Oxford: John Wiley & Sons.

Tagliamonte, Sali, Shana Poplack & Ejike Eze. 1997. Plural marking patterns in Nigerian Pidgin English. *Journal of Pidgin and Creole Languages* 12(1). 103–129.

Trudgill, Peter. 1972. Sex, covert prestige, and linguistic change in the urban British English of Norwich. *Language in Society* 1. 175–195.

Chapter 4

The structure and function of Dangme adverbs and adverbials

Yvette Djabakie Asamoah
Ghana Institute of Languages

Richard Ayertey Lawer
Ho Technical University, Ghana

The primary objective of this paper is to unearth the details of what constitutes adverbs in Dangme using the prototype theory. The paper also sheds light on the morphological and structural properties of Dangme adverbs and adverbials. Analyses of the various structures used revealed that Dangme expresses adverbial concepts through the use of core adverbs and quite a large number of ideophonic and reduplicated elements. The morphology and lexical semantics of these three elements place them as integral members that make up the class. It was found that contrary to other Kwa languages, Dangme does not derive adverbs by affixation, but rather by the process of reduplication. It was also observed that Dangme uses adverbials which include nominals with temporal and locative functions, postpositional phrases, and compounded adverbials. Regarding their syntactic properties, Dangme adverbs were discovered to have shared properties; however, for their adverbial counterparts, their structural and distributional properties are influenced mainly by the semantic interpretation they add to the sentences they occur in and the constituents that they scope over.

1 Introduction

Although a great wealth of studies on African languages has targeted word classes, adverbs have not received as much attention as other lexical categories. Nevertheless, we cannot overlook the substantive findings of the few works on adverbs which show them to have intriguing structural properties (Lusekelo 2010,

Sakyi 2013, Saah & Agbedor 2004, Tabe 2015). In Kwa languages, for instance, it has been revealed that adverbs have unique morphology and distributional qualities. Most of the Kwa languages have core adverbs; they are single-word, monomorphemic lexemes. A lot of the adverbs are very adjective-like in nature, with a great number of them often used both adjectivally and adverbially without any morphological markings (cf. Awuku 2011; Otoo 2014; Saah 2004). In terms of derivation of adverbs, these languages differ considerably. Ewe, for instance, derives adverbs from nouns with the suffix *-tɔe* (cf. Ameka 1991: 55; Dzameshie 1998: ex. 3). Ewe also uses nominalized phrasal verbs to express adverbial functions by attaching the suffixes *-e* or *-i* to the base (Dzameshie 1998, Saah & Agbedor 2004). Meanwhile, Akan presents a different case. It has no morphological means of deriving adverbs from other lexical categories. Aside from the use of a few simple lexical adverbs such as *bɔkɔɔ* 'softly,' *basaa* 'haphazardly,' the language uses ideophones and postpositional phrases as adverbs (Saah 2004: 52).

However, there has not been much discussion about Dangme adverbs. [1] The two works that present some information about Dangme adverbs do not shed light on the nature of this lexical category. Kropp Dakubu's (1987) introductory survey of the language emphasizes that indeed Dangme has adverbs that express aspect, but does not give a thorough analysis about the nature of the adverbs. The study by Abakah et al. (2010) looked at reduplication in Akan, Dangme and Gurene and to some extent discussed the formation of adverbs through the process of reduplication. The prime objectives for this study are two-fold: to discover the various structures used to express adverbial concepts in Dangme and to examine the morphology of Dangme adverbs.

2 Fundamental issues on adverbs

Generally, adverbs are classified from two major perspectives: the function of modifying verbs and propositions, and the morphosyntactic properties that mark the class. Linguistic scholars of the first approach hold that adverbs should be

---

[1]Dangme is a Kwa language of the Niger-Congo language family spoken in Ghana in some parts of the Eastern, Greater Accra and Volta regions. The language has two major dialects, namely coastal and inland dialects (Kropp Dakubu 1988: 94). Dangme shares borders with Ga, Hill Guang (Ɔkere and Lɛtɛ), Akan and Ewe. It is a register tone language with three contrasting basic tones, high, low and mid, which have lexical as well as grammatical functions. The Dangme verbal system is aorist in nature and therefore does not make any distinction regarding the time frames of events. However, the language uses aspectual markers and adverbial expressions to depict grammatical properties such as habitual, progressive and future events (Kropp Dakubu 1988).

identified and defined in terms of their function because every structure has at least a semantic function and can therefore only be understood and explained on the basis of its discourse and function (Curme 1935, DeLancey 2001, Haspelmath 2001, Trask 1993). Therefore, among these linguists, adverbs are usually considered as grammatical adjuncts of verbs or propositions which express semantic notions such as time, manner, place, instrument or circumstance (Hengeveld 1992, Paolo & Ricca 1994). They are said to play a prevalent role in the modification of almost all other lexical categories as well as the clause itself. However, not all these "optional modifying phrases" of verbs and propositions are adverbs. Many of them exhibit varying structures with entirely different morphosyntactic properties. Therefore, for some other linguists, adverbs should be defined and classified on the basis of their structural properties, syntactic behaviour and distribution of the class (Cinque 1999, 2004, Ernst 2001, Van Valin & LaPolla 1997). According to them, even though it is function that provides the primary defining characteristic of adverbs, the easiest way to an initial understanding of the class is usually through their structure and lexical morphology (Huddleston 1988: 31). Thus, within this approach, an adverb is identified as adjoining elements of a verb or a proposition. They are regarded as "syntactically dispensable" lexemes structurally classified into lexical, derived adverbs and other complex constructions with corresponding meanings and function of adverbs (Geuder 2002; Paolo & Ricca 1994: 290). For its morphological structure, the category in several languages consists of simple as well as derived structures formed through the processes of compounding, affixation and all forms of reduplication. Lefebvre & Brousseau (2002) for example, in their description of Fongbe adverbs found them to be "an eclectic class with a few monosyllabic adverbs" as well as "…morphologically-complex adverbs, reduplicated forms, unanalysable bisyllabic forms, and frozen phrases, all of which can be used adverbially." This generalization about Fongbe adverbs, (a language spoken in Benin, Nigeria, Togo, Ghana and Gabon) is a very important premise to this study because it gives us a glimpse of the likely structures that may be encountered. However, there is a considerable degree of overlap between the two approaches as the function and structural properties of adverbs are intertwined. Therefore, many researchers also look at adverbs in terms of their internal structure, morphology and lexical meaning. In this paper, we examine the structure of Dangme adverbs focusing on their nature within the context of their syntactic function.

3 The prototype theory and Dangme adverbs

The Dangme adverb class was identified and classified using prototype theory, which is a functional and highly effective model for word class studies (Givón 2001, Rosch 1978, Taylor 2009). Prototype theory was originally a product of cognitive psychology proposed by Rosch (1978) in her work on the internal structure of categories, but is currently used in various fields. Prototype theory was developed to fill the gaps of the classical and componential analysis models of categorization where features are regarded as essential and binary, so that if an entity does not have the complete set of features, it fails to be a member of the category (Givón 2001, Rosch 1978, Taylor 2009). Thus, entities were strictly categorized in terms of the set of necessary and sufficient features that were associated with the category. This raised many problems as most members of a category tend not to possess all the defining features of the category. In prototype theory, categorization is more of a graded categorization where some members of the class are more central than others (Taylor 2009). Here, it is not a matter of whether an entity possesses all the supposed attributes of a category or not, but to what extent the entity is closely related to other prototypical members. This is not to say that peripheral members are not perfect examples of the group, as they are also representative of a category. For this reason, the theory is used in the classification of lexical classes. This is because these lexical categories often cannot be defined by means of a strict set of criteria as they generally tend to show a family resemblance structure instead (Givón 2001). Therefore, using the prototype theory in this study helps us account for the heterogeneous class the Dangme adverb system exhibits. It also aids our understanding of the central as well as peripheral items that make up the Dangme adverb class.[2]

Croft (2013) asserts that although word classes and other syntactic structures form a significant set of language universals, they tend to be language-specific and construction-specific. Hence, it is necessary to ascertain whether adverbs and adverbials exist in Dangme and the kind of properties that are associated with them. In this paper, we define an adverb in Dangme as any single lexical item that traditionally functions as a modifier of a verb, an adjective, another adverb or a clause. Consider the following examples in (1). In Dangme, adverbs modify verbs, adjectives, other adverbs, or the entire clause. This is shown in (1a), where the Dangme adverb *mɔbɔmɔbɔ* 'sorrowfully' modifies the verb as it gives information about how the action denoted by the verb *lá* 'sing' was performed.

---

[2]This data was obtained from a collection of excerpts from available grammars and other written texts such as the Language Guide and an array of novels and primers published by the Bureau of Ghana Languages.

In (1b), the adverb *sàmínyá* 'really/very' modifies the adjective by showing the level or degree of the adjective *fɛ̀ú* 'beautiful' associated with the noun *jókúɛ́ yòyó* 'girl.' The adverb *nítsɛ̄* 'very' in (1c) modifies the adverb *má* 'early' by giving information about the degree to which the adverb applies. And in (1d), the expression *pɛ̀pɛ̀ɛ̀pɛ̀* 'thoroughly' modifies the entire clause because it depicts the speaker's judgement about an event denoted by the clause.

(1) a. yó-ɔ́   lá   **mɔ̀bɔ̀mɔ̀bɔ̀**.
     woman-DEF sing.PST sorrowfully
     'The woman sang sorrowfully.'

   b. jókúɛ́ yòyó   nɔ̀ hē ngɛ̀ fɛ̀ú   **sàmínyá**.
     child female DEM self be.at beautiful really/very
     'This girl is really/very beautiful.'

   c. wà   títsɛ̄-ɛ̄   bá-á   má **nítsɛ̄**.
     1PL.POSS teacher-DEF come-HAB early very
     'Our teacher comes very early.'

   d. è-péé-ɔ̀   wē   mì   nɔ́ fɛ̀ɛ̀ nɔ́ **pɛ̀pɛ̀ɛ̀pɛ̀**.
     3SG.SBJ-do-HAB house POST everything thoroughly
     'He/she does all his/her house chores thoroughly.'

Thus, a look at the Dangme adverbs in (1) shows that the basic features associated with a prototypical adverb include:

  1. Its function of modifying a verb, an adjective, another adverb or an entire clause.

  2. It must be a monomorphemic, single lexical element

  3. It must have an inherent adverbial meaning. Inherent adverbial meaning here means that these words do not perform any other function like nominal, verbal or adjectival except an adverbial function.

This set of possible core features is what ties together the members of the adverb category in Dangme. Therefore, an adverb which bears all the features outlined above is said to be central or prototypical and the members which often may not possess all the features are known as less central or peripheral members. The Dangme adverbs were found to be associated with other syntactic structures which are different with respect to their form. These structures were classified as adverbials. We define an adverbial in Dangme as any syntactic constituent

that can function as an adverb in a sentence. In the constructions in (2), the expressions in italics also modify a verb, an adjective, another adverb and an entire clause just like adverbs. However, as signaled in the examples, all of them show a different morphosyntactic structure from adverbs as they are larger syntactic units. Therefore, we can conclude that in Dangme, both adverbs and adverbials have the same semantics of denoting manner, time, degree, frequency etc. even though they are different in terms of their forms.

(2) a. yó-ɔ́   lá   **kɛ́ jé mɔ̀tù   ngmlɛ́ nyɔ̀ngmá kɛ̄   káké kɛ̄**
         woman-DEF sing.PST from morning hour   ten   CONJ one CONJ
         bà   sù   píàní   nɔ̀.
         come.PST reach afternoon DEM
         'The woman sang from eleven in the morning till this afternoon.'

     b. jókúɛ́ yòyó   nɔ̀   hē pè fɛ̀ú   **nɛ̄   è-wò**
         child female DEM self do beautiful when 3SG.SBJ-wear.PST
         è-blóónyà   tádè-ɛ́.
         3SG.POSS-Christmas dress-DEF
         'This girl looked beautiful when she wore her Christmas dress.'

     c. wà   títsɛ̄-ɛ̄   bá-á   má **dáá Hɔ̀gbí.**
         1PL.POSS teacher-DEF come-HAB early every Monday
         'Our teacher comes early every Monday.'

     d. è-péé-ɔ̀   wē   mì   nɔ́ fɛ̀ɛ̀ nɔ́.
         3SG.SBJ-do-HAB house POST everything
         lókó è-yà-á   súkúú.
         before 3SG.SBJ-go-HAB school
         'He/she does all his/her house chores before he/she goes to school.'

4 The structure of Dangme adverbs

This section elaborates on the structure of the Dangme adverbs. The discussion encompasses core adverbs, ideophonic adverbs and reduplicated adverbs.

4.1 Core adverbs

Dangme has a number of core adverbs. These are prototypical adverbs which consist of single-word items. They have an inherent adverbial function. Structurally, core adverbs are not derived from other word categories. In Table 1 below is an inventory of some of these core adverbs in Dangme.

Table 1: Core adverbs

Adverb	Gloss	Adverb	Gloss
má	'early'	ékóhú	'again'
pám	'suddenly'	lōlō	'yet'
blɛ̀ūū	'slowly'	kpàmìsáá	'often'
píɔ̀ɔ̀	'now'	ékómē̄	'perhaps'
tɔ̀ɔ̀	'far'	lókóó	'very far'
kpóó	'calmly/quietly'	tsítsà	'far'
hlūū	'a long time'	tũ̄ũ	'darkly'
dáá	'every time/always'	tii	'straightly/directly'
mómó	'already'	dĩ̄	'silently'
pɛ́	'indeed/exactly/just'		

As shown in Table 1, the core adverbs are strictly distinguished by their non-derived quality. They are monomorphemic, not formed through the concatenation of morphemes. Their meanings are not aggregations of the meanings of constituent parts. In terms of their structural distribution, most of these adverbs are verb-modifying adverbs that scope over the verb phrases in sentences. The core adverbs are syntactically restricted in terms of their position in sentences. They cannot occur in either pre-clause position or pre-verbal position. They always occur after the verb phrase in sentences as shown in (3) below.

(3) a. hìɛ́   pũ-ɔ̀   jè   kpò **má**.
       Yesterday sun-DEF leave.PST out early
       'Yesterday, the sun rose early.'

    b. Uede té   sī   **pám**   nē̄   è-yà   tsɔ́mì-ɛ́.
       Dede rise.PST GROUND[3] suddenly CONJ 3SG.SBJ-go.PST errand-DEF
       'Dede stood up suddenly and went for the errand.'

    c. bímɔ̀yó-ɔ̀ fó yā   **hlúú**.
       Baby-DET cry.PST a.very.long.time
       'The baby cried for a long time.'

    d. * Dede **pám**   té   sī   nē̄   è-yà   tsɔ́mì-ɛ́.
       Dede suddenly rise.PST GROUND CONJ 3SG.SBJ-go.PST errand-DEF
       ('Dede stood up suddenly and went for the errand.')

e. * bímɔ̀yó-ɔ́ **hlúú** fó yā.
baby-DET a.very.long.time cry.PST
('The baby cried for a long time.')

In the examples (3d) and (3e), we observe that the sentences have become ungrammatical because we have changed the position of the core adverbs from a post-verbal to a pre-verbal position.

4.2 Reduplicated adverbs

The Dangme adverb class[4] also has derived adverbs which are created through the process of reduplication of adjectives and nouns. For instance, the adverbs *jéhājéhā* 'yearly' and *ésɔ̀ésɔ̀* 'quickly' are formed from the repetition of the lexical forms *jéhā* 'year' and *ésɔ̀* 'quick.'

4.2.1 Adverbs derived from reduplication of adjectives

Most derived adverbs in Dangme are formed from the reduplication of adjectives. Here, there is some form of repetition of the base or part of it to form another word. Consider the following examples.

(4) a. jókúέ yòyó nɔ̀ hē ngē̄ **ésɔ̀** pè.
child female DEM self be.at quick very/really
'This girl is very quick/fast.'

b. nyὲ pò gbὲjὲgbὲ-ὲ mī **ésɔ̀ésɔ̀**.
2PL.SBJ cut.PST street-DEF POST quickly
'You should cross the street quickly.'

In (4) above, the adjective *ésɔ̀* 'quick' is reduplicated to form the adverb *ésɔ̀ésɔ̀* 'quickly.' Many other examples of adverbs formed through reduplication of adjectives are outlined in Table 2 below.

As shown in Table 2, the adjectives in the base are repeated twice to form adverbs. Adjective bases such as *bàsàà* 'haphazard' and *ésɔ̀* 'quick' form the adverbs *bàsàbàsà* 'haphazardly' and *ésɔ̀ésɔ̀* 'quickly' when reduplicated.

---

[3]The word *si* is a "ground" which represents an abstract locative site or orientation from the natural state of rest where you are unconscious to the state where you are awake and conscious and aware of what surrounds you. The ground marker often co-occurs with verbs of movement which normally have a located site or position and a locating entity.

[4]Dangme adverb class used here in this paper refers to both adverbs and adverbials.

Table 2: Reduplicated Adverbs formed from Adjectives

Base form (Adj)	Gloss	Reduplicated form (Adv)	Gloss
mɔ̀bɔ̀	'sad'	mɔ̀bɔ̀mɔ̀bɔ̀	'sorrowfully'
kùnyàà	'extreme'	kùnyàkùnyà	'extremely'
bɔ̀ɔ́	'small'	bɔ̀ɔ́bɔ̀ɔ́	'gradually'
fálíí	'neat'	fálífálí	'neatly'
nyɔ́ngɔ́	'small'	nyɔ́ngɔ́nyɔ́ngɔ́ɔ́	'in small pieces'
àgbò	'big'	àgbò àgbò	'in big pieces'
tsétsɛ́	'smart'	tsétsɛ́ɛ́tsɛ́	'smartly'
bàsàà	'haphazard'	bàsàbàsà	'haphazardly'
kéklé	'first'	kéklééklé	'firstly'

4.2.2 Adverbs derived from reduplication of nouns

In the same way, some Dangme adverbs are formed from the reduplication of nouns. In the creation of these adverbs, the nouns undergo total reduplication to form the adverbs as demonstrated in the constructions in (5).

(5) a. **gbɔ̀kúɛ̀** màá bā   píɔ̀ sɔ̀ nɔ̄.
       evening FUT come right.now
       'Evening will come right now.'

   b. Dédé bá-á   híɔ̀ **gbɔ̀kúɛ̀-gbɔ̀kúɛ̀**.
       Dédé come-HAB here every.evening
       'Dede comes here every evening.'

   c. à-jé-ɔ́   **jéhā** éhé sīsī   ngɛ́.
       3SG.SUBJ-begin-HAB year new under be.at
       Jɔ̀né   kéklé līgbī-ɔ̄   nɔ̄.
       January first   day-DET POST
       'A new year begins on the 1st of January.'

   d. Klò-lí yé-ɔ́   Ngmāyēmī[5] **jéhā-jéhā**.
       Klo-PL eat-HAB Ngmāyēmī   yearly/annually
       'Krobos celebrate Ngmāyēmī yearly/annually.'

In the above example, the adverbs *gbɔ̀kùɛ̀ gbɔ̀kùɛ̀* 'every evening' in (5b) and *jéhā jéhā* 'yearly' in (5d) are formed by reduplicating the nouns *gbɔ̀kùɛ̀* 'evening'

---

[5]Ngmāyēmī festival is a harvest festival celebrated by the people of Manya Krobo.

and *jéhā* 'year' in (5a) and (5c) respectively. More examples of reduplicated adverbs formed from nouns are listed below in Table 3.

Table 3: Adverbs formed from reduplicated Nouns

Base form (Noun)	Gloss	Reduplicated form (Adv)	Gloss
ótsí	'week'	ótsíótsí	'weekly'
píàní	'afternoon'	píànípíàní	'every afternoon'
gbéyè	'fear'	gbéyègbéyè	'fearfully'
bè	'time'	bèbèèbè	'already'
nyɔ̀	'night'	nyɔ̀nyɔ̀ɔ̀nyɔ̀	'every night'
dákā	'box'	dákādákā	'in boxes'
kákē	'one'	kákááká	'one each'
ényɔ̀	'two'	ényɔ̀ényɔ̀	'two each'
nyũ	'water'	nyũnyũ	'watery'
tɛ́	'stone'	tɛ́tɛ́	'stony'/'lumpy'
ngò	'salt'	ngòngò	'salty'

As outlined in Table 3, the single nominal bases *píàní* 'afternoon,' *bè* 'time' and *dákā* 'box' are repeated to derive the adverbs *píànípíàní* 'every afternoon,' *bèbèèbè* 'already' and *dákādákā* 'in boxes' respectively. However, the final three reduplicated forms resulting in *nyũnyũ* 'watery,' *tɛ́tɛ́* 'stony/lumpy' and *ngõngõ* 'salty' become adjectives when reduplicated. These adjectives are used as adverbs with the only change being the constituents they modify in sentences. For instance, a look at the sentences in (6) below shows the use of an adjective adverbially. In (6a), the adjective *tɛ́tɛ́* is a modifier of the noun *ma ku* 'banku' and gives a descriptive nature about the noun. However, in (6b), *tɛ́tɛ́* is used as an adverb to describe how the noun *ma ku* 'banku' was made. Thus, it can be seen that Dangme, like Akan and Ga (Kwa languages) does not show any strict morphological distinction between adjectives and adverbs. To differentiate between the two cases, one must consider the meanings they add to sentences as well as the type of structures they tend to modify.

(6)   a.   Nàkó há   mī   má kú **tɛ́tɛ́**   kō.
           Nako give.PST 1SG.OBJ banku lumpy INDEF
           'Nako gave me some lumpy banku.'

b. Nàkó tsi   má kú-ɔ̀   **tɛ́tɛ́**.
Nako make.PST banku-DEF lumpy
'Nako made the banku lumpy.'

4.3 Ideophonic adverbs

Some Dangme adverbs are ideophonic in nature. This attribute of adverbs in Dangme has also been found in many other Kwa languages (cf. Bodomo 2006, Saah 2004, Dzameshie 1998). In Dangme, ideophonic expressions are usually used to describe the distinctive movement, sound and nature associated with the action denoted by the verbs in sentences. Examples of ideophonic adverbs in Dangme are given below in Table 4.

Table 4: Ideophonic adverbs

Adverb	Gloss
pàtàpàtà	'restlessly'
kúmákúmá	'eagerly'
tsílɔ́tsílɔ́	'in bits/bit by bit'
lílíílí	'grumpily'
glɔ́glɔ́	'to drink swiftly'
hwéhwééhwé	'safely'
ngmángmángmá	'sourly'
mɛ́mɛ́ɛ́mɛ́	'devotedly'
tutúútú	'exactly'
vévéévé	'meticulously'
wìkìwìkì	'violently'

The above-mentioned items are all ideophonic adverbs in Dangme and usually have multiple syllables. Even though they seem like reduplicated forms, their semantics tell the opposite. In cases such as (7) and (8), we see a striking difference between these ideophonic adverbs and other reduplicated forms.

(7) a. gbé bí-ɛ̄   húà   **pàtàpàtà**.
     dog child-DET struggle.PST restlessly
     'The puppy struggled restlessly.'

     b. wà   há   nyūmū-ɔ̄ nyù nɛ̀ è-nù   **kúmákúmá**.
     PL.SUBJ give.PST man-DEF water CONJ 3SG.SUBJ-drink eagerly
     'We offered the man water and he drank it eagerly.'

As can be observed in the sentences in (7) above, *pàtàpàtà* 'restlessly' and *kúmákúmá* 'eagerly' are ideophonic adverbs used to modify the verb phrases in the two sentences. As with reduplicated forms, the initial bases of these adverbs appear to have been reduplicated. However, they cannot be said to have undergone any form of reduplication. This is because there are no corresponding unreduplicated bases. Thus, ideophonic adverbs do not display meanings that have meaning components that can be traced to the supposed base. Though the words *pàtàpàtà* 'restlessly' and *kúmákúmá* 'eagerly' seem like reduplicated forms derived from the initial bases *pàtà* and *kúmá,* they are not. This is because the supposed bases are not meaningful words in the language. Now, let's consider the examples in (8):

(8)   a.   Darley màmí   há   mī   òtìm   **bɔ̀dɔ̀ɔ̀** kō.
            Darley mother give.PST 1SG.OBJ kenkey soft   INDEF
            'Darley's mother gave me a soft kenkey.'

      b.   Darley tsí   má   kú-ɔ́[6] **bɔ̀dɔ̀bɔ̀dɔ̀**.
            Darley make.PST banku-DEF soft
            'Darley made the banku food very soft.'

      c.   míní nɛ́   ò-péé   **dɔ̀léé** mwɔ̀nɛ̀ɔ̀.
            why that 2SG.SUBJ-do dull   today
            'Why are you dull today?'

      d.   Dèdé kàné   blɛ̀fótā-ɔ̀-mɛ̄   **dɔ̀lédɔ̀lé**.
            Dédé count.PST pineapple-DEF-PL lazily
            'Dede lazily counted the pineapples.'

The sentences in (8) show a clear case of derived adverbs which have undergone reduplication. The initial bases: *bɔ̀dɔ̀ɔ̀* 'soft' and *dɔ̀léé* 'lazy/dull' are repeated to form the adverbs *bɔ̀dɔ̀bɔ̀dɔ̀* 'softly' and *dɔ̀lédɔ̀lé* 'lazily' in (8b) and (8c) which are meaningful words in the language. Again, the resultant reduplicated forms depict meanings different from the meanings expressed by the single bases from which they were derived. Aside the ideophonic adverbs with multisyllabic structures, Dangme also has ideophonic adverbs which are monosyllabic. This kind of adverb tends to indicate how an action is performed by imitating the sound, movement and/or other qualities associated with the action denoted by the verb phrase. In observing Akan ideophonic manner adverbs, Saah (2004: 54) refers to them as onomatopoeic expressions which describe the verb phrase with respect

---

[6]*Má kú* is the Dangme name for banku, a Ghanaian dish prepared from fermented corn and cassava dough.

to manner, colour, smell, action, state or intensity. In the examples below, we provide some of the Dangme ideophonic adverbs that evoke the sense of manner.

(9) a. ngmɛ̄-ɛ̄   nɔ̃   sī   **kùm**.
         nut-DEF fall.PST GROUND IDEO
         'The coconut fell *kùm*.'

     b. páyà-à   nɔ̃   sī   **tìm**.
         pear-DEF fall.PST GROUND IDEO
         'The pear fell *tìm*.'

     c. yò-ɔ̀   nɔ̃   sī   **tòó**.
         woman-DEF fall.PST GROUND IDEO
         'The woman fell *tòó*.'

     d. nyūmū-ɔ̄ nù   nyū-ɔ̄   **mìì**.
         man-DEF drink.PST water-DEF IDEO
         'The man drank the water *mii*.'

     e. kɔ̀hìɔ̀-ɔ̀   fìà   **hùù**.
         wind-DEF blow.PST IDEO
         'The wind blew *hùù*.'

     f. tsō-ɔ̄   kũ   **kɛ́**.
         stick-DEF break.PST IDEO
         'The stick broke *kɛ́*.'

     g. lā-ā   tsò   **bòm**.
         fire-DEF burn.PST IDEO
         'The fire burnt *bòm*.'

     h. tsō-ɔ̄   kũ   **kùá**.
         tree-DEF break.PST IDEO
         'The tree broke *kùá*.'

     i. krámáá   pɛ̀   **bɔ̃́ɔ̃́**.
         machine-DEF sound.PST IDEO
         'The machine sounded *bɔ̃ɔ̃*.'

From the sentences in example (9), we observe that the ideophonic adverbs occur after the verbs in their respective sentences. These adverbs show the manner in which the actions denoted the verbs in the various sentences are carried out. For instance, the ideophones *hùù* and *bɔ̃ɔ̃* indicate the manner in which the wind blew and the machine sounded in their respective sentences in (9e) and (9i).

5 The structure of Dangme adverbials

In addition to adverbs, Dangme has other ways it expresses the adverbial meanings. The language uses adverbials which are any syntactic constituent that can function as an adverb in a canonical sentence and can modify the predicate or the proposition. These adverbials can be sub-grouped into two: adverbials with temporal functions and adverbials with locative functions.

5.1 Dangme adverbials with temporal functions

Many of the adverbials in Dangme have temporal meaning, such as those in (10).

(10) a. yòmóyó-ɔ̀   gbó   **Hɔ̀gbì màkɛ̃-ɛ̀.**
          old.woman-DEF die.PST Monday dawn-DEF
          'The old woman died on Monday dawn.'

      b. è-bā   híɔ̃ **līgmī Pɛ̀plɛ̀gbì-ɛ.**
          3SG.SBJ-come.PST here last   Tuesday-DEF
          'He/she came here last Tuesday.'

      c. zūgbãtsɛ̄-ɛ̄   màá bā híɔ̃ **Màjá   nyàgbè.**
          land owner-DEF FUT come here December end
          'The land owner will come here at the end of December.'

      d. wà   màá yà híɛ́   sìní **mwɔ́nɔ̄ gbɔ̀kúɛ̀ ngmlɛ̀ kpàànyɔ̀.**
          1PL.SBJ FUT go watch movie today night   hour   eight
          'We will be going to watch movies at 8 o'clock tonight.'

      e. **àmáné bɔ̀   mī sè-ɔ̀**   wà kpàlé   wà   sè bā
          news   narrate POST back-DEF 1PL return.PST 1PL.POSS back come.PST
          wē   mì.
          home POST
          'After delivering the news, we returned home.'

In the constructions in (10), we see that all the expressions in bold have temporal functions which indicate when the actions depicted by the verbs were performed. Dangme adverbials have a wider distribution than adverbs. They can occur at both the sentence-initial and sentence-final positions and can as well be fronted for focus. However, they generate an ill-formed sentence when placed in the pre-verbal position as illustrated in (11d) below.

(11) a. Darley bā   híↄ̀ **līgmī Pɛ̀plɛ̀gbì-ɛ̀**.
          Darley come.PST here last   Tuesday-DEF
          'Darley came here last Tuesday.'

     b. **līgmī Pɛ̀plɛ̀gbì-ɛ**   Darley bā   híↄ̀.
          last   Tuesday-DEF Darley come.PST here
          'Last Tuesday, Darley came here.'

     c. **līgmī Pɛ̀plɛ̀gbì-ɛ̀**   né Darley bā   híↄ̀.
          last   Tuesday-DEF FM Darley come.PST here
          'It was last Tuesday that Darley came here.'

     d. * Darley **līgmī Pɛ̀plɛ̀gbì-ɛ̀**   bā   híↄ̀.
           Darley last   Tuesday-DEF come.PST here

One other syntactic property that is associated with this group of Dangme ad-
verbials is that they are usually headed by nominals (lexical time nouns) such as
jéhā nↄ̀ 'this year' and subordinators like *bènɛ́* 'when' and *lókó* 'before.' In (12a),
blóónyà 'Christmas' expresses its original function as an argument of the verb
'celebrate' However, in (12b) the nominal *blóónyà* 'Christmas' is used as a tempo-
ral adverb to express the time the action specified by the verb will be completed.
It can be observed that even though here the nominal *blóónyà* 'Christmas' takes
up an adverbial function, the form remains the same.

(12) a. wà-yé-ↄ́   **blóónyà** dáá jéhā.
          3PL-eat-HAB Christmas every year
          'We celebrate Christmas every year.'

     b. à-màá   gbè tsū   nↄ̀ nyà **blóónyà**.
          3SG.SBJ-FUT finish building DEM end Christmas
          'They will complete the building in Christmas.'

5.2 Dangme adverbials with locative functions

Speakers of Dangme also make extensive use of adverbials with locative func-
tions. In Dangme, these adverbials are expressed by postpositional phrases and
compounded adverbs. Consider the following Dangme examples in (13):

(13) a. kↄ̀dú   júàlī-hī bá-á   à-kpè-ↄ̀   ngɛ̄.
          banana seller-PL come-HAB 3PL.POSS-meeting-DEF be.at
          Àsɛ̀sɛ̀wá júà   nↄ̄.
          Àsɛ̀sɛ̀wá market POST
          'The banana sellers have their meeting at the Àsɛ̀sɛ̀wá market.'

b. è-nyὲ   nà   kùngwɔ́ bí-ɔ́ ngɛ̄ **tsū-ɔ̄**   **mì**.
   3SG.POSS-mother see.PST chicken-DET be.at room-DET POST
   'His/Her mother saw the chicken in the room.'

c. gòdòtsɛ̄   kō   húɔ́-ɔ́   sī   ngɛ̄ **pósɔ̄fìsì-ὲ**   **sὲ**.
   mad person INDEF sleep-PROG GROUND be.at post office.DEF back
   'A mad person is sleeping behind the post office.'

In (13), the postpositional phrases *Àsὲsὲwá júà nɔ̄* 'at the Àsὲsὲwá market,' *tsūɔ̄ mì* 'in the room' and *pósɔ̄fìsì-ὲ sὲ* 'behind the post office' are used to indicate where the action denoted by the verb took place. We can also infer from the sentences above that these adverbials are usually preceded by the locative verb *ngɛ* 'be at.' *ngɛ* is a locative verb that precedes the postpositional phrases. Together with the postpositional phrases, they denote the position of an entity with respect to a spatial location or the ordinary sense of place. In Dangme, postpositional phrases cannot occur alone to express the location of an entity.

For their distribution, they either occur at the sentence-final or sentence-initial position, but are not likely to surface in the pre-verbal position in a sentence as illustrated by the sentences in (14). A look at the sentences in (14b) and (14c) show that Dangme adverbials with locative functions can freely be fronted for focus with or without a focus marker *ne*. In addition, when these adverbials occur at the sentence-initial position, they come with the topic marker *-ɔ* which is realized differently depending on its preceding sound.

(14) a. Nàkò nà   klàlá dúkù-ɔ̀   ngɛ̄ **òkplɔ́-ɔ́** **sīsī**.
        Nàkò see.PST handkerchief-DEF be.at table-DEF POST
        'Nàkò found the handkerchief under the table.'

   b. ngɛ̄ **òkplɔ́-ɔ́** **sīsī-ɛ̄**   Nàkò nà   klàlá dúkù-ɔ̀.
      be.at table-DEF POST-TOP Nàkò see.PST handkerchief-DEF
      'It was under the table that Nàkò found the handkerchief.'

   c. ngɛ̄ **òkplɔ́-ɔ́** **sīsī-ɛ̄**   nē Nàkò nà   klàlá dúkù-ɔ̀.
      be.at table-DEF POST-TOP FM Nàkò see.PST handkerchief-DEF
      'It was under the table that Nako found the handkerchief.'

   d. *nàkò ngɛ̄ **òkplɔ́-ɔ́** **sīsī** è-nà   klàlá dúkù-ɔ̀.
       Nàkò be.at table-DEF POST 3SG.SBJ-see.PST handkerchief-DEF

5.2.1 Compounded adverbials

There are locative adverbials that are formed through the process of compounding. This category of adverbials usually consists of a noun and a post postposition

(N+P) which denotes spatio-temporal and locational relations. Usually, the noun occurs as the left-hand constituent, while the postposition occurs as the right-hand constituent. The sentences in (15) below illustrate the adverbial functions of the compounded adverbials.

(15) a. pláfō-hī   fúú là-á   à-hē   ngɛ̄ **yōkú** **nɔ̄**.
     executioner-ᴘʟ many hide-ʜᴀʙ 3ᴘʟ.ᴘᴏss-body be.at mountain top
     'Many executioners hide themselves on mountain tops.'

   b. pànɔ̄yàlī-hī nɛ́   mā   ngɛ̄ **pàhém**.
     fisherman-ᴘʟ see.ɴᴇɢ herring be.at river.face
     'Fishermen do not find herrings on the surface of the river.'

   c. yò   nɔ̀kɔ́tɔ́má dú   wē ngɛ̄ **kpō** **nɔ̄**.
     woman adult   bathe ɴᴇɢ be.at outside top
     'Elderly women do not bathe on the compound.'

   d. *yò   nɔ̀kɔ́tɔ́má dú   wē ngɛ̄ **kpō**.
     woman adult   bathe ɴᴇɢ be.at outside

The example above illustrates instances where Noun-Postposition compounds are used to perform an adverbial function. Based on the data, we realize that this group of derived adverbs is used to express the location or physical space where the action was performed. An interesting observation about this group of derived adverbs is that their structure seems to have some similarity with that of postpositional phrases. Like Noun-Postposition compounds, postpositional phrases also consist of nouns and postpositions. The difference between the two is that the (N+P) compounds are lexicalized and have unpredictable meanings, whereas postpositional phrases have compositional meanings. Neither the word *kpō* 'outside' nor *nɔ̄* 'top' mean 'compound' on their own; this meaning only arises when the two words are compounded. This is demonstrated in (15c) and (15d). As shown in (16), Noun-Postposition compounds require both parts to co-occur before generating meaning whereas items in postpositional phrases are not required to do so. Oftentimes, the noun or noun phrase of the postpositional phrase may refer to the entity as illustrated in the example below. In this case, *yoku* and *pa* still refer to 'mountain' and 'river' even without postpositions.

(16) a. pláfō-hī   fúú làá   à-hē   ngɛ̄ **yōkú-ɔ́**
     executioner-ᴘʟ many hide.ᴘsᴛ 3ᴘʟ.ᴘᴏss-body be.at mountain-ᴅᴇꜰ
     sē.
     back
     'Many executioners hid themselves behind the mountain.'

b. mā   fúú   ngɛ̄ **pà-à**   **mì**.
    herring many be.at river-DEF inside
    'There are many herrings in the river.'

Further examples of these adverbials are presented in Table 5 below. From Table 5, we realise that the compound words are nouns, but they function in sentences as adverbials.

Table 5: Adverbials formed through compounding

Base 1	Gloss	Base 2	Gloss	Compound	Meaning
pà	river	nõ	top	pànõ	surface of the river
pà	river	mì	inside	pàmì	interior of the river
tsō	tree	nõ	top	tsō nõ	treetop
pà	river	sè	back	pàsè	across the river
tsũ	house	mì	inside	tsũ̀mì	room
pà	river	nyã	mouth	pãnyã	river bank
pà	river	túé	ear	pãtùé	riverside
mã	town	nyã	mouth	mã̀nyã̀	suburb
blɔ̄	road	se	back	blɔ̄sè	across the street
wīē	house	nyà	mouth	wīē nyà	front of the house
mà	town	sè	back	màsè	the far side of the town
túé	ear	sè	back	túé sè	behind the ear
tɛ́	stone	nō	top	tɛ́nò	on the rock

6 Conclusion

As the above discussion has demonstrated, Dangme adverbs are a heterogeneous class. They include core adverbs, reduplicated adverbs and ideophonic adverbs. The core adverbs in Dangme are monomorphemic in nature and have an intrinsic adverbial meaning. Most core adverbs in Dangme are verb-modifying adverbs which have scope over the verb phrase. As such, they are very restricted and may not be placed in various positions an adverb may occur in. They only occur after the verb. Ideophonic adverbs form another category of words that belong to the Dangme adverb class. These adverbs consist of monosyllabic and multisyllabic structures. Structurally, ideophonic adverbs with multiple syllables appear

as if they have been reduplicated. However, they cannot be, as there are no corresponding unreduplicated bases. In Dangme, ideophonic adverbs describe the manner of the actions denoted by the verb by expressing the distinctive sounds and movements that accompany the actions. In terms of adverbial derivation, Dangme creates adverbs through reduplication of adjectival and nominal roots.

Besides adverbs, Dangme has other syntactic structures which are used as adverbs. These include adverbials with temporal and locative functions. Dangme locative adverbials are nominal in character. They can occur at both the sentence-initial and sentence final positions and can also be fronted for focus. Nonetheless, they may generate ungrammatical sentences when they appear in the pre-verbal position. Concerning the structure of locative adverbials, they are expressed by postpositional phrases and compounds. All these adverbials are usually preceded by the locative verb *ngɛ* 'be.at.' The postpositional phrases can only surface in sentence-initial and sentence-final positions. Being placed in any other position may generate ill-formed sentences. In cases where they are fronted for focus, they are realized with or without focus markers. However, they are normally followed by a topic marker which is realized differently based on its preceding sound. For the compound adverbs, they made up of nouns and postpositions with spatio-temporal and locational relations. The difference between these adverbials and the postpositional phrases is because the (N+P) compounds become lexicalized when formed and may have a non-compositional meaning.

Against these analyses, it can be said that Dangme uses both adverbs and adverbials in expressing adverbial concepts. The descriptions of the various structures discussed suggest that the class of Dangme adverbs is an open class centered on three categories of words: core adverbs, ideophonic adverbs and reduplicated adverbs. The Dangme adverb class has both monosyllabic and multisyllabic members. Quite a large number of the adverbs with multiple syllables are formed through the process of reduplication which is different from the way adverbs are formed in other Kwa languages. Dangme adverbs (core adverbs, reduplicated and ideophonic adverbs) were discovered to have shared properties in terms of their structural distribution. Their morphological properties and lexical semantics allow them to be characterized as important adverbs in the Dangme language. Dangme adverbials on the other hand consist of temporal and locative nominals, postpositional phrases and compounded adverbials. The structural and distributional properties of Dangme adverbials are influenced by the kind of meaning they add to the sentence and the constituents they scope over. The members of each of the categories share a fixed set of syntactic properties that make them unique.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions:

FM	Focus marker	SBJ	Subject
HAB	Habitual	IDEO	Ideophone
POST	Postposition		

Acknowledgements

References

Abakah, Emmanuel N., Regina O. Caesar & James A. Ababila. 2010. Tonomorphology of reduplication in Akan, Dangme and Gurune. In Mary E. Kropp Dakubu & E. Kweku Osam (eds.), *Studies in the languages of the Volta Basin*, vol. 6 (Proceedings of the Annual Colloquium of the Legon-Trondheim Linguistics Project), 121–140. Legon: University of Ghana.

Ameka, Felix K. 1991. *Ewe: Its grammatical constructions and illocutionary devices.* Australian National University. (Doctoral dissertation).

Awuku, Ameyo S. 2011. Adverbial placement in Ewe: A Role and Reference Grammar. *Legon Journal of the Humanities* 22. 127–175.

Bodomo, Adams. 2006. The structure of ideophones in African and Asian languages: The case of Dagaare and Cantonese. In John Mugane, John P. Hutchinson & Dee A. Worman (eds.), *Selected Proceedings of the 35th Annual Conference on African Linguistics*, 203–213. Somerville, MA: Cascadilla Proceedings Project.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective* (Oxford Studies in Comparative Syntax). New York: Oxford University Press.

Cinque, Guglielmo. 2004. Issues in adverbial syntax. *Lingua* 114(6). 683–710.

Croft, William. 2013. Radical construction grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 211–32. Oxford: Oxford University Press.

Curme, George O. 1935. *A grammar of the English Language. Vol. II: Parts of speech and accidence.* Boston: D.C. Heath.

DeLancey, Scott. 2001. *Lectures on functional syntax.* Notes for the Summer School of the Linguistic Society of America held at the University of California at Santa Barbara. University of Oregon.

Dzameshie, Alex K. 1998. *The structure and functions of adverbials in Ewe.* Legon: Department of Linguistics.

Ernst, Thomas. 2001. *The syntax of adjuncts.* Cambridge: Cambridge University Press.

Geuder, Wilhelm. 2002. *Oriented adverbs: Issues in the lexical semantics of event adverbs.* Universität Konstanz. (Doctoral dissertation).

Givón, Talmy. 2001. *Syntax: A functional-typological introduction*, vol. I. Amsterdam: Benjamins.

Haspelmath, Martin. 2001. Word classes/parts of speech. In Neil Smelser & Paul Baltes (eds.), *Encyclopedia of the Social and Behavioural Sciences*, 16538–16545. Oxford: Pergamon.

Hengeveld, Kees. 1992. Parts of speech. In Michael Fortescue, Peter Harder & Lars Kristoffersen (eds.), *Layered Structure and Reference in a Functional Perspective: Papers from the Functional Grammar Conference, Copenhagen, 1990* (Pragmatics & Beyond New Series 23), 29–55. Amsterdam: Benjamins.

Huddleston, Rodney D. 1988. *English grammar: An outline.* Cambridge; New York: Cambridge University Press.

Kropp Dakubu, Mary E. 1987. *The Dangme language: An introductory survey.* Basingstoke, England/Accra-North, Ghana: Macmillan; Unimax Publishers.

Kropp Dakubu, Mary E. (ed.). 1988. *The languages of Ghana.* London; Accra: Macmillian; Unimax Publishers Limited.

Lefebvre, Claire & Anne-Marie Brousseau. 2002. *A grammar of Fongbe:* Berlin: De Gruyter Mouton.

Lusekelo, Amani. 2010. Adverbs as a word category in Kinyakyusa. *Journal of National Council of Less Commonly Taught Languages* 8. 60–88.

Otoo, Yvonne A. 2014. *Expressing property concepts in Ga.* University of Ghana, Legon. (MA thesis).

Paolo, Ramat & Davide Ricca. 1994. Prototypical adverbs: On the scalarity/radiality of the notion of ADVERB. *Rivista di linguistica* 6. 289–326.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara Lloyd (eds.), *Cognition and categorization*, 27–48. Hillsdale, NJ: Lawrence Erlbaum.

Saah, Kofi K. 2004. A survey of Akan adverbs and adverbials. *Journal of West African Languages* 31(2). 47–71.

Saah, Kofi K. & Paul K. Agbedor. 2004. Adverbs in Akan and Ewe: A comparative study. In Mary E. Kropp Dakubu & E. Kweku Osam (eds.), *Studies in the languages of the Volta Basin*, vol. 2 (Proceedings of the Annual Colloquium of the Legon-Trondheim Linguistics Project), 202–216. Accra: University of Ghana.

Sakyi, Joana Portia. 2013. Modal adverbs and predicative modal adjectives in Akan. In Olanike Ola Orie & Karen W. Sanders (eds.), *Selected Proceedings of the 43rd Annual Conference on African Linguistics*, 188–195. Somerville, MA: Cascadilla Proceedings Project.

Tabe, Florence A. E. 2015. Adverbs in Kenyang. *International Journal of Linguistics and Communication* 3(1). 112–133.

Taylor, John R. 2009. *Linguistic categorization.* Oxford: Oxford University Press.

Trask, Robert L. 1993. *A dictionary of grammatical terms in linguistics.* London & New York: Routledge.

Van Valin, Robert D. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning and function.* Cambridge: Cambridge University Press.

Chapter 5

Phonological variation in Kusaal: A synchronic dialectological study

Samuel A. Asitanga
University of Education, Winneba

Anthony Agoswin Musah
University of Education, Winneba

Samuel Alhassan Issah
University of Education, Winneba

The aim of this chapter is to discuss, analyze, and motivate patterns of phonological variation in two dialects of Kusaal, a Mabia language spoken in north-eastern Ghana where, ostensibly, the Toende dialect is more innovative than the Agole dialect. The paper adopts a synchronic dialectological approach to investigate the observed patterns of phonological variation in the dialects. It assumes a qualitative dimension with primary data sourced through interviews, elicitations and native intuitions. Alternations involving consonants and vowels are discussed. Devoicing of stops is widespread word-finally, while debuccalization of /s/ occurs in post-vocalic position. In addition, certain vowel sequences undergo deletion or coalescence in the Toende dialect.

1 Introduction

Language variation is an inherent property of any human language, with dialects which can be identified and analyzed both synchronically and diachronically (Labov 1990). Similarly, Wardhaugh (2010) notes that languages all over the world are not spoken uniformly due to people's different social and regional

backgrounds. The two regional dialects of Kusaal are not exempt from these assertions, wherein we observe pertinent variations at the phonological level of the Toende and Agole dialects. Kusaal is a Mabia (Gur) language spoken in northeastern Ghana. The language is geographically split into two by the White Volta River, forming two mutually intelligible dialects with distinct phonological patterns. The phonology of the dialects varies significantly in terms of segment alternations, which are quite obvious, such that it is possible to determine the regional affiliation of the speaker from a casual observation of their speech.

This chapter thus sets out to examine the observed synchronic variations using generative dialectological rules. It seeks to describe the segment alternations that occur in the language which result in these pronunciation differences. The study assumes a synchronic approach to dialect studies, where the descriptions of the dialects are based on how they are spoken today without recourse to their historical antecedents.

1.1 Brief linguistic profile of Kusaal

Kusaal has nine phonemic vowels, /i ɪ e ɛ a ɔ o ʊ u/, which have their long correlates as /i: ɪ: e: ɛ: a: ɔ: o: ʊ: u:/. These vowels can be divided into Advanced Tongue Root [+ATR] /i e o u/ and non-Advanced Tongue Root [-ATR] /ɪ ɛ a ɔ ʊ/. In addition, Kusaal has five nasal vowels /ĩ ɛ̃ ɔ̃ ʊ̃ ã/ which are all -ATR. The language has eight vowel sequences: [ɪa], [ɪʊ], [aɪ], [aʊ], [ʊa], [ʊɔ], [ɔɪ] and [ʊɔɪ], which are all produced with a retracted tongue root. These will be referred to as *sequential vowels*. Vowels produced with an advanced tongue root [+ATR] do not occur in sequence in Kusaal. The lengthening and sequencing of vowels are distinctive in the two dialects of Kusaal. For instance, while long /e:/ and /o:/ are not preferred in Agole (except in loanwords), they are prevalent in Toende. On the other hand, while all the sequential vowels are observed in Agole, they are restricted in Toende. Sequential vowels that arise in Toende are /aʊ/, /ɔɪ/, /aɪ/ and /ɪʊ/, which are often observed in:

(1)   a.   some noun roots + class suffix: /bã+ʊ̃k/ [bã ʊ̃k] 'shoulder' and/da:+ʊk/ [da:ʊk] 'wood'

b.   Loanwords: *ajɔpɔɪ* 'seven', *bakɔɪ* 'a week', *wɪ:ʊk* 'red', *awaɪ* 'nine'

c.   Interjections: *ajaɪ* 'of course' *waɪ:* 'wow' (Niggli 2014: 39).

As will be shown in our subsequent discussions, such sequential vowels in Agole are adapted by means of deletion or coalescence with compensatory lengthening in Toende.

Furthermore, Kusaal vowels are also categorized according to tongue region, such as front /i ɪ e ɛ/, back /u ʊ o ɔ/, and central /a/. As observed cross-linguistically, and particularly so for the Mabia group of languages, all the front vowels are non-round while the back ones are round. The vowels are also grouped based on the height of the tongue, such as high /i ɪ u ʊ/, low /a/ and mid /e o ɛ ɔ/. These phonemic inventories are present in both the Agole and Toende dialects of Kusaal (Musah 2018, Niggli 2014). Table 1 contextualizes these facts.

Table 1: The Kusaal vowel system

	Oral			Nasal		
	Front	Central	Back	Front	Central	Back
+ATR	i		u			
−ATR	ɪ		ʊ	ĩ		ʊ̃
+ATR	e		o			
−ATR	ɛ		ɔ	ɛ̃		ɔ̃
+ATR						
−ATR		a			ã	

As regards the consonantal inventory of Kusaal, we count 22 phonemic consonants. The consonants correspond to seven places of articulation and six manners of articulation with voiced and voiceless members at most articulatory points. The Toende and Agole dialects show these phonemic consonant distributions in their respective phonologies (Niggli 2014, Abubakari 2018, Musah 2018).[1] The data in Table 2 presents a consonantal chart depicting all the phonemic consonants in Kusaal for both dialects.

1.2 Phonological variation

Gaskell & Marslen-Wilson (1996) define phonological variation as systematic variations occurring within conjunctions of speech sounds which are triggered by varied phonological processes in different dialects of the same language. Roh (2004) also intimates that phonological variation occurs when a single underlying form in a language is mapped onto multiple outputs. Wolfram & Schilling

---

[1]Note that this paper concludes that [r] is an allophone of /d/ rather a separate phoneme in §4.1.2, so the trill is not listed in the phoneme inventory.

Table 2: Phonemic Consonants of Kusaal

	Bilabial	Labio-dental	Alveolar	Alveo-palatal	Palatal	Labial-palatal	Velar	Labial-velar	Glottal
Stops	p b		t d				k g	k͡p g͡b	ʔ
Fricatives		f v	s z	ʝ					h
Nasals	m		n		ɲ		ŋ		
Approximants					j			w	
Lateral			l						

(2016) posit that phonological patterns can be indicative of regional dialect differences. They add that if a person has a good listening ear for language variation, one can pinpoint a speaker's regional affiliation with considerable accuracy based solely on the pronunciations of lexical items. Evidence drawn from Kusaal data attest to the fact that phonological variation in the dialects is equally obvious, such that even a cursory observation of conversations could straight away help determine the regional background of the speaker.

Gaskell & Marslen-Wilson (1996) analyze place of assimilation variation within some regional dialects of English and find that this kind of variation is usually seen in word boundaries, where a previous consonant adapts the place of articulation of the following segments in some regions, while a similar process is not seen in other regions in English. They note that this affects only coronals such as /t/, /d/ and /n/ when they are followed by non-coronals such as the labials /p/, /b/ and /m/ or the velars /k/, /g/ and /ŋ/. They also note that place assimilation is asymmetric in English, such that non-coronal segments cannot assimilate preceding coronals. For instance, a phrase such as [wɪkɪd pɹæŋk] 'wicked prank' is realized as [wɪkɪb præŋk] whiles 'black tie' [blæk taɪ] is not produced as [blæt taɪ] but as [blæk taɪ] because place assimilation is asymmetric in English (Gaskell & Marshen-Wilson, 1996: 145). Likewise, Mishra & Bali (2011) present a comparative analysis of phonological variation in Hindi dialects and observe that the prevailing cause for the variation in the dialects is vowel quality. According to them, while Awadhi shows allophonic free variation between [eː, oː] and [jɑː, wɑː], as in [djɑːkhɑu] ~ [deːkhɑu] (which can be shortened when the consonant /k/ is lengthened, e.g., [ekɑu] ~ [eːkɑu]), they are in complementary distribution with

the more common /i/ and /u/ in the Bagheli dialect [dustana] ~ [dostana] 'friend-ship'. Similarly, they maintain that high vowels in Bundeli tend to be lower in the other dialects. For instance, [bahota denõ se] in Bundeli is heard as [bahʊta dɪnõ se] 'from many days' in the other dialects of Hindi (Mishra & Bali 2011: 1392). We observe similar processes in Kusaal and will elaborate on these in the data analysis section.

In addition to the foregoing, Eze (2019: 60-61), in an investigation of linguistic variation in Umunze, a dialect of Igbo, establishes segment substitution as the most pronounced instance of phonological variation from standard Igbo. She observes that Umunze speakers use the close back vowel [ṳ] in place of the close front vowel [i] in forms like [dṳ] instead of [di] for 'is', and the voiced labio-dental fricative [v] in place of the voiced bilabial plosive [b] in forms such as [vuː] for [buː] 'to carry'. The velar sound [ɣ], written <gh>, in Umunze, is also the representation of the voiceless labiodental fricative [f] in Igbo. For instance, *oghe* in Umunze is *ofe* in Igbo for 'soup'. The voiced lateral consonant [l] in Umunze is substituted for the nasal alveolar sound [n] in a word like *chileke* instead of *chineke* in Igbo for 'God'. Her data shows that segmental substitutions are not systematic, as each segment could be used to substitute many segments in different words. Her data on Umunze and Igbo are similar to the case of Kusaal, as segment alternations mark the core trigger of phonological variations in the language. For instance, all voiced plosives [b d g] in word-final position in Agole are altered and replaced by their voiceless counterparts [p t k] in Toende.

2 Theoretical framework

This paper hinges on generative dialectology, which aligns with generative grammar and is set within the theory of generative phonology as proposed in *The Sound Pattern of English* (SPE) by Chomsky & Halle (1968). Generative dialectology asserts that since generative phonology accounts for surface forms that are different from underlying forms in one variety, it could also be used to account for the differences in different varieties (Abubakar 1982, Chambers & Trudgill 2004). According to these studies, since dialects are more or less from a uniform language, it is possible to show that they can, for the most part, be described in terms of a common set of underlying forms.

The theory of generative dialectology is guided by the principles of identifying underlying forms based upon which lexical forms are listed in the lexicon, then applying phonological rules to the underlying forms to convert them to surface forms and, ultimately, into their actual pronunciations. (Chambers & Trudgill

2004: 39) state that "in particular, forms involved in alternations of various kinds appear in the lexicon as only one form, the others being the result of the application of rules". The theory proceeds on the premise that a single underlying form can be postulated for related dialects, where the dialects differ based on (a) the phonological rule that applies to the underlying form; (b) the environment in which the rules apply; and/or (c) the order in which the rules apply. Generative dialectologists' concerns are thus the identification of the underlying forms. The remaining facts then derive from their interest in generating phonological rules and in using these rules to generate formulae to account for variations in related dialects.

3 Methodology

We follow the methods and procedures dialectologists use in identifying, describing and presenting dialect differences in line with synchronic dialectological perspectives in this paper. The study assumes a qualitative approach with primary data sourced through interviews, word list elicitations and native speaker intuitions. While the regions of the two dialects, eastern and western Bawku, constitute the research sites of the study, five participants from five communities in each of the dialect areas were consulted. The data was collected by administering questionnaires to participants and their responses captured and stored on an audio recorder. These were then transcribed and analyzed based on the theory of generative dialectology, whereby we first identified and postulated underlying forms for the two dialects while noting observed variations. Following from this, we determined the systematicity in the variations by applying phonological rules to the derived underlying forms.

In line with the theory of generative dialectology, dialectologists identify underlying forms, apply systematic phonological rule(s) to the underlying forms and derive variations from them to account for the differences in related dialects (Al-Hindawi & Al-Aadili 2018; Abubakar 1982; Chambers & Trudgill 2004). There are several methods for selecting underlying forms for dialect studies. Abubakar (1982), for instance, posits that early dialectologists described dialect differences in an ad hoc manner by arbitrarily taking forms from one dialect as the base form and deriving other forms from them for related dialect(s). In analyzing the dialects of modern Faroese, O'Neil (1963) uses this approach in the selection of underlying forms. According to Abubakar (1982), however, the arbitrary selection system was observed to sometimes not be reliable or accurate, thereby giving rise to the notion that underlying forms should be more abstract and independent. This opinion is buttressed by Thomas (1967), who affirms that one must

select forms that are more or less abstract, widespread and independent of the dialects under study. Variations are then derived from these forms and postulated for the dialects by applying phonological rules to the abstract forms.

Other methods used to determine underlying forms include dominance and usage (Goldstein & Iglesias 2001) and the historical antecedence of dialects (Newton 1972), and make apparent the fact that dialectologists do not entirely agree on one unique acceptable criterion for the identification and selection of so-called underlying forms. Abubakar (1982: 30) however, notes that "whichever method one adopts to establish the underlying forms, a claim is made that a generative treatment of dialect differences will formalise the essential fact about dialects: that they have much in common but still have some differences".

In this paper we employ underlying forms that are closer to the Agole dialect than the Toende dialect. We show with empirical evidence that most vowel and consonant changes have occurred in the Toende dialect, making it more divergent from Agole. In the next section, we present and analyze the data.

4 Segment alternation

Segment alternation is a phonological process that allows speakers of a language to alter or modify a segment or a group of segments due to some phonotactic constraints (Katamba 1989). Speakers of the Toende dialect of Kusaal apply this process to both consonants and vowels, mostly in word-medial and word-final positions. Even though this phenomenon is often observed in social dialects, where different social factors such as education, gender, social class, and age may influence the choice of one variable over another, as noted by Labov (1990), it is also observed in regional dialects, as is the case for the present study.

The consonant alternations that are observed in Toende include word-final devoicing (§4.1.1), [d]/[r] alternations (§4.1.2), and debuccalization of the voiceless alveolar fricative /s/ (§4.1.3).

In addition, segment alternations involving vowels of Kusaal are quite pervasive. This arises mostly where a number of sequential vowels in Kusaal are realized as single vowels in Toende, as shown in §4.2. These segment alternations underscore a clear distinction between the two regional dialects.[2]

---

[2]Tone is not indicated in the transcriptions as there are sometimes minor differences between the dialects, which might distract from the segmental alternations that are the focus of the paper.

4.1 Consonant alternation

Consonant alternation separates the two regional dialects of Kusaal, where voiced obstruents in word-final positions are maintained in Agole but are neutralized (devoiced) in Toende in the same environment.

4.1.1 Devoicing

The data in (2) illustrate the distribution of /k/ and /g/ in Kusaal. In Agole there is a contrast between /k/ and /g/ in final position, as shown by comparing (2a-b) with (2c-d). In Toende, however, there is only [k] word-finally. Where Agole has [g], Toende has [k], as in (2c-i). This results in pronunciation discrepancies in the two regional varieties. The /g/ is assumed to be underlying but is devoiced in Toende. The data in (2j-m) shows that that the voiced velar /g/ is realized as [g] in non-final positions in both dialects. UR refers to the underlying representation.

(2)		UR	Agole	Toende	
	a.	/kʊk/	kʊk	kʊk	'chair'
	b.	/mak/	mak	mak	'measure'
	c.	/zug/	zug	zuk	'head'
	d.	/dʊg/	dʊg	dʊk	'pot'
	e.	/dɔːg/	dɔːg	dɔːk	'room'
	f.	/lɛːg/	lɛːg	lɛːk	'to dig'
	g.	/bʊʔɔg/	bʊʔɔg	bʊʔɔk	'valley'
	h.	/lɔdʊg/	lɔdʊg	lɔrʊk	'corner'
	i.	/dadʊg/	dadʊg	darʊk	'ladder'
	j.	/zigi/	zigi	zigi	'gravels'
	k.	/dagɔbʊg/	dagɔbʊg	dagɔbʊk	'left hand'
	l.	/gɪdɪma/	gɪdɪma	gɪrɪma	'respect'
	m.	/gãdɪg/	gãdɪg	gãrɪk	'to respond'

Similarly to the data with velar stops, a final voiced bilabial stop /b/ is realized as a voiceless bilabial stop [p] in Toende (3a-f). We note however that a similar occurrence does not apply to the voiced obstruent [b] at word-initial and word-medial positions in Kusaal (3g-j). There is no observed final [p] in Agole, but [p] can occur in other positions (3k-l), showing that there is a /p/ vs. /b/ contrast in both dialects.

(3) | | UR | Agole | Toende | |
|-----|-----------|------------|---------|------------|
| a. | /sãb/ | sãb | sãp | 'abundance' |
| b. | /sɛb/ | sɛb | sɛp | 'to squat' |
| c. | /lɔb/ | lɔb | lɔp | 'to throw' |
| d. | /ɔb/ | ɔb | ɔp | 'to chew' |
| e. | /mɛːb/ | mɛːb | mɛːp | 'building' |
| f. | /dɔːb/ | dɔːb | dɔːp | 'climbing' |
| g. | /zaba/ | zaba | zaba | 'conflict' |
| h. | /sabɪl/ | sabɪl | sabɪl | 'black' |
| i. | /bãŋ/ | bãŋ | bãŋ | 'ring' |
| j. | /bɛdɪgʊ/ | bɛdɪgʊ | bɛrɪgʊ | 'plenty' |
| k. | /paŋ/ | paŋ | paŋ | 'strength' |
| l. | /pʊpʊːm/ | pʊpʊːm | pʊpʊːm | 'foam' |

Agole shows a contrast between the voiced alveolar stop /d/ and the voiceless alveolar stop /t/ in word-final position (4a-l). However, Toende only has the voiceless alveolar stop [t] in word-final position. In addition, another instance of apparent dialect variation in Kusaal is also shown where the voiced alveolar plosive /d/ is optionally realized as the alveolar trill [r] word-finally in Agole, but still realized as [t] in Toende. This fact is reported in the literature on Kusaal, where it is established that the voiced alveolar plosive [d] and the trill [r] are free variants in word-final position in Agole, which is not the case in Toende (Musah et al. 2013, Musah 2018, Niggli 2014).

(4) | | UR | Agole | Toende | |
|-----|-----------|----------------|---------|--------------|
| a. | /mat/ | mat | mat | 'wet.IDEO' |
| b. | /kat/ | kat | kat | 'to chase' |
| c. | /gbɛd/ | gbɛd ~ gbɛr | gbɛt | 'thigh' |
| d. | /kʊkɔd/ | kʊkɔd ~ kʊkɔr | kʊkɔt | 'throat' |
| c. | /daːd/ | daːd ~ daːr | daːt | 'wood' |
| d. | /bɔːd/ | bɔːd ~ bɔːr | bɔːt | 'to want' |
| e. | /bʊːd/ | bʊːd ~ bʊːr | bʊːt | 'to vindicate' |
| f. | /vãːd/ | vãːd ~ vãːr | vãːt | 'leaves' |
| g. | /dɔːd/ | dɔːd ~ dɔːr | dɔːt | 'dawadawa fruit' |
| j. | /kpaːd/ | kpaːd ~ kpaːr | kpaːt | 'farmer' |
| k. | /nɔːd/ | nɔːd ~ nɔːr | nɔːt | 'mouth' |
| l. | /jʊʔʊd/ | jʊʔʊd ~ juʔur | jʊʔʊt | 'name' |

A rule to account for the trill will be provided in §4.1.2.

There is no voicing alternation for sonorants that occur in word-final, position as shown below in (5).

(5) | | **UR** | **Agole** | **Toende** | |
|---|------|-------|--------|---|
| a. | /saːn/ | saːn | saːn | 'stranger' |
| b. | /puːm/ | puːm | puːm | 'flower' |
| c. | /bã́ŋ/ | bã́ŋ | bã́ŋ | 'ring' |
| d. | /bul/ | bul | bul | 'to germinate' |

Labial-velar stops do not appear in word-final position, so they are excluded from the voicing alternation. Voiceless and voiced fricatives contrast word-initially in words like [saːm] 'to march' and [zaːm] 'evening' or [fã́ːd] 'to save' vs. [vã́ːd] 'leaves' (Agole dialect) (Musah 2018). However, only voiceless fricatives are found word-finally in both dialects, and there are no alternations. As this paper focuses on differences between the two dialects, we do not explore this further.

From the foregoing illustrations, it is apparent that there is a rule of word-final devoicing that applies to voiced stops in Toende. The rule does not also apply to stops in word-initial or medial positions.

(6) Devoicing

$$\begin{bmatrix} \text{-son} \\ \text{-cont} \end{bmatrix} \rightarrow \text{[-voice]} \: / \: \_\#$$ (Toende)

Word-final devoicing has been studied systematically and extensively across many languages in the world (see Dinnsen 1985, Charles-Luce 1985, Slowiaczek & Dinnsen 1985, and Slowiaczek & Szymanska 1989 for some examples in German, Dutch, Polish, Catalan, and Turkish among others). Dinnsen (1985: 266) notes that the rule involves the merger of voiced and voiceless obstruents in favour of the voiceless at word-final positions. According to Charles-Luce (1985: 309), the word-final devoicing rule has been formulated to account for (i) the voice alternation between medial voiced obstruents and final voiceless obstruents, and (ii) the presumed absence of a voice contrast word-finally. The present study does not investigate alternations between word forms within a dialect, but focuses only on word-final devoicing as a difference between Toende and Agole.

4.1.2 [d] ~ [r] alternation

As discussed in §4.1.1, the alveolar trill [r] and the voiced alveolar plosive [d] are free variants in word-final position in Agole. It is, however, prudent to state that they do not substitute for each other in all phonological environments in

Kusaal, and Toende has a different distribution. In word initial position, only [d] occurs. [dadʊg] 'ladder' cannot be pronounced as [*radʊg] or [daːn] 'owner' be pronounced as [*raːn]. Furthermore, the alveolar trill [r] does not begin words in either dialect of Kusaal (Musah et al. 2013).

However, in word-medial position, [d] is found in Agole, but [r] (or [ɾ]) in Toende. Data in (7) show the differences.

(7)		**UR**	**Agole**	**Toende**	
	a.	/bɛdɪgʊ/	bɛdɪgʊ	bɛrɪgʊ	'plenty'
	b.	/bidibiŋ/	bidibiŋ	biribiŋ	'boy'
	c.	/bidikin/	budikin	birɪkin	'noble'
	d.	/lɔdʊg/	lɔdʊg	lɔrʊk	'corner'
	e.	/ɛdʊg/	ɛdʊg	ɛrʊk	'anxiety'
	f.	/fada/	fada	fara	'hardships'
	g.	/fɛdɪg/	fɛdɪg	fɛrɪk	'turn'
	h.	/pʊdʊg/	pʊdʊg	pʊrʊk	'share'
	i.	/jaːdɪm/	jaːdɪm	jaːrɪm	'salt'

This alternation only occurs if the /d/ is between two vowels. Any consonant sequences have [d]. For example, [jadda] 'faith' cannot be pronounced as *[jarra] or [tɪndãʔãn] 'a dry mud' be pronounced as *[tɪnrãʔãn]. As [r] is always a variant of /d/ and never contrasts with it in other positions, we conclude that they are not separate phonemes. The two rules are formulated in (8).

(8) Trilling
    a. /d/ → [r] / _# (optional)   (Agole)
    b. /d/ → [r] / V_V   (Toende)

As defined by the phonological rules, the /d/ to [r] variation rule applies in both dialects, but it is optional in Agole word-finally and obligatory in Toende between vowels in word-medial position.[3]

4.1.3 Debuccalization

Another salient phonological variation in the language is observed in the debuccalization of /s/ to [h] in word-medial and word-final positions in the Toende dialect, where the Agole dialect has [s]. Here, the voiceless alveolar fricative /s/ loses its original place of articulation and becomes [h] in Toende when following a vowel (Niggli 2014: 11, Hudu 2018). The data in (9) provide evidence of this

---

[3]In word-medial position, it is pronounced as [r] or [ɾ].

trend in alternation. While the examples in (9a-e) highlight /s/ being realized as [h] in word-medial position in Toende, examples (9f-j) show the debuccalization of /s/ in word-final position.

(9)

	UR	Agole	Toende	
a.	/bɪʔɪsɪm/	bɪʔɪsɪm	bɪʔɪhɪm	'breast milk'
b.	/bʊʔɔsʊg/	bʊʔɔsʊg	bɔʔɔhʊk	'question'
c.	/faːsɪm/	faːsɪm	faːham	'swollen'
d.	/ɲɔsɪg/	ɲɔsɪg	ɲɔhʊk	'to miss (a target)'
e.	/kãsɪd/	kãsɪr	kãhãt	'hot weather'
f.	/tɛʔɛs/	tɛʔɛs	tɛʔɛh	'to think'
g.	/ɛbɪs/	ɛbɪs	ɛbɪh	'to scratch'
h.	/ɛːs/	ɛːs	ɛːh	'to wipe'
i.	/dɪʔɪs/	dɪʔɪs	dɪʔɪh	'to press'
j.	/ʊɔːs/	ʊɔːs	ɔːh	'to warm up'

It is pertinent to state that when the voiceless alveolar fricative /s/ occurs in word-initial position, it does not debuccalize in Toende, as shown in (10a-d). Furthermore, /s/ and /h/ contrast in word-initial position (10e), although underlying /h/ does not occur in other positions in either dialect.

(10)

	UR	Agole	Toende	
a.	/sugudu/	sugudu	suguru	'peace'
b.	/saman/	saman	saman	'compound'
c.	/sɔːd/	sɔːd	sɔːt	'liver'
d.	/sɪːg/	sɪːg	sɪːk	'spirit'
e.	/hali/	hali	hali	'a lot, very much, greatly'

The phonological rule can be written as follows:

(11) Debuccalization

/s/ → [h] / V__   (Toende)

The rule is applicable in the Toende data, where the voiceless alveolar fricative /s/ in the underlying form debuccalizes into the glottalic fricative [h] in word-medial and word-final positions following vowels.

It is also noted that the glottal fricative [h] and the stop /ʔ/ are phonologically transparent in Kusaal, and allow progressive spreading of vowel features across them. Examples in (9c-e) show that when debuccalization occurs, all features except [high] spread progressively across [h], including the feature [+nasal]. This

is an instance of translaryngeal harmony (Steriade 1986). We do not provide a formal rule for this variation, but note that debuccalization must apply prior to the harmony.

Closely related to the foregoing is the fact that debuccalization does not apply in compounded forms where /s/ is the onset of the second word or morpheme. Morphologically, the voiceless alveolar fricative /s/ can occur word-medially in Toende when the word in question arises from compounded forms. For example, in /daʊ + saan/ → [dasan] 'young man' and /zug + sʊŋ/ → [zusʊŋ] 'luck', /s/ does not debuccalize but remains the same for Toende as well as Agole. Similarly, when the low-central vowel /a/ is used as a prefix to a base beginning with /s/, the fricative does not change in Toende. This is mostly shown in nominal items where the prefix /a-/ functions as a nominalizer in Kusaal. Examples include /a-sibi/ → [asibi] 'Mr Saturday', /a-sɪda/ → [asɪra] 'Mr True', /a-saman/ → [asaman] 'Mr Compound' and /a-sɛr/ → [asɛt] 'Mr Wall-gecko'. where the underlying forms and surface representations remain the same for both dialects of the language, except in Toende where /d/ is realized as [r] in word-medial position. Ostensibly, this could be reflective of root-initial faithfulness in Kusaal, where /s/ is preserved in both dialects because it occurs in root-initial position. Similar observations are made by Ahn (2000a,b), Lee (2000) on root-faithfulness in English phonology.

Hudu (2018: 214) observes a similar process of debuccalization in Dagbani, a related Mabia language, and notes that the process targets coronals and dorsals, making them glottals in Dagbani. The data in (12) illustrate the phenomenon of /s/ → [h] alternation in Dagbani.

(12)   /s/ to [h] alternation in Dagbani (Hudu 2018: 214)

	UR	Dagbani	
a.	/má:sɪlɪ/	[máhɪlɪ]	'cool weather after rain'
b.	/nὲ:-sɪ̀/	[nɛ-hɪ̀]	'awaken-pl.'
c.	/mó:sɪ/	[mɔhɪ]	'become reddish'
d.	/ánáːsɪ/	[ánáhɪ]	'four'
e.	/bìsím/	[bìhím]	'milk'
f.	/bíː-sí/	[bí-hí]	'children'
g.	/boːsɪ/	[bɔhɪ]	'ask'

From the data, it is apparent that the /s/ to [h] debuccalization in Dagbani is similar to that of the Toende dialect of Kusaal. However, while the glottalic fricative [h] is said to occur only as an allophone of /s/ in Dagbani, they are separate phonemes in Kusaal (Hudu 2018: 207; Niggli 2014; Musah 2018). Also, as

the data portray, debuccalization triggers shortening of preceding long vowels in Dagbani, whereas the preceding long vowels are not shortened in Kusaal. For instance, while /má:sɨlɨ/ is [máhɨlɨ] and /áná:sɨ/ is [ánáhɨ] for 'cold weather' and 'four' respectively in Dagbani, /ka:sʊg/ and /du:sug/ are realized as [ka:hʊk] and [du:huk] for 'crying' and 'cleaning' respectively in the Toende dialect of Kusaal.

4.2 Vowel alternation

There are several vowel alternations in Kusaal which constitute phonological variation between the two dialects. This is observed where sequential vowels in the underlying representation are maintained in Agole but undergo a process of vowel deletion or coalescence in Toende. There are eight types of sequential vowels in the language ([ɪa], [ɪʊ], [aɪ], [aʊ], [ʊa], [ʊɔ], [ɔɪ] [ʊɔɪ]) and alternations are observed with those that have a high vowel followed by a non-high vowel: [ɪa], [ʊa] [ʊɔ] [ʊɔɪ]. While some scholars of Kusaal describe sequential vowels as diphthongs and triphthongs (see Musah et al. 2013: 14, Musah 2018: 57, Abubakari 2018: 38, and Niggli 2014: 39), the current paper refers to them as sequential vowels (SVs) (see Adongo 2018 for a similar observation in Gurɛnɛ). Each of the SVs can be optionally bisected by the glottal stop /ʔ/, thereby reshaping the word, with the /ʔ/ forming the onset of a new syllable in the word. For instance, /sɪak/ 'agree' and /fʊɔɪ/ 'remove', which are CVVC and CVVV respectively, could be reshaped as [sɪʔal] 'to meet' and [sʊʔɔɪ] 'own', to assume forms as CV.CVC and CV.CVV respectively (Musah 2018; Asitanga 2021). Since diphthongs and triphthongs are assumed to be inseparable, we argue that they are sequential vowels rather than diphthongs and triphthongs.

The Agole and Toende dialects of Kusaal differ in their phonology with respect to vowel sequence alternations in the language. The variety of Toende spoken in Ghana does not accept certain vowel sequences within morpheme boundaries, except those that are morphologically conditioned such that a noun root + class suffix could produce a diphthong, some loanwords and interjections, as identified by Niggli (2014) and explained in (1). Due to this restriction, some SVs are always adapted by means of either vowel deletion or coalescence in Toende.

4.2.1 Vowel deletion

The vowel deletion process occurs with the nasal SV /ɪã/, while the coalescence process occurs with the oral SV /ɪa/. Consider the data in (13). Deletion makes the SV /ɪã/ in the UR surface in Toende as [ã]. Examples (13g-h) demonstrate that the deletion does not occur with sequences of oral vowels /ɪa/.

(13)   **UR**   **Agole**   **Toende**
    a.   /dɪ̃ã̃ʔãd/   dɪ̃ã̃ʔãd   dã̃ʔãt   'dirt'
    b.   /kpɪãk/   kpɪãk   kpãk   'to economise'
    c.   /pɪ̃ãʔã/   pɪ̃ãʔã   pã̃ʔã   'to speak'
    d.   /tɪãg/   tɪãg   tãk   'to massage'
    e.   /zɪãg/   zɪãg   zãk   'to wither off'
    f.   /ɲɪãg/   ɲɪãg   ɲãk   'stimulus'
    g.   /lɪabʊg/   lɪabʊg   lɛːbʊk   'disturbance'
    h.   /sɪak/   sɪak   sɛk   'fit'

The phonological rule in (14) states that vowels delete before nasal vowels. This makes the nasal SV /ɪã/ becomes [ã] in Toende. There are no such alternations on the other nasal sequential vowels.

(14)   Vowel deletion

$$V \rightarrow \varnothing \ / \ \_\_ \begin{bmatrix} V \\ +\text{nasal} \end{bmatrix}$$ (Toende)

The [+nasal] feature triggers the deletion process, as examples (13g-h) that do not have nasal vowels show a different process rather than deletion. The preservation of [ã] in Toende is also motivated by its degree of sonority, which is higher than the front high vowel /ɪ/. Niggli (2014: 47) also notes that in Kusaal as well "when one vowel of a diphthong is to be deleted, it is the high vowel, not a low or mid vowel".

4.2.2 Coalescence

As hinted above, the SV /ɪa/ becomes [ɛ] in Toende. Here, the SV undergoes a process of coalescence together with compensatory lengthening, preserving the moraic value or timing positions of the input vowels. Under coalescence, the [-back] value of the first vowel is preserved, and the [-high] value of the second vowel is maintained, producing a front mid vowel.

(15)   **UR**   **Agole**   **Toende**
    a.   /dabɪam/   dabɪam   dabɛːm   'fear'
    b.   /fɪam/   fɪam   fɛːm   'freedom'
    c.   /tɪaŋ/   tɪaŋ   tɛːŋ   'beard'
    d.   /bɪal/   bɪal   bɛːl   'naked'
    e.   /pɪan/   pɪan   pɛːn   'a type of cloth'
    f.   /bɪa/   bɪa   bɛː   'to go astray'
    g.   /fɪa/   fɪa   fɛː   'to blame'
    h.   /wɪas/   wɪas   wɛːh   'to analyze'

Coalescence affects not only /ɪa/ to [ɛː] but also /ʊa/ to [ɔː], /ʊɔ/ to [ɔː], and /ʊɔɪ/ to [ɔː] in Toende. Under coalescence, the [+back, +round] values of the first vowel are preserved, and the [-high] value of the second vowel is maintained, producing a back rounded mid vowel.

(16)

	UR	Agole	Toende	
a.	/jʊal/	jʊal	jɔːl	'to babysit'
b.	/sʊas/	sʊas	sɔːh	'to startle'
c.	/vʊaŋ/	vʊaŋ	vɔːŋ	'a cotton tree'
d.	/ʊas/	ʊas	ɔːh	'to warm up'
e.	/zʊal/	zʊal	zɔːl	'to perch'
f.	/bʊɔlʊg/	bʊɔlʊg	bɔːlʊk	'calling'
g.	/jʊɔlɪm/	jʊɔlɪm	jɔːlʊm	'later on'
h.	/vʊɔl/	vʊɔl	vɔːl	'whistle'
i.	/kʊɔsʊg/	kʊɔsʊg	kɔːhʊk	'selling'
j.	/sʊɔl/	sʊɔl	sɔːl	'advantage of'
k.	/lʊɔɪ/	lʊɔɪ	lɔː	'take some'
l.	/bʊɔɪ/	bʊɔɪ	bɔː	'to pour'
m.	/vʊɔɪ/	vʊɔɪ	vɔː	'to uproot'
n.	/dʊɔɪ/	dʊɔɛ	dɔː	'to get up'
o.	/fʊɔɪ/	fʊɔɪ	fɔː	'to remove'

In each of these vowel alternations, the SVs coalesce into a long [ɛː], and [ɔː], preserving the bimoraic length of the SVs in the underlying representation.[4] Furthermore, coalescence occurs in both final and non-final syllables. The vowel alternations are formalized by the rule in (17).

(17)   Coalescence

$$\begin{bmatrix} V \\ \alpha back \\ \alpha round \\ +high \end{bmatrix} \begin{bmatrix} V \\ -high \end{bmatrix} (V) \rightarrow \begin{bmatrix} Vː \\ \alpha back \\ \alpha round \\ -high \\ -low \end{bmatrix} \qquad \text{(Toende)}$$

This rule employs alpha notation, which states that a sequence of a high vowel and a non-high vowel (or two vowels) becomes a long mid vowel of the same backness and rounding as the first vowel. The rule is context-free, reflecting the

---

[4]Sequences of three vowels occur in open syllables, but are reduced to a long vowel with double vowel length, not triple vowel length.

fact that it applies in a variety of environments: in open syllables, and when followed by sonorants or fricatives.

If the following consonant is a voiceless velar or glottal stop, however, the vowel length is not maintained and a short vowel results. This is shown with /ɪa/ → [ɛ] in (18a-f) and /ʊa/ → [ɔ] in (18g-k) when followed by either [k] or [ʔ].

(18)

	UR	Agole	Toende	
a.	/sɪak/	sɪak	sɛk	'enough'
b.	/wɪak/	wɪak	wɛk	'to hatch'
c.	/tɪak/	tɪak	tɛk	'to exchange'
d.	/kpɪak/	kpɪak	kpɛk	'to restrain'
e.	/vɪak/	vɪak	vɛk	'to be burnt'
f.	/kpɪaʔa/	kpɪaʔa	kpɛʔɛ	'neighbor'
g.	/bʊak/	bʊak	bɔk	'to cut open'
h.	/kʊak/	kʊak	kɔk	'to hug'
i.	/lʊak/	lʊak	lɔk	'to elude'
j.	/mʊak/	mʊak	mɔk	'to suck'
k.	/sʊak/	sʊak	sɔk	'a type of fishing equipment'

This process is not observed in Agole, which clearly differentiates the two dialects in the pronunciation of words with such phonological make-ups.

Instead of proposing a second coalescence rule that is context-specific to velars and glottal stops and produces a short vowel, we propose that vowels are shortened before voiceless stops in Toende. The rule does not apply before voiceless fricatives, so the feature [-cont] is included.

(19) Vowel shortening

$$V: \rightarrow V \ / \ \_\_ \begin{bmatrix} -\text{son} \\ -\text{cont} \\ -\text{voice} \end{bmatrix} \qquad \text{(Toende)}$$

Neither dialect has long vowels preceding underlying voiceless stops, so the vowel shortening rule accounts for this phonotactic restriction as well. Toende does have long vowels preceding devoiced stops, as shown in (20). Furthermore, long vowels derived from vowel coalescence can appear before devoiced stops (20g-h). This shows that the language treats voiceless stops and devoiced stops differently.

(20)   **UR**   **Agole**   **Toende**
   a.   /dɔːg/   dɔːg   dɔːk   'room'
   b.   /lɛːg/   lɛːg   lɛːk   'to dig'
   c.   /daːd/   daːd ~ daːr   daːt   'wood'
   d.   /bɔːd/   bɔːd ~ bɔːr   bɔːt   'to want'
   e.   /mɛːb/   mɛːb   mɛːp   'building'
   f.   /dɔːb/   dɔːb   dɔːp   'climbing'
   g.   /lɪab/   lɪab   lɛːp   'to court'
   h.   /kʊɔb/   kʊɔb   kɔːp   'farming'

If devoicing is ordered after vowel shortening, this pattern can be accounted for. However, there are apparently no examples of oral vowel sequences before voiced velar stops to compare directly with the data in (18) before voiceless velar stops; predictions about rule ordering are made based on the data in (20) for voiced labials. This is discussed in the next section.

4.3 Rule ordering

It is imperative to state that the rules are logically ordered in Toende. First, the vowel deletion rule applies only before nasal vowels, whereas the vowel coalescence rule applies to vowels in general. One could add the feature [-nasal] to the coalescence rule, or one could order the vowel deletion rule first, as is done here. This ensures that the oral-nasal vowel sequence is repaired before coalescence applies. Second, vowel shortening follows vowel coalescence because vowel coalescence creates the long mid vowel that is then shortened. This is shown in (21). SR indicates surface representation.

(21)   UR   /tɹak/   /tɹãg/
    Deletion   –   tãg
    Coalescence   tɛːk   –
    Shortening   tɛk   –
    Devoicing   –   tãk
    SR   [tɛk]   [tãk]
             'to exchange'   'to massage'

Finally, devoicing follows vowel shortening to ensure that shortening only occurs before underlying voiceless stops, not those that have been devoiced. Sample derivations of this interaction are shown in (22). The vowel is shortened with /tɹak/ as the /k/ is underlying. But there is no shortening for /dɔːg/ or /lɪab/ because the stop is voiced at that point in the derivation. The devoicing applies following the shortening.

(22) UR   /tɪak/   /dɔːg/   /lɪab/
      Deletion   –   –   –
      Coalescence   tɛːk   –   lɛːb
      Shortening   tɛk   –   –
      Devoicing   –   dɔːk   lɛːp
      SR   [tɛk]   [dɔːk]   [lɛːp]
                    'to exchange' 'room'   'to court'

The rules of debuccalization and trilling are not crucially ordered with respect to the other rules as they does not interact with them. Shortening only applies before stops, and both [s] and [h] are fricatives. Debuccalization must occur prior to translaryngeal harmony, as it creates the context for the harmony rule; however, we do not formalize this. Trilling optionally applies word-finally in Agole, and obligatorily word-medially in Toende, so there is no interaction with word-final devoicing in Toende.

5 Conclusion

This chapter examined phonological variation in Kusaal from a synchronic dialectological perspective. It showed that segment alternation is a pertinent trigger of phonological variation in the two dialects of Kusaal. The study revealed that while voiced obstruents are observed in word-final positions in Agole, their voicing features are neutralized in Toende, because Toende restricts voiced obstruents in word-final position, creating obvious phonological disparities between the dialects. Similarly, while sequential vowels such as /ɪa/, /aʊ/, /ʊɔ/ and /ʊɔɪ/ are barred in Toende, they are allowed in Agole. As a result, a casual observation of how the speakers use the language could adduce accurate results of whether the speaker speaks the Agole or Toende Kusaal variety. This paper therefore explored several of the alternations that cause variations in the phonology of Kusaal, including alternations at the consonantal and vocalic levels. At the consonantal level, prominent alterations include those that result in word-final devoicing of stops as well as the debuccalization of /s/ to [h] following vowels in Toende. As regards the vocalic alternations, we find several instances of dialectal variations between Agole and Toende in instances where high vowel – non-high vowel sequences become long mid vowels in Toende. These are important markers of dialect variations. Finally, we showed that shortening of vowels due to coalescence only occurs before underlyingly voiceless consonants, not those that are devoiced, necessitating rule ordering.

Samuel A. Asitanga, Anthony Agoswin Musah & Samuel Alhassan Issah

Acknowledgements

References

Abubakar, Abdulhamid. 1982. *Generative phonology and dialect variation: A study of Hausa dialects.* University of London. (Doctoral dissertation).

Abubakari, Hasiyatu. 2018. *Aspects of Kusaal grammar: The syntax-information structure interface.* University of Vienna. (Doctoral dissertation).

Adongo, Helen Atipoka. 2018. *Aspect of Gurenɛ phonology.* University of Ghana, Legon. (Doctoral dissertation).

Ahn, Sang-Cheol. 2000a. Consonantal variation and root faithfulness in affixation. In Vida Samiian (ed.), *Proceedings of the Western Conference on Linguistics*, vol. 12, 9–22. Fresno: Department of Linguistics, California State University.

Ahn, Sang-Cheol. 2000b. Root-faithfulness and coalescence in English prefixation. In Hitoshi Iida (ed.), *Proceedings of the 2000 Seoul International Conference on Language and Computation*, 133–143.

Asitanga, Samuel A. 2021. *A phonological and lexical analysis of dialect variation in Kusaal.* University of Education, Winneba. (MPhil Thesis).

Chambers, Jack K. & Peter Trudgill. 2004. *Dialectology.* 2nd ed. (Cambridge textbooks in linguistics). Cambridge: Cambridge University Press.

Charles-Luce, Jan. 1985. Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics* 13. 309–324.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.

Dinnsen, Daniel A. 1985. A re-examination of phonological neutralization. *Journal of Linguistics* 21. 265–279.

Eze, Patricia. 2019. Dialectal variation: A critical study of Umunze Igbo. *Journal of Language and Linguistics* 6. 51–66.

Gaskell, Gareth M. & William D. Marslen-Wilson. 1996. Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 22. 144–158.

Goldstein, Brian A. & Aquiles Iglesias. 2001. The effects of dialect on phonological analysis: Evidence from Spanish-speaking children. *American Journal of Speech-Language Pathology* 10. 394–406.

Al-Hindawi, Fareed Hameed H. & Nesaem Mehdi Al-Aadili. 2018. The basic tenets of generative phonology. *Journal of Education and Practice* 9. 7–16.

Hudu, Fusheini. 2018. Asymmetries in the phonological behaviour of Dagbani place features: Implications for markedness. *Legon Journal of the Humanities* 29(2). 197–240.

Katamba, Francis. 1989. *An introduction to phonology*. New York: Longman Group UK Limited.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2). 205–254.

Lee, Jae-Young. 2000. *The root-affix faithfulness metaconstraint and English phonology*. Paper presented at spring meeting of the Linguistic Association of Korea.

Mishra, Diwakar & Kalika Bali. 2011. A comparative phonological study of the dialects of Hindi. In *17th International Congress of Phonetic Sciences, ICPhS 2011, Hong Kong, China, August 17-21, 2011*, 1390–1393.

Musah, Anthony Agoswin. 2018. *A grammar of Kusaal: A Mabia (Gur) language of Northern Ghana*. Berlin: Peter Lang.

Musah, Anthony Agoswin, Tony Naden & Michael Awimbilla. 2013. *Handbook of Kusaal orthography: A reading and writing guide*. Tamale: GILLBT Press.

Newton, Brian. 1972. *The generative interpretation of dialects: A study of modern Greek phonology*. Cambridge: Cambridge University Press.

Niggli, Urs. 2014. *The structure of Burkina Faso Kusaal (draft edition)*. Ouagadougou, Burkina Faso: SIL Burkina Faso.

O'Neil, Wayne A. 1963. The dialects of Modern Faroese: A preliminary report. *Orbis* XII(2). 393–397.

Roh, KeunYoung. 2004. English dialects and optimality theory: Prominent feature hypothesis. *SNU Working Papers in English Language and Linguistics* 3. 1–19.

Slowiaczek, Louisa M. & Daniel A. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13. 325–341.

Slowiaczek, Louisa M. & Helena J. Szymanska. 1989. Perception of word-final devoicing in Polish. *Journal of Phonetics* 17. 205–212.

Steriade, Donca. 1986. Locality conditions and feature geometry. *Proceedings of the North East Linguistics Society* 17. 595–618.

Thomas, Alan A. 1967. Generative phonology in dialectology. *Transactions of the Philological Society* 66. 179–203.

Wardhaugh, Ronald. 2010. *An introduction to sociolinguistics*. 6th edn. Oxford: Wiley-Blackwell.

Wolfram, Walt & Natalie Schilling. 2016. *American English: Dialects and variation*. 3rd edn. (Language in Society). Oxford: Wiley Blackwell.

Chapter 6

The syntax of negation in Anii

Frances Blanchette

Pennsylvania State University

Deborah Morton

Pennsylvania State University

This paper presents original field work data from the Bassila variety of Anii. We provide the first description of the Anii negation system and its interactions with indefinite noun phrases. We further analyze this apparently bipartite system as one in which a semantically negative preverbal particle obligatorily co-occurs with a postverbal negative focus marker, which does not itself contribute a semantic negation. We discuss the implications of this for our broader understanding of bipartite negation and the typology of negation systems more generally, suggesting a close relationship between focus and negation in systems such as the one found in Anii.

1 Introduction

In this paper we describe and analyze the syntax of negation in the Bassila dialect of Anii, a Ghana-Togo Mountain (possibly Kwa) language spoken on the border between Togo and Benin in West Africa. While previous studies (e.g., Heine 1968, Schwarz & Fiedler 2011, Morton 2012, 2014, Fiedler 2021) have explored various aspects of Anii grammar, phonology, and information structure, the negation system has not yet been fully described.

Using data from original field work, we illustrate the basic syntax of the negation system, and show how indefinite noun phrases interact with negation to yield certain types of negative meaning. Moving beyond these basic descriptions, we further show how, while at first glance Anii appears to fit in the category of languages with a bipartite negation system, with two negative particles marking a single semantic negation, its similarity to other bipartite negation systems

such as French may be only apparent. We propose an alternative hypothesis in which the postverbal particle does not directly contribute to the semantic negation of the clause, but rather marks the clause as obligatorily focus-marked. The data and analysis provide new information on this relatively understudied language, which has the potential to inform broad and longstanding questions on the syntax of negation and the nature of focus more generally.

2 Data and methods

The original data presented in this paper come from recorded spontaneous conversations, written texts, and planned elicitation sessions with Anii native speakers. The elicitation data presented here were largely collected through in-depth sessions with 5 speakers, and some of the more complex cases were checked with 3–4 other speakers. There were relatively few cases of disagreement, and if these arose during elicitation sessions, speakers worked them out with each other and consulted elders in the community to resolve confusion. Elicitations were conducted primarily in French, which speakers learned as an academic language. They are all multilingual and report high levels of switching between Anii, French, and other languages in their daily lives. All are literate in Anii, with high levels of metalinguistic awareness. Elicitation sessions therefore involved collaborative transcription between the interviewer and the research participants. For each variable, participants were asked to translate constructed French examples into Anii, or to choose between minimal pairs or triplets of Anii sentences. Methods for conducting the elicitation sessions adhered as closely as possible to the Small N Acceptability judgment Paradigm (SNAP) (Mahowald et al. 2016): 5–10 examples were included for each aspect of negation observed (e.g., 5 fragment answers to *wh*-questions), and examples were presented to at least 5 people.

3 The basics of Anii negation syntax

The basic word order in Anii is SVO, with obligatory agreement with the subject noun-class, as shown in (1).[1,2]

---

[1]For expository purposes, the data in this paper are given in Anii orthography (Zaske & Atti Kalam 2014) unless otherwise noted. The use of orthography allows us to illustrate the relevant information without introducing complex yet orthogonal linguistic information such as the interaction of lexical and grammatical tone in verb stems. More information on the phonetics and phonology of Anii can be found in Morton (2014).

[2]Note that for noun class CL, we follow the noun-class naming conventions in the Anii orthography (Zaske & Atti Kalam 2014), where the noun classes are named with letters.

(1) a-sna   a   kə u-bu.
CL.Ǝ-dog AGR.CL.Ǝ hit CL.E-snake
'The dog hit the snake.'

Like many other negation systems (e.g., French, Jespersen 1917; Gbe, Aboh 2010; among many others; see Bell 2004 for a review), the negation system in Anii can be described as bipartite, with both a preverbal negative marker *k*V and a postverbal negative marker *na* that obligatorily co-occur in a negative sentence (Morton 2014, Schwarz & Fiedler 2011).[3] The example in (2) illustrates how these markers typically appear relative to basic Anii word order, where the subject is a full nominal phrase.[4]

(2) Ba-smprə   kə ba   bʊŋa   wʊɖa ɩkashɩ   na.
CL.Y-women NEG AGR.CL.Y FAR.PST have CL.B.strength NEG
'Women didn't have power.'

The preverbal negative marker *k*V, realized as *kə* in (2), appears after the full nominal subject *basmprə* 'women', and precedes the agreement marker *ba*.

When the subject is not a full nominal phrase, the preverbal negative marker *k*V appears preceding the subject marker. This is illustrated in (3)–(4).

(3) Kə ba   na nyəm   na.
NEG 3PL.IRR IPFV steal.IRR NEG
'One doesn't steal.'

(4) Kə ba   na kiɖe   bʊ-ja   na...
NEG 3PL.IRR IPFV watch.IRR CL.Ʊ-years NEG
'They didn't look at age...'

It should be noted that there are two instances of the particle *na* in (3) and (4): the imperfective marker, which precedes the verb, and the postverbal (and in these cases sentence-final) negative marker. These two instances of *na* are readily distinguishable, however, since they are not homophonous: the imperfective marker is low-toned and the postverbal negative *na* is high-toned.[5]

---

[3]See Morton (2014: 376-377) for an analysis of the constraints on vowel insertion for *k*V.

[4]The marker *bʊŋa* is glossed as far past, but is not a tense marker. For further information on this marker, please see Morton (2014).

[5]There is a *na* with a high tone meaning 'and' or 'with' that is in fact homophonous with the postverbal negative marker, but it is easily distinguished from this marker by its position in the sentence.

Another basic element of Anii negated sentences is the presence of irrealis morphology. This can be observed by comparing the simple declarative in example (5a) with its negated form in (5b). (These data are given in IPA to illustrate relevant tonal differences, setting aside orthogonal surface-level phonetic implementation. See Morton (2014) for more detailed data and discussion.)

(5) a. [ń   sàrà]
1SG.SBJ walk
'I walked.'

b. [k�à má   sàrá   ná]
NEG 1SG.SBJ.IRR walk.IRR NEG
'I did not walk'

Contrasting (5a) with (5b), we see that the negated form has a distinct subject marker and a high tone on the verb. Morton & Blanchette (submitted) analyze these elements as components of a single irrealis morpheme, which is introduced in the head of a modal phrase. (See Cristofaro 2012, de Haan 2012 for a review of irrealis.) In conjunction with the distribution of subjects, this leads us to propose that the preverbal negative marker is introduced in a negative phrase between TP and the irrealis-introducing modal phrase, as in Figure 1.
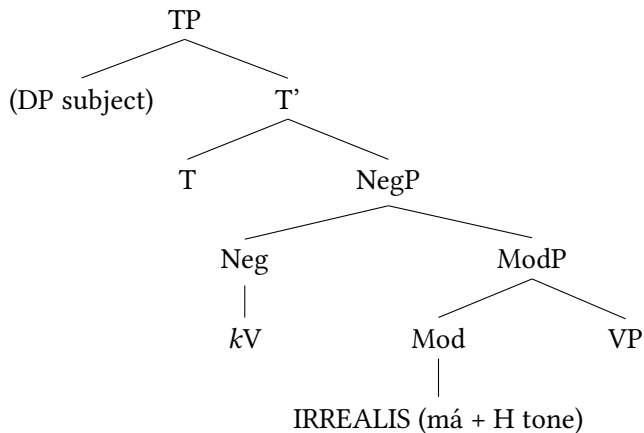
Figure 1: Structure of the Anii Clause

In Figure 1, the preverbal marker *k*V is introduced in the head of a negative phrase (NegP). This NegP is selected by a tense phrase (TP), and it selects for the irrealis-introducing modal phrase. It is this selection relation that Morton &

Blanchette (submitted) argue serves to explain the obligatory co-occurrence of negation and irrealis in Anii sentences. (The Anii system of irrealis marking and Morton and Blanchette's formal analysis of it are complex and it is beyond the scope of this paper to include further detail here.)

4 Indefinite noun phrases and negation

As noted by van der Auwera & Van Alsenoy (2016), while indefinite noun phrases differ crosslinguistically in terms of how they interact with negation, these interactions are often excluded from descriptions of negation systems. However, the behavior of indefinites gives us important insight into how negation systems work. Languages may or may not display Negative Concord, in which morphologically negative noun phrases and negative markers enter into a dependency relation and contribute to the same semantic negation. French (Déprez 1999) is an example of a Negative Concord language:

(6) Je ne   vois rien.
    1SG NEG see nothing
    'I see nothing.'

In example (6), the negative particle *ne* and the noun phrase *rien* 'nothing' contribute to the same semantic negation. Such Negative Concord systems are relatively common across the world's languages, particularly in Romance, though, as van der Auwera & Van Alsenoy (2016) show, they are not the predominant pattern crosslinguistically.

Languages may also have noun phrases that have no overtly negative morphology, but which are sensitive to the polarity of a clause. Such noun phrases, known as Negative Polarity Items, appear to require the presence of a preceding and c-commanding negation or downward entailing operator (Ladusaw 1979). The example in (7) (from Collins et al. 2017: 2) illustrates a Negative Polarity Item construction in Ewegbe:

(7) Kofí *(mé)-kpɔ́ ame   aɖéké o
    Kofi NEG-see   person any   NEG
    'Kofi didn't see anybody.'

In (7), the phrase *áɖéké* 'any' is dependent on the preceding negative marker *mé*, without which the sentence is rendered unacceptable. On the basis of data such as this, forms like *áɖéké* are hypothesized to be Negative Polarity Items that partake in a syntactic dependency with a preceding negation.

Given the possibilities for interactions between noun phrases and negation, it is relevant to ask whether dependencies such as Negative Concord and Negative Polarity are present in Anii. The results of our research suggest that Anii has neither a Negative Concord nor a Negative Polarity system, and it also does not have negative quantifiers that independently contribute negative meaning to a clause (e.g., English *nothing*). Instead, phrasal negative meanings are derived through interactions with preverbal *k*V and a set of non-negative indefinites.

Table 1 contains an inventory of common indefinite noun phrases, and provides information about their morphological composition. Many of the forms include the morpheme *ɖən* 'one', which can also be used freely as an indefinite article.[6] [7]

Table 1: Composition of Anii indefinites

Form	Root(s)	Translation
akoɖən	*ako* 'thing', ***ɖən* 'one'**	'something'
arɛɖən	*arɛ* 'person', ***ɖən* 'one**	'someone'
gaɖən	***ɖən* 'one'**	'something'
gaɖu gaɖən	*ɖu* 'place', ***ɖən* 'one'**	'somewhere'
ʋɖən	***ɖən* 'one'**	'someone'
ʋwor	*wor* 'neighbor'	'someone' (friend)
gɪrɛpi	*gɪpi* 'seed', *arɛ* 'person'	'big/strong person'

Anii indefinite noun phrases such as those in Table 1 may appear under negation, and they may also appear in sentences with no preceding negation or other downward entailing operator, as in the examples in (8).

---

[6]While it appears that the initial morpheme in Anii indefinites is a noun-class marker, this may be only apparent. For example, if *akoɖən* and *gaɖən* both mean 'something', it is not clear why they do not share an initial morpheme. On the other hand, when used as an indefinite article in an analytic noun phrase, the morpheme *ɖən* co-occurs with the class marker of the noun it describes. We set this issue aside here.

[7]While we ultimately conclude that Anii does not have a general Negative Concord or a Negative Polarity system, it is possible that the form *gɪrɛpi* 'big/strong person', which derives from a root with opposite meaning (*gɪpi* 'seed', i.e., something or someone very small), evolved into its current meaning through interactions with the negation system. Our research further suggests that synchronically this form may have some polarity sensitivity. Since this appears to be a peripheral item that does not reflect the central tendencies of the Anii negation system and its interactions with indefinites, we set this matter aside for future research.

(8) a. K'   a   shee   amʊ   akoɖən   na.
         NEG 3SG.SBJ.IRR give.IRR 1SG.OBJ something NEG
         'She did not give me anything.'
         Lit. 'She did not give me something.'

     b. A   shee amʊ   akoɖən.
         3SG.SBJ give 1SG.OBJ something
         'She gave me something.'

In (8a) the phrase *akoɖən*, 'something' is preceded and c-commanded by the negative particle *k*V, and the prose translation reflects a meaning equivalent to the English Negative Polarity Item 'anything'. However, as (8b) shows, this same phrase can also be used with no preceding negation. Comparing Anii examples like (8) with Ewegbe examples such as (7), we can conclude that unlike Ewegbe, Anii indefinites are not sensitive to the polarity of the clause.

Another context type relevant to our understanding of interactions between noun phrases and the negation system is negative fragment answers to *wh*-questions. Such contexts are often used as diagnostics for polarity sensitivity, as illustrated by the following English example:

(9) What did they say? Nothing./*Anything.

In example (9), the term *anything* is unacceptable as an answer to the *wh*-question because it is not preceded by a c-commanding negation or downward entailing operator, a reflection of its polarity sensitivity. The term *nothing*, on the other hand, is perfectly acceptable in such contexts and independently contributes the meaning of a negative quantifier, in the absence of any other negation marking.

Anii indefinites can also be used as fragment answers to *wh*-questions. Since they are not polarity sensitive and have the semantics of non-negative indefinites, in order to contribute a negative meaning in these contexts, they must be accompanied by another element that negates them. Interestingly, the particle that surfaces in such cases is neither preverbal *k*V nor postverbal *na*, but rather *baa*, as shown in (10).

(10) Context: What did they say?

      a. *(baa) akoɖən
          baa   something
          'nothing'

b. (*baa) fɔɩ
    baa   zero
    'nothing'

Contrasting (10a) with (10b), we see that the indefinite *akoḍən* 'something' re-
quires the particle *baa* to trigger the negative meaning 'nothing', whereas non-
indefinite *fɔɩ* 'zero' is incompatible with *baa*, since it inherently contributes the
meaning of 'nothing'. At first glance such examples suggest that *baa* is contribut-
ing a semantic negation. However, independent evidence suggests that *baa* may
not actually be negative. Consider the examples in (11).

(11)   a. Context: Describing a trip from Bassila to Cotonou (the biggest city
           in Benin)
           Gɩ   ta   kpa **baa** m-pa   m-paŋa   njɩ,
           1PL.SBJ if/when arrive baa CL.Ɛ-village AGR.CL.Ɛ-each CL.Ɛ.FOC
           aŋ   tɩ   taba, ba-ḍe   ba   tɩ   lə   amʊ ka
           1SG.EMPH IPFV ask   CL.Y-these AGR.CL.Y IPFV say 1SG.OBJ POSS
           n-nyɩḍa
           CL.Ɛ-name
           'When we would arrive at any village, I would ask and they would
           tell me its name.'
       b. Context: Describing mango season
           Ʊ   ta   kpa **baa** ŋka na, ɩ-maŋgo   ba   ḍa
           2PL.SBJ if/when arrive baa where FOC CL.W-mango AGR.CL.W be.there
           ḍa
           be.there
           'When you went anywhere, there would be mangoes.'

In examples (11a) and (11b), *baa* is not behaving as a negation marker. Instead, it
seems to be acting in concert with another element, *mpaŋa* 'each' in (11a) and *ŋka
na* 'where' in (11b), to refer to a set of elements: the set of all villages in (11a), and
the set of all places in (11b). Moreover, speakers agree that the *baa* that appears in
examples such as those in (11) is the same *baa* that appears in negative fragment
answers such as those in (10).

    Given the data above, we might say that the particle *baa* behaves like a non-
polarity sensitive version of English *any*, which can also take on the meaning of
a universal quantifier (as in free choice *any*, Dayal 2004), a negative quantifier (as
in negative polarity item constructions) or a non-negative indefinite (e.g., under
conditionals) (Collins & Postal 2014). Providing the precise syntax and semantics

of *baa* is beyond the scope of this paper. For now we merely observe that it acts as a negator of indefinites in fragment answers to yield the semantics of a negative quantifier. Future investigation of this particle can build on the description of indefinites provided in this paper, and has the potential to inform theories of ellipsis and quantification more generally.

Scope relations are another point of potential crosslinguistic variation for interactions between negation and indefinites (Matthewson 1998). In Anii, nonnegative indefinites in matrix subject position may take wide or narrow scope with respect to negation, as shown in (12).

(12) a. Context: A woman washes her clothes, puts them in a basket, and hangs them out to dry, but accidentally leaves one washed item in the basket. When she discovers the item she neglected to hang, she says:

Akoḍən   k'   ɪ   kʊ   na.
something NEG 3SG.INANIM be.dry NEG

'Something is not dry.'

   b. Context: A woman washes her clothes and puts them in a basket, but forgets to hang them out to dry. When she discovers them later, she says:

Akoḍən   k'   ɪ   kʊ   na.
something NEG 3SG.INANIM be.dry NEG

'Nothing is dry.'

In (12a), the indefinite *akoḍən* 'something' precedes the negation and takes surface scope, yielding a meaning in which there exists something that is not dry (∃ > ¬). In (12b), the indefinite still precedes the negation, but the context yields an inverse scope reading in which the negation, and not the indefinite, takes wide scope (¬ > ∃). In this case, the sentence is more readily translated as equivalent to an English (or a French) sentence with a negative subject ('nothing').

Indefinites in Anii are also used in constructions that would be stated as *there*-existentials in languages like English and French. The example in (13) contains the indefinite *gaḍu gaḍən* 'somewhere' in subject position.

(13) Ga-ḍu   ga-ḍən ka ga   ḍa   a-shee   sukuru k'
CL.C-place CL.C-one NEG AGR.CL.C be.there.IRR INORD-give school POSS

ʊ-pi   a-cə   tɪ   lee sukuru na
CL.A-child INORD-go finally do school NEG

'There was no place for the schoolchildren to go study.'
Lit: 'Somewhere was not there for the schoolchildren to go study.'

Example (13) is interpreted similarly to (12b), with the negation taking wide scope over the indefinite *ga-ɖu ga-ɖən* 'somewhere', with a translation roughly equivalent to 'nowhere'.

5 The position of the postverbal marker

We now provide an initial description of the syntax of the postverbal marker *na*. Example (14) contains two negative sentences with adverbial adjuncts (*guyo laŋ* 'on the tree' in the first sentence, and *atəŋ* 'on the ground' in the second). In both sentences, the particle *na* follows the adjunct.

(14) Ka a   sə   [gʊ-yo laŋ] na, ka a   sə   [atəŋ]
      NEG 3SG.SBJ sit.IRR CL.Ɛ-tree on   NEG NEG 3SG.SBJ sit.IRR on.the.ground
      na.
      NEG
      'He didn't sit on a tree, he didn't sit on the ground.'

Unlike in (14), in cases where an adverbial contributes a durational meaning, *na* may precede it. Durational adverbial phrases tend to be phonologically heavy in Anii, as with the phrase *halɩ a-ŋɔrɔ a-pə́ləmə* 'for the whole month' in example (15):[8]:

(15) Kali k' a   nyem na k' a   shɩŋ jɩ   ʊ-jɩʊ
      Kali NEG 3SG.SBJ.IRR drink.IRR and NEG 3SG.IRR either eat.IRR CL.E-food
      na [halɩ a-ŋɔrɔ   a-pə́ləmə].
      NEG since CL.A-month AGR.CL.A-entire
      'Kali didn't drink or eat [for the whole month].'

In the interpretation of (15), the adverb is outside of the scope of negation: there was a whole month that was such that Kali did not eat or drink. This suggests that the position of the adverbial could be motivated by scope concerns. However, the following data illustrate that this is not the case:

(16) a. Arɛɖən ka kɔɔ tʊr   gʊ-bɔ   ŋgʊ-ɖe na bʊ-ja
          no-one NEG again read.IRR CL.Ɛ-paper CL.Ɛ-that NEG CL.Ʊ-years
          bʊ-nyɩʊ.
          AGR.CL.Ʊ-two
          'No one has read that newspaper for two years.'

---

[8]This example is particularly interesting since it shows two *k*V markers with one *na*, and involves VP coordination. These phenomena are discussed in more detail in Section 6 below.

b. Arɛɖən ka kɔɔ tʊr   gʊ-bɔ   ŋgʊ-ɖe bʊ-ja
   no-one NEG again read.IRR CL.Ɛ-paper CL.Ɛ-that CL.Ʊ-years
   bʊ-nyɯ   na.
   AGR.CL.Ʊ-two NEG
   'No one has read that newspaper for two years.'

While example (16a), with *na* preceding the adverbial, is the preferred variant, speakers also allow for the order in (16b). In both cases, the adverbial is interpreted outside the scope of negation: there have been two years during which no one read the newspaper. This suggests that the appearance of *na* preceding durational adverbs is not motivated by scope. Below in Section 6 we propose a syntactic analysis that is compatible with the assumption that this positioning is instead due to the phonological heaviness of the durational adverbial phrase.

We now turn to the position of the postverbal particle in questions. Polar questions in Anii obligatorily occur with the clause-final question particle *aa*. When negated, *na* immediately precedes the question particle:

(17) a. Ʊ ce Kʊtɔnʊ aa?
         2SG go Cotonou Q
         'Did you go to Cotonou?'

      b. K' á   ce   Kʊtɔnʊ na aa?
         NEG 2SG.IRR go.IRR Cotonou NEG Q
         'Did you not go to Cotonou?'

Examples such as (17) illustrate how Anii shows some behaviors that align it with some of the geographically proximate and likely (though not closely) related Gbe languages as described in Aboh (2010). In these languages, a series of particles marking elements generally understood to be part of the C-domain, such as topic, focus, and question markers (Rizzi 1997), cluster together toward the end of the clause. In languages with a postverbal negative particle, including those with an apparently bipartite negation system similar to Anii, the postverbal particle also immediately precedes these elements. Example (18), taken from Aboh (2010: 123), illustrates this in Ewegbe:

(18) Kòfí mé-xlẽ   àgbàlẽ ò   à?
      Kofi NEG-read book   NEG Q
      'Didn't Kofi read a book?'

The structure of Anii example (17b) appears identical to Ewegbe example (18): the postverbal negative marker immediately precedes the sentence-final question

particle. Aboh (2010: 131) takes data such as this to support the conclusion that the sentence final marker serves a pragmatic function akin to "speech act modality". (See Biberauer 2009, 2015 for a related analysis of Afrikaans, and Beyer 2009 for an analysis of links between focus and postverbal negation in a number of West African languages.) Below in Section 6, we extend this idea to Anii, proposing that, like in Ewegbe, the Anii postverbal marker is part of the C-domain, where it serves to focalize the negated proposition.

But before presenting our analysis, we illustrate the distribution of post-verbal *na* in multi-clausal sentences, beginning with a coordinate structure:

(19)   a.   Kə   ma   ŋə   am-a-ŋana   na, (na) kə   ma   shɪŋ
            NEG 1SG.IRR see.IRR 1SG.POSS-CL.A-mother NEG (and) NEG 1SG.IRR nor

            kɔɔ   rəŋə   am-pal   ʊ-wələ   na.
            again hear.IRR 1SG.POSS.CL.A-younger.sibling CL.E-voice NEG

            'I didn't see my mother or hear my younger sibling.'

       b.   Munifatu   k'   a   jɪ   ʊ-jɪʊ   na, (na) k'   a
            Mounifatou NEG 3SG.IRR eat.IRR CL.E-food NEG (and) NEG 3SG.IRR

            raʊ   a-nyanʊ   na.
            wash.IRR CL.T-dishes NEG

            'Mounifatou did not eat, nor did she wash the dishes'

The coordination in (19) occurs at the clausal level, and each clause contains a negated proposition. In such cases, the syntax of negation is the same as in non-coordinated structures, with a preverbal *k*V and a postverbal *na* marking each propositional negation.

Subordinate clauses can be marked in a variety of ways. The example in (20) illustrates a negated proposition with a subordinate clause introduced by the complementizer *wàà*.

(20)   Context: A bird flew up to God to ask for rain to end a drought. God responds:

       Aɖe   ka   a   jɔ   wàà ga-tna   ka-gʊ-jarɪ   gʊ
       3SG.DEF NEG 3SG.IRR know.IRR that CL.C-earth POSS-CL.Ɛ-king AGR.CL.Ɛ

       tɪ   lige   aɖe   laŋ na.
       IPFV remember 3SG.OBJ on NEG

       'He (God) did not know that the king of the earth remembered him.'

In example (20) postverbal *na* occurs sentence finally, following the subordinated clause. The preverbal negative marker precedes the matrix verb, and the negation takes matrix scope: the act of knowing (as opposed to remembering) is negated.

When an embedded clause introduced by *wàà* is negated, as in (21), postverbal *na* again occurs at the end of the clause.

(21)   Context: Nouhoum is telling a story about how he was not let into school at the proper age because he was too short, so his father had to intervene:

    Ama-wɛɛ   aḍe   tɪ   lə am-ɪ-mɛɛtr   wàà k

    1SG.POSS-CL.B.father 3SG.DEF IPFV say 1SG.POSS-CL.W-teachers that NEG

    a   lee   ʊ-pi   na

    3SG.IRR do.IRR CL.A-child NEG

    'My father told my teachers: he is not a child.'

Although the negation takes embedded scope in (21), we are unable to determine whether *na* is contained within the embedded clause. Structures with the clause subordination marker *ma* serve to inform this question.[9] In such structures, exemplified in (22), *ma* marks the end of the subordinate clause, and the beginning of the clause may also be marked with a high tone (written orthographically only on third-person pronouns). These clauses typically occur sentence-initially, as in the following non-negated example:

(22)   Context: You have a tradition of eating mangoes when you are angry.

    N cam ʊ-nyana   ma, n   kəm ɪ-maŋgɔ

    1SG hold CL.E-anger SUB 1SG.SBJ suck CL.W-mangoes

    'Since I was angry, I ate mangoes.'

In (22), the end of the clause interpreted as 'since I was angry' is marked by *ma*, which indicates its subordinate meaning relative to the matrix clause: the speaker's state of being angry happened subordinate to the eating.

When subordinate clauses with *ma* are negated, the particle *na* immediately precedes this subordination marker. Example (23) is the negated version of (22):

(23)   Context: You have a tradition of eating mangoes when you are happy.

    kə ma   cam   ʊ-nyana   na ma, n   kəm ɪ-maŋgɔ

    NEG 1SG.IRR hold.IRR CL.E-anger NEG SUB 1SG.SBJ suck CL.W-mangoes

    'Since I was not angry, I ate mangoes.'

---

[9]This *ma* is low-toned, which is different from the *ma* that is the first-person singular subject in irrealis clauses (including negative sentences). The subject marker *ma* has a high tone. This tone difference is not marked orthographically because the difference is already clear from the position in the sentence.

In (23), postverbal *na* resides at the final edge of the negated clause (a proposition), but before the clause subordination marker *ma* which links it to the matrix clause.

A similar pattern is found in relative clauses:

(24) Context: Hakimou is telling a story about how he used to play at his friend's house as a child. His friend is Fulani, a different ethnic group.

  a-rɛ   n-ḍee   ka cee a   yɔ   amʊ   na ma,
  CL.Ǝ-person AGR.CL.Ǝ-REL NEG PERF 3SG.IRR know.IRR 1SG.OBJ NEG SUB

  a   na yɔ   yɔ   wàà n   lee a-fəlandɪja   na
  3SG.SBJ IPFV know know that 1SG.SBJ do/be CL.Ǝ-Fulani.person CL.Ǝ.FOC

  'A person who didn't know me, they would think that I was a Fulani person.'

In (24), the negation scopes within the relative clause in which it resides, and *na* precedes the clause final subordination marker *ma*.

Examples such as (23) and (24) might be taken to suggest that, in general, *na* precedes the clause subordination marker. However, consider the following example:

(25) Context: Malookia is telling a story about his middle school exams. At first, the school authorities said that everyone had passed the exams, but later they announced that half those who passed actually did not, and would be returned to middle school rather than continuing to the next school level. Malookia says "But me, I knew my name would be there on the passing list." Then he continues:[10]

  Amʊ-n-nyɪḍa   kɪ   n   ḍa   m-baḍee bá
  1SG.POSS-CL.F-name NEG AGR.CL.F be.there.IRR CL.F-REL 3PL.SBJ.SUB

  kɔɔ   na   ma ma na.
  return with PL   SUB NEG

  'My name would not be there with those that were returned.'

---

[10]Note that the word for 'name', *nnyɪḍa*, is noun class F, so the subject here is agreeing with that word, the *n* here is low-toned, where the *n* meaning 'I' is high-toned. Note also that there is a *ma* that is a plural marker. This marker is low-toned, like the clause subordination marker, and is used to mark plurality on verbs in some cases. We have checked with multiple native speakers as to which *ma* is which, and we are clear that the clause subordination marker is the second one.

In (25), *na* follows the clause subordination marker. Note that in this case, the relative clause contains the non-negated proposition that some (unnamed) group of people returned something (which we know to be names from the sentence context). The negated proposition in this example is instead the matrix proposition: the act of being there is negated.

Synthesizing these data, in cases where there is more than one proposition, the location of *na* coincides with the proposition over which the negation takes its scope. We therefore conclude that *na* appears at the edge of the proposition that is being negated.

We have shown that *na* precedes phonologically heavy adverbials, question particles, and sometimes the clause subordination marker *ma*, but otherwise occurs clause finally. Examination of multiclausal contexts suggests that *na* marks the final edge of the proposition over which negation takes its scope. This fact will be relevant to our analysis, which groups *na* together with elements that mark discourse-related functions.

6 Bipartite negation and focus marking

Since Jespersen (1917), bipartite negation systems have been understood to take part in a diachronic cycle in which negative meaning is gradually transferred from one negative particle to another. French, in which the preverbal negator *ne* and the postverbal negator *pas* together mark a single sentence negation, is a prototypical example of this. The Jespersen cycle begins with a single negator (*ne* in French) as the only element contributing semantic negation. A second negative element is then added (*pas* in French), and it is largely accepted that its purpose is to reinforce the first, which has undergone semantic weakening. The system then passes through a stage in which both negators are obligatory, contributing to the same semantic negation. Eventually the semantic negation is transferred to the second negator, and the initial one gradually disappears from the language.

Grouping Anii with such bipartite negation systems would lead to the conclusion that both preverbal *k*V and postverbal *na* are contributing to negating the sentence, since both are obligatory. However, there is independent evidence to suggest that postverbal *na*, and Anii more generally, might not participate in the Jespersen Cycle, and that the obligatory nature of this particle instead reflects a particular pragmatic status for negative propositions. (See Biberauer 2009, 2015 for a similar conclusion for Afrikaans, based on different types of evidence.)

Consider again the example in (15), repeated here as (26).

(26) Kali k' a   nyem   na k' a   shɪŋ jɪ   ʊ-jʊ
     Kali NEG 3SG.SBJ.IRR drink.IRR and NEG 3SG.IRR nor eat.IRR CL.E-food
     na [halɪ a-ŋɔrɔ   a-pə́ləmə].
     NEG since CL.A-month AGR.CL.A-entire
     'Kali didn't drink or eat [for the whole month].'

As noted above, (26) contains a phonologically heavy adverbial phrase preceded by *na.* This example is also a coordinate structure, with the coordinated verb phrases headed by *nyem* 'drink' and *jɪ* 'eat' sharing a subject *Kali.* The semantic interpretation of this sentence contains two negations: it is not the case that Kali ate, and it is not the case that Kali drank. Note, however, that these two semantic negations are marked by two *k*V particles, each modifying one of the verbs, but only a single *na.* See Aboh (2010) for similar data from Gbe.

The following example follows a similar pattern:

(27) ʊɖən   k' a   na rəŋə   k' akoɖən   gɪtenshile na
     someone NEG 3SG.SBJ.IRR IPFV listen.IRR NEG something evening NEG
     'No one listens to nothing in the evening.'
     (='Everyone listens to something.')

Example (27) consists of a single clause with two semantic negations: it is not the case that there exists someone who does not listen to music in the evening.[11] The sentence above contains two semantic negations in a single clause, which is logically equivalent to an affirmative. Crucially, this is again achieved by the inclusion of two instances of preverbal *k*V, but only one postverbal *na.*

---

[11]In example (27), the particle *k*V appears to attach directly to the indefinite *akoɖən*, suggesting that it takes its scope directly over this indefinite, as opposed to at the sentential level. This is unusual, since constituent negation—negation that scopes below the propositional level—is not common in Anii. This can be seen in the following example, where something that would commonly be expressed with constituent negation in English requires a full clause in Anii:

(i) N   ɖoo   kə ma   kara   ama-pooloo na
    1SG.SBJ go.out NEG 1SG.SBJ.IRR wear.IRR 1SG.POSS-coat NEG
    'I went outside with no coat.'
    (='I went outside, I didn't wear my coat.')

Example (27) suggests that negation can attach directly to indefinites within a proposition to contribute the meaning of a negative quantifier. We set aside the question of how such instances of constituent negation should be represented and the extent of their productivity for future research.

The same phenomenon occurs in biclausal (28). This example contains an instance of preverbal *k*V in both the matrix and the embedded clause, but only one *na*:

(28) Kə ma   lə/faŋa   wàà Kofi k' a   jɩ
      NEG 1SG.SBJ.IRR say.IRR/think.IRR that Kofi NEG 3SG.SBJ.IRR eat.IRR
      akoɖən   na
      something NEG
      'I didn't say/think that Kofi didn't eat something/anything.'

In cases of subordination containing a relative clause, again both the matrix and the subordinate clause are semantically negated by *k*V:

(29) Kə ma   sɔlɔ   ɩ-film   mba-ɖee kə ba   wʊɖa
      NEG 1SG.SBJ.IRR love.IRR CL.W-film CL.W-REL NEG 3PL.SBJ.IRR have.IRR
      kokoroko mba-ɖee n   yɔ   ma na.
      CL.B.hero CL.W-REL 1SG.SBJ know SUB NEG
      'I don't like films that don't have heroes who I know.'

Examples such as these suggest that *k*V is the true semantic negator in the clause, since more than one *na* is not necessary to express more than one semantic negation. This raises a number of interesting questions regarding the role of the postverbal negative particle. Given the strong obligatory nature of this particle, it does not appear to be weakening in any way (as in, e.g., French preverbal *ne*).

We would like to suggest instead, on both empirical and theoretical grounds, that the particle *na* is a negative focus marker. The obligatory occurrence of this negative focus marker results from a general requirement in Anii that negative sentences be focalized. This hypothesis follows and extends the analysis in Aboh (2010: 131), where Fongbe is hypothesized to have a peripheral clause-typing negation marker whose main contribution to sentence meaning is at the discourse-pragmatic level. We extend this analysis to suggest that the discourse-pragmatic function of the peripheral marker in Anii, and possibly other languages, is focal in nature. Before providing further theoretical support for this analysis, we illustrate some data that point toward a more general focal status of *na*.

There are a number of focus-marking strategies in Anii, including particles, fronting, changes in aspect marking, and clefting and nominalization of verbs. When focus particles are used, often in combination with fronting, they typically indicate object focus (Schwarz & Fiedler 2011, Morton 2014). Consider the examples in (30).

(30)   a.   Context: A friend and I see a child put something in her mouth. I ask
        my friend if the child ate a piece of meat.

        Gɩ-ca   jɩ   a   jɩ
        CL.Đ-bean FOC.CL.Đ 3SG.SBJ eat

        'She ate a BEAN.'

     b.   Context: A friend and I see a child eating. I ask my friend if the child
        ate rice.

        A-ca   nɩ   a   jɩ
        CL.T-bean FOC.CL.T 3SG.SBJ eat

        'She ate BEANS.'

In (30a), the singular object *gɩca* 'bean' is focused by the particle *jɩ*, and in (30b)
the plural object *aca* 'beans' is focused by *nɩ*. The focus particles are morpholog-
ically distinct because they must agree in noun class with the noun they modify.

    The examples in (31) (modified from Schwarz & Fiedler 2011) show that in cases
where the subject noun class is unclear, the focus particle surfaces as *na*. In such
cases, the focus particle has a high tone and is thus fully homophonous with the
na in negative sentences:

(31)   a.   maakɔ na ʊ-pi   a-nyɩʊtaja   a   pɩ   na   afal
        what   FOC CL.A-child AGR.CL.A-second AGR.CL.A come with home

        'What did the second child bring home?'

     b.   ŋkəŋ na a   təŋ ʊ-pur   ʊ-ŋono gʊ-yá
        there FOC 3SG.SBJ send CL.A-child.POSS CL.A-old CL.Ɛ-market

        'That's when she sent her oldest child to the market.'

In (31a) the particle *na* serves to focus the *wh*-phrase in object position, and in
(31b) it focuses a pronominal element (*ŋkəŋ*). Importantly, neither of these ele-
ments has a noun class. This suggests that *na* may serve as a general or default
focus marker. While this remains to be tested more explicitly, we take it as an
initial clue toward understanding the role of *na* in negative sentences.

    Given these data, we would like to suggest that, in addition to serving as a
general focus marker in non-negative contexts, the particle *na* acts as a sort of
negative focus marker in negative sentences. Under this analysis, the presence
of *na* is not triggered by the need for a semantic negator (contra what would be
expected under a Jespersonian analysis). Instead, we propose that its presence
is motivated by the inherently focal status of negation in Anii. Its obligatory na-
ture can then be understood in terms of a requirement that negative propositions

be marked as focused. Under this hypothesis, it is possible that the negative focus marker is in fact underlyingly the same as the homophonous general focus marker that appears in (31). We leave further investigation of this possibility to future research.

The analysis suggested above raises the question of why negation might be obligatorily focused in Anii. To begin to address this question, we first consider what is meant by the term focus. As pointed out by López (2009), standard definitions of focus such as "new information" are relatively broad, and may fail to make precise predictions regarding what should be focused in a given language. López proposes instead that the traditional notion of focus should in fact be construed in terms of the related notion of contrast. Under his definition, contrast opens up a variable and simultaneously provides a value for that variable. For example, in (30b), the term *aca* 'beans' is part of a contrast domain of all the things the child might have eaten, and focusing (or contrasting) the item leads to its selection from that set.

In extending this analysis to English negation, Blanchette (2013) proposes that negation is inherently a form of contrast. The contrast domain for negation consists of the affirmative and negated forms of the negative-marked proposition, and the negated form is selected. In the context of the Anii data, it is possible that what we are seeing is a lexification of this inherently contrastive meaning of negation.

With respect to syntax, if *na* is a negative focus marker, then we would expect it to reside in the C-domain. Since this domain is peripheral, occurring at the beginning of a structure, this raises the question of why the marker surfaces sentence-finally. To explain this, we follow Aboh (2010) and others in hypothesizing that the entire Anii proposition undergoes fronting to a specifier position, yielding the surface order seen in negative sentences. While Aboh (2010) hypothesizes that in Gbe languages, such fronting occurs into the specifier of a NegP within this domain (as part of a successive pied-piping operation), we propose that Anii negated propositions are fronted into the specifier of a Focus Phrase (FocP), as shown in Figure 2 (which is example (3) repeated).

In the structure in Figure 2, the entire proposition *kə ba na nyəm* is fronted into the specifier of a focus phrase, yielding the sentence-final position of *na*.[12]

Recall from Section 5 that while most adverbials precede the sentence final particle in negated sentences, durational adverbials, which are phonologically heavy,

---

[12]We leave open for future research whether a successive pied-piping process occurs, such as the one proposed in Aboh (2010). This may be needed for questions, in which the question particle also occurs in a sentence-final position, and other similar sentence types.
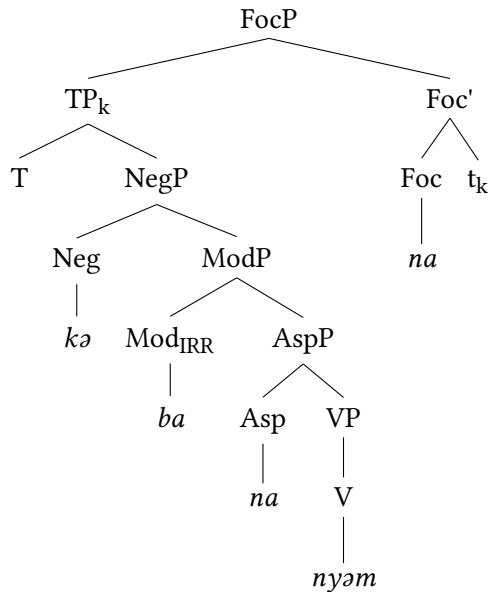
Figure 2: Tree for *kə ba na nyəm na*, 'One doesn't steal.'. The *ba* here is an irrealis subject marker, and may actually originate lower in the tree, though that is not crucial to the structures discussed here. For more discussion of the syntax of irrealis, please see Morton & Blanchette (submitted). This *na* is an imperfective marker.

tend to follow this particle. The fact that these phonologically heavy elements behave differently from other elements with similar syntactic status provides some additional support for the movement analysis we have presented here. Specifically, these can be analyzed as involving an adjunction structure in which the heavy adverbial is stranded in sentence-final position following raising of the focused negated proposition. The relevant structure is illustrated in Figure 3 (the sentence is example (15) repeated).

7 Summary and discussion

In this paper we have provided the first detailed description of the Anii negation system, including interactions between negation and indefinites. We have shown that, with respect to the preverbal negative marker, Anii places negation following TP. We briefly touched upon the relationship between negation and the notion of irrealis, which obligatorily co-occur in Anii, and pointed toward

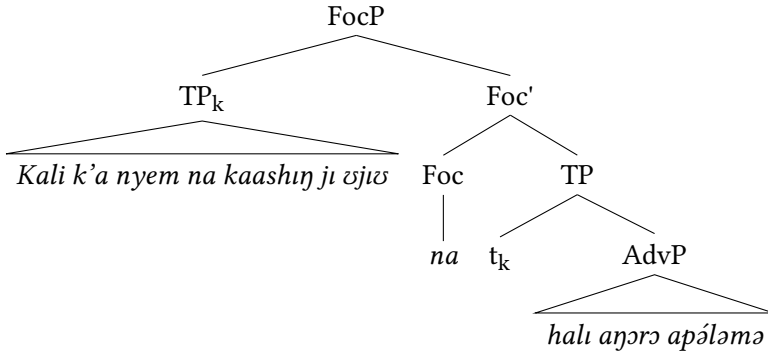Figure 3: Tree for the sentence *Kali k' a nyem na kaashɪŋ jɪ ʊ-jʊ na halɪ aŋɔrɔ apɔ́ləmə*, 'Kali didn't drink or eat for a whole month.'

an analysis of this in Morton & Blanchette (submitted) in which the irrealis morpheme reflects an underlying Modal Phrase selected by negation.

Moving beyond the basic architecture of Anii negative sentences and their interactions with indefinites, we further suggested that the apparent bipartite nature of the negation system in Anii may be a reflection of the inherent focal or contrastive status of propositional negation. We hypothesized that the postverbal marker *na* lexifies this inherent contrastive status, which is obligatorily marked in Anii. With respect to its syntax, this analysis places Anii in line with other apparently bipartite systems such as Ewegbe (Aboh 2010) and Afrikaans (Biberauer 2015), in which the postverbal marker is assumed to be generated in a peripheral position within the C-domain. One factor that distinguishes our proposal from the one in Aboh (2010) is that we do not analyze the postverbal particle as contributing to the semantic negation of the clause. Instead, we propose that its contribution is pragmatic in nature, and that its obligatory presence is due to the focal status of negation.

The hypothesis that the postverbal particle in Anii negative sentences is focal in nature leads to a number of questions. On the empirical side, one question is raised by biclausal sentences with a semantic negation in each clause and two *k*Vs but only one *na*. Under our proposal, if the sentence contains two negated propositions, then we might expect both propositions to be obligatorily focus-marked, leading to two occurrences of *na*. It is possible that, since focus is a discourse-level function, a single *na* is all that is needed to meet the requirement that negative propositions be focused within the discourse. Further discussion of this, however, is beyond the scope of this paper. To explore this question further, we would need to examine the semantics and pragmatics of more biclausal

sentences, along with the specific pragmatics of focus and the distribution of additional discourse-related particles in Anii.

Another question that our analysis gives rise to is whether there are other systems that have been analyzed along Jespersonian lines which could alternatively be analyzed as obligatorily focused negation. One aspect of the Anii system that distinguishes it from traditionally Jespersonian systems like French is the existence of focus particles. Anii has a complex system of morphological focus marking, with multiple focus particles whose distribution is determined by the syntactic properties of the focused constituent (Schwarz & Fiedler 2011). On the other hand, French does not have a focus particle, but instead marks focus through strategies such as clefting and prosody (Karssenberg 2017). An additional difference between Anii and French (and some other systems) pertains to how negative and indefinite noun phrases interact with negation: As illustrated in Section 4, French has a Negative Concord system while Anii does not. It is possible that these differences between Anii and French may be reflective of a more general typological distinction between truly bipartite systems and those in which the bipartite nature of the negative system is only apparent. That the Anii data we have presented in this paper lead to such questions and potential connections serves as an illustration of the potential this understudied language has to inform linguistic theory.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

CL	class marker	POL	polarity marker
EMPH	emphatic	REL	relativizer
INANIM	inanimate	SUB	clause subordination marker
INORD	infinitive-type marker meaning 'in order to'		

Acknowledgements

References

Aboh, Enoch O. 2010. C-type negation markers on the right edge. In Enoch Aboh & James Essegbey (eds.), *Topics in Kwa syntax*, 109–139. Dordrecht: Springer.

Bell, Arthur James. 2004. *Bipartite negation and the fine structure of the negative phrase.* Ithaca, NY: Cornell University Press.

Beyer, Klaus. 2009. Double negation-marking. In Norbert Cyffer, Erwin Ebermann & Georg Ziegelmeyer (eds.), *Negation patterns in West African languages and beyond*, 205–222. Amsterdam: John Benjamins.

Biberauer, Theresa. 2009. Jespersen off course: The case of contemporary Afrikaans negation. In Ellie van Gelderen (ed.), *Cyclical change*, 12–58. Amsterdam: John Benjamins.

Biberauer, Theresa. 2015. Nie sommer nie: Sociohistorical and formal comparative considerations in the rise and maintenance of the modern Afrikaans negation system. *Stellenbosch Papers in Linguistics Plus* 47. 129–174.

Blanchette, Frances. 2013. Negative concord in English. *Linguistic Variation* 13(1). 1–47.

Collins, Chris & Paul M. Postal. 2014. *Classical NEG raising: An essay on the syntax of negation* (Linguistic Inquiry Monographs, vol. 67). Cambridge, MA: MIT press.

Collins, Chris, Paul M. Postal & Elvis Yevudey. 2017. Negative polarity items in Ewe. *Journal of Linguistics* 54(2). 331–365.

Cristofaro, Sonia. 2012. Descriptive notions vs. Grammatical categories: Unrealized states of affairs and 'irrealis'. *Language Sciences* 34(2). 131–146.

Dayal, Veneeta. 2004. The universal force of free choice. *Linguistic Variation Yearbook* 4(1). 5–40.

de Haan, Ferdinand. 2012. Irrealis: Fact or fiction? *Language Sciences* 34(2). 107–130.

Déprez, Viviane. 1999. The roots of negative concord in French and French-based creoles. In Michel DeGraff (ed.), *Language creation and language change: Creolization, diachrony and development*, 375–428. Cambridge, MA: MIT press.

Fiedler, Ines. 2021. The gender system of Anii. *STUF-Language Typology and Universals* 74(2). 369–385.

Heine, Bernd. 1968. A preliminary survey of the noun classes of Basila. *Journal of African Languages* 7(1). 1–13.

Jespersen, Otto. 1917. *Negation in English and other languages*, vol. 1. Copenhagen: AF Høst.

Karssenberg, Lena. 2017. French il y a clefts, existential sentences and the Focus-Marking Hypothesis. *Journal of French Language Studies* 27(3). 405–430.

Ladusaw, William A. 1979. *Polarity sensitivity as inherent scope relations*. Austin: The University of Texas at Austin. (Doctoral dissertation).

López, Luis. 2009. *A derivational syntax for information structure* (Oxford Studies in Theoretical Linguistics, vol. 23). USA: Oxford University Press, USA.

Mahowald, Kyle, Jeremy Hartman, Peter Graff & Edward Gibson. 2016. SNAP judgments: A small N acceptability paradigm (SNAP) for linguistic acceptability judgments. *Language* 92(3). 619–635.

Matthewson, Lisa. 1998. On the interpretation of wide-scope indefinites. *Natural Language Semantics* 7(1). 79–134.

Morton, Deborah. 2012. [ATR] harmony in an eleven vowel language: The case of Anii. In Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison & Tristan M. Purvis (eds.), *Selected Proceedings of the 42nd Annual Conference on African Linguistics*, 70–78. Cascadilla Proceedings Project.

Morton, Deborah. 2014. *The temporal and aspectual semantics and verbal tonology of Gisida Anii*. The Ohio State University. (Doctoral dissertation).

Morton, Deborah & Frances Blanchette. Submitted. Irrealis as modality: Evidence from Gɪsɪɖa Anii. *Journal of Linguistics*.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Springer, Kluwer Academic Publishers.

Schwarz, Anne & Ines Fiedler. 2011. *Linguistic fieldnotes III: Information structure in Gur and Kwa languages*. Potsdam: Repository of the Potsdam University.

van der Auwera, Johan & Lauren Van Alsenoy. 2016. On the typology of negative concord. *Studies in Language* 40(3). 473–512.

Zaske, Stefanie & Hakimou Atti Kalam. 2014. *Ecriture de la langue Anii*. Bassila: SIL. 1–4.

Chapter 7

Metaphors and Metonyms of *ti* and *yi* 'HEAD' expressions in Nzema and Dangme

Regina Oforiwah Caesar
University of Education, Winneba

Mohammed Yakub
University of Education, Winneba

Raymond Teye Akrobettoe
University of Education, Winneba

The study examines how *ti* and *yi* 'head' is conceptualised physically and metaphorically in Nzema and Dangme, two Kwa languages of Ghana. It highlights the similarities and differences based on socio-cultural experiences and conceptualisation patterns of the Nzema and Dangme people. In this study, we position our scholarship on metaphors and metonyms of 'head' among the Nzema and the Dangme as culturally constructed. In Nzema and Dangme, the head serves as a container for conceptualisations of intelligence and wickedness. Phenomena like fortune, insanity and death are construed in terms of cognitive representation of the head. Hardwork, unity, trust, and transparency are virtues that are communicated through head-related expressions in Nzema and Dangme.

1 Introduction

Nzema is a Kwa (Niger-Congo) language spoken mainly in the Western Region of Ghana and some parts of Côte d'Ivoire (Aboagye 1992). Nzema also refers to the speakers of the language. Their total population, according to the census (Ghana

Statistical Service 2021), is 342,090. Nzemaland is bordered to the West by Côte d'Ivoire, to the East by the Ahanta, to the North by the Aowin (or Anyi) and Wassa, and to the South by the Gulf of Guinea. Nzema is studied from Basic to the Tertiary levels of education in Ghana. The basic word order of Nzema follows the Subject-Verb-Object (SVO) pattern (Kwaw 2008: 10).

Dangme is also a Kwa language, spoken in the Greater Accra Region and some parts of the Eastern and Volta Regions of Ghana (Kropp Dakubu 1987). Pockets of speakers are also found in Nɔtse and Gatsi in Togoland (Ameka & Kropp Dakubu 2008). Dangme has seven dialects, namely Adaa, Nugo, Gbugblaa, Osudoku, Sɛ, Manya Klo and Yilɔ Klo. Like Nzema, Dangme has Subject-Verb-Object (SVO) basic word order in a canonical clause structure (Lawer 2017). The language has three level tones (high, mid, low). These tones perform lexical and grammatical functions (Caesar 2016). Even though the Nzema and Dangme people are far apart in terms of their geographical locations in Ghana, it is assumed that their conceptualisations about human body parts (head) would be closely related, since both are Kwa languages with similar basic structural patterns. This serves as one of the motivations for undertaking this study.

In the present study, we examine the conceptual interpretations of expressions that incorporate *ti* and *yi* 'head', as they collocate with different lexical items (mainly adjectives, verbs and adpositions) in specific contexts to give same and/or different meanings in Nzema and Dangme. We attempt to unravel the metaphoric and metonymic relationships that exist in the examples examined. Head expressions that prevail in only one of these languages are also explored to show the differences in the conceptualisations of 'head' in Nzema and Dangme. These research questions guided the study:

1. What positive and negative experiences are conceptualised and described in Nzema and Dangme through head-related expressions?

2. To what extent are Dangme head-related expressions different from or similar to that of Nzema?

3. What are the common lexical items that collocate with *ti* and *yi* 'head' to achieve various interpretations?

This paper is organised into sections. Beyond the introduction, §2 discusses the notion of embodiment cognition and §3 presents a review of the interrelationship between metaphor and metonymy. §4 provides orientation about the theoretical frameworks adopted to underpin the analysis of data. The next section, §5, describes the methodology employed in carrying out the study, followed

by the presentation and analysis of data in §6. In §7, we present a discussion of our findings in light of existing literature on head conceptualisations, and we conclude the study in §8.

2 Embodiment and cognition

In a broader sense, Rohrer (2007) perceives the *embodiment hypothesis* as "the claim that human physical, cognition and social embodiment ground our conceptual linguistic systems" (cited in Ochieng 2022: 53). According to Kraska-Szlenk (2020a: 2), "in neuroscience and cognitive psychology, the strong connection between the human mind and body is known as *embodiment cognition*". The notion of *embodiment*, as scholars such as Yu (2004) and Gibbs (2005) point out, refers to the bodily basis of human perceptions about the environment, and also structures our conceptual system. Evans & Green (2006) argue that:

> The nature of conceptual organisation arises from bodily experiences, so part of what makes conceptual structure meaningful is the bodily experience, with which it is associated. (Evans & Green 2006: 157)

Evans and Green's position provides an insight that body-parts also play a fundamental role in our conceptual system and understanding of our natural and social environment. Corroborating this observation, Yu (2020) states as follows:

> The human body, along with bodily experiences, is a salient source domain for conceptual mappings onto the more abstract target domains such as human cognition, emotion, disposition, and so on. Body part terms found in metaphoric usages in language constitute linguistic manifestations of underlying conceptual metaphors. (Yu 2020: 12-13)

Research on embodiment has received attention from scholars across the world, especially from a cognitive-semantic perspective. Studies such as Yu (2004) on Chinese, Kraska-Szlenk (2014) on Swahili, Agyekum (2015, 2016, 2018) on Akan, Abdul-Hakim & Dogbey (2019) on Kasem, Bisilki (2019) and Bisilki & Yakpo (2021) on Likpakpaln, among others, have examined body-part expressions from a conceptual viewpoint. Several works have concentrated specifically on the conceptualisation of HEAD across languages. Among them are Luo (2020) on Chinese, Kóczy (2019) on Hungarian, Gaby & Bradley (2019) on Yanyuwa, Will (2020) on Hausa, Sime (2020) on Amharic, Al-Saleh et al. (2020) on English and Spanish, Bisilki (2022) on Likpakpaln, and Ochieng (2022) on Dholuo, among others. The

previous studies commonly found out that the metonymic conceptualisation of HEAD[1] can be extended to describe feelings and processes that are experienced within the head. In these studies, HEAD FOR PERSON metonymy is seemingly pervasive. The metonymic extensions can also relate to concepts that are external to humans, such as LOCATION, UNIT OF MEASUREMENT, TOPIC, TITLE, and HEADING. The analysis of the metaphors of HEAD showed that it is primarily connected to REASONING and in some cases to EMOTION and MEMORY.

3 Metaphor and metonymy

Metaphor and metonymy, among other tropes, are attributed to cognitive processes (Kövecses 2006, Cruse 2011, Peña-Cervel & Ibañez 2022). Metaphors, for instance, as Lakoff & Johnson (1980) contend, are not simply poetic parts of speech, or obscure idiomatic expressions, but rather part of human entire cognitive reasoning through which they are able to understand and talk about the social world in which they live. Gibbs et al. (2004) opine that metaphor is not merely a linguistic rhetoric figure, but an aspect of cognition which constitutes a fundamental part of people's ordinary thought, reason, and imagination. Writing on the cognitive aspects of metaphor, Steen (2011) asserts that:

> Metaphor has turned out to be a conceptual mechanism, a 'figure of thought', by which specific and operational knowledge about more concrete phenomena and experience is projected onto a wide range of more abstract ones. (Steen 2011: 28)

In line with the above postulations, Yu (2020: 13) notes that "metaphor is primarily a matter of thought. It is how we think and reason". Metaphors are conceptually grounded on physical and mental experiences and metaphorical conceptual systems in the mind (Lakoff & Johnson 1980, Lee 2005, Agyekum 2015, 2018). A striking feature of metaphor is that it involves understanding an abstract concept (target domain) based on another, which is rather concrete (source domain), and this understanding is based on the perceived similarities between two concepts or entities (Lakoff & Johnson 1980, Lakoff 1993, Gibbs 2008, Steen 2011, Diaz-Vera 2022).

Metonymy, on the other hand, is a *stand-for relationship*, which establishes connection within a single domain (Dawson-Ahmoah & Wonkyi 2021). Metonymy

---

[1]By convention, conceptual metaphors are written in capital letters.

is not necessarily anchored in perceived similarities, but rather spatial, temporal, or conceptual contiguity within the same cognitive model (Agyekum 2018: 24, Ochieng 2022: 52). In metonymic relations, as Kövecses (2006: 99) avers, one conceptual entity, the vehicle, provides mental access to another conceptual entity (thing, event, property), the target, within the same frame or cognitive domain. Thus, Kövecses describes metonymy as *within-domain mapping*. As Ochieng (2022: 52) points out, there are two basic metonymic mappings: PART-FOR-WHOLE and WHOLE FOR PART. Agyekum (2018) indicates that:

> When we talk of any part of the body as being part of the entire body or personality, we are having a partonomic (*part-and-whole relation*) view between the body and its parts, and these are terms of *metonym*. (Agyekum 2018: 24-25)

In this study, we show that figurative expressions involving the human head in Nzema and Dangme are grounded in metaphorical and metonymic abstractions. This seeks to enhance our understanding of the way these two cultures perceive their world based on embodiment and cognition.

4 Theoretical framework

The paper draws largely on Conceptual Metaphor Theory (CMT) (Lakoff & Johnson 1980), which seeks to primarily explain that our cognitive faculty, and for that matter our conceptual systems play a vital role in experiencing and understanding our day-to-day realities. Another theory that is considered effective to underpin this study is the Cultural Conceptualisations and Language framework, which also deals with cultural schemas and features of language that have cultural basis. Some orientations of these two frameworks adopted are presented in Sections §4.1 and §4.2 respectively.

4.1 Conceptual Metaphor Theory

Metaphor was noted for its artistic function, and a kind of "decoration" that is added to an ordinary plain language, perhaps to achieve speech embellishment. In contrast to the traditional views of metaphor, Lakoff & Johnson (1980: 5) assert that "the essence of metaphor is understanding and experiencing one kind of thing in terms of another". This is to say that, in contemporary studies, metaphors have been described beyond mere linguistic and communication ornaments; they are seen as "a mechanism" of thought and reason (Lakoff &

Johnson 1980, Lakoff & Turner 1989, Lakoff 1993). Conceptual metaphor theorists see metaphor as a mental process that transfers images from one concept to another. Conceptual metaphors typically structure more abstract concepts (target) in terms of more concrete/physical (source) concept whereby we are able to understand unfamiliar concepts better (Kövecses 2002). Semino (2008) notes that conceptual metaphors are systematic sets of correspondence, or "mappings" across conceptual domains, whereby a "target" domain is partly structured in terms of a different "source" domain. This underscores that certain mental mappings are essential to influence our thinking and imagining in everyday life. Take for instance, the conceptual metaphor LOVE IS A JOURNEY. Here, Lakoff (1993) presents LOVE as abstract (target domain) and JOURNEY as concrete (source domain). Lakoff (1993: 224) argues that one can conceptualise the LOVERS as TRAVELERS, and the LOVE RELATIONSHIP as A VEHICLE. He makes his point clearer by stating:

> Just as significant life events are special cases of events, so events in a love relationship are special cases of life events. Thus, the LOVE IS A JOURNEY metaphor inherits the structure of the LIFE IS A JOURNEY metaphor. What is special about the LOVE IS A JOURNEY metaphor is that there are two lovers who are travelers and that the love relationship is a vehicle. The rest of the mapping is a consequence of inheriting the LIFE IS A JOURNEY metaphor. Because the lovers are in the same vehicle, they have common destinations, that is, common life goals. Relationship difficulties are impediments to travel (Lakoff 1993: 223).

According to the Conceptual Metaphor Theory, unfamiliar and delineated concepts in our experiences can only be grasped by means of other concepts that we understand in clearer terms (Agyekum 2018: 13). The body part "head" provides a conceptual basis that can be projected metaphorically to explain and describe other abstract concepts in both Nzema and Dangme communication. Thus, adopting CMT as a theoretical lens to guide this paper is considered effective and appropriate.

4.2 Cultural conceptualizations

Cultural Conceptual Metaphor is a facet of Cultural Linguistics, a sub-branch of linguistics which explores the relationship between language, culture, and conceptualisations (Sharifian 2011, Sharifian 2017). The early notion of Cultural Linguistics may be attributed to the works of Langacker (1994) and Palmer (1996),

who seem to make a common observation that cognitive linguistics recognises and embraces "cultural knowledge" as a significant facet of grammar. In this regard, Palmer (1996) argues that cognitive linguistics can be applied to the study of language and culture; where Palmer makes use of the term *imagery*, perhaps to make reference to cognitive concepts. Following this, Sharifian (2011) elucidated the concept of Cultural Linguistics from a multidisciplinary perspective, using the term *Cultural Conceptualizations*; which he notes enables members of a cultural group to think in one mind (see also Sharifian 2003). According to Sharifian (2017: 2), "Cultural Linguistics engages with features of human languages that encode or instantiate culturally constructed conceptualisations encompassing the whole range of human experience". This suggests that many features of human languages are entrenched in cultural conceptualisations. Cultural Linguistics, thus, is responsible for exploring features of language that have cultural basis (Sharifian 2014). It employs three analytical tools such as "cultural schema", "cultural category" and "cultural-conceptual metaphor". Sharifian (2014) avers that the meaning of many lexical items of human languages lends itself best to cognitive schemas that are abstracted from cultural experiences. Metaphorical expressions are pervasive in ordinary language; they are not just used for artistic purposes. These everyday metaphors reveal cognitive and cultural conceptions of the world (Jackendoff & Aaron 1991).

Among the Nzema and Dangme people, as this paper reveals, head related expressions encode both positive and negative human experiences that are culturally constructed. To this end, our analysis further takes insight from Cultural Conceptual Metaphor, a facet of the Cultural Linguistics framework to help provide adequate description of how the Nzema and Dangme conceptualise the human head in their day-to-day communication.

5 Methodology

Data for the study were obtained from both primary and secondary sources. A total of 30 respondents were involved in the primary data collection. Twenty-four (24) of them were Level 300 students of the University of Education, Winneba. They comprise 12 Dangme and 12 Nzema students. The remaining six respondents were elderly native speakers of Dangme and Nzema, three from each language community. Their ages ranged from 60 to 75 years. The elderly people were purposely selected because of their rich linguistic and socio-cultural knowledge in the two languages. We tasked the students to write expressions that relate to "the head" in Dangme and Nzema. As native speakers, we also relied on our intuitions to generate additional data for the study. The secondary

data were obtained from Dangme and Nzema literary materials: *Dangme Abɛ gbi* (Accam 1972) and *Nyamenle asa ɛnlomboɛ ne* (Soboh-Blay 2013) respectively. Through semi-structured interviews, we cross-checked the data with the elderly respondents and audio-recorded the interview sessions, which were transcribed and translated for the analysis. In all, fifty (50) expressions related to the head were assembled, but thirty-five (35) were purposively selected for the analysis. The data were coded, categorised, and thematised for the analysis. In §6, the head-related expressions in both languages are provided, followed by morpheme-by-morpheme interlineal glossing and translations in English, accompanied by analysis.

6 Analysis of Nzema and Dangme head-related expressions

This section deals with the data presentation and analysis in the light of metaphor and metonymy. We agree with Peña-Cervel & Ibañez (2022) who argue that metaphor (conventionally formulated as A is B) is a different trope than metonymy (A for B), and that the two must not be conflated, although they can co-occur. In our analysis, therefore, we attempt to explicate metaphoric expressions differently from metonymic expressions related to "head". The analysis falls under two major headings: (i) head-related expressions that encode positive experiences (§6.1) and (ii) head-related expressions that encode negative/unpleasant experiences (§6.2). Each category has various sub-headings. We provide the examples in Nzema and Dangme and present the interlinear glossing as well as the English translation.

6.1 Head expressions that encode positive experiences

The data demonstrate that head-related expressions among the Nzema and the Dangme encode pleasant concepts and experiences, including luck, intelligence, good intentions, and comfort. We take the discussion up by considering these thematic areas.

6.1.1 Expressing luck/fortune

Both Nzema and Dangme speakers consider GOOD HEAD metonymically as FORTUNE. The expressions in (1) illustrate the phenomenon:

(1)   'HEAD' for FORTUNE Metonymy

    a. Nzema
        ɛ   ti   ɛ-yɛ   boɛ
        2SG.POSS head PFV-make well
        'Your head has been good'. (You are fortunate)

    b. Dangme
        E   yi   hi
        2SG.POSS head good
        'Your head has been good'. (You are fortunate)

In both examples, (1a) and (1b), GOOD HEAD is used metonymically to stand for FORTUNE. As we see in these expressions, only the body part 'head' is modified by the value adjective good; however, GOOD HEAD stands for all other things that work out successfully for a person. For instance, if a woman is able to bear children after marriage, her fortunate situation is not described by mentioning the stomach/womb or the vagina that are directly involved in the process of child bearing/delivery. Rather, the HEAD would be mentioned as in the following examples:

(2)   'HEAD' for FORTUNE Metonymy

    a. Nzema
        ɛ   ti   ɛ-yɛ   boɛ kɛ wɔ-nyia m-ralɛ wɔwo   la
        2SG.POSS head PFV-be good that 2SG-born PL-child 2SG-born PART
        'Your head has been good for having gotten children'. (You are fortunate for bearing children)

    b. Dangme
        O   yi   hi   kaa o fɔ bi-mɛ
        2SG.POSS head be.good that 2SG born child-PL
        'Your head has been good for having gotten children'. (You are fortunate for bearing children)

Besides, if a person is able to escape from any danger or accident, the head is used to express his/her fortunate situation.

Moreover, in Nzema and Dangme cultures, and for that matter according to their "anatomic conceptualisations", the human head is perceived as the most "significant" part of the body, as it accommodates certain vital organs such as eyes, mouth, ears, and brain. Though the heart, the kidney and other internal organs cannot be taken for granted, the brain remains crucial because whoever

loses his brain becomes dysfunctional in every respect. This "super important" brain is situated within the head, which means that the head keeps something good. Therefore, describing a person as having 'good head' implies that all good (positive) things work out successfully for the person. In other words, the person is said to be lucky. Agyekum (2018) shows that the Akan also have a similar conception, where LUCK is construed as GOOD HEAD and UNLUCKY is seen as BAD HEAD, where bad head implies that one has a bad omen (Agyekum 2018: 44). He uses the expressions in (3a) and (3b) to exemplify lucky and unlucky respectively.

(3) Akan

    a. Ne   tiri   yɛ
       3sg.poss head cop.good
       'His head is good'. (He is lucky)

    b. Ne   tiri   n-yɛ
       3sg.poss head neg-cop.good
       'His head is not good'. (He is not lucky) (Agyekum 2018: 44)

6.1.2 Expressing intelligence

In discussing intelligence and/or unintelligence among the Nzema and Dangme, the mention of 'head' is key. As we have hinted earlier, the head contains the brain/mind, which is also associated with human intellectual activities. Since the head harbours the brain, the CONTAINER and the CONTAINED relationship (see Agyekum 2018: 44) is noticeable under this category. Consider the Nzema and Dangme examples in (4a) and (4b) respectively:

(4) 'HEAD' for INTELLIGENCE/THINKING Metonymy

    a. Nzema
       Ɔ-lɛ   etile
       3sg-poss head
       'He/she has a head'. (He/she is intelligent)

    b. Dangme
       E   ngɛ yi
       3sg poss head
       'He/she has a head'. (He/she is intelligent)

In examples (4a) and (4b), we find that both Nzema and Dangme speakers perceive the head as the centre for reasoning. In these constructions, head is a complement of the subject pronouns Ɔ and *E* 'he/she' respectively. Here, head is a possessed entity, which presents us with the metaphor BEING INTELECTUALLY INCLINED is POSSESSING A HEAD. These expressions show the metonymic use of head for a part of it which relates to thinking.[2] Here, we see a WHOLE-FOR-PART metonymy. Instead of mentioning the brain, which is a portion of the elements contained in the head, and which is used for thinking, the entire head is rather mentioned, leading to the conceptualisation and interpretation of intelligence. Since the brain resides in the head, whoever has a head is considered as "having a brain", which implies that the person has intellectual capabilities. Other possible expressions that describe intelligence deal with the concept of "maturity". Dawson-Ahmoah & Wonkyi (2021: 101) indicate that maturity is an "innate trait" that is demonstrated in deeds, which is believed to reside in the mind. The Nzema and Dangme share this perception, and so they akin maturity to intelligence as can be seen in (5) below:

(5)   MATURE BRAIN is INTELLIGENCE Metaphor

    a. Nzema
       Ye   adwenle ɛ-di
       3SG.POSS brain   PFV-grow
       'His/her brain is mature'. (He/she is intelligent)

    b. Dangme
       E   juɛmi nya   wa
       3SG.POSS brain mouth be.grown
       'His/her brain is mature'. (He/she is intelligent)

The examples in (5a) and (5b) both make specific mention of BRAIN. In these examples, BRAIN is described as GROWN or MATURE to express a person's intellectual capability. The verb 'grow' is significant to the utterance interpretation as it expresses increment/expansion in size or height. When a tree, mahogany, for instance, is fully grown/matured, the trunk expands, and its roots can penetrate the soil and reach very far. In a similar sense, a 'mature brain' enables its owner to think far, so as to make reasonable decisions. This "far-reaching phenomenon" is what foregrounds the metaphoric correlation between MATURE BRAIN and INTELLIGENCE.

The next examples capture the notion of "openness" and "deepness" to conceptualise intelligence.

---

[2]We are grateful to an anonymous reviewer for pointing out to us this important observation.

(6) OPEN/DEEP BRAIN is INTELLIGENCE Metaphor

    a. Nzema
        Ye   adwenle nu   ɛ-buke
        3SG.POSS brain   inside PFV-open
        'His/her brain inside is deep'. (He/she is intelligent)

    b. Dangme
        E   juɛmi mi   kuɔ
        3SG.POSS brain inside deep
        'His/her brain inside is deep'. (He/she is intelligent)

In (6), we notice the conceptual metaphor OPEN-MINDEDNESS is INTELLI-GENCE. In this example, intelligence is measured in terms of the openness of one's mind, whose interpretation is enhanced by the use of the postposition *nu* or *mi* 'inside'. The metaphoric abstraction and correlation between open-mindedness and intelligence is triggered by the fact that a widely open entity or space covers a broader or vast area. So, if one's brain/mind is open, it implies that the person's mind, which represents wisdom/intelligence, perhaps has no boundaries, and for that matter the person has vast knowledge (see also Dawson-Ahmoah & Wonkyi 2021). The Dangme expression in (6b) is somewhat in line with the Nzema version in (6a). As a widely open entity can be loaded with a lot of materials or goods, a deep entity, say a hole, can also be loaded with many goods. In (6b), therefore, intelligence is measured in terms of the depth of one's brain. A DEEP BRAIN, as the Dangme expression suggests, is able to contain a lot of WISDOM, which arrives at the metaphors DEEP BRAIN is INTELLIGENCE and SHALLOW BRAIN is UNINTELLIGENCE. In (6a) and (6b), the relevant noun 'brain' occurs in a dependent position in a postpositional construction. The post-positional phrases *nu ɛbuke* 'inside open' and *mi kuɔ* 'inside deep' contribute to describing the state of affairs of the BRAIN in relation to intelligence.

6.1.3 Expressing good intentions and positive mindset

In both Nzema and Dangme cultures, white colour provides conceptual basis to talk about positive mindset and wholeheartedness. As opposed to black colour, which is linked to evil-mindedness as would be highlighted later in this paper, white colour symbolises victory and purity. Using the metaphor of WHITE COL-OUR (PURITY) to express POSITIVE MINDSET, however, is strictly associated with the internal part of the head as seen in example (7):

(7)   PURE HEAD for GOOD INTENTIONS Metonymy

   a.   Nzema

  Ɔ ti akunlu le fufule

  3sg.poss head inside   cop white

  'His/her head inside is white'. (He/she has good intentions)

   b.   Dangme

  E yi mi ngɛ hei

  3sg.poss head inside cop clear/clean

  'His/her head inside is pure'. (He/she has good intentions)

In these examples, we observe that the Nzema and Dangme expressions incorporate 'white' and 'pure' respectively to communicate that one has good intentions or thinks positively. Both expressions make use of the postposition 'inside'; implying that the assumed clarity/whiteness, which is linked to positive thinking (one's mindset), is situated within the head. Therefore, PURE HEAD is a WHOLE-FOR-PART metonymy for GOOD INTENTIONS. In Nzema and Dangme, clear is expressed by white, and the semantic motivation for the extension of white to clear is that the person has nothing to hide. So, if there is whiteness or purity in a person's head, it means that they do not harbour negativities in their minds. In other words, such people easily forgive and forget. It also implies that they are sincere, not envious or jealous, and always wish that other people would make progress in their endeavours.

6.1.4 Expressing comfort

As part of their cultural conceptualisations, the Nzema and the Dangme liken a state of comfort to "sweetness", which is believed to exist in the head of a less-troubled individual. The expressions in (8) came up as part of the data in this regard.

(8)   SWEETNESS IN THE HEAD for COMFORT Metonymy

   a.   Nzema

  ɔ ti akunlu yɛ ye fɛ

  3sg.poss head inside   has 3sg sweet

  'There is sweetness inside his/her head'. (He/she is comfortable)

   b.   Dangme

  E yi mi ngɔɔ lɛ

  3sg.poss head inside be.sweet 3sg

  'There is sweetness inside his/her head'. (He/she is comfortable)

In examples (8a) and (8b), there is a WHOLE-FOR-PART metonymy in the sense that being comfortable in life deals with one's psychological state, which relates to "stable (undisturbed) mindedness" that resides in the head as a whole.[3] Both expressions imply that a person has no worries, or that their troubles are highly controllable. Any consumable that is sweet, people enjoy it with comfort. However, in consuming something bitter, the consumer feels uncomfortable. Therefore, sweetness existing in a person's head implies that they have a "limited burden", if not completely free from troubles, and for that matter they enjoy life.

6.2 Head expressions that encode negative/unpleasant experiences

Some unpleasant experiences such as unintelligence (lack of academic excellence), insanity, wickedness and death are expressed via the conceptualisation of the human head among the Nzema and Dangme. We discuss these phenomena in the following sub-sections.

6.2.1 Expressing unintelligence

The expressions in (9) can imply lack of academic excellence. Both (9a) and (9b) capture the concept of INCOMPLETENESS to describe one's lack of adequate academic excellence, such as:

(9) INCOMPLETE HEAD for UNINTELLIGENCE Metonymy

    a. Nzema
        ɛ   ti   le   ebule
        2SG.POSS head COP half
        'Your head is incomplete'. (You are unintelligent)

    b. Dangme
        O   yi   su-i
        2SG.POSS head up.to.number-NEG
        'Your head is incomplete'. (You are unintelligent).

In (9a) and (9b), the Nzema and Dangme examples respectively employ *ebule*, 'half' and *su-i*, 'not up to the number' as quantifying adjectives which collocate with 'head' to conceptualise and express unintelligence. In (9a), while the Nzema

---

[3]Here, one also sees the conceptual metaphors, COMFORT is SWEETNESS and DISCOMFORT is BITTERNESS.

perceive an unintelligent person as possessing 'half head', the Dangme conceptualise an unintelligent person as someone who possesses a head which is 'not up to number', in other words, the person's head is not 'full' as in (9b). Here, the brain, which relates to thinking, is perceived as insufficient (not up to its maximum quantity that can foster one's reasoning). In the examples, however, the 'whole head' is mentioned and described as incomplete. This reveals a WHOLE-FOR-PART metonymy to conceptualise unintelligence.

In (10), we present examples in Nzema and Dangme that use the phenomenon of death to express unintelligence.

(10) DEAD HEAD for UNINTELLIGENCE Metonymy

    a. Nzema
       Ɛ   ti   ɛ-wu
       2SG.POSS head PFV-die
       'Your head is dead.' (You are unintelligent).

    b. Dangme
       O   yi   tɔ   gbo
       2SG.POSS head container die.PFV
       'Your head is dead.' (You are unintelligent)

In example (10), HEAD (the container) rather stands for BRAIN (the contained), and the entity for reasoning. When a person dies, all their senses become dysfunctional; not responding to any stimuli (Yakub & Agyekum 2022). Just as an animate entity does not perform any activity when they are dead, so the brain, represented by the head, is described as 'dead' when a person is unable to think critically to achieve academic excellence. Here, the brain is seen as not able to receive information and process it accurately for a desired output.

The following examples in (11) and (12) are further expressions that encode unintelligence, which are also peculiar to the Dangme people. The expressions incorporate 'rotten head' and 'finished head' respectively. These also present a metonymic use of HEAD, as it stands for the BRAIN that is actually associated with intellect. Take the examples as follows:

(11) ROTTEN HEAD for UNINTELLIGENCE Metonymy[4]
      Dangme
      O   yi   mi   sa
      2SG.POSS head inside be.rot.PFV
      'Your head inside is rotten'. (You are unintelligent)

(12) FINISHED HEAD for UNINTELLIGENCE Metonymy
      Dangme
      O   yi   ta
      2SG.POSS head finish.PFV
      'Your head is finished'. (You are unintelligent)

In (11), the Dangme see an unintelligent person as someone who possesses a head with rotten content. The use of *mi* 'inside' (of the head) can enhance our understanding that reference is being made to the brain, since its habitat is within the head. This metonymic conceptualisation is prudent in the sense that it is the human head that possesses the brain used to think about the world around us. Thus, to be intelligent, according to the Dangme, one must have the head containing the brain and other components which must be alive, active, and functioning. The expression in (12) implies that one's reasoning ability, especially for academic purpose, is non-existent.

6.2.2 Expressing senselessness

Unlike the foregoing examples in Section 6.2.1 which focus on the lack of academic excellence, the next examples in (13) represent the generic expressions that describe senselessness. Consider these expressions:

(13) HEADLESS for SENSELESS Metonymy

   a. Nzema
       Ɛ-n-lɛ   etile
       2SG-NEG–have head
       'You are headless'. (You do not reason)

---

[4]A rotten entity can be likened to a dead entity, since they both share the semantic attribute, -life. Therefore, examples (11) and (12) are analogous in terms of semantic relevance. Although these examples are largely metonymic, one also observes the metaphors, DEAD/ROTTEN HEAD (BRAIN) is UNINTELLIGENCE and ALIVE HEAD (BRAIN) is INTELLIGENCE.

b. Dangme

O   be   yi
2SG NEG.have head
'You are headless'. (You do not reason)

The expressions in (13a) and (13b) also explain how the Nzema and Dangme conceptualise the phenomenon of SENSELESSNESS. It is evident in the two expressions that both societies use HEADLESSNESS to express one's irrational or unreasonable act. The expressions imply that to reason, one must possess the HEAD, which is used metonymically to represent the BRAIN. Here, the bare noun, HEAD is used, and this is relevant for the interpretation; suggesting that a person's mind does not exist. The head contains the schemas which is used to reason or think about the world around us, which also helps in perceiving and responding to complex situations. So, to express senselessness, the Nzema and Dangme see a person's head to be absent. In this way, the individual cannot use the schemas to organise his or her thought to make meanings of concepts.

6.2.3 Expressing insanity

Absolute madness is attributed to "damaged" head. The Nzema and Dangme share this conception and respectively use *zɛkye* and *puɛ* 'to spoil' in describing insanity, as can be seen in (14):

(14)   DAMAGED/SPOILT HEAD for MADNESS Metonymy

a. Nzema

Ɔ   ti   ɛ-zɛkye
3SG.POSS head PFV-spoil
'His/her head is spoilt'. (S/he is insane)

b. Dangme

E   yi   puɛ
3SG.POSS head spoil-PFV
'His/her head is spoilt'. (S/he is insane)

Unlike in (11), where the Dangme use ROTTEN HEAD/(BRAIN) to conceptualise UNINTELLIGENCE, (14a) and (14b) capture SPOILT HEAD to metonymically express INSANITY. Only a part of the head, BRAIN (CONTAINED) relates to thinking and what can become dysfunctional, but the entire HEAD (CONTAINER), is used in the expressions to imply madness. Generally, any spoilt entity goes bad and loses its value. Based on this conceptualisation, the Nzema and

Dangme describe one who is insane as having a head that has dysfunctional and useless content.

6.2.4 Expressing wickedness

The concept of MISCHIEVOUSNESS is well construed among the Nzema and Dangme by linking it with DARKNESS, which is believed to exist in a person's head. Consider the following expressions in (15a) and (15b):

(15)   DARKNESS IN THE HEAD is WICKEDNESS Metaphor

    a. Nzema
        Ɔ   ti   akunli le   awozinli
        2sg.poss head inside cop darkness
        'There is darkness in your head'. (You are wicked)

    b. Dangme
        O   yi   mi   wo   diblii
        2sg.poss head inside wear/keep darkness
        'There is darkness in your head'. (You are wicked)

The Nzema and Dangme examples in (15a) and (15b) respectively use the colour adjectives *awozinli* and *diblii* 'dark' to indicate negative mindedness. In these expressions, DARKNESS IN THE HEAD is mapped onto 'WICKEDNESS'. As we have hinted earlier, the Nzema and Dangme also associate black colour, and for that matter darkness, with bad/unfortunate occurrences and situations such as death and bereavement. "Darkness in the mind", however, is specifically attributed to evil deeds and thoughts. The correlation between darkness in the head and wickedness is triggered by the negative attributes that both the Nzema and Dangme societies associate with darkness. They see wicked people as not possessing 'purity' in their heads (minds). This implies that such wicked people can even murder their fellow humans mercilessly, including undertaking other brutalities. The statement in (16) below is another Dangme expression that describes wickedness. It employs *ti* 'be thick' to fully arrive at the interpretation that a person's mind is not 'transparent', and that many evil thoughts can be hidden behind their minds which one would hardly become aware of. The expression says:

(16)   THICKNESS IN THE HEAD is WICKEDNESS Metaphor
       Dangme
       O   yi   mi   ti
       2SG.POSS head inside be.thick
       'Your head inside is thick'. (You are wicked)

   Here, THICKNESS IN THE HEAD is metaphorically mapped onto WICKED-NESS. This also means that the Dangme perceive the contents of the head to be THIN for a person to be ascribed with the positive attribute, KINDNESS. However, for a person to be seen as WICKED, then the contents in their head must be THICK.

6.2.5 Head expressions concerning death

As a fear-based concept, the subject of DEATH is usually not discussed freely in many cultures without resorting to cover-up strategies like euphemisms to tone-down the shock, discomfort and embarrassment that might be caused (Yakub & Agyekum 2022). This is common among both the Nzema and Dangme people. Both have "roundabout" means to express death. The examples in (17) and (18) especially show how the Nzema employ metaphorical euphemisms related to *ti* 'head' to escape from the explicit pronouncement of death. Take the examples as follows:

(17)   TO HIT ONE'S HEAD ON THE GROUND is TO BE CONQUERED BY
       DEATH Metaphor
       Nzema
       Ɔ   ti   ɛ-bɔ   aze
       3SG.POSS head PFV-hit floor
       'His/her head has hit the ground'. (He/she is dead)

(18)   SHORT HEADEDNESS is WIDOWHOOD Metaphor
       Nzema
       Ɔ   ti   le   sinli
       3SG.POSS head COP short
       'His/her head is short/reduced'. (He/she is a widow/widower)

   In the context of death-discourse among the Nzema, examples (17) and (18) are specific head-related expressions that are used to avoid overt pronouncement of death. In (17), the conceptual metaphors, DEATH is DOWN and LIFE is UP, are

noticeable. The Nzema perceive the phenomenon of death as a struggle/war between life (being alive) and death (being dead). So, if one's head is said to have hit the ground, in the context of discussing death, it implies that the individual's life is conquered by death (see also Yakub & Agyekum 2022: 138). This is one of the subtle ways to report the demise of a person in Nzema language and culture. In (18), we find that both a widow and a widower are described figuratively as people who are "short-headed" or that they have "reduced head". In this example, the conceptual metaphor, WIDOWHOOD is REDUCED HEAD is evidenced. Married couples are expected to put heads together to plan for their present and future lives. Therefore, "short/reduced-headedness", according to the Nzema cultural conceptualisation, implies that the widow has lost an additional 'head', which represents her husband's wisdom and knowledge through which he contributed significantly during their decision making among other things. The same is conceived of a man who loses his wife by death. Put simply, the Nzema see people who lose their spouses as possessing "reduced or shortened" heads because a 'head' that supports their decision making has passed on. Hence, one also observes HEAD for PERSON metonymy in (18).

In (19), the Dangme expression related to *yi* 'head' to rather talk about the cessation of descent line is also provided.

(19)   FINISHED HEAD for END OF LINEAGE Metonymy
        Dangme
        Aku   yi   ta
        Aku.POSS head PFV.finish
        'Aku's head has finished'. (Aku has passed on without giving birth)

The Dangme talk about the discontinuity of a descent line by employing a metonymic expression as in (19), where 'head' is used for the lineage. Specifically, Dangme speakers use this expression to imply that an adult passed on without giving birth to any child, and so they have left to the spiritual underworld without leaving children behind. Therefore, from Dangme cultural perspective, everything about such a person is finished, especially, no descendants would ever emerge from that person.

6.3 Proverbial expressions related to *ti* and *yi* 'head'

The data further revealed that both Nzema and Dangme speakers have conventionalised proverbs which incorporate *ti* and *yi* 'head' respectively to communicate and inculcate positive values and virtues such as unity and cooperation

(collective decision marking), proper time management and hard work, among others. Examples of these proverbs are discussed in the following sub-sections.

6.3.1 Proverbs that express cooperation and collective decision making

Speakers of Nzema and those of Dangme think that any decision making that involves the collective wisdom of people is mostly fruitful, as opposed to one that is championed by an individual. Thus, both cultures have proverbs that encourage collaboration and collectiveness, especially with regard to decision making as we notice in (20):

(20)   HEAD for PERSON Metonymy

    a. Nzema
        Etile ko le   namule
        Head one cop cottage
        'Single head is cottage'.

    b. Dangme
        Yi   kake yɛ   dami
        Head one go.neg meet
        'Single head does not attend a meeting'.

Metonymically, both examples use HEAD to stand for A PERSON. The proverbial expressions point out that "a single person does not deliberate" to yield a fruitful outcome. The Nzema example in (20a) also metaphorically compares individual decision making with "living in a cottage". Life in the cottage is quite unpleasant in the sense that all social amenities that make life in the urban cities enjoyable are completely absent at the cottage. This notion is transferred to conceptualise individual decision making as portrayed in (20a). A decision taken by 'single head' also lacks many brains on board, and so it is likely to lack successful accomplishment and/or implementation. In this case, people may not realise any "enjoyable" outcome of such a decision. The Dangme proverb in (20b) communicates similar sense as the Nzema version. It buttresses the fact that individualism must be discarded in human endeavours, especially in terms of decision making. Crucially, both proverbs seek to promote unity and cooperation among members.

6.3.2 Assigning duties to rightful individuals

The following proverbial expressions in (21) also reveal PART-FOR-WHOLE metonymy, where we are cautioned not to assign duties to wrongful persons when

the people who deserve to undertake the activities are available. The examples are:

(21) HEAD/KNEE for PERSON Metonymy

    a. Nzema
       Etile de aze a egyakɛ nangonloma ɛ-n-zoa kyɛlɛ
       Head sit down PART leg knee EMPH-NEG-wear cap
       'The head must not be ignored for the knee to rather wear a cap'.

    b. Dangme
       Yi hi si nɛ nakutso bu pɛɛ
       Head sit.NEG down CONJ knee wear cap
       'The head must not be ignored for the knee to rather wear a cap'.

The analogy of the knee in these proverbs, as Bobuafor (2021: 205) suggests based on parallel proverbs she provides in Tafi and Ewe, could be due to the common practice where one removes their hat and hangs it on their knee. The advisory content of these proverbs (Bobuafor 2021: 205) is that "one should let appropriate person do the things that are required". A hat/cap is worn on the head, thus when the head is readily available, no other part of the body must wear the hat. Since it could be disastrous for people who lack adequate knowledge and expertise to embark on certain "skill-based" activities, these proverbs inform us to always allow experts to do their job.

6.3.3 Proverbs about unpredictability of intentions

The Nzema and the Dangme have common proverbs which communicate that a person's intentions cannot be known unless they disclose their own agenda. Such proverbs advise people to take precautions when dealing with their fellows. They convey the message that human beings must not be fully trusted because one cannot determine what others harbour in their mind. The proverbs capture what we have referred to in this paper as *opposing metaphoric link* between pawpaw and the human head to drive home such advisory message. Consider the examples in (22):

(22) HEAD not as PAWPAW Metaphor

    a. Nzema
       Etile ɛ-n-le kɛ kpakpa
       head EMPH-NEG-COP like pawpaw
       'The head is not like pawpaw'.

    b. Dangme

        Yi   be   kaa gɔ

        head have.NEG like pawpaw

        'The head is not like pawpaw'.

Though pawpaw and the human head may share some similarities, for example, in terms of shape (roundness), some other features make the two entities distinct. For instance, pawpaw is a soft edible fruit that can easily be cut (opened) to know its content. The human head, however, is not easily opened unless one aims at killing his fellow. So, the Nzema and Dangme employ these proverbs to inform members of the society to "fear" mankind, since we cannot be certain whether one thinks positively or negatively. If a person is up to doing good or evil, no one can easily detect or read from their mind.

6.3.4 Proverb about hard-work and proper retirement preparation

Hard-work and perseverance is an essential virtue every society cherishes. Indolence is frowned at because it hinders socio-economic development and breeds vices, including theft and murder (Ibrahim et al. 2022). Example (23) is a proverb in Nzema that repudiates laziness. It points to the benefit of hard-work and the essence of proper time management in terms of retirement preparation and planning. The proverb says:

(23)   BLACK/GREY HEAD for AGE Metonymy

        Nzema

        Wɔ-an-zoa   etile bile zo a   ɛ-soa   egyɛne zo

        2SG-NEG-carry head black top PART 2SG-carry grey   top

        'If you refuse to carry your load on black head, you would carry it on grey head'.

In this proverb, the Nzema conceptualise and liken BLACK HEAD to YOUTHFUL AGE, a period in which one has enough strength to work hard, and when he/she is expected to do so. The GREY HEAD also represents OLD AGE, a period when a person is expected to retire from any active service and relax comfortably at home, reaping what they had sown. The proverb encourages industriousness; doing hard-work when a person is capable (at their youthful age), else, they would suffer when they are not capable (at old age).

7 Discussion of findings

Adopting the Cultural Conceptual Metaphor Theory, the study examined how *ti* and *yi* 'the human head' is conceptualised physically and metaphorically among the Nzema and Dangme speakers. Lexical categories such as adjectives, verbs, adverbs and adpositions in both languages were noted for collocating with *ti* and *yi* 'head' to arrive at various contextual interpretations. The paper showed that both Nzema and Dangme have a rich variety of expressions related to the HEAD, which serves as the CONTAINER of the BRAIN/MIND for conceptualisations of GOOD LUCK, INTELLIGENCE, POSITIVE THINKING, and COMFORT. This study has shown that Nzema and Dangme see the HEAD as a locus of intellect and reasoning/thinking. This observation corroborates the findings of several studies such as Agyekum (2018) in Akan, Gwarzo (2015) and Will (2020) in Hausa, Kóczy (2019) in Hungarian, Bisilki (2022) in Likpakpaln, and Ochieng (2022) in Dholuo. Kraska-Szlenk (2014: 119) argues that "because of its upper location and distinctive features such as hair, colour and style, and facial features, the head stands out as a prominent part of the physical appearance of a person which triggers a cross-linguistically common metonymy: HEAD FOR PERSON. In this study, we found metonymic expressions in Nzema and Dangme, where the HEAD stands for A PERSON, an observation that also supports the findings of studies by Kraska-Szlenk (2020b: 145) and Ochieng (2022: 57). As Kraska-Szlenk (2014: 16) indicates, HEAD can refer to a part of the head, which is the fragment covered with hair. In line with this observation, Will (2020) and Ochieng (2022) also found HEAD FOR HAIR (WHOLE-FOR-PART) metonymy in Hausa and Dholuo as illustrated in (24) and (25) respectively; however, it is important to note that examples like these were not evident in our data.

(24) Hausa
      Ta   sa   a-ka   aske ma-ta   kai kwal
      3SG.F.PFV cause 4-PFV.REL shave DAT-3FSG head being.hairless
      'She had her hair shaved completely'. (Will 2020: 160).

(25) Dholuo
      Mary o-dhi   suko wi-ye
      Mary PFV-go plait head-POSS.3SG
      'Mary has gone to plait her hair'. (Ochieng 2022: 57).

Our data showed that both Nzema and Dangme speakers resort to "colour metaphors", specifically *white* and *black* in relation to the internal part of the head to describe POSITIVE THINKING and CRUELTY respectively. This is also

prevalent in Akan conceptualisation of HEAD (Agyekum, 2018). Almost all bad deeds (criminal acts) take place under the cover of darkness, and this seems to be universal, hence the metaphors WHITE/CLEAR/LIGHT is VIRTUE and DARK-NESS is VICE (see also Agyekum 2018: 42). Another intriguing finding is that, while Nzema and Dangme speakers conceptualise DARKNESS IN THE HEAD as WICKEDNESS (see example 8 above), Hausa speakers associate DARKNESS IN THE HEAD with IGNORANCE/LACK OF KNOWLEDGE (Will 2020: 175). In Hungarian also (Kóczy 2019: 239), DARKNESS IN THE HEAD relates to DULL-NESS/INABILITY TO UNDERSTAND SOMETHING. We identified head related expressions in Nzema that point to death and widowhood (see (17) and (18)). The Dangme example in (19), with regard to death, rather expresses the end of descent line. These observations, however, seem to be rare in previous scholarship on the figurative conceptualisation of HEAD. Parallel proverbial expressions related to HEAD in Nzema and Dangme were noted to express the sense of collaboration and cooperation, hard-work and proper time management, assigning duties to capable and experienced people among others. These expressions seem to be un-common in the literature explored in this study, except the work of Bobuafor (2021) on Tafi and Ewe proverbs.

8 Conclusion

The paper showed that both Nzema and Dangme speakers use head-related ex-pressions to describe pleasant and unpleasant experiences such as fortune, in-telligence, positive mindset, wickedness, insanity, unintelligence, and death. We also found that speakers of both languages have close similarities in their cultural conceptual systems in terms of head-related expressions. This, as we have hinted earlier in §1, may partly be due to the fact that Nzema and Dangme are both Kwa languages, and that both are spoken in the Southern part of Ghana. Both lan-guages appear to exhibit similar syntactic and semantic properties, hence the use of verbs like *open*, *grow* and the adposition *inside* to describe intelligence, that is perceived to be situated in the HEAD. We assume that the apparent linguistic similarities may account for why both languages use the colour adjectives *white* and *black* to modify HEAD when expressing positive mindset and wickedness respectively. The study concludes that, largely, the Nzema and the Dangme have similar tropes related to HEAD, and that both groups perceive their world based on embodiment and socio-cultural experiences.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions:

CONJ	conjunction	EMPH	emphatic particle
PART	particle		

References

Abdul-Hakim, Umar & Emmanuel Dogbey. 2019. Kasem metaphoric extension of *yi* 'eye' and *yuu* 'head' expressions. *International Journal of Literature and Arts* 7(1). 26–31.

Aboagye, Paul Alfred Kwesi. 1992. *Nzema Aneε ne Anwo Mgbanyidwεkε [The History of Nzema language]*. Accra: Bureau of Ghana Languages.

Accam, T. N. 1972. *Dangme Abε Gbi (Dangme proverbs)*. Accra: Bureau of Ghana Languages.

Agyekum, Kofi. 2015. *Ani* 'eye' metaphorical expressions in Akan. *Jounal of West African Languages* 42(1). 1–29.

Agyekum, Kofi. 2016. Bodily state and metaphors related to *ho* 'body' in Akan. *Metaphor and the Social World* 6(2). 326–344.

Agyekum, Kofi. 2018. *Akan body parts expressions: Cognitive semantics and pragmatic approach*. Accra: Adwinsa Publications Ltd.

Ameka, Felix K. & Mary E. Kropp Dakubu (eds.). 2008. *Aspect and modality in Kwa languages* (Studies in Language Companion Series 100). Amsterdam: John Benjamins.

Bisilki, Abraham Kwesi. 2022. Grammaticalised and non-grammaticalised conceptualisations of *liyi* 'head' in Likpakpaln (Konkomba). *Language & Communication* 83. 61–76.

Bisilki, Abraham Kwesi & Kofi Yakpo. 2021. 'The heart has caught me': Anger metaphors in Likpakpaln (Konkomba). *Sociolinguistic Studies* 15(1). 65–89.

Bisilki, Abraham Kwesi. 2019. Body-part terms in Likpakpaln. *Journal of West African Languages* 46(2). 146–163.

Bobuafor, Mercy. 2021. Cultural values and the pragmatic significance of proverbial sayings in Tafi and Ewe. *Journal of Pragmatics* 178. 192–207.

Caesar, Regina Oforiwah. 2016. An analysis of interrogative constructions in Dangme. *Journal of Literature, Languages and Linguistics* 21. 35–49.

Cruse, Alan. 2011. *Meaning in language: An introduction to semantics and pragmatics 3rd edition*. Oxford: Oxford University Press.

Dawson-Ahmoah, Grace N. & Patrick N. Wonkyi. 2021. Conceptualising MATU-RITY in the Mfantse dialect of Akan. *Sociolinguistic Studies* 15(1). 91–110.

Diaz-Vera, Javier. E. 2022. Soft heart and hard souls: The multiple texture of Old English feelings and emotions. *Cognitive Linguistic Studies* 9(1). 128–151.

Evans, Vyvyan. & Melanie Green. 2006. *Cognitive Linguistics: An introduction.* Edinburgh: Edinburgh University Press.

Gaby, Alice & John Bradley. 2019. *Wulaya* 'head' in Yanyuwa. In Iwona Kraska-Szlenk (ed.), *Embodiment in cross-linguistic studies: The 'Head'* (Brill's Studies in Language, Cognition and Culture 20), 263–272. Leiden: Brill.

Ghana Statistical Service. 2021. *2021 Population and Housing Census preliminary report.* Accra. http://www.census2021.statsghana.gov.gh/.

Gibbs, Raymond. W. 2005. *Embodiment and cognitive science.* Cambridge: Cambridge University Press.

Gibbs, Raymond. W. 2008. *The Cambridge Handbook of metaphor and thought.* New York: Cambridge University Press.

Gibbs, Raymond. W., Paula Lenz Costa Lima & Edson Francozo. 2004. Metaphor is grounded on embodied experience. *Journal of Pragmatics* 36. 1189–1210.

Gwarzo, Yusif Ahmed. 2015. *A study of metaphoric and metonymic expressions of body part terms hand and head in Hausa.* University of Malaya. (MA thesis).

Ibrahim, Rahinatu T., John Nyame, Mohammed Yakub, Kwasi Adomako & Esther Nana A. Wiafe-Akenten. 2022. Some Nzema philosophical concepts portrayed in their proverbs. *Ghana Journal of Languages, Linguistics and Literature* 1(1). 25–37.

Jackendoff, Ray & David H. Aaron. 1991. Review of *More Than Cool Reason: A Field Guide to Poetic Metaphor* by George Lakoff and Mark Turner. *Language* 67(2). 320–338.

Kóczy, Judit Baranyiné. 2019. 'He cracked his head feverishly': Conceptualizations of HEAD and THINKING in Hungarian. In Iwona Kraska-Szlenk (ed.), *Embodiment in cross-linguistic studies: The 'Head'* (Brill's Studies in Language, Cognition and Culture 20), 219–244. Leiden: Brill.

Kövecses, Zoltan. 2006. *Language, mind and culture: A practical introduction.* Oxford: Oxford University Press.

Kraska-Szlenk, Iwona. 2014. *Semantics of body part terms: General trends and a case study of Swahili.* Munchen: Lincom Europa.

Kraska-Szlenk, Iwona (ed.). 2020a. *Body part terms in conceptualisation and language usage* (Cognitive Linguistic Studies in Cultural Contexts 12). Amsterdam: John Benjamins.

Kraska-Szlenk, Iwona. 2020b. Metonymic extensions of the body part 'head' in mental and social domains. In Iwona Kraska-Szlenk (ed.), *Embodiment in cross-linguistic studies: The 'Head'* (Brill's Studies in Language, Cognition and Culture 20), 136–154. Amsterdam: John Benjamins.

Kropp Dakubu, Mary E. 1987. *The Dangme language.* London: Macmillan Publishers.

Kwaw, Francis Ehoma. 2008. *Nzema Aneɛmɛla Tagyee Ne (The Grammar of Nzema).* Ayinase: Atwe Royal Publishers.

Kövecses, Zoltan. 2002. *Metaphor: A practical approach.* Oxford: Oxford University Press.

Lakoff, George. 1993. The contemporary theory of metaphor. In A. Ortony (ed.), *Metaphor and thought*, 202–251. Cambridge: Cambridge University Press.

Lakoff, George & Mark Johnson. 1980. *Metaphors We Live by.* Chicago: University of Chicago Press.

Lakoff, George & Mark Turner. 1989. *More than cool reason: A field guide to poetic metaphor.* Chicago: University of Chicago Press.

Langacker, Ronald. W. 1994. Culture, cognition and grammar. In Martin Pütz (ed.), *Language contact and languages conflict*, 25–53. Amsterdam: John Benjamins.

Lawer, Richard Ayertey. 2017. *Compounding in Dangme.* University of Ghana, Legon. (MA thesis).

Lee, David. 2005. *Cognitive Linguistics: An introduction.* Oxford: Oxford University Press.

Luo, Yongxian. 2020. 'Head' as a link of embodiment in Chinese. In Iwona Kraska-Szlenk (ed.), *Embodiment in cross-linguistic studies: The 'Head'* (Brill's Studies in Language, Cognition and Culture 20), 99–123. Amsterdam: John Benjamins.

Ochieng, Joseph Jaoko. 2022. Conceptualisation of Head in Dholuo. *Studies in African Languages and Cultures* 56. 49–80.

Palmer, Gary. B. 1996. *Toward a theory of cultural linguistics.* Austin: University of Texas Press.

Peña-Cervel, María Sandra & Francisco Jose Ruiz de Mendoza Ibañez. 2022. *Figuring out figuration: A cognitive linguistic account.* Amsterdam: John Benjamins.

Rohrer, Tim. 2007. The body in space: Dimensions of embodiment. In Jordan Zlatev Tom Ziemke & Roslyn M. Frank (eds.), *Embodiment*, 339–378. Berlin: De Gruyter Mouton.

Al-Saleh, Tasneem, Jihad Al-Shuaibi & Ronad Al Momani. 2020. A cognitive analysis of head and heart metaphors in English and Spanish. *International Journal of Arabic-English Studies (IJAES)* 20(2). 115–132.

Semino, Elena. 2008. *Metaphor in discourse.* Cambridge: Cambridge University Press.

Sharifian, Farzad. 2003. On cultural conceptualisations. *Journal of Cognition and Culture* 3(3). 187–207.

Sharifian, Farzad. 2011. *Cultural conceptualisations and language: Theoretical framework and applications*. Amsterdam: John Benjamins.

Sharifian, Farzad. 2014. Conceptual metaphor in intercultural communication between speakers of Aboriginal English and Australian English. In Andreas Musolff, Fiona MacArthur & Giulio Pagani (eds.), *Metaphor and intercultural communication*, 118–129. London: Bloomsbury Publishing.

Sharifian, Farzad. 2017. *Cultural linguistics: Cultural conceptualisations and language*. Amsterdam: John Benjamins.

Sime, Abinet. 2020. Semantics of Amharic *ras* 'head'. In Iwona Kraska-Szlenk (ed.), *Embodiment in cross-linguistic studies The 'Head'* (Brill's Studies in Language, Cognition and Culture 20), 183–204. Leiden: John Benjamins.

Soboh-Blay, Alfred. 2013. *Nyamenle Asa Ɛnlomboɛ Ne (The Heavy hand of God)*. Accra: Bureau of Ghana Languages.

Steen, Gerard. 2011. The contemporary theory of metaphor—now new and improved. *Review of Cognitive Linguistics* 9(1). 26–64.

Will, Izabela. 2020. The conceptualisation of HEAD among the Hausa based on verbal and nonverbal representations. In Iwona Kraska-Szlenk (ed.), *Embodiment in cross-linguistic studies: The 'Head'* (Brill's Studies in Language, Cognition and Culture 20), 157–182. Amsterdam: John Benjamins.

Yakub, Mohammed & Kofi Agyekum. 2022. Metaphorical euphemisms in death-discourse among the Nzema. *Studies in African Languages and Cultures* 56. 127–153.

Yu, Ning. 2004. The eye for sight and mind. *Journal of Pragmatics* 36. 663–686.

Yu, Ning. 2020. Linguistic embodiment in linguistic experience: A corpus-based study. In Iwona Kraska-Szlenk (ed.), *Body part terms in conceptualisation and language usage: Cognitive linguistic studies in cultural contexts*, 11–30. Amsterdam: John Benjamins.

Chapter 8

Left-sided vs. right-sided phonology of labial-velars

Michael Cahill

SIL International

Labial-velars are shown to be phonologically single units by diagnostics such as duration and their patterning in syllable structure. However, they also exhibit processes in which the active feature is [dorsal] on their left side, but [labial] on their right. This is shown by partial nasal assimilation as ŋKP and KPm, as well as other processes. Several phonological models are found to be inadequate to explain the range of these behaviors. Standard or enhanced versions of Feature Geometry and Articulatory Phonology can account for the phonological sidedness behaviors noted here, but not with other behaviors, and issues still remain in combining the phonology with the phonetics.

1 Introduction

The term "labial-velars" refers to $\widehat{kp}$, $\widehat{gb}$, $\widehat{ŋm}$ (and modifications thereof such as $^{ŋ}\widehat{gb}$) with approximately simultaneous labial and velar articulations. In this paper, unless referring to a specific type of labial-velar, these will be generalized under the label KP.

The "approximately simultaneous" label for the labial and velar articulations is an appropriate designation in many contexts – they largely overlap. But there is additional systematic, cross-linguistic detail that is foundational to this paper. In every case in which this issue has been examined, the velar articulation slightly precedes the labial, and the labial persists slightly longer. For example, this large but partial overlap is clearly seen in the electromagnetic articulography measurements in Figure 1 for Ewe [ewe] (Maddieson 1993). In this study, metal pellets

were glued to the lips and tongue dorsum, and a metal-detecting sensor produced the positional readouts. Note that the "k" articulation slightly precedes the "p" articulation, and the "p" articulation persists after the "k" has finished.
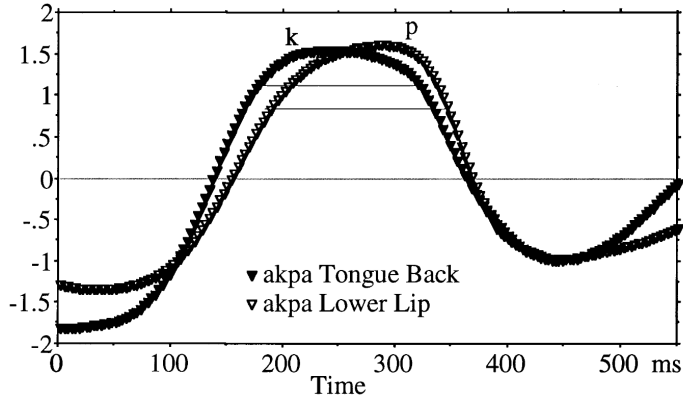

Figure 1: Coordination of lower lip and tongue back movements in the Ewe word *ákpá* 'too much'. Y-axis is vertical displacement, normalized scale, mean of ten tokens aligned at release; horizontal lines indicate the likely duration of actual contact of the articulator. (Figure 6 from Maddieson (1993))

Spectrographic evidence from Leggbo [agb] in Figure 2 also shows a velar onset and labial release (note the "velar pinch" of F2 and F3 going into the consonant). Spectrograms of intervocalic KP are also presented for Dedua ([ded]) and Efik ([efi]) in Ladefoged & Maddieson (1996: 336-37) and others in Connell (1994), showing a velar onset and a labial release. So, a KP largely, but not totally, overlaps velar and labial articulations.

Besides the phonetic evidence of gestural overlap, KPs exhibit a variety of patterns indicating they are units rather than sequences:

- They occur in languages which have only unambiguous single consonant syllable onsets (CV, CVN, CVV, CVVN). No consonant clusters occur word-initially in many of the languages cited in this paper, yet KP does occur word-initially.

- Their duration is much closer to single stops than to clusters. Consonant clusters typically have 1.5-2 times the durations of single segments (Ladefoged & Maddieson 1996: 333). The duration of labial-velars is slightly longer than simple stops (Yoruba gb/b = 132/128 ms), but does not approach
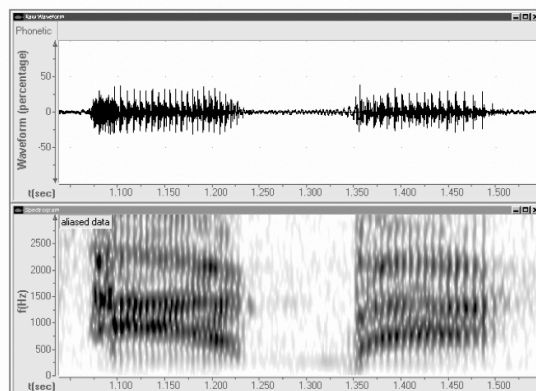
Figure 2: Spectrogram of [ag͡ba] from Leggbo (Nigeria) (recording courtesy of Julie Larson)

the duration of consonant clusters. They also report similar data for Ewe, as Connell (1994) does for Igbo, and Demolin (1991) for Mangbetu. In a direct comparison, Ladefoged and Maddieson also show by spectrographic evidence that Eggon's [kp] consonant *cluster* is longer than the [k͡p] *unit*, though no numerical measurements are given.[1]

- Turning to specific language data, in Ewe reduplication, KPs also act as single segments. A word with an initial consonant cluster has a reduplicant with only the first consonant, including KP, e.g. *fle* 'to buy', *fe-flee* 'bought' vs. *kplo* 'to lead', *kpo-kplo* 'leading' (Ansre 1963).

- In Kaanse [gna], /kp/ becomes totally (not partially) voiced after a nasal, e.g. *sànɪ kpógorò* 'sheep-shelter', *sʉ́ŋ gbógorò* 'chicken-shelter' (Showalter pc).

---

[1]Gouskova & Stanton (2021: 183) "doubt a universal correlation between segmenthood and duration". However, the claim in this paper is limited to labial-velars, not the wider variety of sounds ("universal") they consider. More crucially, they do not consider the evidence here, and most particularly, they note that "the most straightforward evidence for a duration/segmenthood link would have to come from languages that contrast complex segments with same-phone clusters" (Gouskova & Stanton 2021: 184). This is precisely what the Eggon example above presents.

- In Mano [mev], /gb/ has an allophone [ŋm] before a nasalized vowel; that is, the voiced labial-velar is completely (not partially) nasalized (Welmers 1973: 47)[2]

This multiplicity of patterns indicates that KP are phonological units rather than sequences.[3] However, other processes are sensitive to edge effects, and the presence of both "unit phonology" and "edge phonology" indicates phonology operating on different levels. Nasal place assimilation, which can occur with nasals either preceding or following KP, is but one process that shows sensitivity to the edges of KP, not the whole segment.

2 Left-sided phonology

2.1 Nasal place assimilation

A nasal preceding KP, whether as prenasalization, a distinct morpheme, or within a morpheme, is reliably transcribed as *ŋKP* or *^{ŋ}KP* as a product of assimilation or an independent phoneme in over 85 documented languages for which I have data (Appendices A-B), while a total assimilation as *ŋmKP* or *^{ŋm}KP* is attested in at least 56 languages (Appendices C-D). There is sometimes uncertainty as to which of these two options is correct, but some investigators (e.g. Boyeldieu 2006) have differentiated *ŋKP* from *ŋmKP* in different languages (Bagiru and Ngiti), so these researchers, at least, are not only aware of the difference, but deliberately record these as different.

As Clements & Rialland (2008: 42) note:

> …in homorganic nasal–stop sequences, it is the dorsal feature that typically spreads to the preceding nasal, yielding [ŋmgb] or [ŋgb].

---

[2]A reviewer points out that the Kaanse and Mano examples have an alternate explanation, that the KP could be a cluster, not a unitary segment. As noted on the previous page, syllable structure inventories in these languages make this alternative untenable; /kp/ and /gb/ occur word-initially, but no unambiguous consonant clusters do. Also for these languages, Kaanse has codas only of nasal and glottal stop (Showalter 1997), and Mano does not have any stop consonant clusters (Khachaturyan 2015, 2018).

[3]Because of the inherent contradiction in the features associated with labial and velar articulations at the time, Jacobsen et al. (1953) claimed that labial-velars were an extreme form of consonant *cluster*. As Anderson (1976: 20) notes, such a claim is "counter to all previous treatments and (what is more to the point) quite at variance with the phonetic and phonological properties" of labial-velars, such as those noted here.

I regard the few transcriptions recorded as *mKP* as dubious, which I address in this section.

The two types of well-documented nasal place assimilation are thus limited to a total place assimilation *ŋmKP*, or a partial place assimilation *ŋKP*. It is the latter, illustrating sensitivity to the left edge of KP, that we are primarily concerned with here.

Ryder (1987) cites several languages in which a nasal segment assimilates to KP as [ŋ] (as well as languages assimilating as [ŋm]). Partial nasal place assimilation is exemplified below for a few languages. See Appendix A for a fuller list of 20 languages illustrating this phonological process.

(1) Gã [gaa] (Ryder 1987)
  a. ŋgbek   'my child'
  b. ŋkpai   'my cheeks'
  c. taaŋkpee 'sisal'

(2) Dagaari [dga] (Kennedy 1966, personal data)
  a. kpàŋkpàŋ 'upper arm'
  b. gbáŋgbáŋ 'noon'

(3) Vagla [vag] (Crouch & Smiles 1966)
  a. tʃaŋkpalŋa 'antelope'
  b. sʌŋgbo   'baboon'

(4) Mono [mnh] (Olson 2005)
  a. ŋgba   'be many'
  b. kéŋgbā   'alone'

Olson (2005: 33) notes the pronunciation as [ŋg͡b], though written orthographically as <ngb>

(5) Gonja [gjn] (Painter 1970)
  gbìŋgbìŋ   'big'

Painter (1970: 36) notes that "when /gb/ is preceded by a syllabic nasal this nasal has a velar, not a bilabial or labial-velar articulation".

Many languages also exhibit a prenasalized KP. Similar to the pattern across syllables above, this is often realized as ^{ŋm}KP, but also occurs as ^{ŋ}KP, as in (6) from Bongo. Kilpatrick (1985: 8) notes, "The prenasalization just has velar closure, rather than both labial and velar closure".

(6)   Bongo [bot] (Kilpatrick 1985)
    /ᵑgb/:   ᵑgbáyá   'corn'

The dozens of languages which have ᵑKP as a phoneme include Ambele [ael] (Nganganu Kenfac 2001), Kako [kkj] (Ernst 1996), Yango [yng] (Bostoen & Donzo 2013), Mündü [muh] (Jeffrey & Polley 1981), Avokaya [avu] (Callinan 1981), Logo [log] (Goyvaerts 1983). See Appendix B for a list of 66 such languages.

The patterns [ŋKP] and [ŋ͡mKP] are well-documented (see Appendix C for 10 languages showing assimilation as [ŋ͡mKP], and Appendix D for 46 languages showing /ŋmKP/ as a phoneme). In contrast, the actual likelihood of an [mKP] can be uncertain, for several reasons.

First, the literature that reports assimilation as [m] does so in a way that indicates the writer has never considered the possibility of [ŋm]. A labial closure is observed, and it is assumed that that is all there is to it.

Second, in some languages, the orthographic convention does not match phonetic or phonological reality. Yoruba uses orthographic <p> to represent /kp/ (Folarin 1987, inter alia) and orthographic <m> for [ŋm].

Third, the different nasals which are possible before KP are not always easy to distinguish by ear alone. Besides personal experience, we see that this issue was noted almost a century ago. Ward (1933) writes concerning Efik [efi]:

> It is, however, extremely difficult to *hear* which is being said without *seeing* the presence or absence of lip-articulation... There are some words in which *m* has been written and others in which *ŋ* occurs. *mkpa*, death; *ŋkpɔ*, thing. It is probable that both articulations are made at the same time, i.e. a labio-velar nasal consonant... (Ward 1933: 10)

Note that Ward indicates the probability of [ŋmkp], but never writes it as such. The difficulty of distinguishing [ŋmkp] and [ŋkp] is exacerbated by their word- and utterance-initial positions, as these positions have no vocalic transition into the nasal. For Efik more recently, (Welmers 1968: xii) observes that a nasal before /kp/ in Efik is pronounced "with simultaneous closure at the lips and with the back of the tongue", i.e. [ŋm], but Welmers (1973: 47) comments that "For some unknown reason, in the usual orthography of Efik, *mkp* is written in some cases but *ŋkp* in others". Both Cook (1969) and Ohala & Ohala (1993) elucidate this by noting that the nasal assimilating to /kp/ manifests itself as either [ŋ] or [ŋm] in Efik.

Another case of alleged [mkp] occurs in Bikele, also called Kol [biw], cited as having /mkp/ and /mgb/ in Begne (1980: 30-33). However, the more recent

Henson (2007) records every nasal before KP as [ŋ], and notes the possibility of them being [ŋm] (Henson, pc). Furthermore, Begne notes that the number of words with either labial-velar is very limited, around a dozen, and these are all borrowed, generally from Ewondo [ewo]. The cognate Ewondo words in two dialects are shown below (data courtesy of Steven Bird, pc).

(7) | | **Bikele (Yaoundé)** | **Ewondo (Mbalmayo)** | **Ewondo** | gloss |
|---|---|---|---|---|
| a. | mkpálá | ə́ŋ̀kpálá | ŋ̀mkpálá | 'playful' |
| b. | mkpámág | ə́ŋ̀kpámán | ŋ̀mkpámán | 'new' |
| c. | mkpeg | ə́ŋ̀kpɔ̂k | ŋ̀mkpâk | 'favorite co-wife' |
| d. | mgba | ə́ŋ̀gbà | ŋ̀mgbà | 'friendliness' |

Interestingly, the Ewondo data, in which particular care was taken in the nasal transcription, shows [ŋkp] in one dialect and [ŋmkp] in the other. A borrowed word is typically changed to fit the receptor language's phonology. But the fact that neither the source language nor any other languages clearly attest [mkp] makes the report of Bikele /mkp/ dubious.

Finally, some publications which record /mKP/ are corrected at a later time. For example, Boyd (1997) lists /mkp/ as a phoneme. Upon query, she responded (pc, 2023) "Yes, it is /ŋmkp/ or perhaps more correctly ŋᵐkp. Sorry for the 'shorthand'".

As the above examples show, transcriptions of *mKP* are uncertain at best.

To sum up, a nasal preceding KP may be totally assimilated to the place of KP as *ŋmKP*, but it is also common to have a partial place assimilation, and if so, this yields *ŋkp*, not *mkp*. The nasal assimilates to the left edge of KP, the velar.

2.2 Other left-sided phonology

In Kɔnni [kma] (Cahill 2007a), vowel epenthesis occurs between segmental morphemes with differing place values (e.g. r-k, r-b, b-ŋ, and b-kp), but not between morphemes with the same place (e.g. r-t, b-b, n-r, and g-kp):

(8)   Kɔnni
   a.   /b-kp/:   /kɔb-kpɪŋ/   →   kɔ̀b-ì-kpɪ́ˈíŋ   'big bone'
   b.   /g-kp/:   /hɔg-kpɪŋ/   →   hɔ̀k-kpɪ́ˈíŋ   'big woman'

Note that in (8a) the labial /b/ preceding the KP is treated as a different place than that KP, but in (8b) the velar /g/ preceding the KP is treated as the same place as the KP. This process, involving sounds relating to the left side of KP,

treats KP as velar. This pattern would predict that for a KP-C sequence, a vowel would epenthesize if C is velar, but Kɔnni does not end morphemes with KP.

It is anticipated that more research will uncover more cases of "left-sided phonology", but with KP often limited to morpheme-initial or word-initial positions, the required environment for these is not as common as for the right side, and many phonology sketches do not go into the detail needed to document either of these.

3 Right-sided phonology

3.1 Nasal place assimilation

Nasals occurring after KP are not nearly as common as those preceding KP, with even so well-informed a writer as Ohala (1993: 690), among others, not showing awareness that they exist. But the documented cases mirror those discussed in §2; i.e., *KPŋm* and *KPm* occur, but not *KPŋ*.

The *KPŋm* pattern has few documented cases, listed below and in Appendix F. For the Kuta dialect of Gwari [gbr], Hyman & Magaji (1970) cite phonetic syllables [pᵐa], [tⁿa], [kᵑa], [kpᵑᵐa], but give no actual words which contain them.

Mada [mda] of Nigeria (Price 1989) has an unusual syllable type: a stop followed by a syllabic nasal, e.g. the middle syllable in [kpa.kŋ.ki] 'tree stump'. When a nasal follows KP, there is total place assimilation as KPŋm (the posited underlying /m/ below is arbitrary, and could just as well be posited as /KPŋm/):

(9)   Mada
     a.   /kpm̀/   [kpŋm̀]   'kapok tree'
     b.   /gbm̄/   [gbŋm̄]   'canoe'

Konabere [bbo] (Phil Davison, pc) also has a number of words with syllabic nasals as syllable peak (tone unmarked).

(10) Konabere
     a.   g͡bŋ͡m   'black'
     b.   k͡pŋ͡m   'war'

Other languages show partial nasal place assimilation on the right side of the labial-velar (Appendix E). Especially relevant to this discussion, the Tyebaara Senoufo language [sef], (Mills 1984: 94) shows partial nasal place assimilation as KPm:

(11) Tyebarra Senoufo
 a.   kpmɔ́:   'to beat'
 b.   n�<̀>-gbmɔ́:   'herb doctor'

Finally, the Gwari language [gbr] (Rosendall 1992) exhibits *both* patterns, with nasals on either side of a KP, in at least the Giri dialect. This shows that the partial nasal place assimilation is dependent on whether the nasal precedes or follows the KP. In (12c) particularly, we see both in a single word.[4]

(12) Gwari
 a.   tʃĩŋkpè   'stool'
 b.   kpmàmí   'okra'
 c.   wʲédʒíŋgbmà 'dark'

3.2 Other right-sided phonology

Nafaanra [nfr] (Jordan 1980) has not only syllabic and plain nasals (13a-b), but also post-oralized nasals (13c-d), a relatively rare phenomenon. The post-oralized labial-velar releases into a labial (ŋm^b), not a velar (13d). This process, involving sounds *following* KP, treats KP as labial.

(13) Nafaanra
 a.   nthó:sì 'tomato' c.   n^{d}ú: 'to climb'
 b.   mãnã   'nose'   d.   ŋm^{b}a 'him'

Parallel to the two possibilities of nasal place assimilation (ŋg͡b and ŋ͡mg͡b), one would predict that there would be cases in which a post-oralized nasal labial-velar releases into a full labial-velar, that is, not only ŋm^b as above, but also ŋm$^{g͡b}$. At this point, I am not aware of such cases, probably because post-oralized nasals are rare in the world's languages,[5] and the intersection of this probability with that of the frequency of labial-velars results in a lower probability still.

In Ejagham [etu], Watters 1981, /i/ becomes [ɨ] following either a labial OR labial-velar when the vowel precedes a velar (14a,b). Here, the right side of labial-velars patterns with labials. Note that if the /i/ precedes a velar but follows "any

---

[4]At least some non-African languages also illustrate some of these same patterns. The Yeletnye language of Papua New Guinea [yle] also illustrates nasals both before and after a KP, e.g. [ŋ͡mg͡ba:] 'constrict', [k͡pŋ͡mĩ:] 'coconut'. However, these are total nasal place assimilation, not partial (Henderson 1995: 8). The Nambo [ncm] language, also of Papua New Guinea, has a phonemic prenasalized KP, which shows up as the partial assimilation ŋg͡b, as in /jɐ̃ŋg͡b/ 'bag' (Kashima 2021).

[5]See Wetzels & Nevins (2018) for a discussion of post-oralized nasals vs. prenasalized stops.

consonant other than a labial or labial-velar" (Watters 1981: 39), the /i/ is realized as [ɪ], as in (14c-e).

(14) Ejagham
    a. /é-bîg/   →   [éβîk]   'it is enough'
    b. /ɔ̀-kpígì/ → [ɔ̀kpɨ́ɣì] 'you turned'
    c. /à-ríg/   →   [àrík]   'ropes'
    d. /ɔ̀-síŋ/   →   [ɔ̀síŋ]   'mangoes'
    e. /ɔ̀-kîg/   →   [ɔ̀kʰîk]   'cheek' (Watters, pc)

In the Mande language Dan (Santa) [daf], it is reported that before nasalized vowels, the labial components of both /kp/ and /gb/ "tend to be realized as [m]" (Bearth & Zemp 1967). They give the examples in (15) (the "1" superscripts mark high tone). The nasality of the vowel spreads only to the labial portion, the right side, of KP:

(15) Dan
    a. /gbã^1/ [gmã^1] 'leg'
    b. /kpã^1/ [kmã^1] 'basement'

Since labiality is often (but not always) associated with rounding, one might predict the possibility of a process rounding a vowel following a KP. If so, this would also occur following a plain labial. I am not aware of any such processes, but in a somewhat related process in some languages, KP is consistently labialized, e.g. Williamson (1965: 19) notes that in the Kolokuma dialect of Ijo, kp and gb are "both produced with rounded lips".

In summary, we see that labial-velars exhibit a mixture of identities. For some processes, they act as units with no discernable internal phonological structure. But with other phonological processes, they exhibit a sensitivity to the left and right edges of KP, with the left edge acting as velar, and the right edge acting as labial. How can these left- vs. right-sided processes be formalized, with KPs composed of both [dorsal] and [labial] features? The challenge of how to account for this bifurcate nature is the topic of the remainder of this paper.

4 Phonological approaches

How a few major relevant phonological theories have interacted with labial-velars is summarized in this section. As we shall see, while some models explain some phenomena quite well, no theory covers all the phonological patterns of labial-velars that have been presented in the previous two sections.

4.1 SPE

Chomsky & Halle (1968)'s *Sound Pattern of English* (SPE) has largely been superseded, but it did specifically address labial-velars. In this framework, features were unordered and binary. The pertinent question became whether labial-velars "are labials with extreme velarization or velars with extreme rounding" (Chomsky & Halle 1968: 311). Anderson (1976: 21) expands on this concept, noting that the two articulations have opposite values of the [anterior] feature, which is impossible for a single segment. Consequently, labial-velars were required to have either one place or the other as primary, and the other as secondary. Even with phonetically identical exponents, it was a language-specific matter which place was primary.

The partial nasal place assimilation as ŋKP discussed in §2.1 was the evidence that led Chomsky & Halle (1968) and Anderson (1976), though not as strongly in Anderson (1981), to propose velar as the primary place of articulation for KP in some languages. But we have seen that this is a consequence of the left-sided phonology of KP. They do not discuss the existence of [ŋmKP], which, by the same reasoning, should show a labial-velar as having *both* places of articulation as primary. Also, the previously cited case of Gwari in (12), in which a nasal assimilates as velar to the left of KP and labial to the right (*wʲédʒíŋgbmà* 'dark'), shows that nasal assimilation cannot be the defining factor in determining primary place.

Besides leaning on nasal place assimilation as evidence, the SPE approach was supported by Anderson (1976) largely on distributional grounds, nicknamed the "filling the gaps" criterion. For example, Limba has /k/ and /gb/, but has no /g/, so /gb/ "fills the gap" as a velar, with secondary labialization. If another language has /kʷ/ but no /p/, then /kp/ fills the gap as a labial, with secondary velarization. This criterion thus wholly depended on criteria external to the segment itself, rather than phonetic nor phonological characteristics inherent to the segment. Applying the two criteria for a language also can give opposing results. Efik is a language for which Anderson (1976) posits labial-velars as labial as primary because of distributional patterns, but notes later (Anderson 1981: 499) that the nasal assimilation pattern points to velar as primary.

4.2 Feature Geometry

The standard Feature Geometry model (FG, Clements & Hume 1995) accounts for some labial-velar phonology, but not the left- and right-sided phonology described above. A specific modification to FG developed below would be con-

sistent with these directional patterns, but then becomes problematic for other labial-velar phonology.

In Feature Geometry, the [dorsal] and [labial] features are in separate tiers, and thus are specifically unordered (Clements & Hume 1995: 249, 253), since only features or elements in the same tier are ordered (unlike the elements in syntactic trees). This unordering means that features in separate tiers are simultaneous, and such features overlap phonetically. The configuration in Figure 3a below accounts nicely for the phonological unity of KP. However, because the features are unordered, the partial nasal place assimilation to one feature but not the other is purely arbitrary, and thus the predominance of ŋKP and KPm is not predicted.
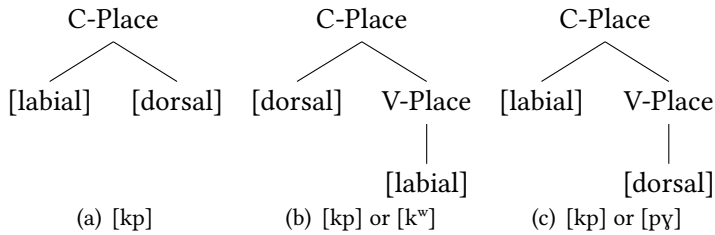
Figure 3: Possible labial-velar geometries (partial representation)

Alternative representations are Figure 3b,c and these correspond somewhat conceptually to the SPE model in that either [dorsal] or [labial] is the primary feature. Since the existence of a particular feature does not imply the degree of closure of that feature, Figure 3b,c are ambiguous in the phonetic segment they represent. Cahill (1998) proposed Figure 3c as a universal configuration for labial-velars, with rules of assimilation to V-Place to account for ŋKP, and assimilation to C-Place to account for ŋmKP. This also accounts for the unusual Dagbani pattern in which a labial-velar becomes a labial-coronal before a front vowel (Cahill 2007b):

Figure 4: Dagbani assimilation rule: $\widehat{KPI} \rightarrow \widehat{TPI}$

But though there are several phonological processes, such as co-occurrence restrictions and neutralization patterns, that indicate the phonological prominence of [labial] for KP (cf. Cahill 2006), the configuration in Figure 3c does not account for KPm, or the other edge effects in (1–15). Connell (1998-1999) specifically asserts that labial-velars are problematic for Feature Geometry, largely because of the type of asymmetries discussed here.
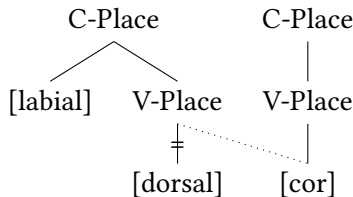
The central issue of this study is the fact that labial-velars act both as single units and as complex sounds, with significant differences in processes sensitive to their left side and right side. Other classes of segments – contours – have somewhat of the same issue, but we will see that labial-velars have significant differences from these. Contours include prenasalized and postnasalized stops, which contain both [+nasal] and [-nasal] in the same segment, and affricates, which contain both [+continuant] and [-continuant] in the same segment. For affricates, the ordering of the [continuant] features is predictable, with [-continuant] always preceding [+continuant]. Partly for this reason, Lombardi (1990) proposes "that affricates are composed of [-cont] and [+cont] specifications which are unordered at underlying representation and throughout the phonological derivation, although they are ordered phonetically". See similar discussion of this lack of underlying ordering of [±cont] in van de Weijer (1996).

Prenasalised and postnasalized stops are unlike affricates in that their parts are phonologically ordered, shown by the existence of contrast between prenasalized segments with the [+nas][-nas] order and postnasalized segments with a crucial [-nas][+nas] order. See further discussion of these in an aperture model, but with [nasal] as privative in Steriade (1993).

To deal with the directionality issues, van de Weijer (1996) proposes a two-root analysis in the Dependency Phonology framework (as in Figure 5 for clicks, as well as prenasalized stops, postnasalized stops, affricates, and labial-coronals, though without a specific application to labial-velars). For all these, the two root nodes are connected to a single timing position.

The pre-/postnasalized stops and affricates have only partial relevance to labial-velars, since articulations of these do not overlap, as those of labial-velars do. Ulfsbjorninn (2021) has a similar proposal within Element Theory, with ordered root nodes for labial-velars, linked to a single timing position. This differs from Van de Weijer's proposal, which deals with contour segments, not complex segments.

However, the basic idea of a segment that can be represented both as a unit and also with different featural structure on the left and right sides in FG is worth a closer examination, though we will ultimately see it has critical weaknesses. van de Weijer (1996: 65) notes:
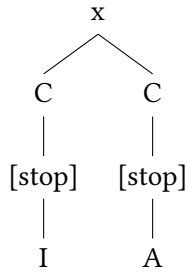
Figure 5: Representation of click reproduced from van de Weijer (1996: 199)

In Clements & Keyser (1983), a case is made for representing long consonants and vowels with two positions on the timing tier, both connected to the same root node on the melodic tier. The opposite situation, two root nodes connected to a single timing position, is also predicted to be a well-formed phonological representation.

This connection of two featural nodes to a single timing position also has a parallel in tone systems of the world. Yip (1989) solved the problem of contour tones spreading as units by proposing another node in the autosegmental representation. To represent a contour tone, at least in the Asian languages, Yip proposed the configuration in Figure 6.
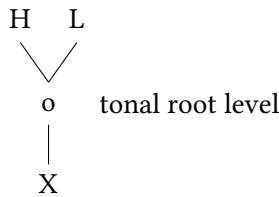
Figure 6: Contour tone representation from Yip (1989)

The rule for spreading a HL contour tone as a contour spreads the "tonal root level" node intermediate between the TBU and the H and L tones, and carries them both along. Could it be possible that the same type of reasoning could be applied to the unitary nature yet differing phonologies of the left and right sides of labial-velars? As Clements & Hume (1995: 259) wrote, "...any feature or feature set that assimilates as a unit must constitute a node on an independent tier of its own".
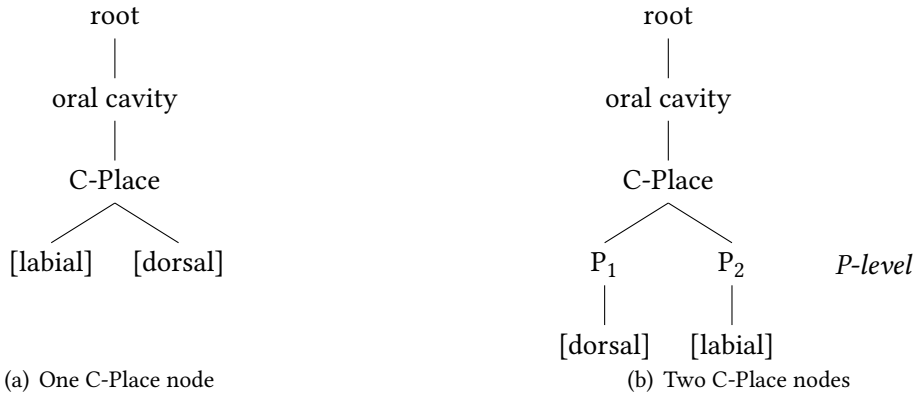
Figure 7: One vs. two C-Place nodes for labial-velars

In Figure 7a, the figure represents a somewhat fuller representation of Figure 3a, a common way that KP has been presented in Feature Geometry, with the [labial] and [dorsal] features attached to the same timing unit, overlapping, but unordered. The representation in Figure 7b depicts a new node level, provisionally labeled "P", between the C-Place node and the terminal features. The [dorsal] and [labial] attached to two separate P nodes. The nodes P_1 and P_2 are crucially ordered, with [dorsal] preceding [labial] here.

This ordered configuration of Figure 7b can now account for both the unitary nature of KP and its phonological edge effects. Total nasal assimilation will assimilate the place of a nasal on either side of the KP to the C-place node, yielding ŋmKP or KPŋm. Partial nasal place assimilation on the left will yield the correct ŋKP by associating the place of the nasal to the P_1 node, and partial nasal place assimilation on the right will yield the correct KPm by associating the place of the nasal to the P_2 node.

A significant drawback of the configuration in Figure 7b is that how to interpret this is not obvious, in light of the relative phonetic timing of the [dorsal] and [labial] features, as gestures. In Figure 7a, the features are attached to a single C-Place and this to a single timing unit, implying simultaneous articulations, which largely agrees with the phonetics presented in Figure 1. However, in Figure 7b the P_1 and P_2 nodes are sequential and non-overlapping, just as the H and L tones were sequential in Figure 6. Thus, the configuration in Figure 7b, though consistent with the left-and right-sided phonological patterns of KP discussed here, is not consistent with the articulatory phonetics of KP.

However, while prenasalization displays a relatively sharp boundary between [+nas] and [-nas], the phonological falling tone HL does not consist of a level H that drops instantaneously to a level L. Rather, it moves smoothly from one articulation to the next. So, might the sequential features [dorsal] [labial] be interpreted phonetically as 1) moving from one place to the other, with 2) significant overlap in articulation? These are two separate issues. Both tone and KP shift from one place to the other. But while the [labial] and [velar] places overlap, the H and L tones do not overlap, but have a brief transition that is neither H nor L. The overlap problem remains for KP.

While the issue of phonetic overlap is problematic, the Dagbani process in Figure 4 KP → TP / _front vowels, repeated below as Figure 8 and expressed in terms of the configuration in Figure 7, is actually incompatible with this.
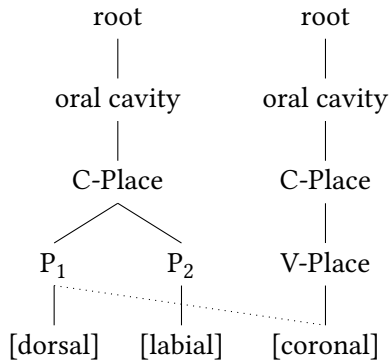
Figure 8: Dagbani assimilation rule: $\widehat{\text{KP}}$I →$\widehat{\text{TP}}$I (modified)

This is firstly because the [coronal] feature of the vowel displaces and delinks the [dorsal] feature of KP. But the ordering of features in Figure 7b means the KP [dorsal] is not adjacent to the vocalic [coronal]; the [coronal] would have to cross lines to associate to the mother node of [dorsal] which it displaces. Secondly, unlike the well-formed rule in Figure 4, [coronal] above is shared by two unlike nodes, and it is not at all clear how this ill-formedness could be repaired.

At this point, it appears that while different KP phenomena can be captured quite neatly by different versions of FG, no version of FG can be proposed as universal to account for all the phonology of labial-velars.

4.3 Feature Class Theory

Constraints in various instantiations of Optimality Theory can be formulated to describe the phenomena above and other KP phenomena, but would describe

surface patterns without providing a principled and non-arbitrary account.

The Feature Class Theory (FCT) variety of Optimality Theory described in (Padgett 1995, 2002) does away with the nodes and organization of FG in favor of direct reference to the features. With respect to labial-velars, Padgett has dealt specifically with both partial and total nasal assimilation on the left side of KP. He accounts for ŋKP as in Figure 9, with the nasal ([+son]) linking directly to the [dorsal] feature rather than a C-Place node.

Partial NPA to [ɡb] in FCT

[+son]   [-son]

[dor]   [lab]
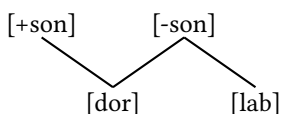
Figure 9: Partial place assimilation, yielding [ŋɡb] Padgett (1995: 153)

However, since the [dor] and [lab] features are unordered, they could just as easily be represented as [lab] [dor]. In that case, the result would yield the undesirable mKP. Also, this approach, exactly like the basic Feature Geometry discussed above, does not and cannot distinguish between this and the mirror image right-sided nasal place assimilation of KPm.

4.4 Articulatory Phonology

Nasal place assimilation and the other directional phenomena cited above may be more amenable to a phonetically-sensitive approach rather than to an abstract phonological one. Articulatory Phonology (AP, Browman & Goldstein 1986, 1989, 1990, 1992) makes reference to primitive units of gestures and their temporal coordination as speech proceeds, as represented in the "gestural score" of Figure 10, with each gesture having both a location and degree of closure.

Such a gestural score makes reference to timing in a more fine-grained degree than to models discussed above. The coordination of gestures is expressed as degree of temporal overlap. The differing nasal assimilation patterns can be represented straightforwardly, as in the scores below, using the "box notation" of Browman & Goldstein (1989, 1990, 1992).

Here, the relevant articulators start with VELUM, with a "wide" articulation indicating it is open, i.e. there is airflow through the nose – the nasal part of the utterance. The T.BODY articulator, as "closed velar", indicates the tongue dorsum is firmly against the roof of the mouth, and lasts for the entire pronunciation except for a small portion at the end. The timing of these two gestures agrees

with Maddieson's electromagnetic articulography graph in Figure 1. The LIPS are closed for the latter part of the pronunciation. Note that when the velum closes, that is when the lips close, thus providing a demarcation between the nasal *ŋ* and the KP: partial nasal place assimilation.

The difference between *ŋK* and *ŋmKP* is simply that the LIPS gesture is extended into the VELUM gesture, as in Figure 11. The difference between Figure 10 and Figure 11 is more gradient than categorical, and thus may exhibit variability due to presently unknown factors. If so, this may help explain why Efik speakers are observed to pronounce both *ŋKP* and *ŋmKP*, as previously noted in §2.1 by Cook (1969) and by Ohala & Ohala (1993).

Similar representations can be generated for *KPm* and *KPŋm*, being basically mirror images of the above.

Interestingly, AP representation may give a clue as to what is problematic about the unattested [mKP]. If [mKP] represents the phonetics, then first a labial nasal is articulated, then the labiality ceases in favor of a velar articulation, but then returns again later in the articulation. It is represented as follows.

In Figure 12, we see that the labial gesture would be interrupted briefly in the articulation which is first [m], then becomes non-labial for the velar part of [KP], then immediately returns to labial for the bulk of [KP]. This rapid on/off/on setting of *bilabial* is a complex gesture and less likely to occur than a simpler one.

Note that if the gap between the two bilabial gestures was erased, producing one bilabial gesture, then we would have [**m**KP], but with the velar gesture K completely hidden.

I will not attempt a gestural score for all of the phenomena in §2 and §3. Some should be straightforward in a gestural account, as with the Nafaanra labial-velar nasal's oral release [ŋm^{b}a] in (13).

However, it is not at all clear from previous AP literature how interaction of the consonant KP and vowels is to be handled. Vocalic epenthesis in the Kɔnni /kɔb-kpɪŋ/ → kɔb-ì-kpɪ́ˈíŋ in (8) seems like it should be amenable to a gestural account, but probably additional machinery would be required within the Browman & Goldstein (1989, 1990, 1992) approach. More challenging yet is the centralization of a high vowel when between a labial and a velar in Ejagham in (14). It is possible that a version of AP could describe it, but it is not obvious how. A more in-depth exploration will have to wait for another occasion.

5 Q-theory

Q-theory, as expounded by Inkelas & Shih (2017) and Shih & Inkelas (2019), offers a quantized extension of Articulatory Phonology, with each segment Q com-
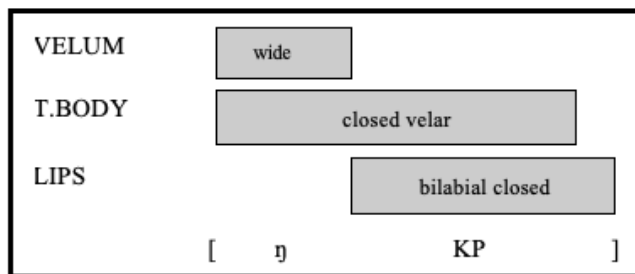
Figure 10: Gestural score for *ŋKP*
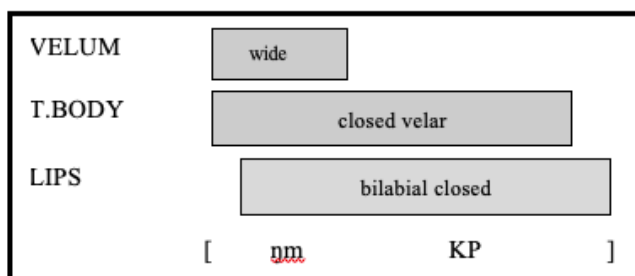
Figure 11: Gestural score for *ŋmKP*
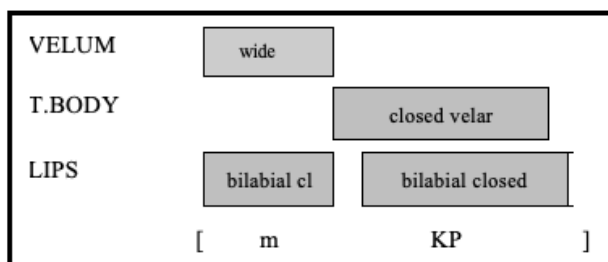
Figure 12: Gestural score for unattested *mKP*

posed of subsegments q_1, q_2, q_3. These correspond to the onset transition, main target, and release of a segment. Crucially for this discussion, each subsegment is *featurally uniform*, with no further internal structure or divisions. This representation has generally been paired with the Agreement by Correspondence framework of Rose & Walker (2004) *inter alia*, but here we focus only on the representational aspects of Q-theory.

It might be thought that q_1 and q_3, being at the edges of a segment, offer a promising approach to the left- and right-sided labial-velar phonology noted in this paper. The left-hand phonology would be a q to q_1 correspondence, and the right-hand phonology would be a q_3 to q correspondence.

However, Inkelas & Shih (2017: 1) specifically note:

> "Contour segments possess *distinct phases sequenced in time*; this crucial sequencing differentiates them from doubly articulated segments, such as labiovelars, in which distinct gestures are (nearly) simultaneous" (my added emphasis).

Q-theory excels at addressing such true contour segments, such as affricates and prenasalized stops, which have clear boundaries between their sub-parts. In Inkelas & Shih (2017), they give cases of the Kiyaka prenasalized unit stop [ŋg] being represented by the q subunits (ŋ g g), and the tone on the Changzhi vowel vowel ə in [təʔ$_{213}$] is represented by the q subunits (ə$_2$ ə$_1$ ə$_3$). Each of these q subunits has features which do not overlap with an adjacent q. Labial-velars differ from these by having a majority of their articulatory target consist of *two* distinct major features overlapping: [labial] and [dorsal]. This would be represented in Q-theory by KP (k kp p), with the subsegment q_2 being the two-unit "kp". This two-unit overlap is not allowed in Q-theory, though it was in Articulatory Phonology.

Thus the problem for Q-theory applying here is not at the edges, but in the putative central q_2 subsegment of a labial-velar. This reinforces the notion that labial-velars cannot be treated as contour segments, but that their nature as *complex* segments requires a different approach.

6 Conclusions and further research

The examples of left-sided and right-sided phonology across languages are not numerous, but do illustrate a clear pattern. Differential nasal assimilation is the most often reported phenomenon, possibly because it is the most easily observable one. Phenomena such as the Kɔnni and Ejagham vowel patterns require

detailed investigation, which is not available for many of the languages listed in the Appendices. But the cross-linguistic pattern is that labial-velars do reveal phonological processes sensitive to a velar configuration on the left, and a labial one on the right, but never the reverse.

At this point, we see that no current phonological model is able to capture the totality of labial-velar phonology. Not all models have even attempted to include labial-velars, and those which have done, have generally only referred to the bare existence of KP, not to the phonological patterns noted in this and other works. Labial-velars are complex segments, neither simple nor contours, and details of their phonology have been largely unexplored. It remains to be seen if, in fact, it is possible to incorporate the totality of the phonetic and phonological facts of labial-velars into a single model.

Acknowledgements

Appendix A Assimilation as [ŋKP]

	Language [ISO]	Sample	Gloss	Reference
1.	Adioukrou [adj]	tóŋkpó	'daba'	Kaul (2006)
2.	Aizi, Tiagbamrin [ahi]	atɪŋgbra	'bouteille'	Herault (1971)
3.	Anufo (Chakosi) [cko]	ŋgbɛ̃	'empty'	Stanford & Stanford (1970)
4.	Birifor [biv]	kpaŋkpan	'upper arm'	Kuch (1993)
5.	Bongo [bot]	ᵑgbáyá	'corn'	Kilpatrick (1985)
6.	Chumburung [ncu]	ŋ̀kpìnɔ̀	'chests'	Price (1975)
7.	Dagaari [dga]	gbáŋbáŋ	'noon'	Kennedy (1966)
8.	Dɛg [mzw]	dàŋgbàlá	'walking stick'	Crouch & Herbert (2003)
9.	Gã [gaa]	taaŋkpee	'sisal'	Ryder (1987)
10.	Gangam [gng]	ūsɛ̀ŋ̀gbéŋ́l	'dog'	Reimer (2022)
11.	Gonja [gjn]	gbìŋgbìŋ	'big'	Painter (1970)
12.	Gwari [gbr]	tʃìŋkpè	'stool'	Rosendall (1992)
13.	Gbaya-Mbodomo [gmm]	líŋkpòŋ	'vine for swinging'	Boyd (1997)
14.	Konkomba [xon]	ŋ̄gbéèm	'full'	Steele & Weed (1966)
15.	Kɔnni [kma]	tìŋgbáŋ	'floor'	Cahill (2007a)
16.	Kusaal [kus]	nɪŋgbɔŋ	'skin'	Spratt & Spratt (1968)
17.	Mbembe, Cross River [mfn]	kpenaŋkpen	'every, each'	Barnwell (1969)
18.	Mono [mnh]	kéŋgbā	'alone'	Olson (2005)
19.	Ncam (Bassar)[bud]	ŋ́-gbàñ	'skin'	Cox (1998)
20.	Vagla [vag]	tʃaŋkpalŋa	'antelope'	Crouch & Smiles (1966)

Appendix B /ŋKP/ as independent phoneme

	Language [ISO]	Reference
1.	Ambele [ael]	Nganganu Kenfac (2001)
2.	Avokaya [avu]	Callinan (1981)
3.	Bagiro (Furu) [fuu]	Boyeldieu (2006)
4.	Baka [bkc]	Léonard (2009)
5.	Balanta-Ganja (Fjaa, Fca) [bit]	N'Diaye-Corréard (1970)
6.	Bali [bcp]	Grégoire (2003)
7.	Banda, Mid-Southern (Yakpa) [bjo]	Cloarec-Heiss (1978)
8.	Banda, South Central (Ngbugu) [lnl]	Cloarec-Heiss (1978)
9.	Banda, West Central [bbp]	Cloarec-Heiss (1978)
10.	Banda-Bambari (Linda) [liy]	Cloarec-Heiss (1978)
11.	Banda-Yangere [yaj]	Moñino (1988)
12.	Bangando [bgf]	Baron (1995)
13.	Bangba [bbe]	Boone (1995)
14.	Bekwel (Bekwil) [bkw]	Phillips (2009)
15.	Birri [bvq]	Santandrea (1966)
16.	Bofi [bff]	Moñino (1995)
17.	Bulu [bum]	Yanes & Moise (1987)
18.	Digo (Chidigo) [dig]	Nicolle (2013)
19.	Ding [diz]	Muluwa & Bostoen (2015)
20.	Dongo ('Dongo-ko) [doo]	Moñino (1988)
21.	Esimbi [ags]	Stallcup (1980)
22.	Fang [fan]	Medjo Mvé (1997)
23.	Gbanu [gbv]	Moñino (1995)
24.	Gbanziri (Gbanzili) [gbg]	Bostoen & Donzo (2013)
25.	Gbaya (Kresh, Kreish) [krs]	Boyeldieu (2006)
26.	Gbaya Southwest [gso]	Moñino (1995)
27.	Gbaya-Mbodomo [gmm]	Boyd (1997)
28.	Gobu (Gubu, Gabu) [gox]	Cloarec-Heiss (1978)
29.	Gola [gol]	Koroma (1994)
30.	Indri [idr]	Santandrea (1969)
31.	Jula, Odienne (Wojenaka) [jod]	Derive (1983)
32.	Kako [kkj]	Ernst (1996)
33.	Kare (Kali) [kbn]	Elders (2006)
34.	Kol (Bikele) [biw]	Henson (2007)

	Language [ISO]	Reference
35.	Kpagua [kuw]	Cloarec-Heiss (1978)
36.	Kpatiri (Kpatili, Gbayi) [kym]	Boyd (1988)
37.	Kuo [xuo]	Elders (2006)
38.	Kyoli (Chori, Cori) [cry]	Dihoff (1976)
39.	Lele [lln]	Frajzyngier (2001)
40.	Lendu [led]	Boyeldieu (2006)
41.	Logo [log]	Goyvaerts (1983)
42.	Lutos [ndy]	Olson (2013)
43.	Mangbetu [mdj]	Larochette (1958)
44.	Mayogo [mdm]	McCord (1989)
45.	Mbandja (Mbanza) [zmz]	Cloarec-Heiss (1978)
46.	Mbum [mdd]	Elders (2006)
47.	Mono [mnh]	Olson (2005)
48.	Mono [mru]	Elders (2006)
49.	Mündü [muh]	Jeffrey & Polley (1981)
50.	Ndai (Galke, Pormi) [gke]	Elders (2006)
51.	Ngbaka Ma'bo [nbm]	Thomas (1963)
52.	Ngbaka Manza [ngg]	Selezilo (2006)
53.	Ngbandi, Northern [ngb]	Bostoen & Donzo (2013)
54.	Ngbandi, Southern [nbw]	Bostoen & Donzo (2013)
55.	Ngombe [ngc]	Grégoire (2003)
56.	Ngundu [nue]	Cloarec-Heiss (1978)
57.	Nzakara [nzk]	Santandrea (1965)
58.	Pagibete [pae]	Reeder (1998)
59.	Pande [bkj]	Murrell (2022)
60.	Sango [sag]	Samarin (1967)
61.	Sere [swf]	Moñino (1988)
62.	Togoyo [tgy]	Santandrea (1969)
63.	Wumboko (Mboko) [bqm]	Mutaka & Ebobissé (1996/7)
64.	Yakoma [yky]	Moñino (1988)
65.	Yango [yng]	Bostoen & Donzo (2013)
66.	Zande [zne]	Bostoen & Donzo (2013)

Appendix C Assimilation as [ŋ͡mKP]

	Language [ISO]	Sample	Gloss	Reference
1.	Agni [any]	ŋ̀mgbàfwɛ̂	'jeune homme'	Ouattara (2006)
2.	Dan [daf]	ŋm gbe	'my arm'	Bearth & Zemp (1967)
3.	Efutop [ofu]	ŋm̀-kpìb	'ant, tailor'	Crabb (1965)
4.	Ejagham [etu]	ŋ̀m-gbɛ̀	'leopard'	Watters (1981)
5.	Eton [eto]	ŋ́mkpôŋ	'pumpkin leaves'	Van de Velde (2008)
6.	Gɪdɪrɛ (Adele) [ade]	ŋm̀-kpá	'life'	Kleiner (1989)
7.	Kpelle [xpe]	ŋm-gbiŋ	'myself'	Welmers (1962)
8.	Nkonya [nko]	ŋ̀m-kpàà	'paths'	Peacock (2011)
9.	Samue [wbf]	kpɛŋmgba	'gall bladder'	Ouattara (2015)
10.	Yoruba [yor]	o ŋm gbo	'he is hearing'	Bamgbose (1969)

Appendix D /ŋ͡mg͡b/ as independent phoneme

(Bangolan alone also has /ŋ͡mk͡p/)

	Language [ISO]	Reference
1.	Baka [bdh]	Persson (2004)
2.	Banda-Ndele (Banda-Tangbago) [bfl]	Moñino (1988)
3.	Bangolan [bgj]	Mbah (2003)
4.	Belanda Bor [bxb]	Gilley (2004)
5.	Belanda Viri (Viri) [bvi]	Bilal (2004)
6.	Beli (Jur Beli) [blm]	Stirtz (2014)
7.	Bhogoto [bdt]	Boyd (2015)
8.	Bila [bip]	Kutsch Lojenga (2003)
9.	Bongo [bot]	Persson (2004)
10.	Buwal [bhs]	Viljoen (2009)
11.	Bwa (Benge, Bua-Yewu) [bww]	De Wit (2020)
12.	Cuvok [cuv]	Dadak (2021)
13.	Daba [dbq]	Lienhard & Giger (2009)
14.	Eloyi [afo]	Armstrong (1969)
15.	Etulo [utr]	Armstrong (1969)
16.	Gavar [gou]	Viljoen (2009)

	Language [ISO]	Reference
17.	Gbaya-Bossangoa (Gbeya) [gbp]	Samarin (1966)
18.	Iceve-Maci (Icheve) [bec]	Cox (2013)
18.	Ipulo [ass]	Tuinstra (2015)
20.	Jur Modo (Mödö) [bex]	Persson (2004)
21.	Kakwa [keo]	Onziga & Gilley (2012)
22.	Keliko [kbo]	Kilpatrick (2004)
23.	Kisi, Southern [kss]	Childs (1992)
24.	Komo [kmw]	Thomas (1982)
25.	Kuwaa [blh]	Marchese (1984)
26.	Kwakum (Bakoum) [kwu]	Hare (2018)
27.	Langbashe (Langbasi) [lna]	Cloarec-Heiss (1978)
28.	Lika (Liko) [lik]	De Wit (2008)
29.	Lulubo (Olu'bo) [lul]	Kilpatrick (2004)
30.	Ma'di [mhi]	Kilpatrick (2004)
31.	Mambila [mcu]	Connell (1998-1999)
32.	Mbo [zmw]	Rasmussen (2015)
33.	Mbudum [xmd]	Dadak (2014)
34.	Morokodo [mgc]	Persson (2004)
35.	Moru [mgd]	Kilpatrick (2004)
36.	Ndogo [ndz]	Bilal (2004)
37.	Ngbaka [nga]	Maes (1959)
38.	Ngiti [niy]	Boyeldieu (2006)
39.	Nzakambay (Touboro) [nzy]	Mbanji (1996)
40.	Omi (Kaliko-Omi, Omiti) [omi]	Bradley (2004)
41.	Suga (Nizaa) [sgi]	Kjelsvik (2002)
42.	Suma [sqm]	Bradshaw (1995)
43.	Tiv [tiv]	Kropp Dakubu (1980)
44.	Tuki [bag]	Clements & Rialland (2008)
45.	Twendi (Cambap) [twn]	Connell (2002)
46.	Yaka (Aka) [axk]	Duke (2001)

Appendix E [KPm]

	Language [ISO]	Sample	Gloss	Reference
1.	Gwari [gbr]	kpmàmí	'okra'	Rosendall (1992)
2.	Tyebaara Senufo [sef]	kpmɔ́:	'to beat'	Mills (1984)

Appendix F [KPŋ͡m]

	Language [ISO]	Sample	Gloss	Reference
1.	Konabere [bbo]	g͡bŋ͡m	'black'	P. Davison (pc)
2.	Mada [mda]	kpŋm̀	'kapok tree'	Price (1989)
3.	Kuta Gwari [gbr]	kp^{ŋm}	no specific words given	Hyman & Magaji (1970)

References

Anderson, Stephen R. 1976. On the description of multiply-articulated consonants. *Journal of Phonetics* 4. 17–27.

Anderson, Stephen R. 1981. Why phonology isn't "natural". *Linguistic Inquiry* 12. 493–539.

Ansre, Gilbert. 1963. Reduplication in Ewe. *Journal of African Languages* 2. 128–132.

Armstrong, Robert. G. 1969. Iboid. In John Bendor-Samuel (ed.), *The Niger-Congo languages*, 323–336. Lanham, MD: University Press of America/SIL.

Bamgbose, Ayo. 1969. Yoruba. In Elizabeth Dunstan (ed.), *Twelve Nigerian languages*, 163–172. New York: Africana Publishing Corporation.

Barnwell, Katharine G. L. 1969. *A grammatical description of Mbembe (Adun dialect), a Cross River language*. University of London. (Doctoral dissertation).

Baron, Gerard. 1995. *Dictionnaire ʼbangando-francais*. Ms.

Bearth, Thomas & Hugo Zemp. 1967. The phonology of Dan (Santa). *Journal of African Languages* 6. 9–29.

Begne, Leopold Prosper, II. 1980. *The Phonology of Bikele, A Cameroonian language*. Illinois Institute of Technology. (Doctoral dissertation).

Bilal, Clement Murba Wau. 2004. The Adamawa-Ubangi languages in Sudan. *Occasional Papers in the Study of Sudanese Languages* 9. 175–184.

Boone, Douglas. 1995. Bangba survey report. In Gert de Wit (ed.), *Compendium of survey reports : Non-Bantu languages*, vol. 2. Language Survey Department Summer Institute of Linguistics Eastern Zaire Group.

Bostoen, Koen & Jean-Pierre Donzo. 2013. Bantu-Ubangi language contact and the origin of labial-velar stops in Lingombe (Bantu, C41, DRC). *Diachronica* 30(4). 435–468.

Boyd, Raymond. 1988. Le kpatiri ou gbāyī, une nouvelle lange du groupe ngbandi. In Yves Moñino (ed.), *Lexique comparatif des langues oubanguiennes*, 35–49. Paris: Geuthner.

Boyd, Virginia L. 1997. *A phonology and grammar of Mbódɔ̀mɔ̀*. University of Texas at Arlington. (MA thesis).

Boyd, Virginia L. 2015. *Précis d'orthographe Bhogoto*. Yaoundé: SIL Cameroon.

Boyeldieu, Pascal. 2006. Reflexes of a labiovelar series in Central Sudanic. In Al-Amin Abu-Manga, Leona Gilley & Anne Storch (eds.), *Insights into Nilo-Saharan language, history and culture. Proceedings of the 9th Nilo-Saharan linguistics colloquium, Institute of African and Asian Studies, University of Khartoum, 16-19 February 2004* (Nilo-Saharan 23), 129–151. Cologne: Rüdiger Köppe Verlag.

Bradley, David. 2004. *Draft omiti phonology statement*. Unpublished ms.

Bradshaw, Mary. 1995. Tone on verbs in Suma. In Akinbiyi Akinlabi (ed.), *Theoretical approaches to African linguistics*, 255–271. Trenton, NJ: Africa World Press.

Browman, Catherine P. & Louis Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3. 219–252.

Browman, Catherine P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6. 201–251.

Browman, Catherine P. & Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In John Kingston & Mary E. Beckman (eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech*, 341–376. Cambridge: Cambridge University Press.

Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49. 155–180.

Cahill, Michael. 1998. Nasal assimilation and labiovelar geometry. In Ian Maddieson & Thomas J. Hinnebusch (eds.), *Language history and linguistic descrip-*

tion in Africa. Trends in African linguistics 2. (Papers presented at the 26th ACAL, UCLA, 1995), 127–136. Trenton, NJ: Africa World Press.

Cahill, Michael. 2006. *Labial-velars are labial*. Paper presented at the Linguistic Society of America Annual meeting, Albuquerque, NM, Jan. 5-8, 2006.

Cahill, Michael. 2007a. *Aspects of the morphology and phonology of Kɔnni*. Dallas, TX: SIL International.

Cahill, Michael. 2007b. *The phonetics and phonology of Dagbani labial-velars*. Paper presented at the Linguistic Society of America Annual meeting, Anaheim, CA, Jan. 4-7, 2007.

Callinan, Lynne. 1981. A preliminary study of Avokaya phonemes. *Occasional Papers in the Study of Sudanese Languages* 1. 64–73.

Childs, George Tucker. 1992. Language typology and reconstruction: The prenasalized stops of Kisi. *Studies in African Linguistics* 23. 65–80.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Clements, George N. & Elizabeth V. Hume. 1995. The internal organization of speech sounds. In John Goldsmith (ed.), *A handbook of phonological theory*, 245–306. Oxford: Blackwell.

Clements, George N. & Samuel J. Keyser. 1983. *CV phonology - A generative approach to the syllable*. Cambridge, MA: MIT Press.

Clements, George N. & Annie Rialland. 2008. Africa as a phonological area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 36–85. Cambridge: Cambridge University Press.

Cloarec-Heiss, France. 1978. *Étude préliminaire à une dialectologie banda*. Raymond Boyd & France Cloarec-Heiss (eds.) (Bibliothèque de la SELAF 65). Paris: SELAF. 11–32.

Connell, Bruce. 1994. The structure of labial-velar stops. *Journal of Phonetics* 22. 441–476.

Connell, Bruce. 2002. Aspects of the phonetics of Cambap. *Studies in African Linguistics* 31. 179–210.

Connell, Bruce. 1998-1999. Feature geometry and the formation of labial-velars: A reply to Mutaka and Ebobissé. *Journal of West African Linguistics* XXVI. 17–32.

Cook, Thomas. 1969. Efik. In Elizabeth Dunstan (ed.), *Twelve Nigerian languages*, 35–46. New York: Africana Publishing Corporation.

Cox, Bruce. 2013. *A phonological overview of Iceve-Maci*. ms, SIL Cameroon. Available at. http://www.sil.org/africa/cameroun/index.html.

Cox, Monica. 1998. *Description grammaticale du ncam (bassar)*. Ecole Pratique des Hautes Etudes, CNRS. (Doctoral dissertation).

Crabb, David. W. 1965. *Ekoid Bantu languages of Ogoja, part 1.* West Africa Language Monographs 4. Cambridge University Press.

Crouch, Marjorie & Patricia Herbert. 2003. *The phonology of Dɛg.* Collected Field Notes Series No. 24. The Institute of African Studies, University of Ghana, Legon.

Crouch, Marjorie & Nancy Smiles. 1966. *The phonology of Vagla.* Collected Field Notes Series No. 4. The Institute of African Studies, University of Ghana, Legon.

Dadak, Ndokobai. 2014. *Esquisse phonologique de la langue mbudum.* Unpublished ms. Yaoundé: CABTAL.

Dadak, Ndokobai. 2021. *Grammaire cuvok: Langue tchadique centrale du cameroun.* Leiden University. (Doctoral dissertation).

De Wit, Gerrit. 2008. *Lika phonologie.* SIL Electronic Working Papers 2010-002. https://www.sil.org/resources/publications/entry/7818.

De Wit, Gerrit. 2020. *Boa-Yewu, a Bantu language with a seven-vowel system and ATR vowel harmony.* Ms, SIL International.

Demolin, Didier. 1991. Les consonnes labio-vélaires du mangbetu. *Pholia* 6. 85–105.

Derive, Marie-José. 1983. *Etude comparée des parlers manding ivoiriens.* Abidjan, Côte d'Ivoire: Université d'Abidjan, Institut de linguistique appliquée.

Dihoff, Ivan R. 1976. *Aspects of the tonal structure of Chori.* University of Wisconsin-Madison. (Doctoral dissertation).

Duke, Daniel J. 2001. *Aka as a contact language: Sociolinguistic and grammatical evidence.* University of Texas at Arlington. (MA thesis).

Elders, Stefan. 2006. Issues in comparative Kebi-Benue. *Africana Linguistica* 12. 37–88. DOI: 10.3406/aflin.2006.962.

Ernst, Urs. 1996. *Alphabet et orthographe du kakɔ (kakɔ-est).* Yaoundé: SIL Cameroon.

Folarin, Antonia Yetunde. 1987. *Lexical phonology of Yoruba nouns and verbs.* University of Kansas (doctoral dissertation).

Frajzyngier, Zygmunt. 2001. *A grammar of Lele* (Stanford Monographs in African Languages). Stanford, CA: CSLI Publications.

Gilley, Leoma. 2004. The Lwoian family. *Occasional Papers in the Study of Sudanese Languages* 9. 165–174.

Gouskova, Maria & Juliet Stanton. 2021. Learning complex segments. *Language* 97.1. 151–193.

Goyvaerts, Didier L. 1983. Some aspects of Logo phonology and morphology. In M. Lionel Bender (ed.), *Nilo-Saharan language studies*, 272–280. Lansing, MI: African Studies Center, Michigan State University.

Grégoire, Claire. 2003. The Bantu languages of the forest. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 349–370. New York: Routledge.

Hare, Stacey. 2018. *Tone in Kwakum (A91) with an application to orthography.* Graduate Institute of Applied Linguistics, Dallas. (MA thesis).

Henderson, James. 1995. *Phonology and grammar of Yele, Papua New Guinea.*

Henson, Bonnie. 2007. *The phonology and Morphosyntax of Kol.* University of California, Berkeley. (Doctoral dissertation).

Herault, Georges. 1971. *L'aïzi: Esquisse phonologique et enquête lexicale* (Documents Linguistiques, XX). Abidjan: Institut de Linguistique Appliquee, Unversité d'Abidjan.

Hyman, Larry M. & Daniel J. Magaji. 1970. *Essentials of Gwari grammar.* Occasional Publication No. 27, Institute of African Studies, University of Ibadan.

Inkelas, Sharon & Stephanie Shih. 2017. Looking into segments. In Karen Jesney, Charlie O'Hara, Caitlin Smith & Rachel Walker (eds.), *Proceedings of the 2016 Annual meeting on phonology.* Linguistic Society of America. DOI: 10.3765/amp.v4i0.3996.

Jacobsen, Roman, Gunnar Fant & Morris Halle. 1953. *Preliminaries to speech analysis.* Cambridge: MIT Press.

Jeffrey, Dorothea & Linda Polley. 1981. Phonology and morphophonemics in Mündü. *Occasional Papers in the Study of Sudanese Languages* 1. 1–42.

Jordan, Dean. 1980. *The phonology of Nafaara.* Collected Field Notes Series No. 17. The Institute of African Studies, University of Ghana, Legon.

Kashima, Eric. 2021. The phonetics of Nmbo (Nɐmbo) with some comments on its phonology (Yam family; Morehead district). In Kate L. Lindsey & Dineke Schokkin (eds.), *Phonetic fieldwork in Southern New Guinea* (Language Documentation & Conservation Special Publication 24), 53–75. Mānoa: University of Hawaiʻi.

Kaul, Guy. 2006. Morphophonologie l'adioukrou. In Will Leben & Firmin Ahoua (eds.), *Morphohonologie des langues kwa de Côte d'Ivoire*, 77–102. Cologne: Rüdiger Köppe Verlag.

Kennedy, Jack. 1966. *The phonology of Dagaari* (Collected Field Notes Series No. 6). Legon: The Institute of African Studies, University of Ghana.

Khachaturyan, Maria. 2015. Grammaire du mano. *Mandenkan* 54.

Khachaturyan, Maria. 2018. A sketch of dialectal variation in Mano. *Mandenkan* 59. 31–56.

Kilpatrick, Eileen. 1985. Bongo phonology. *Occasional Papers in the Study of Sudanese Languages* 4. 1–62.

Kilpatrick, Eileen. 2004. Orthographies of Moru-Ma'di languages. *Occasional Papers in the Study of Sudanese Languages* 9. 85–92.

Kjelsvik, Bjørghild. 2002. *Verb chains in Nizaa.* University of Oslo. (Doctoral dissertation).

Kleiner, Renate. 1989. *Phonology notes of Gɪdɪrɛ, the language of the Adele.* Unpublished ms., GILLBT, Tamale, Ghana.

Koroma, Regine. 1994. *Die Morphosyntax des Gola* (Afrikanistische Monographien 4). Köln: Institut für Afrikanistik, Universität zu Köln.

Kropp Dakubu, Mary E. 1980. *West African language data sheets*, vol. III. West African Linguistic Society.

Kuch, Lawrence. 1993. *The phonology of Birifor.* Collected Field Notes Series No. 21. The Institute of African Studies, University of Ghana, Legon.

Kutsch Lojenga, Constance. 2003. Bila (D32). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 450–474. London, New York: Routledge.

Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the world's languages.* Cambridge, MA: Blackwell Publishers.

Larochette, Joe. 1958. *Grammaires des dialectes Mangbetu et medje suivi d'un manuel de conversation et d'un lexique* (Annales du Musée Royal du Congo Belge, Sciences de l'Homme Linguistique Volume 18). Tervuren, Belgique: Musée royal du Congo belge.

Léonard, Yves. 2009. *Orthographe baka.* Yaounde: SIL Cameroon.

Lienhard, Ruth & Martha Giger. 2009. *Document de reference pour la standardisation de la langue daba.* Yaoundé: SIL Cameroon.

Lombardi, Linda. 1990. The nonlinear organization of the affricate. *Natural Language & Linguistic Theory* 8.3. 375–425.

Maddieson, Ian. 1993. Investigating Ewe articulations with electromagnetic articulography. *Forshungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München* 31. 181–214.

Maes, Védaste. 1959. *Dictionnaire Ngbaka-francais-neerlandais.* Tervuren: Commissie voor Afrikaanse Taalkunde / la Commission de Linguistique Africaine.

Marchese, Lynell. 1984. *Atlas linguistique kru.* Third edition. Abidjan: Université d'Abidjan, Institute de Linguistique Appliquee.

Mbah, Njeck Mathaus. 2003. *Phonological sketch of Bangolan.* Yaoundé: NACALCO. https://www.sil.org/resources/archives/47360.

Mbanji, Bawe Ernest. 1996. *Alfabe nzakambay (alphabet chart of Nzakambay).* SIL Chad.

McCord, Michael S. 1989. *Acoustical and autosegmental analyis of the Mayogo vowel system.* University of Texas at Arlington. (MA thesis).

Medjo Mvé, Pither. 1997. *Essai sur la phonologie panchronique des parlers Fang du Gabon et ses implications historiques.* Université Lumière-Lyon 2. (Doctoral dissertation).

Mills, Elizabeth. 1984. *Senoufo phonology, discourse to syllable*. Dallas, TX: Summer Institute of Linguistics.

Moñino, Yves. 1988. *Lexique comparatif des langues oubanguiennes*. Paris: Geuthner.

Moñino, Yves. 1995. *Le proto-gbaya: Essai de linguistique comparative historique sur vingt-et-une langues d'Afrique centrale*. (SELAF, 357.) Paris: Peeters.

Muluwa, Joseph Koni & Koen Bostoen. 2015. *Lexique comparé des langues bantu du Kwilu (République démocratique du Congo)* (Grammatical Analyses of African Languages Volume 48). Cologne: Rüdiger Köppe Verlag.

Murrell, Paul. 2022. *A phonological overview of Bhogongo*. Ms., Association Centrafricaine pour la Traduction de la Bible et l'Alphabétisation (ACATBA). Bangui, Central African Republic.

Mutaka, Ngessimo M. & Carl Ebobissé. 1996/7. The formation of labial-velars in Sawabantu: Evidence for feature geometry. *Journal of West African Languages* XXVI. 3–14.

N'Diaye-Corréard, Geneviéve. 1970. Le systéme des classes du Fca, ou Balante (dialecte Ganja). *Journal of African Languages* 8-2. 102–119.

Nganganu Kenfac, Lucy. 2001. *The phonology of Ambele*. Université de Yaoundé. (Doctoral dissertation).

Nicolle, Steve. 2013. *A grammar of Digo*. Dallas, TX: SIL International.

Ohala, John J. 1993. The relation between phonetics and phonology. In William J. Hardcastle & John Laver (eds.), *The handbook of phonetic sciences*, 674–694. Malden, MA: Blackwell Publishers.

Ohala, John J. & Manjari Ohala. 1993. The phonetics of nasal phonology: Theorems and data. In Marie K. Huffman & Rena A. Krakow (eds.), *Nasals, nasalization, and the velum*, 225–249. San Diego, CA: Academic Press.

Olson, Kenneth. 2005. *The phonology of Mono* (SIL and UTA Publications in Linguistics 140). Dallas, TX: SIL International & the University of Texas at Arlington.

Olson, Kenneth. 2013. *A sketch of Lutos phonology*. Unpublished ms. http://www.sil.org/resources/archives/60749.

Onziga, Yuga Juma & Leoma Gilley. 2012. Phonology of Kakuwâ (Kakwa). *Occasional Papers in the Study of Sudanese Languages* 10. 1–15.

Ouattara, Irma. 2006. Morphophonologie de l'agni-ano. In Will Leben & Firmin Ahoua (eds.), *Morphohonologie des langues Kwa de Côte d'Ivoire*, 77–102. Cologne: Rüdiger Köppe Verlag.

Ouattara, Virpi. 2015. *A phonological and tonal analysis of Samue using Optimality Theory*. University of Turku. (Doctoral dissertation).

Padgett, Jaye. 1995. *Partial class behavior and Nasal place assimilation.* Proceedings of the Southwestern Optimality Theory Workshop. Coyote Working Papers in Linguistics. University of Arizona, Tucson, pp. 145-183.

Padgett, Jaye. 2002. Feature classes in phonology. *Language* 78.1. 81–110.

Painter, Colin. 1970. *Gonja: A phonological and grammatical study.* Bloomington: Indiana University.

Peacock, Wesley. 2011. *The phonology of Nkonya.* Collected Field Notes Series No. 27. The Institute of African Studies, University of Ghana, Legon.

Persson, Janet. 2004. Notes on the phonology of Jur Mödö. *Occasional Papers in the Study of Sudanese Languages* 1. Juba: Summer Institute of Linguistics, Institute of Regional Languages, and the University of Juba, 43–54.

Phillips, John. 2009. *Guide d'orthographe bɛkwyel.* ms. SIL-Congo.

Price, Norman. 1975. *The phonology of Nchimburu.* Collected Field Notes Series No. 14. The Institute of African Studies, University of Ghana, Legon.

Price, Norman. 1989. *Notes on Mada phonology* (Language Data, Africa Series, Publication 23). Dallas, TX: Summer Institute of Linguistics.

Rasmussen, Kent. 2015. *A first look at Mbo [zmw] (D.R. Congo).* Paper presented at the 46th Annual Conference on African Linguistics, University of Oregon (Eugene), March 26-28, 2015.

Reeder, JeDene. 1998. *Pagibete, a Northern Bantu borderlands language: A grammatical sketch.* University of Texas at Arlington. (MA thesis).

Reimer, Jeannette. 2022. *The segmental phonology and tonology of Dye (Gangam) and its verb classification system* (SIL eBooks 80). Dallas: SIL. https://www.sil.org/resources/publications/entry/9477.

Rose, Sharon & Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80.3. 475–531.

Rosendall, Heidi James. 1992. *A phonological study of the Gwari lects.* Dallas, TX: Summer Institute of Linguistics.

Ryder, Mary Ellen. 1987. An autosegmental treatment of Nasal assimilation to labial-velars. In Barbara Need, Eric Schiller & Anna Bosch (eds.), *Papers from the 23rd annual regional meeting of the Chicago linguistic society: Part two. Parasession on autosegmental and metrical phonology*, 253–265. Chicago: Chicago Linguistic Society.

Samarin, Willam J. 1967. *A grammar of Sango.* The Hague: Mouton.

Samarin, William J. 1966. *The Gbeya language.* Berkeley & Los Angeles, CA: University of California Press.

Santandrea, Stefano. 1965. *Languages of the Banda and Zande groups.* Naples: Istituto Universitario Orientale.

Santandrea, Stefano. 1966. The Birri language: Brief elementary notes. *Afrika und Übersee, Bund XLIX:* 81-105. 196–234.

Santandrea, Stefano. 1969. *Note grammaticali e lessicali sul gruppo Feroge e sul Mundu.* Naples: Istituto Universitario Orientale.

Selezilo, Apollinaire. 2006. *Description sytematique du ngbaka-manza de Bogangolo in Centrafrique.* University of Yaoundé I. (Doctoral dissertation).

Shih, Stephanie & Sharon Inkelas. 2019. Autosegmental aims in Surface-optimizing phonology. *Linguistic Inquiry* 50.1. 137–196.

Showalter, Stuart. 1997. Coup de glotte, nasalité, et schèmes syllabiques en kaansa. *Gur Papers/Cahiers Voltaiques No* 2. 137–146.

Spratt, David & Nancy Spratt. 1968. *The phonology of Kusal* (Collected Field Notes Series 10). Legon: The Institute of African Studies, University of Ghana.

Stallcup, Kenneth L. 1980. Noun classes in Esimbi. In Larry Hyman (ed.), *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8), 139–153. Los Angeles: University of Southern California.

Stanford, Ronald & Lyn Stanford. 1970. *The phonology and grammar of Chakosi.* Collected Field Notes Series No. 11. The Institute of African Studies, University of Ghana, Legon.

Steele, Mary & Gretchen Weed. 1966. *The phonology of Konkomba.* Collected Field Notes Series No. 3. The Institute of African Studies, University of Ghana, Legon.

Steriade, Donca. 1993. Closure, release, and nasal contours. In Marie K. Huffman & Rena A. Krakow (eds.), *Nasals, nasalization, and the velum*, 401–470. San Diego, CA: Academic Press.

Stirtz, Timothy. 2014. 'Bëlï phonology, tone and orthography. *SIL Electronic Working Papers* 2014-02. http://www.sil.org/resources/publications/entry/5601.

Thomas, Jacqueline M. C. 1963. *Le parler Ngbaka de Bokanga.* Paris: Mouton & Co.

Thomas, John Paul. 1982. *A morphophonology of Komo: Non-tonal phenomena.* University of North Dakota. (MA thesis).

Tuinstra, Anna. 2015. *Ipulo orthography guide.* Yaoundé: SIL Cameroon.

Ulfsbjorninn, Shanti. 2021. Labiovelars and the labial-velar hypothesis: Phonological headedness in Bare Element Geometry. *Glossa: A journal of general linguistics* 1. 111.

Van de Velde, Mark. 2008. *A Grammar of Eton.* New York: Mouton de Gruyter.

van de Weijer, Jeroen M. 1996. *Segmental structure and complex segments.* Berlin, Boston: Max Niemeyer Verlag.

Viljoen, Melanie. 2009. *Précis d'orthographe pour la langue gavar, edition préliminaire.* Yaoundé: SIL Cameroon.

Ward, Ida C. 1933. *The phonetic and tonal structure of Efik.* Cambridge: W. Heffer & Sons, Ltd.

Watters, John. 1981. *A phonology and morphology of Ejagham: With notes on dialect variation.* UCLA. (Doctoral dissertation).

Welmers, William E. 1962. The phonology of Kpelle. *Journal of African Languages* 1. 69–93.

Welmers, William E. 1968. *Jukun of Wukari and Jukun of Takum* (Occasional Publication No. 16). Ibadan: Institute of African Studies, University of Ibadan.

Welmers, William E. 1973. *African language structures.* Berkeley & Los Angeles, CA: University of California Press.

Wetzels, W. Leo & Andrew Nevins. 2018. Prenasalized and postoralized consonants: The diverse functions of enhancement. *Language* 94.4. 834–836.

Williamson, Kay. 1965. *A grammar of the Kolokuma dialect of Ijo* (West African Language Monographs 2). Cambridge: Cambridge University Press.

Yanes, Serge & Eyinga Essam Moise. 1987. *Dictionnaire boulou-francais francais-boulou avec grammaire.* Editions "P. Monti" Sangmelima, Saronno, Italy.

Yip, Moira. 1989. Contour tones. *Phonology* 6.1. 149–174.

Chapter 9

Post-verbal focus by *wəli* in Kenyan Maay

Chiara Di Maio
Syracuse University

Haley Muth
Syracuse University

This paper provides an overview of the construction involving the function word *wəli* in Kenyan Maay. To our knowledge, there is yet to be discussion of this construction in the literature on the language. Our goal is to provide an initial account of its use and function in order to better understand its role in the language. Its characteristics and distribution will also be compared to its apparent analog *wáxa(a)* in Somali. We argue that *wəli*, like *wáxa(a)*, is implicated in cataphoric focus, requiring displacement of a focused element to a position to the right of the language's Verb Complex.

1 Introduction

The topic of this paper is a specific construction used in Kenyan Maay that is introduced by the function word *wəli*. Our goal is to provide an initial account of its use and function. In order to better understand its role in the language, its characteristics and distribution will be compared to its apparent analog *wáxa(a)*[1] in Somali. In Somali, *wáxa(a)*, according to Svolacchia et al. (1995), is a cataphoric focus marker that appears before the Verb Complex and focuses an element following it. In other words, the focus marker *wáxa(a)* is a cataphor co-referent

---

[1]The length of the final vowel varies by dialect.

with a later expression which is in focus. The configuration of the *wəli* construction and its syntactic effects, as will be demonstrated below, are very similar to *wáxa(a)*, and we illustrate that the word *wəli*, like *wáxa(a)*, places a post-verbal constituent into narrow focus. However, from the data we have collected, we find some nuanced differences in the form and function of *wəli* as compared to *wáxa(a)*. In this paper, we will first give a concise overview of Kenyan Maay in §2. Following this in §3 we will demonstrate that the function of the *wəli* construction is to place elements in narrow focus. Then, in §4, we will explain the configuration of the *wəli* construction, the various elements which it can place in focus, and the contexts in which it tends to appear. In §5, we will compare *wəli*'s usage with another focus marker in Kenyan Maay with different properties. In §6, we will briefly describe sentences where two focus markers occur in Kenyan Maay. Finally, in §7, we will discuss avenues of future research by examining studied phenomena of Somali focus markers and drawing some comparisons with what we have established for the *wəli* construction in Kenyan Maay. We will also elaborate on the possibility of *wəli*'s involvement in broad focus and provide some concluding remarks.

2 Kenyan Maay

Kenyan Maay (ISO: ymm) is a Lowland East Cushitic language spoken in southern Somalia, parts of Kenya, and by speakers in the diaspora in the U.S., Europe, and elsewhere. Kenyan Maay and Somali are members of the Omo-Tana subgroup (see Figure 1).

Lowland East Cushitic
Southern — Saho 'Afar
Nuclear — Transversal
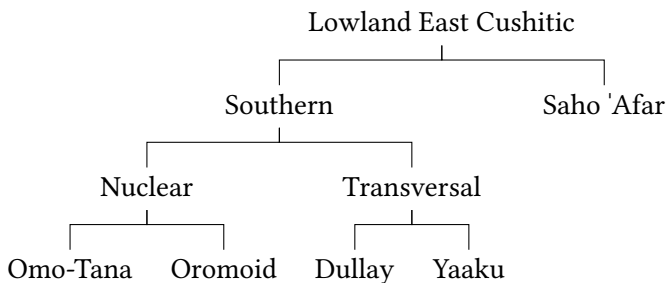Omo-Tana — Oromoid — Dullay — Yaaku

Figure 1: Lowland East Cushitic tree adapted from Tosco (2000: 108)

The data included in this paper were collected from two female speakers from Dadaab, Kenya (see Figure 2). Data from the first speaker, Habiba Noor, was collected primarily during a field methods course at Syracuse University, as well

as through additional elicitation sessions conducted after the completion of the course. The first speaker's parents are from Bu'ale, Somalia, and she grew up in Dadaab until age 15, after which she came to the US. The second speaker, Wilo Matan, also moved to the U.S. when she was 15, but her parents are from Baidoa, Somalia. Maay remains the primary language of the household for both speakers. We will refer to them as Speaker 1 and Speaker 2, respectively, where a distinction between the two is needed. We refer to their dialect as Kenyan Maay to distinguish it from other Maay varieties discussed in the small literature on the language (Biber 1982; Comfort & Paster 2009; Paster 2006, Paster 2010, Paster 2018; Saeed 1982a). Kenyan Maay is also the subject of a recent master's thesis by Smith (2022). Locations of dialects covered in these works are shown in Figure 2.
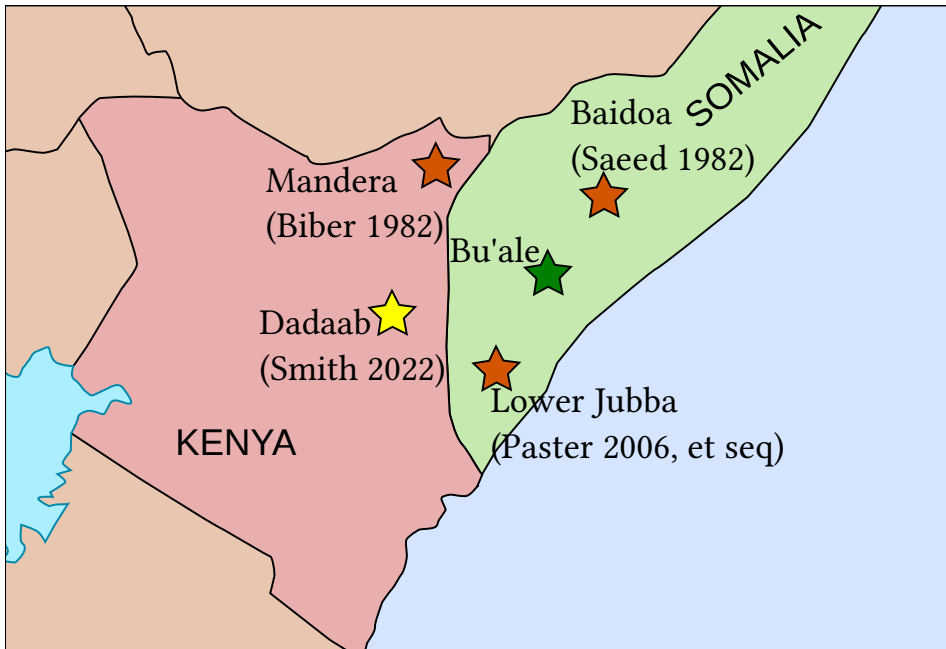
Figure 2: Locations of Maay studies

From what is reported in the published literature, "Central Somali", as discussed in Saeed (1982a), appears to be quite similar to Kenyan Maay. Central Somali, however, is understood to be a variety of Maay spoken around Baidoa (where Speaker 2's parents are from). Smith (2022) refers to this variety as Baydhabo (Baidoa) Maay. The two varieties share the same marker of pre-verbal focus, *jaa* (see Section 5), but Central Somali as described by Saeed employs a different marker of post-verbal focus, *wey ba*. Based on the data provided by Saeed

on *wey ba*, it appears to operate in much the same manner as *wəli* by focusing a post-verbal element. Speaker 1 recognizes *wey ba* but does not use it. Beyond providing a few examples, Saeed (1982a)'s sketch does not delve into much detail about focus in Central Somali.

Maay is closely related to Somali (ISO: som), but the languages are not mutually intelligible (Paster 2018). In §7.1, the post-verbal focus construction of Kenyan Maay will be compared to its Somali analog, with which it shares many similarities. Cushitic languages generally present a basic word order of SOV (Gebert 1986: 45), and this also appears to be the case in Kenyan Maay in pragmatically neutral sentences. Svolacchia & Puglielli (1999) note that Somali has fairly free word order with regard to its arguments, and as we shall see, Kenyan Maay shows evidence of this as well. Both languages make use of a constituent known as the Verb Complex (see Gebert 1986; Green 2021: 255; Puglielli 1981a; Saeed 1999: 163). According to Puglielli (1981a), the Verb Complex includes all the elements that express grammatical relations, namely verbs and adpositions (referred to by Puglielli as prepositions). Furthermore, it can be considered a micro-structure of the entire sentence since it also includes pronominal referents of NPs, which occur outside of the Verb Complex (Puglielli 1981a). In (1) and (2) below, examples in Somali and Kenyan Maay are provided with the Verb Complex in brackets.[2]

(1)   unə́ ʊntə́-ð-ə   [sən   sii-an-ə́ŋ]   isíŋ.[3]   Kenyan Maay
     1PL   food-F-DEF 2PL.OCL give-1PL-PRES 2PL

     'We give food to you.'

(2)   adí-gu   w[-áad   rab-t-aa].   Somali
     2SG-M.DEF.SUBJ[4] DECL-2SG want-2SG-PRES

     'You want it.' (Green 2021: 310)

Kenyan Maay makes use of an inflectional system to mark person, number, gender, and tense. Auxiliary verbs are also used to indicate aspect and mood. Negation is achieved by the addition of the negator *mə* before the predicate, as well as through inflection of the verb in certain instances.

---

[2]We will continue to bracket the verb complex in subsequent examples so that the location of nominal constituents relative to the verb complex is clear.

[3]As there is not yet a standard orthography for Maay, in this paper the IPA is used to provide a broad phonetic transcription for our examples. Morpheme boundaries are indicated by hyphens. It should be noted that an orthography was developed for Mukhtar & Ahmed's 2007 English-Maay dictionary, however that system was unacceptable to our speakers.

[4]Note that we use masculine (M) and feminine (F) agreement in our glosses, which generally equates to what is called K and T series agreement in Green (2021).

3 Focus

3.1 Storyboard

The *wəli* construction in Kenyan Maay consistently entails constituent movement to the right of the Verb Complex. In order to further examine the construction and its function, and on comparison to its analog *wáxa(a)* in Somali, we used a storyboard (from www.totemfieldstoryboards.org by Littell 2010) which was expressly designed to elicit narrow subject and object focus (narrow focus being the focusing of a single constituent, as opposed to broad focus, in which an entire utterance is in focus). The storyboard was used with both speakers, but individually. Speaker 1 used both markers of pre-verbal and post-verbal focus, while Speaker 2 used the marker of pre-verbal focus for the storyboard, but used the marker of post-verbal focus elsewhere. Therefore, we will mostly discuss Speaker 1's responses below. The story involves two animals (a squirrel and a crab) discussing the party they are currently attending. We acted as the squirrel, and would ask our speakers, who played the crab, various questions about who brought what item to the party. For example, we would ask the speakers "What did the snake bring?" and they would respond "The snake brought the drinks". The story was told once for object focus, asking what was brought by each animal, and then again for subject focus, asking who brought each item.

For the first elicited sentence, we provided a potential response in English to demonstrate the task, following which Speaker 1 provided the declarative sentence in (3). In subsequent items, for which we did not provide any prompt, Speaker 1 produced sentences using the *wəli* construction for almost every response – as in (4) – except for one response for which the marker of pre-verbal focus (described in §5) was used.[5]

(3) éj-k-ə   ʊntə́-ð-ə   [hagað-Ø-í].   Kenyan Maay
    dog-M-DEF food-F-DEF bring-3SG.M-PST
    'The dog brought the food.'

(4) éj-k-ə   wəli [hagað-Ø-í]   ʊntə́-ð-ii.   Kenyan Maay
    dog-M-DEF FOC bring-3SG.M-PST food-F-RDEF
    'The dog brought THE FOOD.'

The question which elicited (4) above was *éjkii máj hagaðí?* 'What did the dog bring?'. The element that is expected to be in narrow focus is the object *ʊntə́ðii*

---

[5]In (4) and the remainder of the paper, small caps is used in the translation to indicate which constituent is focused.

'the food'. When the *wəli* construction is used in (4), this element appears after the Verb Complex, whereas in (3) it occurs pre-verbally.

Sentence (5) was elicited in response to the question *áj ʊntɔ́ðii hagaðí?* 'Who brought the food?'. This sentence differs in that the subject is targeted for narrow focus. As expected, the element targeted for focus appears post-verbally. Therefore, whether intended focus is on the subject or object, the element focused by *wəli* appears in post-verbal position.

(5) ʊntɔ́-ð-ii   wəli [hagað-Ø-í]   éj-k-ə.   Kenyan Maay
     food-F-RDEF FOC bring-3SG.M-PST dog-M-DEF
     'THE DOG brought the food.'

We also tested subject focus with plural subjects and different persons, in order to determine whether the use of the *wəli* construction affected subject-verb agreement, since in Somali, verbal agreement can be affected when a subject is in focus (Banti 2011; Green 2021: 308-311; Puglielli 1981b: 13; Saeed 1999: 192; Svolacchia et al. 1995). In (6), verbal agreement in number was maintained despite the post-verbal movement of the subject.

(6) hés-t-ii   wəli [ʃeen-Ø-é-ŋ]   məlai-jaál-k-ə.   Kenyan Maay
     music-F-RDEF FOC bring-3-PST-PL fish-PL-M-DEF
     'THE FISH brought the music.'

However, when testing for different persons, there was some variation between our two speakers. Speaker 1 generally maintains person-marking under subject focus, except for 2SG. Whereas this would normally be realized as *-t-* in the absence of post-verbal subject focus, under subject focus it is realized as *-Ø-* (7), making it identical to 1SG and 3SG masculine. A sentence with a verb that maintains the regular 2SG agreement marking is ungrammatical (8). Note that this reduced agreement occurs in the past tense, but not the present tense.

(7) wəli ʊntɔ́-ð-ə   [ʃéen-Ø-ej]   aðɔ́.   Kenyan Maay
     FOC food-F-DEF bring-2SG-PST 2SG
     'YOU brought the food.'

(8) *wəli ʊntɔ́-ð-ə   [ʃéen-t-ej]   aðɔ́.   Kenyan Maay
     FOC food-F-DEF bring-2SG-PST 2SG
     'YOU brought the food.'

In contrast, when a subject is in focus, Speaker 2 neutralizes all person markers with the exception of 3sg in the feminine form. The paradigms of both speakers for verbal inflection under the subject focus condition in the past tense can be seen in Appendix B (§B).

The storyboard questions required answers in which a single constituent was both new to the discourse and of communicative interest, making it the focus of the sentence. The fact that the *wəli* construction was consistently employed by Speaker 1 while performing the storyboard suggests that it does function as a means of narrow focus. Furthermore, the data indicate that the element which receives this focus is placed in the post-verbal position. With the construction's basic form and function outlined, we turn in the next section to exploring in more detail the types of constituents that can be compelled to move post-verbally under *wəli* focus.

4 *Wəli* constructions

4.1 Basic Configuration

As introduced above, when the function word *wəli* appears pre-verbally, there is an obligatory shift or displacement of an element to post-verbal position. In the absence of *wəli*, the shifted element would be expected to appear pre-verbally in pragmatically neutral contexts. Sentence (9), which has basic SOV word order, has its object *buúggə hoostiísa* 'the book's underside' before the verb *jaalé* 'be located above, be at here'.

(9) qaláŋ-k-ə buúg-g-ə hoos-tiísa [jaal-Ø-é].   Kenyan Maay
     pen-m-def book-m-def under-poss be.at-3sg.m-prs
     'The pen is under the book.'

This can be compared to (10), where the object *buúggə biðiðís* 'the book's left' appears post-verbally. The use of *wəli* always results in some element appearing in post-verbal position.

(10) qaláŋ-k-ə wəli [jaal-Ø-é]   buúg-g-ə biði-ðís.   Kenyan Maay
      pen-m-def foc be.at-3sg.m-prs book-m-def left-poss
      'The pen is to the left of the book.'

However, the focused element does not have to immediately follow the Verb Complex and can be preceded by another NP, as in (11).

(11) wəli [hir-Ø-éj]   elbép-k-ə   dərə́s-k-ə.   Kenyan Maay
FOC close-3SG.M-PST door-M-DEF neighbor-M-DEF
'The neighbor closed the door.'

4.2 Elements focused by *wəli*

The function of the *wəli* construction is to place into focus a variety of different constituents, which includes nominals of many types, and even those modified by a relative clause (which is how Kenyan Maay performs most attribution). Also included are adverbials (which are nominal themselves) and even larger adverbial clauses and complement clauses. The construction does not appear to be able to focus verbs, nor a full Verb Complex, and also cannot focus adjectival participles.

4.2.1 Nominal Constituents and Complement Clauses

Nominals of any argument type can be placed in focus by *wəli*, including subjects and objects of various types. For example, oblique objects in pragmatically neutral sentences typically appear pre-verbally, as in (12). However, when focused by *wəli*, as in (13), they appear after the verb. Note, however, that because the adpositions governing these nominals reside in the Verb Complex, they remain *in situ* even under the focus condition, rather than moving with their nominal object. So, while in (12) the adposition *ən* follows the element it locates, in (13), the nominal moves to post-verbal position while the adposition does not.

(12) usə́   məðərəsə́-ð-ii [ən loɣa-j-é].[6]   Kenyan Maay
3SG.M school-M-RDEF ADP walk-3SG.M-PST
'He walks to the school.'

(13) anə́ jé   wəli [ɪn hogaami-j-é] gurú-ge.   Kenyan Maay
1SG 3SG.F FOC ADP lead-1SG-PST house-POSS
'I lead her to my house.'

Complement clauses can also be focused by the *wəli* construction. Speaker 1 used the *wəli* construction almost invariably for sentences containing com-

---

[6]The glide here is a phonologically-conditioned allomorph of the 1SG and 3SG masculine suffix, which is otherwise Ø. According to Saeed (1982a: 8), VV sequences are not allowed and are repaired by [j] epenthesis.

plement clauses.[7] Sentences containing multiple clauses have *wəli* in the main clause, as in (14).

(14) anə́ wəli [éxr-Ø-ej]   ɪnti   fərás-ɪn-Ø-ə.   Kenyan Maay
      1SG FOC say-1SG-PST COMP be.happy-VBZ-1SG-PRS

      'I said THAT I AM HAPPY.'

As in previous sentences, *wəli* follows the subject and precedes the Verb Complex. In (14), *wəli* is placing into focus the embedded clause *ɪnti fərásɪɪnə* 'that I am happy', which occurs post-verbally.

4.2.2 Adverbials

Most adverbials in Maay are nouns governed by adpositions, which is also true for Somali (Saeed 1999: 124). The distribution of the adverbial nouns, like adverbials in many other languages, is fairly free. However, as nouns, these elements also have the ability to be placed into focus by *wəli*, and when this occurs, their position post-verbally becomes obligatory. This can be seen in (15) where *wəli* places into focus the adverbial *tartiíp* 'quietness', which accordingly follows the verb.

(15) láŋ-k-ii   wəli [ɪn axri-j-é]   tartiíp.   Kenyan Maay
      man-M-RDEF FOC ADP read-3SG.M-PST quietness

      'The man read QUIETLY'. (Lit. with quietness)

4.2.3 Adjectival Participles and Relative Clauses

Like in Somali (Green 2021: 158), Kenyan Maay derives adjectival participles through verbalizing suffixes, e.g., (-*ən* in (17). Unlike adverbs, adjectival participles are unable to be focused on their own via this construction, as attempted in (16). This likely stems from the fact that they appear in relative clauses and are bound by the nominal that governs them. As seen in (17), focusing on the "redness" of the plate formally requires focus on the entire noun phrase.

(16) *ɑðə wəli səháŋ [o   kaðə-t-í]   gʊðʊʊð-ən.   Kenyan Maay
      2SG FOC plate from.1SG take-2SG-PST red-VBZ

      'You take a RED plate from me.'

---

[7]At this point, only one sentence has been elicited in which a complement clause was used but the *wəli* construction was not used. The complement clause in that sentence still appeared in post-verbal position despite the absence of *wəli*.

(17) ɑðə́ wəli [o   kaðə-t-í]   səháŋ gʊðʊʊðə́-ə́n.   Kenyan Maay
     2SG FOC from.1SG take-2SG-PST plate red-VBZ

     'You take A RED PLATE from me.'

In (17), *gʊðʊʊðə́n* is a subject relative clause (lit. 'a plate that is red'). As Saeed
(1999: 214) reports, this also occurs in Somali, where nominal modification fre-
quently involves restrictive relative clauses that are not preceded by a relative
pronoun.

4.3 Common environments

Our elicitations showed a variety of syntactic environments in which *wəli* can oc-
cur. Environments in which a focus construction is likely to occur are sentences
which contain new information. This tendency to focus new information is il-
lustrated in (18) and (19). In the first sentence, Speaker 1 was asked to translate
the English sentence 'I ate the bread'. In this sentence, the information was all
known to the speaker, and there was no morphological focus marking. However,
when the second sentence was elicited, Speaker 1 used the *wəli* construction to
focus the new information (the amount of bread eaten).

(18) anə́ rootʰə́-g-ii   [ʕaam-Ø-í].   Kenyan Maay
     1SG bread-M-RDEF eat-1SG-PST

     'I ate the bread.'

(19) anə́ wəli [ʕaam-Ø-é] lə́mə ʒəp   rootʰə́.   Kenyan Maay
     1SG FOC eat-1SG-PST two   piece bread

     'I ate TWO PIECES OF BREAD.'

It is also the case that *wəli* almost invariably occurs in sentences with com-
plement clauses, as mentioned above in Section 4.2.1. The construction is also
frequently employed in sentences with ditransitive verbs (20) and in sentences
which include adjuncts (21). The correlation between these more complex sen-
tences and the use of the *wəli* construction is related to the fact that these more
complex sentences carry more new information.

(20) malə́ŋ-k-ə   wəli [sii-j-í]   buúg-g-ə   ɲaɲur-t-ə́. Kenyan Maay
     teacher-M-DEF FOC give-3SG.M-PRS book-M-DEF cat-F-DEF

     'The teacher gives a book to THE CAT.'

(21) məlái-k-ə hés-t-ii wəli [ʃeen-Ø-í] halfə́-t-ə.   Kenyan Maay
fish-ᴍ-ᴅᴇꜰ music-ꜰ-ʀᴅᴇꜰ ꜰᴏᴄ bring-3ꜱɢ-ᴘꜱᴛ party-ꜰ-ᴅᴇꜰ

'The fish brought the music at ᴛʜᴇ ᴘᴀʀᴛʏ.'

Apart from elicitations, this construction also frequently occurred when the first speaker was asked to translate the story 'The North Wind and the Sun' (see Appendix A [A]).

5 Pre-verbal focus

In addition to *wəli*, Kenyan Maay also makes use of the marker of pre-verbal focus *jaa*. Very seldom in the course of elicitations, Speaker 1 offered an utterance employing *jaa*, which prompted further inquiry into its use and the ways in which it can be distinguished from *wəli*. It was often the case that *jaa* was provided as an alternative to *wəli*.

(22) e-jaál-k-ii jaa [ʃeen-Ø-é-ŋ] ʊntə́-ð-ii.[8]   Kenyan Maay
dog-ᴘʟ-ᴍ-ʀᴅᴇꜰ ꜰᴏᴄ bring-3-ᴘꜱᴛ-ᴘʟ food-ꜰ-ʀᴅᴇꜰ

'ᴛʜᴇ ᴅᴏɢꜱ brought the food.'

(23) ʊntə́-ð-ii wəli [ʃeen-Ø-é-ŋ] e-jaál-k-ii.   Kenyan Maay
food-ꜰ-ʀᴅᴇꜰ ꜰᴏᴄ bring-3-ᴘꜱᴛ-ᴘʟ dog-ᴘʟ-ᴍ-ʀᴅᴇꜰ

'ᴛʜᴇ ᴅᴏɢꜱ brought the food.'

As seen in (22) and (23), both of which were elicited during the storyboard enactment, the focus markers *jaa* and *wəli* both appear before the Verb Complex, but while *wəli* focuses a constituent following the Verb Complex, *jaa* focus whatever constituent immediately precedes it.

Similar to *wəli*, the word order of sentences with *jaa* is more constrained. In sentences with *jaa*, both the focused constituent and the focus marker must precede the Verb Complex, or the sentence will be ungrammatical, as (24) is below.

(24) *[ən duruk-t-í] tartiíp ɲaɲur-t-ə́ jaa.   Kenyan Maay
ᴀᴅᴘ move-3ꜱɢ-ᴘꜱᴛ quietness cat-ꜰ-ᴅᴇꜰ ꜰᴏᴄ

'ᴛʜᴇ ᴄᴀᴛ moved quietly.' (Lit. with quietness)

---

[8]The reader may note that word order in this sentence is not the unmarked SOV word order. As mentioned above, word order in Kenyan Maay is relatively free. It is unclear as of yet what discourse factors result in various word orders.

6 Sentences with more than one focus marker

In the course of our elicitations with Speaker 1, we also encountered, but did not explore in detail, instances where both the focus markers *wəli* and *jaa* occur in the same sentence. Such examples can be seen in (25) and (26).

(25) éj-k-ii   bəluúgə jaa wəli [ən roor-ø-é]   sɛ'ið.   Kenyan Maay
      dog-M-RDEF blue   FOC FOC ADP run-3SG.M-PRS speed

      'The blue dog runs fast. (Lit. with speed)'

(26) láŋ-k-ii   jaa wəli [ɛxr-Ø-é]   ɪnti fərás-ɪn-Ø-ə.
      man-M-RDEF FOC FOC say-3SG-PRS COMP be.happy-VBZ-3SG.M-PRS

                                                              Kenyan Maay

      'The man says that he is happy.'

Both of the sentences above can be considered complex, as the first sentence contains a relative clause and the second, a complement clause. In (25), *jaa* highlights the blueness of the dog. We asked the speaker questions about both a blue dog and a red dog, which were present as paper models. Therefore, *jaa* appears to be involved in contrastive focus, as the speaker used it to be clear about which of the two dogs she was describing. In this particular instance, *wəli* is involved in corrective focus of the adverbial *sɛ'ið* 'speed'. This sentence was given in response to the question, 'Did the blue dog run slow?' Therefore, Speaker 1 was both contrasting the color of the dog and correcting us on the speed of the dog.

Sentence (26) was received during a series of grammaticality checks with Speaker 1, in which we asked about different ways to form the sentence 'The man says that he is happy'. In (26), *jaa* is used to focus the subject, whereas it may be that *wəli* indicates reported speech. One insight from Speaker 1 was that using *wəli* in these specific instances was more appropriate if the reported speaker ('the man') was not present in the speech situation.

It is possible for two focus markers to occur in the same sentence in Somali as well (Ajello 1995; Tosco 2002; Green 2021: 336). In Somali, both markers of pre-verbal and post-verbal focus can occur in the same sentence, as in (27).

(27) waddam-ó kalé   ayáa wáxaa [la-gú sam-ee-y-eý]   baadhitaan-ó
      country-PL other FOC FOC   ISP-in do-FAC-3SG-PST.RED research-PL
      saliidó ka-lá   duwán.   Somali
      oil   in-with various

      'In other countries, various oil tests have been conducted.' (Green 2021: 336)

In Somali, one focus marker focuses an adverbial clause, and the second focus marker highlights a post-verbal element (Green 2021: 336). According to Tosco (2002: 37), the choice to focus an adverbial is a strategy used when attention is on the "development of the action". Tosco (2002: 39) also claims that the *wáxa(a)* construction in Somali is a cleft involving a relative clause, which is what allows the double focus, as it is only in simple sentences that double focus is disallowed. It would appear that the Somali and Kenyan Maay cases presented above differ, as *jaa* in (25) and (26) focuses the subject and not an adverbial.

Multiple focus marking, while uncommon, has been regularly observed to occur (Krifka 1991). For example, van der Wal & Namyalo (2016) discuss the interaction of two separate focus strategies (one pre-verbal and one post-verbal) in Luganda (Bantu). They claim that because the two focus strategies have different functions (identification and exclusion),[9] they can be combined. Comparison with this analogous case might suggest that the appearance of both markers of pre-verbal and post-verbal focus in (25) and (26) above is allowed because they are being used for separate purposes. However, as can be seen from Speaker 1's use of both *jaa* and *wəli* to enact our storyboard (see 22 and 23), the functions of these two focus markers must overlap as they are being used in the same context.

7 Future research and concluding remarks

7.1 Comparisons with Somali focus markers

This initial description of *wəli* offers some insight into focus marking in Kenyan Maay that extends well beyond the only other discussion of the matter in the literature, namely Saeed (1982a). It sets the stage for further exploration into the topic, which in closely related languages like Somali is arguably one of the most extensively described and analyzed aspect of the language's grammar. In this way, research on Somali focus presents a natural way ahead in terms of both description of Kenyan Maay focus but also in building a microtypological profile of how focus is encoded in these languages. Somali, as is well known, exhibits several morphosyntactic particularities under its "subject focus condition" (Green 2021: 308-311; Puglielli 1981b: 13; Saeed 1999: 192; Svolacchia et al. 1995). This includes prohibition on the use of pronoun clitics, reductions in verb agreement, and also the prohibition of subject marking. In addition, there is a known parallel between subject focus and subject relative clauses. Below, we will provide

---

[9]According to van der Wal & Namyalo (2016: 356), exclusive focus occurs when there is some referent in a set of alternatives to which the predicate does not apply, whereas identificational focus identifies a referent for which a presupposed proposition is true.

some initial observations of similarities and differences between Kenyan Maay and Somali focus markers, in order to lay the foundation for further comparison through continued research.

The clearest and most direct parallel with the *wəli* focus marker in Kenyan Maay is Somali's focus marker *wáxa(a)*. A typical sentence employing this focus marker in Somali is shown in (28). Both *wəli* and *wáxa(a)* are markers of post-verbal focus that introduce the cataphoric focalization of a constituent that follows the Verb Complex.

(28) wáxa [tag-Ø-ay]   Cáli.   Somali
      FOC go-3SG.M-PST.RED Cali

      'CALI went.' (Green 2021: 304)

The two focus markers are also semantically related, as both *wəl* in Kenyan Maay and *wáx* in Somali mean 'thing'. Synchronically, when used otherwise, *wəli* in Kenyan Maay presumably contains the "remote" definite determiner and *wáxa* contains a "basic" definite determiner, with both meaning 'the thing'. Both words can occur in sentences as nouns, rather than focus markers, as in (29) and (30).

(29) aðə́ [kəsaa-s-é]   jé   wəl-i   ɪntə [ən koj-t-í]?   Kenyan Maay
      2SG know-2SG-PST 3SG.F thing-RDEF COMP ADP come-3SG.F-PST

      'Do you know why she came here?'
      (Lit. Do you know the thing that she came here for?)

(30) w-áydin [t-aqaan-n-aa-n]   wáx-[aad doón-ey-s-aa-n].   Somali
      DECL-2PL 2PL-know-2-PRS-PL thing-2PL want-PROG-2-PRS-PL

      'You (PL) know what you want.' (Green 2021: 305)
      (Lit. You know the thing that you want.)

This opens up an interpretation of *wəli* and *wáxa(a)* focus constructions as being clefts, which reflects how they have often been translated. For example, the sentence *ʊntə́ðii wəli hagaðé éjkə* 'THE DOG brought the food' could potentially be translated as 'the thing that brought the food was THE DOG'. This is the view supported by Saeed (1982b), Tosco (1997: 132) (for Tunni), and Tosco (2002). However, Green (2021: 303) points out that his Somali speakers reject cleft readings.

Another notable feature of the *wáxa(a)* focus construction in Somali is that when the subject is in focus, the subject-verb agreement paradigm marked on the verb is reduced (see Andrzejewski 1968; Andrzejewski 1969). Somali verbs show reduced agreement (2SG, 2PL and 3PL are the same as 3SG.M) in subject relative clauses and with focused subjects in main clauses (see relevant paradigms

in Green 2021: 309 and Saeed 1982b: 79-80). As discussed in §3.1, Kenyan Maay maintains full inflection for agreement in number when the subject is in focus. For person (when using the past tense), Speaker 1 maintains full inflection except for 2sG subjects, while Speaker 2 removes all person markers except for 3sG.F subjects.

Finally, *wáxa(a)* and other Somali focus markers have the ability to coalesce with subject pronoun clitics (SPC) (Green 2021: 304). An example of this coalescence is given in (31), in which *wáxa(a)* combines with the 3sG feminine SPC *-ay*.

(31) wáx-ay [taha-sh-ay] ín-ay shaqá-da [qab-a-t-ó]. Somali
    FOC-3sG.F try-3sG.F-PST COMP-3sG.F work-F.DEF do-MID-3sG.F-IRR

    'She tried TO DO THE WORK.' (Green 2021: 304)

SPCs have not been observed in Kenyan Maay. In lieu of SPCs, Kenyan Maay makes regular use of independent personal pronouns. This contrasts with Somali, for which independent personal pronouns are associated with emphasis (Andrzejewski 1961).

When it comes to Kenyan Maay's *jaa*, a direct comparison can be drawn to the behavior of *báa* and *ayáa* in Somali. The use of the former is shown in (32).

(32) macállin-ka báa buugg-ág [ná sii-y-eý]. Somali
    teacher-M.DEF[10] FOC book-PL 1PL.OBJ give-3sG.M-PST.RED

    'THE TEACHER gave books to us.' (Green 2021: 260)

The focus marker *báa* follows the nominal constituent it places into focus, and occurs before the Verb Complex, just as *jaa* does. The focus marker *ayáa* operates in much the same way as *báa*, but its use is limited to a somewhat more formal register, typically absent from spoken Somali and found instead in written forms of the language.

As with *wáxa(a)*, both *báa* and *ayáa* can coalesce with SPCs, with the negative marker *aán* and with the interrogative marker *ma* (Green 2021: 295-303). This has not been observed with *jaa* in Kenyan Maay.

Yet to be explored in Kenyan Maay are the parallels between focus marking and the behavior of subject relative clauses and any evidence of subject marking.

7.2 Broad Focus

While we have mentioned narrow, contrastive, and corrective focus in Kenyan Maay, the reader may wonder about if and how broad focus readings are encoded

in the language. In Somali, broad focus involves the marker of pre-verbal focus *báa* or *ayáa* following the subject and SOV word order (Green 2021: 327). Post-verbal cataphoric focus is not used for broad focus.

To see if this holds true for Kenyan Maay as well, we attempted to elicit the use of *wəli* for broad focus. Various short videos were shown to Speaker 1, the videos having simple content such as a dog chasing a cat. We then asked our speaker *máj də jí?* 'What's up?' or 'What happened?' and recorded her response. This was repeated several times, and each time, she provided a sentence in which the focus marker *jaa* appeared following the subject, as shown in (33). The *wəli* construction was not used.

(33) ɲáɲur-t-ii jaa [kə dʒibi-t-í] éj-k-ii.   Kenyan Maay
      cat-F-RDEF FOC ADP sleep-3SG.F-PST dog-M-RDEF
      'THE CAT slept on the dog.'

Speaker 1's reluctance to use *wəli* may be due to the fact that broad focus in Kenyan Maay, like Somali, is simply not compatible with cataphoric focus.

7.3 Concluding remarks

In this paper, we have achieved the modest goal of describing and beginning to characterize certain aspects of Kenyan Maay focus, particularly as it pertains to the cataphoric focus marker *wəli*. We have illustrated the basics of its behavior, which has not yet been discussed in the literature, and have also discussed the ways that it can be compared to the marker of pre-verbal focus *jaa*, as well as the marker of post-verbal focus *wáxa(a)* in Somali. We have shown that although there are several parallels between focus marking in Somali and Kenyan Maay, some of the known peculiarities of Somali subject focus appear to be absent in Kenyan Maay. Some limitations of our research are that we have only worked with two speakers, and that most of our data comes from elicitations. We hope that future research will fill in some of these gaps and explore how *wəli* is used with a greater variety of speakers and genres.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

jó **wəli** es   karrəhen-∅-é-ŋ qóf-k-ii
3PL FOC REFL agree-3-PST-PL person-M-RDEF
They agreed that THE PERSON

ən   hor gul-éjs-əð-ə
COMP first succeed-FAC-MID-3SG-IRR
THAT FIRST SUCCEEDED

qokkáð-∅-ə   safrɘ́-g-ə   marɘ́-ð-ə
take.off-3SG.M-IRR traveler-M-DEF cloak-F-DEF
IN TAKING OFF THE TRAVELER'S CLOAK

qof-k-ə   kalə ku hɘ́gbəð-ən.
person-M-DEF other than be.stronger-VBZ
is the person (who) is stronger than the other.

kuðímbə dabelɘ́ komfúr **wəli** afuf-t-éj   ənti təbərtíi eh
then   Wind North FOC blow-3SG.F-PST COMP strength be
Then the North Wind blew WITH STRENGTH

lakín marwelbɘ́ səaj-ð-ə   afuf-∅-í
but whole   time-F-DEF blow-3SG.M-PRS
but the whole time (that) he blew

safrá-g-ə   marɘ́-ð-ə   kə   dudum-∅-ɘ́s.
traveler-M-DEF cloak-F-DEF around fold-3SG.M-PRS
the traveler wrapped his cloak around him.

markii-dembə dabelɘ́ komfúr əs   dip-t-í.
time-last   Wind North REFL give.up-3SG.M-PST
At last, the North Wind gave up.

kudəmbə irí-ð-ə   iftin-t-í   sə kulul
then   Sun-F-RDEF shine-3SG.F-PST DISJ warmth
then the Sun shone out warmly

markíibə   safrá-g-ə   **wəli** əs qokað-∅-éj   mará-ð-iis.
immediately traveler-M-DEF FOC REFL take.off-3SG.M-PST cloak-F-POSS
and immediately, the traveler took off HIS CLOAK.

markii kudembə dabelɘ́ komfúr **wəli** karrə-t-éj
at   last   Wind North FOC agree-3SG.F-PST
The North Wind agreed

ənti irí-ð-ə   hə́gbəð-ən-t-ə.
COMP Sun-F-DEF be.stronger-VBZ-3SG.F-PST
THAT THE SUN WAS STRONGER.'

Appendix B Full and reduced verbal agreement under subject focus for the past tense

Table 1: Full and reduced verbal agreement under subject focus for the
past tense in Somali (adapted from Saeed 1984: 83)

	Somali full agreement	Somali reduced agreement
1SG	keen-∅-ay	keen-∅-áy
2SG	keen-t-ay	keen-∅-áy
3SG.M	keen-∅-ay	keen-∅-áy
3SG.F	keen-t-ay	keen-t-áy
1PL	keen-n-ay	keen-n-áy
2PL	keen-t-een	keen-∅-áy
3PL	keen-∅-een	keen-∅-áy

keen 'bring'

Table 2: Full and reduced verbal agreement under subject focus for the
past tense in Kenyan Maay (Speaker 1 and Speaker 2)

	full agreement	Kenyan Maay S1	Kenyan Maay S2
1SG	ʃeen-∅-í	ʃéen-∅-ej	ʃéen-∅-ej
2SG	ʃeen-t-í	ʃéen-∅-ej	ʃéen-∅-ej
3SG.M	ʃeen-∅-í	ʃéen-∅-ej	ʃéen-∅-ej
3SG.F	ʃeen-t-í	ʃéen-t-ej	ʃéen-t-ej
1PL	ʃeen-n-í	ʃéen-n-ej	ʃéen-∅-ej
2PL	ʃeen-t-é-ŋ	ʃeen-t-é-ŋ	ʃéen-∅-ej
3PL	ʃeen-∅-é-ŋ	ʃeen-∅-é-ŋ	ʃéen-∅-ej

ʃeen 'bring'

References

Ajello, Roberto. 1995. La focalizzazione in Somalo. In Roberto Ajello & Saverio Sani (eds.), *Scritti linguistici e filologici in onore di Tristano Bolelli*, 1–28. Pisa: Pacini.

Andrzejewski, Bogumił W. 1961. Notes on the substantive pronouns in Somali. In Malcolm Guthrie (ed.), *Notes on the substantive pronouns in Somali*, 80–99. London: School of Oriental & African Studies, University of London.

Andrzejewski, Bogumił W. 1968. Inflectional characteristics of the so-called 'weak verbs' in Somali. *African Language Studies* 9. 1–51.

Andrzejewski, Bogumił W. 1969. Some observations on hybrid verbs in Somali. *African Language Studies* 10. 47–89.

Banti, Giorgio. 2011. Somali language. In Uhlig Siegbert (ed.), *Encyclopaedia Aethiopica*, vol. 4, 693–696. Wiesbaden: Harrassowitz.

Biber, Douglas. 1982. Accent in the Central Somali nominal system. *Studies in African Linguistics* 13. 1–10.

Comfort, Jade & Mary Paster. 2009. Notes on Lower Jubba Maay. In Masangu Matondo, Fiona Mc Laughlin & Eric Potsdam (eds.), *Selected Proceedings of the 38th Annual Conference on African Linguistics*, 204–216. Somerville, MA: Cascadilla Proceedings Project.

Gebert, Lucyna. 1986. Focus and word order in Somali. *Afrikanistiche Arbeitspapiere* 5. 43–69.

Green, Christopher R. 2021. *Somali Grammar*. Berlin, Boston: De Gruyter Mouton.

Krifka, Manfred. 1991. A compositional semantics for multiple focus constructions. *Semantics and Linguistic Theory* 1. 17–53.

Littell, Patrick. 2010. *Animal Party*. Retrieved on Nov 5, 2021. www . totemfieldstoryboards.org.

Mukhtar, Mohamed Haji & Omar Moalim Ahmed. 2007. *English-Maay dictionary*. London: Adonis & Abbey Publishers.

Paster, Mary. 2006. Aspects of Maay phonology and morphology. *Studies in African Linguistics* 35. 73–120.

Paster, Mary. 2010. Optional multiple plural marking in Maay. In Franz Rainer (ed.), *Current issues in linguistic theory 310: Variation and change in morphology*, 177–192. Amsterdam: John Benjamins.

Paster, Mary. 2018. Gender instability in Maay. In Jason Kandybowicz & Harold Torrence (eds.), *Selected Proceedings of the 45th Annual Conference on African Linguistics*, 205–218. Berlin: Language Science Press.

Puglielli, Annarita. 1981a. La derivazione nominale in somalo. In Annarita Puglielli (ed.), *Aspetti morfologici, lessicali e della focalizzazione*, 1–52. Rome: MAE, Dipartimento per la Cooperazione allo Sviluppo.

Puglielli, Annarita. 1981b. *Studi somali 2: Sintassi della lingua somala.* Annarita Puglielli (ed.). Rome: MAE, Dipartimento per la Cooperazione allo Sviluppo.

Saeed, John I. 1982a. Central Somali: A grammatical outline. *Monographic Journals of the Near East: Afroasiatic Linguistics* 8(2). 77–119.

Saeed, John I. 1982b. *Focus and Topic in Somali.* London University. (Doctoral dissertation).

Saeed, John I. 1984. *The syntax of focus and topic in Somali* (Kuschitische Sprachstudien 3). Hamburg: Buske.

Saeed, John I. 1999. *Somali.* Amsterdam: John Benjamins.

Smith, Katrina. 2022. *The Maay Maay nominal system and its tonology.* Syracuse University. (Master's thesis).

Svolacchia, Marco, Lunella Mereu & Annarita Puglielli. 1995. *Aspects of discourse configurationality in Somali.* Katalin E. Kiss (ed.). Oxford: Oxford University Press. 65–98.

Svolacchia, Marco & Annarita Puglielli. 1999. Somali as a polysynthetic language. In Lunella Mereu (ed.), *Boundaries of Morphology and Syntax*, 99–123. Amsterdam: John Benjamins.

Tosco, Mauro. 1997. *Af Tunni. Grammar, Texts, and Glossary of a Southern Somali Dialect.* Cologne: Rüdiger Köppe Verlag.

Tosco, Mauro. 2000. Cushitic overview. *Journal of Ethiopian Studies* 33(2). 87–121.

Tosco, Mauro. 2002. A whole lotta focusin' goin' on: Information packaging in Somali texts. *Studies in African Linguistics* 31(1-2). 27–53.

van der Wal, Jenneke & Saudah Namyalo. 2016. The interaction of two focus marking strategies in Luganda. In Doris L. Payne, Sara Pacchiarotti & Mokaya Bosire (eds.), *Diversity in African languages: Selected papers from the 46th Annual Conference on African Linguistics*, 355–377. Berlin: Language Science Press.

Chapter 10

Pronouncing PRO in Wolof

Suzana Fong

In Wolof, control clauses differ in whether the embedded PRO subject is pronounced. In some control clauses, the subject is phonologically null, as expected, while it is an overt pronoun in others. The main questions that arise are then: why do control clauses in this language differ in the phonological realization of PRO? Which control theory is compatible with such realization? I suggest that control clauses where the subject is an overt pronoun project a ΣP which 'impedes' movement. Assuming that control is derived by movement, I model the pronounced PRO as the partial residue of movement that has been impeded. Control clauses with a null subject, in turn, are restructured.

1 Introduction

Obligatory control is a phenomenon whereby the subject of an embedded clause, usually nonfinite, is null and coindexed with a matrix argument; the latter can be a subject or an object. The embedded null subject is indicated below as 'PRO'.[1]

(1) a. Sindhu$_1$ tried [PRO$_{1/*2}$ to eat natto].
     b. Lasha convinced Sindhu$_1$ [PRO$_{1/*2}$ to eat natto].

The phonological nullness of PRO is usually obligatory:

(2) a. * Sindhu tried [Anna/she to eat natto].
     b. * Lasha convinced Sindhu [Anna/she to eat natto].

---

[1]For an overview on control phenomena and theories, see Landau 2013.

Based on the phonological nullness of PRO, we can divide control theories in the following way:

(3) *Control theory typology*

    a. Inherent theories: phonological nullness is an inherent property of either PRO or of the control clause.

    b. Derivational theories: the phonological nullness of PRO is acquired during the derivation.

    c. Arbitrary theories: there is no necessary relationship between the syntax and semantics and PRO and its phonological realization. It can be null, but it does not have to be.

In inherent theories, the embedded subject of control clauses is null either because this is a property of the lexical item PRO or because there is no space in such clauses for a subject. In Chomsky (1981), for instance, PRO must be null because this is the only way for this DP to vacuously satisfy the Case Filter.

(4) *Case Filter*
    *NP, where NP has a phonetic matrix but no case.   (Chomsky 1981)

In Wurmbrand (1998) and her subsequent work, control can be obtained via restructuring, a phenomenon whereby embedded nonfinite clauses can have a truncated structure. This truncation can be so extreme that the embedded clause may not accommodate a subject. The phonological nullness of PRO is then trivially caused by the absence of a subject.

(5) *Restructuring analysis of (1a)*
    Sindhu tried [$_{VP}$ to eat natto].

In derivational theories, in turn, PRO does not start out phonologically null. This property is a consequence of some independent process or principle that occurs during the course of the derivation. For the Movement Theory of Control (MTC, Hornstein 1999), there is no PRO per se, nor is there a dedicated control module. Rather, control reduces to raising and the embedded subject of a control clause is null because this is the residue of movement of a DP (the controller) through multiple thematic positions.
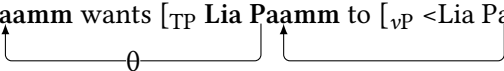
(6) *MTC analysis of (1a)*
    [Sindhu tried [$_{TP}$ <Sindhu> to [$_{vP}$ <Sindhu> eat natto]]]

In other words, for the MTC, the phonological nullness of PRO reduces to the rules that regulate linearization. Notably, the residue of movement is usually null.

A prediction that emerges from this analysis is that, if some independent factor prevents a lower copy from being deleted, the embedded subject in control clauses can be pronounced. Lee (2003) shows that this is the case in copy control in San Lucas Quiaviní Zapotec.[2]

(7) San Lucas Quiaviní Zapotec
 R-cààa'z Lia Paamm [g-ahcnèe Lia Paamm Gye'eihlly].
 HAB-want FEM Pam [IRR-help FEM Pam Mike]
 'Pam wants to help Mike.' (Lee 2003: 62, adapted)

Lee's proposal is that the embedded instance of *Lia Paamm* is a fully pronounced copy of movement.

(8) *Lee's MTC analysis of (7)*
 [**Lia Paamm** wants [$_{TP}$ **Lia Paamm** to [$_{vP}$ <Lia Paamm> help Mike]]]

More generally, then, in derivational theories of control, it is in principle possible for PRO to be pronounced.

Lastly, for arbitrary theories, the phonological nullness of PRO is an accidental property. In principle, nothing in the syntactic derivation of control clauses or in their semantics prevents PRO from being phonologically overt. One example of such a theory is McFadden & Sundaresan (2018), where there is only a minimal pronoun whose behavior as obligatory control PRO, arbitrary PRO, or dropped *pro* depends on the environment where it occurs.[3]

Against this background, we can turn to control in Wolof (Niger-Congo; Senegal). The complement clause of verbs like *jéem* 'try' is headed by a bare verb. The subject of that verb is interpreted as the matrix subject.[4]

(9) a. Xadi jéem-na togg ginaar.
 Xadi try-NA.3SG cook chicken
 'Xadi tried to cook chicken.'

---

[2] See also backwards control (Polinsky & Potsdam 2002).

[3] McFadden & Sundaresan (2018) focus on the syntactic properties of PRO. This theory is *compatible* with PROs with different phonological properties, though I believe further work would be required to predict when PRO is overt.

[4] Unless stated otherwise, all Wolof data were collected by the author in partnership with three consultants, native speakers of Wolof from Dakar, Senegal. All the data presented were checked with the three consultants via online elicitations. I thank S. M. Ndao, A. B. Sow, and S. Sène for their invaluable partnership in this project.

b. Maymuna fas-na   jàng taalif b-i.
Maymuna want-NA.3SG read poem CM.SG-DEF
'Maymuna wants to read the poem.'

As expected from the discussion above, an overt pronoun is prohibited.

(10) a. * Xadi jéem-na **mu**   togg ginaar.
Xadi try-NA.3SG 3SG.SBJ cook chicken
Lit.: 'Xadi tried she to cook chicken.'

b. * Maymuna fas-na   **mu**   jàng taalif b-i.
Maymuna want-NA.3SG 3SG.SBJ read poem CM.SG-DEF
Lit.: 'Maymuna wants she to read the poem.'

Wolof also has constructions where the embedded subject is an overt pronoun, instead of being phonologically null.

(11) a. Dimbali-na-a a-b   xale **mu**   jàng téere b-i.
help-NA-1SG   INDEF-CM.SG child 3SG.SBJ read book CM.SG-DEF
'I helped a child read the book.'

b. Dimbali-na-a a-y   xale **ñu**   jàng téere b-i.
help-NA-1SG   INDEF-CM.PL child 3PL.SBJ read book CM.SG-DEF
'I helped some children read the book.

Such a pronoun is obligatory, at least in the Wolof dialect surveyed here.[5,6]

---

[5]The alternative without a pronoun was either judged outright ungrammatical by the consultants I worked with or were accompanied by comments such as "some speakers talk like this, but this is not standard". However, see data in Dione (2019), where the pronoun is descriptively optional. As we are going to see below, the overtness of the embedded subject in *dimbali* control correlates with other properties (e.g. impossibility of clitic climbing and obligatoriness of resumptive pronoun under *Wh*-movement). The opposite set of properties obtains with subject control, where the embedded subject is null. It would be interesting to check whether these properties dovetail in the same way in Dione's data.

[6]The data presented in this paper would lead one to believe that the relevant distinction is one between subject and object control, the former disallowing overt pronominal subjects and the latter requiring them. However, that this is not the relevant criterion is suggested by the occurrence of a pronoun in interrogative control clauses, where the controller is a subject:

(i) Sàmba ak   Roxaya xam-na-ñu   k-an   la-y-**ñu**-fa   àndal.
Sàmba with Roxaya know-NA-3PL CM.SG-who FOC.OBJ-IPFV-3PL=LOC invite
'Sàmba and Roxaya know who to invite there.'

Due to logistical complications, a full incorporation of interrogative control data into the present analysis is still outstanding.

(12) a. * Dimbali-na-a a-b   xale jàng téere b-i.
          help-NA-1SG   INDEF-CM.SG child read book CM.SG-DEF
          Int.: 'I helped some child read the book.'

      b. * Dimbali-na-a a-y   xale jàng téere b-i.
          help-NA-1SG   INDEF-CM.PL child read book CM.SG-DEF
          Int.: 'I helped some children read the book.'

(13) illustrates the same facts with the verb *yey* 'convince'.

(13) Yey-na-a   Isaa rekk *(mu)   bind a-b   taalif.
      convince-NA-1SG Isaa only *(3SG.SBJ) write INDEF-CM.SG poem
      'I convinced only Isaa to write a poem.'

Given these data, we may ask the following questions:

(14) a. What governs the pronunciation of the subject of control clauses in
          Wolof?

      b. When the controlled subject is pronounced, why is it a pronoun?

      c. Can the pronounced PRO in Wolof help tease apart control theories
          with respect to the phonological properties of PRO (see typology in
          3)?

The questions in (14) can only be asked if sentences like (11) are indeed instances of control. The presence of the overt pronoun in the embedded clause makes them not look like true cases of control, given that PRO is usually phonologically null. In §3, we apply standard tests for control (e.g. *de se* reading, sloppy reading under ellipsis, bound reading) and conclude that this overt pronoun is a bound variable, just like obligatory control PRO. In §4, we bring back the sentences without a pronounced subject in (9b) and (10b) and compare them with clauses with a pronounced PRO. Specifically, we will see how these clauses differ with a pronounced subject with respect to Ā-resumption (i.e. the occurrence of a resumptive pronoun marking a position a phrase Ā-moved from) and clitic climbing: control clauses with a pronounced PRO require Ā-resumption and prohibit clitic climbing, while control clauses with a null PRO have the opposite behavior regarding the same properties. In §5, I propose an analysis where the clausal complement of predicates like *dimbala* 'help' and *yey* 'convince' is a ΣP that impedes movement.

(15) [_CP_ [I helped a child [_ΣP_ **she** to [_vP_ <a child> read the book]]]]

As seen in (15), A-movement to a further θ-position (resulting in control) or Ā-movement across ΣP leaves behind a resumptive pronoun. The resumptive pronoun is a partially pronounced copy in a movement chain (Van Urk 2018).

This analysis captures why a pronounced PRO and Ā-resumption dovetail in the same construction. Furthermore, it rounds out the typology of control as A-movement that is expected from the Copy Theory of Movement.

2 Morphosyntactic properties of the pronounced PRO

Before we investigate the control properties of constructions like (11) above, we examine the basic morphosyntactic properties of the pronoun that occurs in those constructions. Such a pronoun is a subject or nominative pronoun and cannot be accusative, and it is obligatory. It also cannot be replaced with a full DP.

The pronoun that occurs in the constructions like (11) comes from the subject or nominative paradigm (the rightmost column in Table 1).

Table 1: The pronominal system of Wolof (adapted from Zribi-Hertz & Diagne 2002: 29)

	Object clitics	Oblique pronouns	Subject markers
1SG	ma	man	(m)a
2SG	la	yaw	nga/ya
3SG	ko	moom	∅/(m)u
1PL	ñu	ñoom	ñu
2PL	leen	yeen	ngeen/yeen
3PL	leen	ñoom	ñu

This pronoun cannot be replaced with its accusative counterpart.

(16) a. * Dimbali-na-a a-b   xale=**ko**   jàng téere b-i.
          help-NA-1SG   INDEF-CM.SG child=3SG.ACC read book CM.SG-DEF
          Int.: 'I helped some child read the book.'

      b. * Dimbali-na-a a-y   xale=**leen**   jàng téere b-i.
          help-NA-1SG   INDEF-CM.PL child=3PL.ACC read book CM.SG-DEF
          Int.: 'I helped some children read the book.

Lastly, it cannot be replaced with a lexical DP, regardless of whether or not it contains a pronoun coindexed with the matrix antecedent.[7]

(17) a. * Dimbali-na-a a-b   xale **yaay=am**   jàng téere
           help-NA-1SG   INDEF-CM.SG child mother=POSS.3SG read book
           b-i.
           CM.SG-DEF
           Int.: 'I helped some child for his mother to read the book.'

      b. * Dimbali-na-a a-b   xale **Roxaya** jàng téere b-i.
           help-NA-1SG   INDEF-CM.SG child Roxaya read book CM.SG-DEF
           Int.: 'I helped some child for Roxaya to read the book.'

Having surveyed the basic morphosyntatic properties of the pronoun that occurs in the constructions like (11), we now turn to its semantic properties. We shall see that it behaves like a bound variable, a signature property of the subject of control clauses.

3 Bound variable properties of the embedded pronoun

The pronoun that occurs in the subject postion of the clause subcategorized by verbs like *yey* 'convince' and *dimbala* 'help' in Wolof passes several standard tests employed to identify bound variables. The data examined in this section thus support the claim that such pronoun behaves like control PRO despite the fact that it is pronounced. We will compare the behavior of this pronoun with the behavior of dropped subjects in Wolof; the former behaves like a bound variable, while the latter behaves like a free variable.

The first hint that the constructions investigated are instances of control comes from the fact that there is no idiom preservation in *dimbali* sentences.

(18) Sa   jaan wàcc-na.
      POSS.2SG snake descend-NA.3SG
      'Your snake descended.'
      'You did what you had to do.'

---

[7]A reviewer asks whether the pronounced PRO could be replaced with a full copy of the controller. This is indeed a relevant question and is a gap in the data I currently have. I hope to be able to fill this gap in the future.

(19)   a. Isaa dimbali-na   sa   jaan j-i   **mu** wàcc.
           Isaa help-NA.3SG POSS.2SG snake CM/SG-DEF 3SG descend
           'Isaa helped your snake descend.'
           #'Isaa helped you do what you had to do.'

       b. Isaa wax-na   sa   jaan j-i   **mu** wàcc.
           Isaa tell-NA.3SG POSS.2SG snake CM.SG-DEF 3SG descend
           'Isaa told the snake to descend.'
           #'Isaa told you do what you had to do.'

   The pronoun in constructions like (11) systematically contrasts with pronominal subjects of finite clauses. The latter are also unpronounced, though presumably because they are dropped arguments.

   First, obligatory control PRO should be obligatorily coreferent with a local and c-commanding antecedent, which acts as its controller. A dropped subject in Wolof can have its interpretation established in the discourse.

(20)   Bu dee **Mareem**$_k$ moom,   njiit l-i   dafa foog-oon ne   ***pro***$_k$
       BU DEE Mareem   3SG.OBL boss CM.SG-DEF DAFA think-PST COMP 3SG
       dafa-y   xalamal Roxaya.
       DAFA-IPFV praise   Roxaya
       'As for Mareem, the boss thought that she praised Roxaya.'

The pronoun that occurs in *dimbala* sentences, however, cannot.

(21)   # Bu dee **Mareem** moom,   *pro* dimbali-na-a Mbaye **mu**   bind
         BU DEE Mareem 3SG.OBL 1SG help-NA-1SG Mbaye 3SG.SBJ write
         a-b   taalif.
         INDEF-CM.SG poem
         Lit.: 'As for Mareem, I helped Mbaye for her (Mareem) to write a poem.'

Interestingly, one of the consulted speakers offered (22) as a correction to (21). In (22), the matrix subject was dropped to accommodate the reference of the discourse-salient *Mareem*.

(22)   Bu dee **Mareem**$_k$ moom,   ***pro***$_k$ dimbali-na   Mbaye mu bind
       BU DEE Mareem   3SG.OBL 3SG   help-NA.3SG Mbaye 3SG write
       a-b   taalif.
       INDEF-CM.SG poem
       'As for Mareem, she helped Mbaye write a poem.'

Second, the antecedent of a dropped subject can be a higher subject or object.

(23) a. ***pro**_k* Wax-na-a   Mbaye [ne   ***pro**_k* jot-na-a   a-b
          1SG   say-NA-1SG Mbaye [COMP 1SG   receive-NA-1SG INDEF-CM.SG
          leetar].
          letter]
          'I told Mbaye that I received a letter.'

      b. *pro* Wax-na-a   **Mbaye**_k* [ne   ***pro**_k* jot-na   a-b
          1SG say-NA-1SG Mbaye   [COMP 3SG   receive-NA.3SG INDEF-CM.SG
          leetar].
          letter]
          'I told Mbaye that he received a letter.'

Conversely, the antecedent of the embedded pronoun in control sentences with a pronounced PRO must be the matrix object (i.e. it cannot be the matrix subject).[8]

(24) a. * ***pro**_k* Dimbali-na-a Sàmba **ma**_k* togg ginaar   g-i.
          1SG   help-NA-1SG Sàmba 1SG   cook chicken CM.SG-DEF
          Lit.: 'I helped Sàmba for me to cook the chicken.'

      b. *pro* Dimbali-na-a **Sàmba**_k* **mu**_k*   togg ginaar   g-i
          1SG help-NA-1SG Sàmba   3SG.SBJ cook chicken CM.SG-DEF
          'I helped Sàmba cook the chicken.'

Finally, the antecedent must c-command the pronounced pronoun:

(25) a. Dimbali-na-a [_DP rakk-u   Roxaya ak   Faatu]_k **mu**_k*   jàng
          help-NA-1SG [sister-LNK Roxaya with Faatu] 3SG.SBJ read
          téere b-i.
          book CM.SG-DEF
          'I helped [Roxaya and Faatu]'s sister read the book.'

      b. * Dimbali-na-a [_DP rakk-u   Roxaya ak   Faatu_k] **ñu**_k*   jàng
          help-NA-1SG [sister-LNK Roxaya with Faatu   ] 3PL.SBJ read
          téere b-i.
          book CM.SG-DEF
          Int.: 'I helped [Roxaya and Faatu]'s sister, so that Roxaya and Faatu would read the book.'

---

[8]JJ Lim correctly points out that the intended meaning may be itself ill-formed. It is also suggested that *convince* is used instead. I hope to be able to do this in the future.

These data also show that the number property of the antecedent and that of the pronoun must match.

Besides its difference from dropped subjects, pronouns in *dimbala* sentences also behave like bound variables, just like control PRO (Landau 2013). That obligatory control PRO is a bound variable can be diagnosed by properties such as (i) Obligatory coreference, (ii) Obligatory sloppy reading under ellipsis, and (iii) Obligatory *de se* interpretation.

As a bound variable, obligatory control PRO should yield only sloppy readings under VP ellipsis. This is exactly what can be found in Wolof.

(26) Bu dee Isaa moon, wax-na-a Kumba mu   jàng a-b   téere,
     BU DEE Isaa 3SG.OB say-NA-1SG Kumba 3SG.SBJ read INDEF-CM.SG book
     waaye wax-u-ma   Roxaya < mu   jàng a-b   téere >.
     but   say-NEG-1SG.SBJ Roxaya   3SG.SBJ read INDEF-CM.SG book
     'As for/According to Isaa, I told Kumba to read a book, but not Roxaya.'

   a.   I didn't tell Roxaya for her (= Roxaya) to read the book.
   b. * I didn't tell Roxaya for Kumba to read the book.
   c. * I didn't tell Roxaya for Isaa to read the book.

(27) Yey-na-a   sama yaay mu   jënd kër g-u   bees,
     convince-NA-1SG POSS.1SG mother 3SG.SBJ buy house CM.SG-COMP new
     wayee yey-u-ma   sama   baay.
     but   convince-NEG-1SG POSS.1SG father
     'I convinced my mother to buy a new house, but not my father.'
   a.   I didn't convince my father for him to buy a new house.
   b. * I didn't convince my father for my mother to buy a new house.

Furthermore, in attitude contexts, obligatory control PRO should be obligatorily interpreted *de se* (relative to its controller). This is also what can be found in Wolof.

(28) Maryam wax-na   Kadeer mu   dem.
     Maryam say-NA.3SG Kadeer 3SG.SBJ leave
     'Maryam told Kadeer to leave.'
   a. # Maryam is hosting a party. She hears that a certain waiter named Kadeer is being a nuisance. Maryam tells the nearest waiter, "Kadeer has to go." Unbeknownst to her, she's talking to Kadeer.

b. Maryam is hosting a party. She hears that a certain waiter named Kadeer is being a nuisance. Maryam tells Kadeer, "You have to go."

(29) Faatu dafa   yey   Kadeer, mu noppi.
Faatu do.3SG convinced Kadeer 3SG shut.up
'Faatu convinced Kadeer to shut up.'

a. # Suppose Faatu listens with Kadeer to a recording of a speech. The speaker in the recording is Kadeer himself, although he is not aware of that (Kadeer had a cold at the time of the recording so his voice is unrecognizable). After a while, Faatu feels she has had enough and wants to put an end to it. She says, "This dude should shut up." Kadeer agrees.

<div align="right">(context adapted from Landau 2015)</div>

b. Faatu and Kadeer are arguing. Faatu tells Kadeer, "You should shut up." Kadeer agrees.

A free pronoun, contrastively, does not have to be interpreted *de se*:

(30) … Maryam ak   Roxaya wax-na-ñu Kadeer ne   *pro* war-na   jënd
Maryam with Roxaya say-NA-3PL Kadeer COMP 3SG should-NA.SG buy
oto b-u   bees.
car CM.SG-COMP new
'Roxaya and Maryam told Kadeer that he should by a new car.'

a. Maryam and Roxaya work in an office where park spaces are labeled with the car owner's name. The car parked on the space labeled 'Kadeer' is in bad shape. During an office party, they are talking to some worker they don't know. They comment to him, "Kadeer should buy a new car." Embarrassingly, it turns out the person they are talking to is Kadeer himself. As I am recounting this incident to you, I say, …

b. Maryam and Roxaya work selling cars. They see a car pulling up. The car is in bad shape. A guy gets out of the car. It is their friend Kadeer, who came to the car dealership where they work for a quick visit. They tell him: 'You should buy a new car'. As I am recounting this incident to you, I say, …

Finally, a bound reading can also be witnessed with antecedents headed by *only* and *no*.

(31) Wax-u-ma   b-enn   ndongo mu   ñëw ci   baal b-i.
      say-NEG-1SG CM.SG-one student 3SG.SBJ come PREP party CM.SG-DEF
      'I told no student to come to the party.' (Lit.: I didn't tell a student to come
      to the party.)

      # I hate Kadeer. I am throwing a party and am inviting my students to it.
        But I tell each of them that Kadeer is not invited, and therefore if they
        come they shouldn't bring Kadeer along with them. Kadeer was very
        sad because...[9]

(32) Yey-na-a   Isaa rekk mu   bind a-b   taalif.
      convinced-NA-1SG Isaa only 3SG.SBJ write INDEF-CM.SG poem
      'I convinced only Isaa to write a poem.'

      a. # I have three students, Faatu, Kumba, and Isaa. I am trying to get
          them to express themselves in verse, which they are not used to.
          Faatu and Isaa are enthusiastic about acquiring new abilities, so
          they accepted the assignment. Kumba, however, refused to do it.

      b. I have three students, Faatu, Kumba, and Isaa. I am trying to get
          them to express themselves in verse, which they are not used to.
          Isaa is enthusiastic about acquiring new abilities, so he accepted the
          assignment. Faatu and Kumba, however, refused to do it.

In sum, even though there is an overt pronoun in sentences like in (11), these
constructions can be classified as instances of obligatory control. The data ex-
amined in this section show that these pronouns are bound variables, just like
obligatory control PRO in a language like English. In other words, in some control
sentences, Wolof has instances of a pronounced PRO, which occurs as an overt
pronoun in control clauses subcategorized by verbs like *yey* 'convince', *dimbala*
'help', and *wax* 'tell'.

4 The size of control clauses in Wolof

I just established that the pronoun in *dimblali* sentences is a bound variable, just
like obligatory control PRO. An obvious question to ask now is: why does Wolof
have what can be described as an overt PRO. To answer this question, it may be
useful to compare both types of control clauses in Wolof introduced in §1. To
recall, in sentences headed by a verb like *dimbala* 'help', a subject pronoun is

---

[9]Thank you to Itai Bassi for providing the context for this sentence!

obligatory in the embedded clause (33). In contrast, in a sentence headed by a verb like *jéem* 'try,' the same pronoun is prohibited and the embedded subject is necessarily null (34).
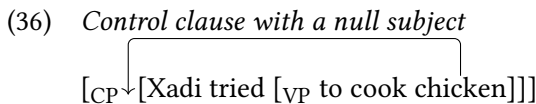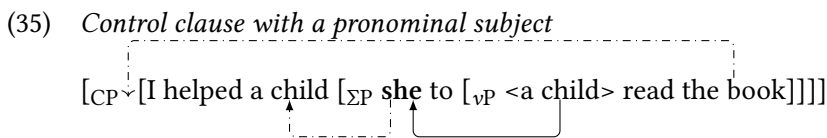
(33) Dimbali-na-a a-b   xale *(**mu**)   jàng téere b-i.
     help-NA-1SG INDEF-CM.SG child *(3SG.SBJ) read book CM.SG-DEF
     'I helped a child read the book.'

(34) Xadi jéem-na   (*****mu**)   togg ginaar.
     Xadi try-NA.3SG (*3SG.SBJ) cook chicken
     'Xadi tried to cook chicken.'

This difference dovetails with other properties, as summarized in Table 2 below. Control clauses with an obligatory pronominal subject require the resumption of an Ā-moved element, while disallowing clitic climbing. Control clauses with an obligatory null subject have exactly the opposite behavior.

Table 2: Dovetailed properties of different types of control clauses

Main verb	Pronounced PRO	Ā-Resumption	Clitic climbing
Dimbala, wax, yey	✓	✓	*
Jéem, fas	*	*	✓

To account for these facts, I propose that control clauses with a pronominal subject and those with a null subject differ in size and, furthermore, that the former makes movement more difficult, though not impossible. An overt pronoun is a correction effect that results from the attempt to cross it.

(35) *Control clause with a pronominal subject*

    [$_{CP}$ [I helped a child [$_{\Sigma P}$ **she** to [$_{vP}$ <a child> read the book]]]]

(36) *Control clause with a null subject*

    [$_{CP}$ [Xadi tried [$_{VP}$ to cook chicken]]]

This analysis of control sentences with a pronounced PRO is inspired by Lee's (2003) analysis of copy raising. This is desirable because there are empirical similarities between Wolof control sentences with a pronounced PRO and copy control in San Lucas Quiaviní Zapotec.

4.1 Clitic climbing

Control clauses in Wolof differ with respect to clitic climbing. In control clauses where the subject is an overt pronoun, a clitic must stay inside the embedded clause. In control clauses with a null subject, it must climb into the matrix clause. The latter observation has already been made by Gowda & Wu (2020) and Martinović (2021). As also observed by Martinović (2021), this difference suggests that subject control clauses can be analyzed in terms of restructuring (Wurmbrand 1998 and subsequent work).

For starters, clitics in Wolof cannot climb into the matrix clause from an embedded finite clause.

(37) a. Gis-na-a   sama   xarit ci   xewam b-i.
         see-NA-1SG POSS.1SG friend PREP wedding CM.SG-DEF
         'I saw my friend at his wedding.'

   b. Gis-na-a=**ko**   ci   xewam b-i.
       see-NA-1SG=OBJ.3SG PREP wedding CM.SG-DEF
       'I saw him at his wedding.'

   c. Mareem xalaat-na   [ne   gis-na-a=**ko**   __ ci   xewam
       Mareem think-NA.3SG [COMP see-NA-1SG=3SG.ACC   PREP wedding
       b-i   ].
       CM.SG-DEF]
       'Mareem thinks I have seen him at his wedding.'

   d. * Mareem xalaat-na=**ko**   [ne   gis-na-a   __ ci   xewam
       Mareem think-NA.3SG=OBJ.3SG [COMP see-NA-1SG   PREP wedding
       b-i   ].
       CM.SG-DEF]
       Int.: 'Mareem thinks that I saw him in his wedding.'

In control clauses with a pronominal subject, the clitic must stay inside the control complement.

(38) a. Kadeer dimbali-na Mareem mu   jënd=**ko**.
         Kadeer help-NA.3SG Mareem 3SG.SBJ buy=3SG.ACC
         'Kadeer helped Mareem buy it.'

   b. * Kadeer dimbali-na=**ko**   Mareem mu   jënd __ .
       Kadeer help-NA.3SG=3SG.ACC Mareem 3SG.SBJ buy
       Int.: 'Kadeer helped Mareem buy it.'

(39) illustrates the same fact for the verb *yey* 'convince'.

(39) Jàngalekat b-i   yey-na{*=**ko**}   ndongo y-i
     teacher   CM.SG-DEF convince-NA.3SG=3SG.ACC student CM.PL-DEF
     ñu   bind{=**ko**}.
     3PL.SBJ write=3SG.ACC
     'The teacher convinced the students to write it.'

However, clitic climbing is obligatory in control clauses with a null subject.

(40) a. Maymuna fas-na   jàng taalif b-i.
          Maymuna want-NA.3SG read poem CM.SG-DEF
          'Maymuna wants to read the poem.'

      b. * Maymuna fas-na   jàng=**ko**.
          Maymuna wantNA.3SG read=3SG.ACC
          Int.: 'Maymuna wants to read it.'

      c. Maymuna fas-na=**ko**   jàng __.
          Maymuna want-NA.3SG=3SG.ACC read
          'Maymuna wants to read it.'

To account for these facts, I propose that control clauses with a null subject are restructured. This has already been proposed by Gowda & Wu (2020) and Martinović (2021). Following Wurmbrand's (1998) definition of restructuring, restructured clauses are severely truncated; that is, they lack functional projections usually found in clauses, including a layer where subjects are base-generated. The fact that the subject is null in restructured clauses thus follows trivially from the fact that there is no subject syntactically represented in restructured clauses. The idea that these clauses are truncated also explains why a clitic can only find an appropriate host in the matrix clause.

By the same reasoning, control clauses with an overt subject cannot be as severely restructured. They must be bigger than clauses with a null subject, so that clitic climbing is blocked.

4.2 Resumptive pronoun with Ā-movement

Another difference between control clauses in Wolof relates to Ā-resumption, the occurrence of a clitic pronoun in the position where some phrase Ā-moves from.

If Ā-movement, instantiated by *Wh*-movement and clefting, proceeds from a control clause with a pronominal subject, a resumptive pronoun occurs obligatorily, marking the position the *Wh*-phrase moved from.

(41)   a.   K-an   la   jàngelekat b-i   dimbali ndongo l-i
            CM.SG-who FOC.OBJ teacher   CM.SG-DEF help   student CM.SG-DEF
            dimbali mu   nataal=*(**ko**)?
            help   3SG.SBJ draw=3SG.ACC
            'Who did the teacher help the student draw?'

     b.   L-an   la   jàngalekat b-i   yey   ndongo
            CM.SG-what FOC.3SG teacher   CM.SG-DEF convince student
            y-i   ñu   bind=*(**ko**)?
            CM.PL-DEF 3PL.SBJ write=3SG.ACC
            'What did the teacher convince the students to write?'

(42)   Ginaar g-i   la   Maymuna dimbali Roxaya mu
      chicken CM.SG-DEF OBJ.FOC.3SG Maymuna help   Roxaya 3SG
      togg*(=**ko**).
      cook*(=3SG.ACC)
      'The chicken, Maymuna helped Roxaya cook.'

In contrast, in control clauses where the subject is null, a resumptive pronoun is prohibited under the same circumstances.

(43)   a.   *K-an   la   Roxaya d-oon   jéem a   nataal=**ko**?
            CM.SG-who FOC.OBJ Roxaya IPFV-PST try   INF draw=3SG.ACC
            Int.: 'Who did Roxaya try to draw?'

     b.   K-an   la   Roxaya d-oon   jéem a   nataal?
            CM.SG-who FOC.OBJ Roxaya IPFV-PST try   INF draw
            'Who did Roxaya try to draw?'

(44)   Ginaar g-i   la   Maymuna fas   yéene togg(*=**ko**).
      chicken CM.SG-DEF OBJ.FOC.3SG Maymuna want   want   cook(*=3SG.ACC)
      'The chicken, Maymuna wanted to cook.'

5 Analysis

In the previous section, I have shown that, in control clauses where the subject is obligatorily pronominal, Ā-resumption is obligatory, while clitic climbing is

banned. In control clauses where the subject is obligatorily null, the opposite state of affairs obtains. A question that arises at this juncture is how to *relate* these properties.

As briefly mentioned above, I propose that control clauses where the subject is null are restructured. The same proposal has already been made by Gowda & Wu (2020) and Martinović (2021).

(45)   Xadi jéem-na   togg ginaar.
       Xadi try-NA.3SG cook chicken
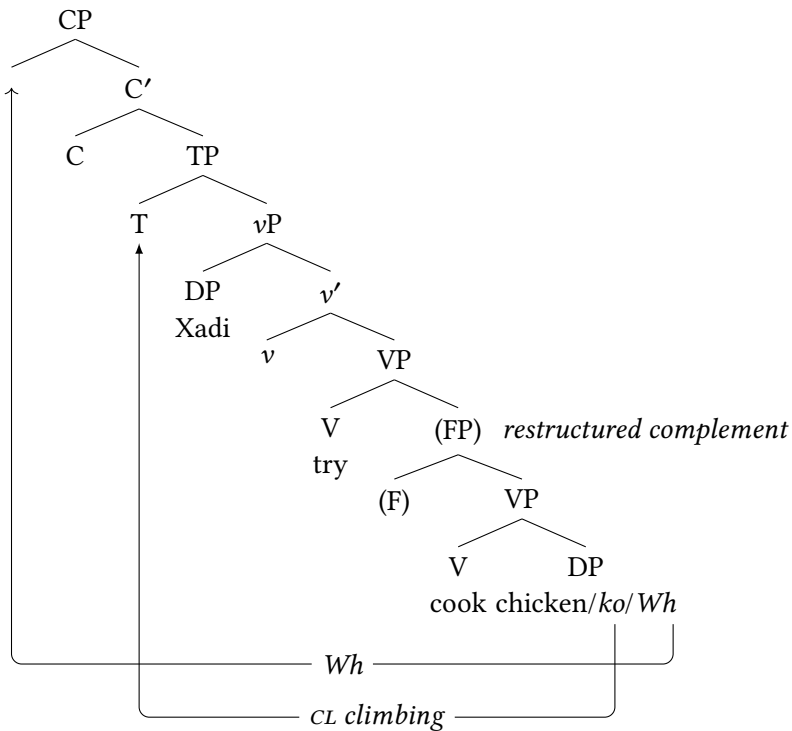   'Xadi tried to cook chicken.'

Figure 1: Representation of sentence (45)

In the representation in Figure 1, the pre-verbal DP *Xadi* is base-generated inside the matrix clause. The embedded clause, being truncated, lacks a subject position. For an analysis of the interpretation of sentences with a restructured nonfinite complement, see e.g. Grano (2015).

Suzana Fong

Clitic climbing is obligatory because the only functional projection that can host the clitic is in the matrix clause. Additionally, there is no Ā-resumption because the embedded clause is so small, it could not impede Ā-movement. The subject is obligatorily null because there is not enough space for a subject.[10]

By contrast, in control sentences where the embedded subject is pronominal, the complement clause is a ΣP, which is bigger than a restructured clause. ΣP is stipulated to impede different types of movement.[11]

(46) Dimbali-na-a a-b   xale **mu**   jàng téere b-i.
      help-NA-1SG   INDEF-CM.SG child 3SG.SBJ read book CM.SG-DEF
      'I helped a child read the book.'

Clitic climbing is not possible because ΣP impedes movement. Likewise, ΣP is not an appropriate host for a clitic, presumably because it is phonologically null. The subject is an overt pronoun as a corrective effect of ΣP impeding A-movement through different thematic positions (Hornstein 1999).[12] Ā-resumption is a corrective effect of the same type: ΣP impedes Ā-movement.[13]

That control clauses with a pronominal subject are bigger is further supported by binding facts. A more deeply embedded pronoun in these clauses can be coindexed with the matrix subject.[14]

---

[10]Alternatively, the embedded clause is restructured and MTC-style A-movement (Hornstein 1999) does not leave any overt residue (Martinović 2021).

[11]For the moment, I do not have a more precise label for the control clauses with a pronounced PRO. It could well be simply a CP. I keep the unspecified ΣP to reflect the current stage of the research, which also lacks a precise formalization for the intended "impediment" of movement resorted to in the present analysis and imposed by ΣP.

[12]Martinović (2021) has already proposed a movement analysis for subject control in Wolof, though the author rejects that the structures analyzed here should be analyzed as control clauses with a pronounced PRO. This cannot be the conclusion taken from §3.

[13]A reviewer correctly asks why ΣP blocks clitic climbing, while permitting, but "impeding" phrasal A and Ā-movement. At the moment, I can offer some speculations to address this important question. It is possible that relevant distinction is between head movement like clitic climbing and phrasal movement like A-movement into a further thematic position and Ā-movement of a DP. *Why* this should be the case, however, is not something I am presently able to answer. Alternatively, it is possible that clitics require a functional projection to be hosted (assuming head movement of the verb into this functional projection) and ΣP is an appropriate host. As such, it could be the case that ΣP does not impede clitic climbing per se; rather, the clitic does not need to climb into the matrix clause because it has found an appropriate host within ΣP.

[14]These data also indicate that what is taken here as a pronoun that is a pronounced PRO is indeed a pronoun and not an agreement prefix. The latter is not expected to be relevant for binding, so it would not help in delimiting the embedded clause as a binding domain. A true pronoun, on the other hand, can be the subject that defines a binding domain.
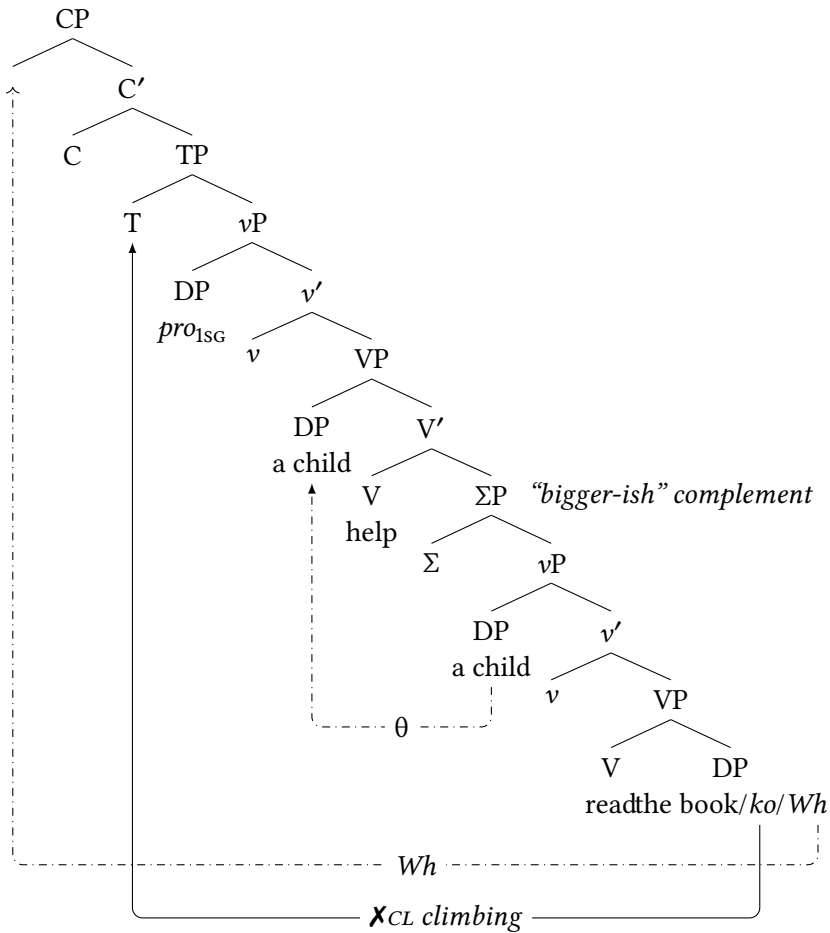
Figure 2: Representation of sentence (46)

(47) a. Maymuna$_k$ wax-na   Roxaya mu xool=ko$_k$.
          Maymuna   say-NA.3SG Roxaya 3SG see=3SG.OBJ
          'Maymuna told Roxaya to look at her.'

     b. Maymuuna$_k$ yey-na   Roxaya mu xool=ko$_k$.
          Maymuna   convince-NA.3SG Roxaya 3SG see=3SG.OBJ
   'Maymuna convinced Roxaya to look at her.'

This interpretive possibility suggests that the embedded clause is a binding domain that excludes the subject. Binding domains, in turn, are usually taken to be

bigger structures which contain a subject and which are impervious to syntactic relationships like government.

5.1 Towards a formalization of pronounced PRO in Wolof

As just mentioned, I propose that there is an overt pronoun as a corrective effect of crossing ΣP, which impedes movement. An analysis of the overtness of PRO as a consequence of the difficulty of A-movement (in the Hornstein 1999 sense) has already been proposed for copy control in San Lucas Quiaviní Zapotec by Lee (2003).

But why exactly is a pronoun pronouced in the embedded clause? I propose that the overt pronoun is a partially pronounced copy (Van Urk 2018). More precisely, I assume that nouns have a complex structure where person features are represented at D. In partially pronounced copies, NP is deleted, but D survives. The exponence of D is a pronoun.

(48)
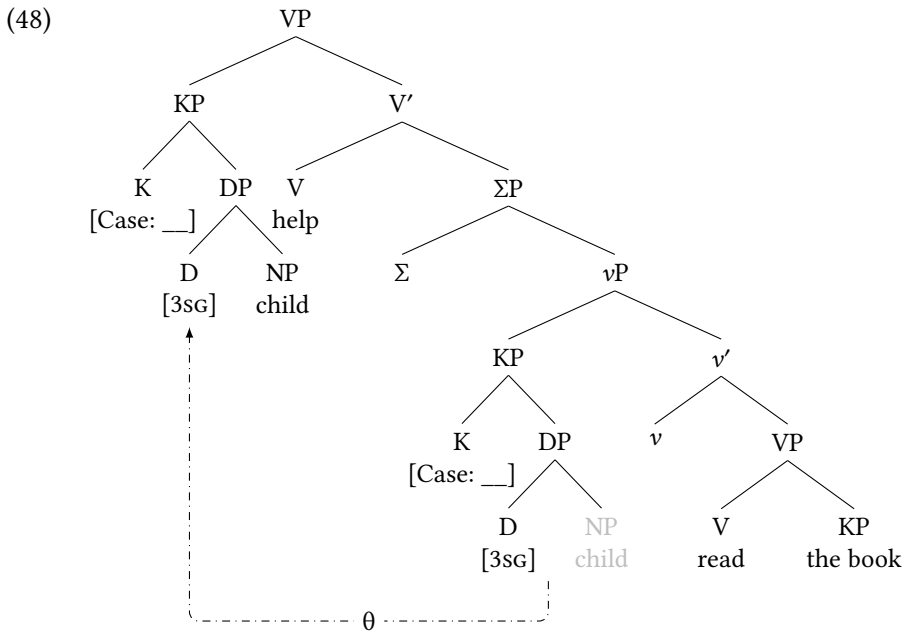
Finally, I assume that [Case: _] in the lower embedded copy of the controller (i.e. PRO) remains unvalued throughout the derivation and is exponed as unmarked nominative case (Preminger 2014). This is why the pronounced PRO in Wolof is a subject/nominative pronoun.

An advantage of the analysis proposed here is that it relates Ā-resumption and the overtness of the PRO in Wolof control sentences with a pronounced PRO: both are resumptive pronouns that emerge as a corrective effect to the movement impediment imposed by ΣP. It seems undesirable to treat the co-occurrence of these properties in the same construction (Wolof control sentences with a pronounced PRO) as coincidental. Likewise, it allows for Wolof control to be related to copy control, as it is found in San Lucas Quiaviní Zapotec. In the analysis proposed here and that proposed by Lee (2003) for San Lucas Quiaviní Zapotec, the pronounced PRO (a subject pronoun in Wolof, a full copy in Zapotec) is the residue of "impeded movement".

6 Discussion and outlook

This project is motivated by the questions in (14). According to the analysis put forward here, there is a difference between control clauses in Wolof because they have different sizes. The PRO in control clauses subcategorized for by verbs like *dimbala* or *yey* is pronounced because it is a residue of movement that has been impeded.

With respect to how the phonological properties of PRO are derived in this analysis, it is of the derivational type, thus further bolstering this category. Specifically, I assumed an MTC (Hornstein 1999 et seq.) framework. This type of theory can account not only for the pronunciation of PRO in Wolof control, but crucially for why it correlates with Ā-resumption. As I show below, the present analysis also rounds out the typology of control as movement and relates it to the typology of Ā-movement.

However, it cannot be the case that PRO is always inherently null, as inherent theories would have it, given control sentences with a pronounced PRO in Wolof. Arbitrary theories, in contrast, do offer some flexibility in the pronunciation of PRO. However, they may fail to capture the correlation between a pronounced PRO and Ā-resumption. A general question that can be asked is why, to the best of our knowledge, PRO is silent in the majority of languages. This is not expected if phonological nullness is an arbitrary property.

Beyond these questions, the analysis put forth here also rounds out the typology of the realization of the subject of control clauses, as expected from the movement theory of control and the copy theory of movement. Starting with Ā-movement, the copy theory of movement predicts the existence of four linearization possibilities:

(49) *Only higher copy is pronounced (English)*

What did Yuwei eat < what > for breakfast?

(50) *Lower copy is pronounced (covert Wh-movement; Mandarin)*

Zhangsan yiwei Lisi mai   le   shenme?
Zhangsan thinks Lisi bought what
'What does Zhangsan think Lisi bought?' (Huang 1983)

(51) *Multiple copy pronunciation (German)*

Wem   glaubst du **wem**   deine Eltern vertrauen?
who.DAT believe you who.DAT your parents trust
'Who do you think your parents trust?' (Pankau 2013)

(52) *Lower copy is partially pronounced (pronoun copying; Dinka)*

Yè **kɔ̂ɔc-kò**   cí̠i   Bôl   **ké** t̠i̠iŋ?
be.3SG people-which PRF.OV Bol.GEN 3PL see.NF
'Which people has Bol seen?' (Van Urk 2018: 12c)

In control derived by movement, the exact same four possibilities can be seen, with Wolof, as analyzed here, being an instance of the partial pronunciation of the lower copy, analogous to the Dinka Ā example (52).

(53) *Only higher copy is pronounced (English)*

Lasha convinced **Sindhu** [< Sindhu > to eat natto].

(54) *Lower copy is pronounced (backwards control; Tsez)*

< kidbā > [**kidbā** ziya   bišra   ] yoqsi.
          [girl.ERG cow.ABS feed.INF] began
'The girl began to feed the cow.' (Polinsky & Potsdam 2002: 2)

(55) *Multiple copy pronunciation (copy control; San Lucas Quiaviní Zapotec)*

R-càaa'z **Lia Paamm** [g-ahcnèe **Lia Paamm** Gye'eihlly].
HAB-want FEM Pam   [IRR-help FEM Pam   Mike   ]
'Pam wants to help Mike.'(Lee 2003: 62, adapted)

(56) *Lower copy is partially pronounced (Wolof)*

Dimbali-na-a **a-b**   **xale mu**   jàng téere b-i.
help-NA-1SG INDEF-CM.SG child 3SG.SBJ read book CM.SG-DEF
'I helped a child read the book.'

In sum, Wolof control sentences with a pronounced PRO are exactly as expected if the copy theory of movement is assumed and combined with the movement theory of control. Wolof control sentences with a pronounced PRO thus provide further support for these theories.

Pronounced PROs have also been documented in Bùlì (Sulemana 2021) and Mandarin Chinese (Li 2021). In both languages, the pronounced PRO coexists with a null counterpart, similarly to what has been shown here in Wolof. Sulemana's and Li's findings and the findings in this paper are strikingly similar, in that the pronounced PRO can be demonstrated to occur in clauses that are bigger than the clauses where the null PRO occurs. This generalization is also supported by the fact thatthe pronounced PRO is associated with focus in Romance languages, Hungarian (Szabolcsi 2009), and Tamil (Sundaresan 2010). Assuming that focus also requires a more complex left periphery, the generalization seems to be that a pronounced PRO correlates with a more complex clausal structure. Future work on the phonological realization of control PRO should take this generalization into account. Likewise, the fact that pronounced PRO is possible, but null in the overwhelming majority of cases, is in need of an explanation.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

CM	class marker	OBL	oblique
LNK	linker	PREP	preposition
NA	sentential particle for neutral sentences (*na*)		

Acknowledgements

to Mouhamadou Deme for putting me in contact with Ndao and thank you to Iz-abela Jordanoska for putting me in contact with Deme. I first became interested in control sentences with a pronounced ᴘʀᴏ when working with L. Touré.

References

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Dione, Cheikh Bamba. 2019. Clause structure, pro-drop and control in Wolof: An LFG/XLE perspective. *Nordic Journal of African Studies* 28(3).

Gowda, Yadav & Danfeng Wu. 2020. *Clitic climbing and linear adjacency in Wolof: Paper presented at the 94th Annual Meeting of the Linguistic Society of America.*

Grano, Thomas. 2015. *Control and restructuring*, vol. 56. Oxford: Oxford University Press.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30(1). 69–96. DOI: 10.1162/002438999553968.

Huang, C-T. James. 1983. *Logical relations in Chinese and the theory of grammar*. Massachusetss Institute of Technology. (Doctoral dissertation).

Landau, Idan. 2013. *Control in generative grammar: A research companion*. Cambridge: Cambridge University Press.

Landau, Idan. 2015. *A two-tiered theory of control*, vol. 71. Cambridge,MA: MIT Press.

Lee, Felicia. 2003. Anaphoric R-expressions as bound variables. *Syntax* 6(1). 84–114. DOI: 10.1111/1467-9612.00057.

Li, Daoxin. 2021. Controlling overt subjects in Mandarin. *Proceedings of the Linguistic Society of America* 6(1). 303–316.

Martinović, Martina. 2021. *Exhaustive control as raising: The case of Wolof.* Ms. https://ling.auf.net/lingbuzz/006072.

McFadden, Thomas & Sandhya Sundaresan. 2018. Reducing pro and PRO to a single source. *The Linguistic Review* 35(3). 463–518. DOI: 10.1515/tlr-2018-0003.

Pankau, Andreas. 2013. *Replacing copies: The syntax of wh-copies in German.* Utrecht University. (Doctoral dissertation).

Polinsky, Maria & Eric Potsdam. 2002. Backward control. *Linguistic Inquiry* 33(2). 245–282. DOI: 10.1162/002438902317406713.

Preminger, Omer. 2014. *Agreement and its failures* (Linguistic Inquiry Monographs 68). Cambridge, MA: MIT Press.

Sulemana, Abdul-Razak. 2021. *Non-finite complementation: A case study of Bùlì.* Massachusetts Institute of Technology. (Doctoral dissertation).

Sundaresan, Sandhya. 2010. A phase-based account of the PRO/anaphor distinction. *Proceedings of ConSOLE XVIII* 1. 19.

Szabolcsi, Anna. 2009. *Overt nominative subjects in infinitival complements cross-linguistically* (NYU Working Papers in Linguistics Spring 2009). New York: New York University.

Van Urk, Coppe. 2018. Pronoun copying in Dinka Bor and the copy theory of movement. *Natural Language & Linguistic Theory* 36(3). 937–990. DOI: 10.1007/s11049-017-9384-x.

Wurmbrand, Susanne. 1998. *Infinitives*. Massachusetts Institute of Technology. (Doctoral dissertation). https://dspace.mit.edu/handle/1721.1/9592.

Zribi-Hertz, Anne & Lamine Diagne. 2002. Clitic placement after syntax: Evidence from Wolof person and locative markers. *Natural Language & Linguistic Theory* 20(4). 823–884. DOI: 10.1023/A:1020494714861.

Chapter 11

Typology of Tigrinya WH-interrogatives

Keffyalew Gebregziabher

CNRS-IKER UMR 5478

Maia Duguine

CNRS-IKER UMR 5478

In this paper, we investigate wh-interrogatives in Tigrinya. We show that Tigrinya at least employs three independent strategies, including the cross-linguistically known wh-in-situ and wh-movement, to express content questions. First, we demonstrate that wh-interrogatives in Tigrinya occur in different syntactic positions, suggesting that the simple parametric dichotomy between wh-in-situ and wh-movement does not explain all the facts in Tigrinya. Then, we run syntactic diagnostics, such as long distance dependency, reconstruction effects, weak crossover and island effects, and show that the three strategies indeed exhibit different sensitivities to the list of syntactic diagnostics. Finally, we examine some interpretive properties, namely presuppositionality and exhaustivity, of the three wh-interrogative strategies and we observe that while the wh-in-situ and wh-movement strategies exhibit no exhaustivity effects, the third strategy appears sensitive to both interpretive properties.

1 Introduction

In this paper, we discuss the nature of wh-interrogatives in Tigrinya, an Ethio-Semitic language mainly spoken in Ethiopia and Eritrea. Tigrinya has a basic SOV word order (Hetzron 1972) in a neutral context (1).

(1) dɨmu ʔančwa səgʷigʷ-a
     cat   mouse chase.GER-3FSG.SBJ
     'A cat chased a mouse.'

Wh-phrases[1] in the language can surface in different syntactic positions of the clause. The examples in (2) show that both wh-questioned subjects and objects can surface in what seems to be their canonical argument position, maintaining the SOV word order.

(2) a. mən nɨ-selam   riʔ-u-wwa?
         who ACC-Selam see.GER-3MSG.SBJ-3FSG.OBJBJ
         'Who saw Selam?'
     b. selam nɨ-mən   riʔ-a?
         selam ACC-who see.GER-3FSG.SBJ
         'Who did Selam see?'

Wh-questioned subjects and objects can also surface to the left of the clause and be followed by the verb. In this case, the basic order of Tigrinya is altered, since the verb is not the final element of the clause: (3a) shows S_{wh}VO, whereas (3b) shows O_{wh}VS.[2]

(3) a. mən riʔ-u-wwa   nɨ-selam?
         who see.GER-3MSG.SBJ-3FSG.OBJ ACC-Selam
         'Who saw Selam?'
     b. nɨ-mən   riʔ-a   selam?
         ACC-who see.GER-3FSG.SBJ Selam
         'Who did Selam see?'

In comparison to the structures in (2), non-neutral orders are possible in which the verb remains in final position. For instance, O_{wh}SV order is possible with a wh-questioned object (4).

(4) nɨ-mən   selam riʔ-a?
     ACC-who Selam see.GER-3FSG.SBJ
     'Who did Selam see?'

---

[1] Some of the basic Tigrinya wh-words include *mən* 'who', *ʔɨntay*, 'what', *ʔabəy* 'where', and *kəməy* 'how'. Some of these words may inflect for case as in *nɨ-mən* ACC-who 'whom' (Gebregziabher 2013).

[2] We mainly provide examples with verbs in the perfective and gerundive aspects as they are functionally the same in Tigrinya. We largely set aside the imperfective forms here.

Wh-questioned adjuncts in the language can also maintain the SOV word order (5) or alter it by moving the wh-phrase to the left-edge of the clause along with the verb (6).

(5) a. selam nɨ-yared   ʔabəy riʔ-a-tto?
         Selam ACC-Yared where see.GER-3FSG.SBJ-3MSG.OBJ
         'Where did Selam see Yared?'

     b. selam nɨ-yared   kəməy riʔ-a-tto?
         Selam ACC-Yared how   see.GER-3FSG.SBJ-3MSG.OBJ
         'How did Selam see Yared?'

(6) a. ʔabəy riʔ-a-tto   selam nɨ-yared?
         where see.GER-3FSG.SBJ-3MSG.OBJ Selam ACC-Yared
         'Where did Selam see Yared?'

     b. kəməy riʔ-a-tto   selam nɨ-yared?
         how   see.GER-3FSG.SBJ-3MSG.OBJ Selam ACC-Yared
         'How did Selam see Yared?'

Finally, wh-questions in which the wh-subject or wh-object surfaces to the right of the verb are unattested (7-8).

(7) *nɨ-selam riʔ-u-wwa   mən?
     ACC-selam see.GER-3MSG.SBJ-3MSG.OBJ who
     'Who saw Selam?'

(8) *selam riʔ-a   nɨ-mən?
     Selam see.GER-3FSG.SBJ ACC-who
     'Who did Selam see?'

In this paper, based on the above facts, we show that Tigrinya has three independent strategies to express wh-interrogatives. The first strategy, illustrated in (2), involves a standard wh-in-situ configuration of the type observed in languages like Mandarin Chinese, Japanese, etc., where the wh-phrase remains in its base-position. The second strategy, illustrated in (3), involves movement of the wh-phrase immediately left-adjacent to the finite verb, followed by T-to-C movement (cf. 4). This is similar to what is observed in typical wh-movement languages, such as English, Hungarian, etc. Finally, the third strategy, illustrated in (4), involves movement of the wh-phrase to the left edge of the main clause, without the verb accompanying it. In order to simplify the discussion (pending their

full description), we refer to these strategies as (i) wh-in-situ, (ii) wh-movement, and (iii) peripheral-wh, respectively.

The remainder of the paper is organized as follows: Section 2 presents diagnostics for A'-movement dependencies for distinguishing the different wh-interrogative strategies under discussion. We demonstrate how the three wh-interrogatives differ with respect to islands, reconstruction and WCO effects. Section 3 looks into the interpretive differences. Finally, Section 4 concludes with some remarks on future endeavours.

2 Diagnosing A'-movement properties

This section explores the syntactic properties of each of the three wh-configurations introduced above, with respect to their A'-properties, by testing long-distance dependencies, reconstruction, weak crossover, and island-sensitivity.

On the one hand, we show that three very different wh-strategies emerge, exhibiting some regularities behind the apparent "free" word order of Tigrinya wh-interrogatives. On the other hand, we show that among these three strategies, the wh-in-situ strategy indeed shows strong similarities with typical wh-in-situ languages, whereas the wh-movement strategy has the properties expected from a language with wh-movement, validating the characterization we made in Section 1 merely based on the observation of word order facts. The third strategy, the peripheral-wh strategy, displays, as we will see, hybrid properties, between a movement and a non-movement construction.

2.1 Long-distance dependency (LDD)

Long-distance dependency refers to a syntactic-semantic relation between a constituent (e.g., a wh-phrase or a pronoun) and a syntactically licensed position (e.g., a gap) in an embedded clause (Ross 1967, Chomsky 1977). In wh-movement languages, an embedded wh-phrase surfaces at the left-edge of the matrix clause (through successive-cyclic movement), exhibiting long-distance dependency with respect to its lower position, as illustrated in (9).[3]

(9)   Who$_i$ does Mary know [that Anna kissed t_i]?

---

[3]Some of the Tigrinya complementizers include *zɨ-*, *kəmzɨ-* and *ʔɨntə-*: *zɨ* is used to introduce relative and nominalized clauses, but *kəmzɨ-*, which also includes *zɨ-*, introduces CP complements. In contrast, *ʔɨntə-* is used to introduce non-finite clauses and indirect questions (see Gebregziabher 2023 for a detailed discussion).

In bona-fide wh-in-situ languages, long-distance dependencies are observable through scopal effects (Huang 1982, Bayer & Cheng 2017, a.o.). This is nicely illustrated by the ambiguous example given in (10) from Mandarin Chinese (hereafter Chinese), where the ambiguity arises due to the narrow and wide scope readings (example from Bayer & Cheng 2017: 4, see also Huang 1982: 254).

(10) Mandarin Chinese
      Bótōng zhīdào Huángróng xǐhuān shéi (?)
      Botong know   Huangrong like   who
      a. 'Botong knows who Huangrong likes.'   (indirect question)
      b. 'Who does Botong know Huangrong likes?'   (matrix question)

In Tigrinya, both standard wh-movement and wh-in-situ configurations exhibit long-distance dependencies (similar to Chinese and English, respectively).

Before we introduce long distance dependencies in Tigrinya, first notice that clausal complementation in Tigrinya features an embedded clause to the left of the matrix verb, introduced by a complementizer prefixed on the embedded finite verb, as illustrated in (11):

(11) yared [selam nɨ-hailu   kəmzɨ-səʕam-ət-to]
      Yared Selam ACC-Hailu COMP-kiss.PFV-3FSG.SBJ-3MSG.OBJ
      fəlit'-u
      know.GER-3MSG.SBJ
      'Yared knew that Selam kissed Hailu.'

Using (11) as a baseline example, we show that (i) the wh-movement strategy, as in English, allows for long-distance dependency across clauses, and (ii) the wh-in-situ strategy, as in a typical wh-in-situ language, such as Chinese, allows wh-phrases to take wide scope.[4]

In conformity to the standard wh-movement strategy, Tigrinya wh-questioned embedded objects can surface to the left of the matrix verb, in a position that is unambiguously outside the embedded clause as shown in (12).

---

[4]In fact, cross-linguistically, this wide scope bearing property is widely attested in a variety of languages with the typical wh-in-situ strategy (see, for example, Ko 2005 on Japanese, Torrence & Kandybowicz 2015 on Krachi, Green & Jaggar 2003 on Hausa, Sabel & Zeller 2006 on Zulu, Bayer & Cheng 2017 on Bangla, a.o.).

(12) nɨ-mənᵢ məsil-u-wwo   yared [selam *t*ᵢ
ACC-who think.GER-3MSG.SBJ Yared Selam
zɨ-səʕam-ət-to]?
COMP-kiss.PFV-3FSG.SBJ-3MSG.OBJ
'Who did Yared think that Selam kissed?'

Consistent with the standard wh-in-situ strategy, Tigrinya wh-questioned embedded objects can remain in their base-generated positon and can have wide scope reading as illustrated in (13).

(13) yared [selam nɨ-mən   kəmzɨ-səʕam-ət]   fəlit'-u?
Yared Selam ACC-who COMP-kiss.PFV-3FSG.SBJ know.GER-3MSG.SBJ
a. 'Yared knew who Selam kissed'
b. 'Who did Yared know that Selam kissed?'

With regard to the third wh-interrogative strategy identified above, the peripheral-wh strategy, it also allows long-distance dependency. Examples in (14)-(15) illustrate how the questioned-object of the embedded verb surfaces unambiguously in the matrix clause (and unlike in example (12) above, the matrix verb does not surface right-adjacent to it). In this case, the embedded verb also remains in its final position (15).

(14) nɨ-mənᵢ yared [*t*ᵢ zɨ-səʕam-ki]   məsil-u-wwo?
ACC-who Yared   COMP-kiss.PFV-2FSG.SBJ think.GER-3MSG.SBJ-3MSG.OBJ
'Who did Yared think that you kissed?'

(15) nɨ-mənᵢ yared [selam *t*ᵢ kəmzɨ-səʕam-ət-to]
ACC-who Yared Selam   COMP-kiss.PFV-3FSG.SBJ-3MSG.OBJ
fəlit'-u?
know.GER-3MSG.SBJ
'Who did Yared know that Selam kissed?'

It is interesting to note that as in (12) above, a mixed pattern is possible, where the embedded verb can surface to the left of its subject (16), apparently due to T-to-C movement.

(16) nɨ-mənᵢ yared [*t*ᵢ kəmzɨ-səʕam-ət-to   selam]
ACC-who Yared   COMP-kiss.PFV-3FSG.SBJ-3MSG.OBJ Selam
fəlit'-u?
know.GER-3FSG.SBJ
'Who did Yared know that Selam kissed?'

Thus, the three Tigrinya wh-interrogative strategies, as expected, exhibit long-distance A'-dependencies.

2.2 Reconstruction

Another diagnostic widely used for A'-movement is *reconstruction* – a phenomenon which refers to the observation that certain syntactic relations which hold in the position at which the constituent attaches before A'-movement, still hold after movement (Chomsky 1981, 1986). For example, Binding Condition A, which requires anaphors to be bound in their local domain, is such a relation.[5] The examples in (17) illustrate the point.

(17)   a.   [Which pictures of herself$_i$] does Mary$_i$ like t_i?
        b.   [Which pictures of each-other$_i$] did [Mary and John]$_i$ like t_i?

In (17), the two wh-phrases, *which pictures of herself* and *which pictures of each other*, as objects of the verb *like*, are associated with two structural positions: a moved/derived position at the beginning of the sentence and a base-merge position before the movement happens. Thus, since the reconstructed constituents contain anaphors, namely the reflexive pronoun *herself* and the reciprocal *each other*, they must be bound by a local antecedent in order to satisfy Condition A after the reconstruction happened. Therefore, the only way to satisfy Condition A is to assume that the wh-phrase, along with the anaphors, have been reconstructed to their base-merge position at LF (see Fox & Nissenbaum 2004 and references cited therein for a range of other alternatives).

In Tigrinya, there are different types of anaphors, including the reciprocal *ħidħid* (Kogan 1997), and reconstruction of wh-phrases is permitted irrespective of whether the verb moves along with the wh-phrase or not. In the baseline sentence in (18), for example, *nənayħidħidom* is a reciprocal anaphor and the plural DP *Selam and Yared* is its antecedent. Because the noun phrase *Selam and Yared* c-commands the anaphor *nənayħidħidom*, Condition A is trivially satisfied, and a bound interpretation is licensed.

(18)   selam-n   yared-n   nənayħidħidom   siʔlitat   fəty-om
        Selam-and Yared-and each.other   pictures   like.GER-3MPL.SBJ
        'Selam and Yared liked pictures of each-other.'

---

[5]For current overview and debate on the topic, see Safir (2004), Bruening & Al Khalaf (2019), a.o.

The examples in (19) and (20) show how wh-phrases reconstruct in both the wh-movement and the peripheral-wh strategies, respectively. Surprisingly, the bound interpretation of the anaphors is also licensed in these examples, despite the fact that the DP, *Selam-n Yared-n*, fails to c-command the anaphor, *nǝnayħi-dħɨdom*. The bound interpretation is, however, expected if the anaphor is interpreted in its reconstructed position, where it is c-commanded by *Selam and Yared* in accordance with the locality requirement (i.e., Principle A of the binding theory). In this regard, the wh-movement strategy is acceptable compared to the peripheral-wh, irrespective of their reconstruction as illustrated by (19) compared to (20).[6]

(19) ?? Ɂayǝnay nǝnayħidħɨdom$_i$ siɁlitat   fǝty-om   [selam-n
      which   each.other   pictures like.GER-3PL.SBJ Selam-and
      yared-n]$_i$?
      Yared-and
      'Which pictures of each-other did Selam & Yared like?' Wh-movement

(20) Ɂayǝnay nǝnayħidħɨdom$_i$ siɁlitat [selam-n   yared-n]$_i$
      which   each.other   pictures Selam-and Yared-and
      fǝty-om?
      like.GER-3MPL.SBJ
      'Which pictures of each-other did Selam and Yared like?'   Peripheral-wh

Thus, if the presence/absence of reconstruction effects (under Binding Condition A) is indicative of movement, then the above examples illustrate differences among the three types of wh-interrogative strategies in Tigrinya.

2.3 Weak crossover

Weak crossover (WCO) refers to the condition where A'-moved constituents cannot dislocate across c-commanding pronouns that they end up binding (Chomsky 1977, Wasow 1979, Chomsky 1981; see also Safir 2017 for a recent overview and extensive discussion). In the literature, it has been assumed that A'-movement is subject to WCO effects because a wh-phrase cannot cross a c-commanding constituent that embeds a co-indexed pronoun (Chomsky 1977, Wasow 1979, Postal 1993, a.o.).

---

[6]Here we are using three degrees of unacceptability judgements offered by the consultants: single question mark (?) refers to sentences that are judged as less/mildly deviant by some speakers, the two question marks (??) for highly deviant, and the asterisk (*) for out-right ill-formed construction for the majority of the speakers.

For example, in English, the movement of wh-phrases induces WCO effects as illustrated in (21). (21) is bad because the trace of the wh-phrase in the object position does not c-command the pronoun that binds it inside the subject.

(21) ?? Who$_i$ does [his$_i$ mother] like t_i?

Similarly, Huang (1982) shows that in-situ wh-expressions in Chinese also trigger WCO effects, despite remaining in an in-situ position (22).

(22) * [Ta$_i$-de mama] xihuan shei$_i$?
       s/he-DE mother like   who
       *Whom$_i$ does his/her$_i$ mother like t_i?

In Tigrinya, some WCO effects are observed among the different wh-interrogative strategies. First notice that Tigrinya possessive pronouns appear as a suffix attached on the possessee (e.g., *ʔaddiʔ-u*/mother-his/'his mother'). In (23), the possessive pronoun -*a* 'her' attaches to the noun *ʔaddə* 'mother' and gets its relevant interpretation by co-indexing with the proper name (i.e., an R-expression) *Selam*, which it does not c-command.

(23) ʔaddiʔ-a$_i$ nɨ-selam$_i$ tɨ-fət-u?
       mother-her ACC-Selam 3-like.IPFV-3FSG.SBJ
       'Her$_i$ mother likes Selam$_i$.'

Thus, using (23) as a base and the assumption that (possessive) pronouns can be dependent on wh-phrases when certain specific structural conditions are met, we test whether the possessive pronoun can be a variable bound by the wh-phrase, i.e., co-indexed with the object *nɨmən* (respecting/avoiding some Binding Principles). If the wh-dependency gives rise to WCO effects, then one can conclude that such dependency is derived by movement; by contrast, if the wh-dependency does not exhibit any WCO effects, presumably movement is not involved.

The Tigrinya wh-movement strategy gives rise to WCO effects, as the deviant example in (24) illustrates.

(24) ?? nɨ-mən$_i$ tɨ-fət-u   ʔaddiʔ-a$_i$?
       ACC-who 3-like.IPFV-3FSG.SBJ mother-her
       'Who$_i$ does [her$_i$ mother] like?'   Wh-movement

WCO effects are also observed in the peripheral-wh construction as the ungrammatical construction in (25) illustrates.

(25)   * nɨ-mən$_i$ ʔaddiʔ-a$_i$ tɨ-fət-u?
       ACC-who mother-her 3-like.IPFV-3FSG.SBJ
       'Who$_i$ does [her$_i$ mother] like?'   Peripheral-wh

Finally, WCO effects are attenuated in in-situ wh-interrogatives as the less-deviant construction in (26) demonstrates.

(26)   ? ʔaddiʔ-a$_i$ nɨ-mən$_i$ tɨ-fət-u?
       mother-her ACC-who 3-like.IPFV-3FSG.SBJ
       'Who$_i$ does [her$_i$ mother] like?'   Wh-in-situ

Thus, the three strategies appear fairly different with respect to WCO: the contrast between (23) and (24-26) is associated to the WCO condition (i.e., a variable cannot be the antecedent of a pronoun or an anaphor that it does not c-command (cf. Reinhart & Reuland 1993). While the ungrammaticality with the wh-in-situ strategy is less severe than the peripheral-wh, the wh-movement strategy is worse than the peripheral-wh. Given that WCO is not a unified phenomenon, the results are not unexpected. Nevertheless, the difference, which essentially lies in degree of marginality or deviance of the questions) is mild and could be due to other factors, not necessarily tied to movement.[7]

2.4 Strong islands

As least since Ross (1967), syntactic islands are considered as standard diagnostics for identifying the presence/absence of wh-movement. In layman terms, islands are a form of "blockade" for certain constituents to move out of certain syntactic configurations. Over the years, syntactic islands— both strong and weak — have been refined to show distinct properties (see Szabolcsi & den Dikken 2003 for an overview), and in what follows we discuss both in distinguishing the three wh-strategies in Tigrinya.

Wh-movement cannot target wh-phrases embedded in "strong" (absolute) islands, such as adjuncts, relative clauses (RCs) modifying an NP, sentential subjects, or coordination (see den Dikken 2018 and references cited therein for a recent discussion). For example, the English sentences given in (27) illustrate how the extraction of wh-phrases from adjunct islands, Complex NP islands and subject islands renders the sentences ungrammatical (see Boeckx 2008: 155, Huang 1982: 497 for more examples).

---

[7]A reviewer wonders whether the three strategies exhibit Strong crossover (SCO) effects; while space precludes us from presenting the examples here, we report that preliminary results show that the wh-movement strategy induces SCO violation, whereas the other two strategies do not.

(27) a. *Which boy$_i$ did Mary laugh [after Sarah kissed t_i]?

　　b. *What$_i$ does John like [the woman who wears t_i]?

　　c. *Who$_i$ do you think [pictures of t_i] would please John?

In contrast, in typical wh-in-situ languages in-situ wh-phrases can be embedded inside syntactic islands.[8] For instance, Huang (1982: 496-8) shows that in Chinese in-situ wh-phrases can surface inside islands while taking matrix scope. The example in (28) illustrates this with a wh-phrase within an RC (from Bayer & Cheng 2017: 5).

(28) Mandarin Chinese

Bótōng xǐhuān shéi xǐe　　de shū?

Botong like　　who write DE book

'For which x, x a person such that Botong likes the book that x wrote?'

There is nonetheless an argument-adjunct asymmetry, whereby unlike in-situ argument wh-phrases, in-situ adjunct wh-phrases display the typical island effects of A'-movement (Huang 1982: 525-527). Compare (29) with (28) (from Bayer & Cheng 2017: 5).

(29) Mandarin Chinese

* Qiáofēng xǐhuān Bótōng wèishénme xǐe　　de shū?

　Qiaofeng like　　Botong why　　　write DE book

Intended: 'For what reason x, such that Qiaofeng like the book that Botong wrote for x?'

Now turning back to Tigrinya, we begin, once again, with the baseline constructions given in (30)-(31). (30) is a model for an adjunct island: the clause embeds a finite adjunct clause, headed by the complementizer *siləzi-* 'since', prefixed onto the verb. (31) is an example of a complex NP, where the object DP is modified by an RC; that is, the noun *məs'ħaf* 'book' is modified by a finite RC, *məħazaʔ-a zi-s'əħaf-o* '(that) her friend wrote'. Finally, (32) and (33) are examples of subject islands represented with a complex subject, both NP and sentential subject, respectively.

---

[8]While this is not an isolated case of Chinese (see, for example, Sabel & Zeller 2006 on Zulu, Abels & Muriungi 2008 on Kiitharaka, Bayer & Cheng 2017 on Bangla, Ko 2005 on Japanese and Korean, a.o.), there are, as an anonymous reviewer points out, some African languages where in-situ wh-phrases inside islands are blocked (see Amaechi & Georgi 2020 on Igbo, Zentz 2016 on Shona, and Torrence & Kandybowicz 2015 on Krachi where wh-in-situ are barred from islands). Notice, however, that many of these African languages involve optional wh-movement unlike typical wh-in-situ languages.

(30) selam [yared məħaza siləzɨ-rəxəb-ə]   təħagʷis-a
      Selam Yared friend   since-find.PFV-3MSG.SBJ be.happy.GER-3FSG.SBJ
      'Selam got excited because Yared found/met a friend.'

(31) selam [məħazaʔ-a zɨ-s'əħaf-o   məs'ħaf] gəziʔ-a
      Selam friend-her REL-write.PFV-3MSG.OBJ book   buy.GER-3FSG.SBJ
      'Selam bought a book her friend wrote.'

(32) a. nay ʔaddiʔ-u   siʔli   nɨ-Yared yə-ħɨgʷis-u-wwo
         NAY mother-his picture ACC-Yared 3-please.GER-3MSG.SBJ-3MSG.OBJ

      b. [sɨʔli   ʔaddiʔ-u] nɨ-Yared yə-ħɨgʷis-u-wwo
         picture mother-his ACC-Yared 3-please.GER-3MSG.SBJ-3MSG.OBJ
         'A picture of his mother pleased Yared'

(33) [yared fərəs mɨ-gzɨʔ-u]   nɨ-selam
      Yared horse NMLZ-buy-3MSG.SBJ ACC-Selam
      ʔəgrim-u-ll-a
      surprise.GER-3MSG.SBJ-APPL-3FSG.OBJ
      'That Yared bought a horse surprised Selam.'

Based on the above background and baseline sentences, we show that the three
Tigrinya strategies are indeed different, as they exhibit different sensitivities for
different islands.

2.4.1 Wh-movement

In Tigrinya, the "wh-movement" strategy is sensitive to island effects, and that
reinforces our initial hypothesis that this strategy, as in other wh-movement lan-
guages, indeed involves movement of the wh-phrase. Evidence comes from the
fact that movement of the wh-phrase out of a strong island (e.g., adjunct, complex
NP or subject island) accompanied by the matrix verb results in ungrammatical
sentences as illustrated in (34-36). This happens irrespective of whether the em-
bedded verb is itself fronted or not. For example, (34) is a typical example of an
adjunct island (marked with brackets for convenience). In this case, movement
of the wh-phrase *nɨmən* 'who(m)' from this strong island position to the left edge
of the matrix clause yields an ungrammatical construction.

(34) Adjunct island

   a. *nɨ-mən   təħagʷis-a   selam [sɨləzɨ-rəxəb-ə
       ACC-who be.happy.GER-3FSG.SBJ Selam since-find.PFV-3MSG.SBJ
       yared]?
       Yared
       'Who did Selam get excited because Yared found?'

   b. *nɨ-mən   təħagʷis-a   selam [yared
       ACC-who be.happy.GER-3FSG.SBJ Selam Yared
       sɨləzɨ-rəxəb-ə]?
       since-find.PFV-3MSG.SBJ
       'Who did Selam get excited because Yared found?'

In the same vein, a bona-fide Complex NP island, given in (35), presents the same result. The sentence is ungrammatical because the wh-phrase *mən* 'who', which originates as part of the RC 'the book that someone wrote', dislocated to the left-edge of the sentence, violates the complex NP constraint (Ross 1967, Bošković 2015). The ungrammaticality is not due to the *that*-trace effect.

(35) Complex NP island

   a. *mən gəziʔ-a   selam [zɨ-s'əħaf-o   məs'ħaf]?
       who buy.GER-3FSG.SBJ Selam REL-write.PFV-3MSG.OBJ book
       '*Who did Selam buy a book that wrote?'

   b. *mən gəziʔ-a   selam [məs'ħaf zɨ-s'əħaf-o]?
       who buy.GER-3FSG.SBJ Selam book   REL-write.PFV-3MSG.OBJ
       '*Who did Selam buy a book that wrote?'

Finally, (36) is an example of a subject island, in which movement out of a complex NP or CP subject — two strong islands — renders the sentences un-grammatical, as expected. Note that the NP subject island has parallel structure with that of possessive constructions in Tigrinya. Tigrinya has two possessive strategies: (i) *nay*-marked with possessor possessee order and (ii) bare (non-*nay*-marked) with possessee possessor order (Gebregziabher 2013). Now compare the base-line examples in (32) with their derivatives in (36a) and (36b). Although both strategies are not always available with relational nouns, in this context, the fact that *picture*-nouns permit multiple complements make the comparison possible, causing the difference in ungrammaticality to arise: the *nay*-marked (a) appears less severe than the non-*nay*-marked.[9]

---

[9]This is consistent with the cross-linguistic view that languages show variation in terms of

(36)   Subject island

    a. ?? nay mən yə-ħugʷis-u-wwo   sɨʔli (nɨ-)Yared?

        NAY who 3-please.GER-3MSG.SBJ-3MSG.OBJ picture ACC-Yared

        'Who did a picture of please Yared?'

    b. * mən yə-ħugʷis-u-wwo   sɨʔli (nɨ-)Yared?

        who 3-please.GER-3MSG.SBJ-3MSG.OBJ picture ACC-Yared

        'Who did a picture of please Yared?'

We also tested sentential subjects because it is natural that sentential subjects yield strong island environments (as Ross 1967 originally observed). The wh-movement strategy shows sentential subject island effects as well (37).

(37)   Sentential subject island

    a. ?? ʔɨntay ʔəgrim-u-ll-a   [yared

        what   surprise.GER-3MSG.SBJ-APPL-3FSG.OBJ Yared

        mɨ-gzɨʔ-u]   nɨ-selam?

        NMLZ-buy-3MSG.SBJ ACC-Selam

        'What that Yared bought surprised Selam?'

    b. * ʔɨntay ʔəgrim-u-ll-a   [mɨ-gzɨʔ-u

        what   surprise.GER-3MSG.SBJ-APPL-3FSG.OBJ NMLZ-buy-3MSG.SBJ

        yared] nɨ-selam?

        Yared ACC-Selam

        'What that Yared bought surprised Selam?'

2.4.2 Peripheral-wh

A striking contrast emerges when we compare wh-movement with the peripheral-wh strategy. First, peripheral-wh, unlike the wh-movement strategy, displays no adjunct island effects (38). This again happens irrespective of whether the verb inverts around the subject or not within the embedded clause.

---

    which complement (e.g., a possessor/theme/agent) they permit to extract out of the complex noun phrase (see Gavruseva 2000 for a cross-linguistic proposal; see also Alexiadou et al. 2007 for an extensive overview).

(38) Adjunct island

    a. *nɨ-mən selam [yared sɨləzɨ-rəxəb-ə]
        ACC-who Selam Yared since-find.PFV-3MSG.SBJ
        təħagʷis-a?
        be.happy.GER-3FSG.SBJ
        'Who did Selam get excited because Yared found?'

    b. *nɨ-mən selam [sɨləzɨ-rəxəb-ə   yared]
        ACC-who Selam since-find.PFV-3MSG.SBJ Yared
        təħagʷis-a?
        be.happy.GER-3FSG.SBJ
        'Who did Selam get excited because Yared found?'

Second, while the wh-movement strategy renders the construction ungrammatical irrespective of the inversion of the verb inside the embedded clause (35), this is not the case with the peripheral-wh (39): No complex NP island is observed with the peripheral-wh only when subject-verb inversion does not take place inside the embedded clause (39b).

(39) Complex NP island

    a. ? ʔɨntay selam [zɨ-s'əħaf-ə   səbʔay] rəxib-a?
        what Selam REL-write.PFV-3MSG.SBJ man   meet.GER-3FSG.SBJ
        'What did Selam meet a man wrote?'

    b. * ʔɨntay selam [səbʔay zɨ-s'əħaf-ə]   rəxib-a?
        what Selam man   REL-write.PFV-3MSG.SBJ meet.GER-3FSG.SBJ
        'What did Selam meet a man wrote?'

Finally, unlike the wh-movement strategy, the peripheral-wh strategy exhibits no subject island effects as the examples in (40a)-(41) show. Note, however, that a difference in ungrammaticality arises with the two nominal structures w.r.t. subject islands: the non-*nay*-marked does not seem to exhibit subject island effects. Compare (40a) with (40b).[10]

---

[10]Recall that possessive constructions in Tigrinya involve two independent strategies: The *nay*-marked is largely reserved for alienable possession, whereas the non-*nay*-marked is used for inalienable ones (see Gebregziabher (2012, 2013) for discussion). Both strategies are possible here because different thematic relationships, namely a possessor, a theme, or an agent, can be established with the head noun *picture*, and in many languages extraction of the possessor is only allowed Alexiadou et al. (2007).

(40)   Subject island

    a. [nay mən siʔli]   (nɨ-)Yared yə-ħəgʷis-u-wwo?
       NAY who picture ACC-Yared 3-please.GER-3MSG.SBJ-3MSG.OBJ

    b. *mən siʔli   (nɨ-)Yared yə-ħəgʷis-u-wwo?
       who picture ACC-Yared 3-please.GER-3MSG.SBJ-3MSG.OBJ
       'Who did a picture of please Yared?'

(41)   Sentential subject island
?ɨntay [yared mɨ-gzɨʔ-u]   nɨ-selam
what Yared NMLZ-buy-3MSG.SBJ ACC-Selam
?əgrim-u-ll-a?
surprise.GER-3MSG.SBJ-APPL-3FSG.OBJ
'What did that Yared bought surprise Selam?'

2.4.3 Wh-in-situ

On the other hand, with the standard wh-in-situ, where the wh-phrase remains in its base-generated position, no strong island effects are observed just as in the case of typical wh-in-situ languages, as the examples in (42)-(45) illustrate.

(42)   Adjunct island
selam [yared nɨ-mən   sɨləzi-rəxəb-ə]   təħagʷis-a?
Selam Yared ACC-who since-find.PFV-3MSG.SBJ be.happy.GER-3FSG.SBJ
Lit. 'Selam got excited because Yared found/met who?'

(43)   Complex NP island
selam [mən zɨ-s'əħaf-o   məs'ħaf] gəziʔ-a?
Selam who REL-write.PFV-3MSG.OBJ book   buy.GER-3FSG.SBJ
Lit. 'Selam bought a book who wrote?'

(44)   Subject island

    a. nay mən siʔli   nɨ-Yared   yə-ħəgʷus-o?
       NAY who picture ACC-Yared 3-please.IPFV-3MSG.SBJ

    b. [siʔli   mən]   nɨ-Yared yə-ħəgʷus-o
       picture mother-his ACC-Yared 3-please.IPFV-3MSG.SBJ-3MSG.OBJ
       Lit. 'A picture of who pleases Yared?'

(45) Sentential subject island
[yared ʔɨntay mɨ-gzɨʔ-u]   nɨ-selam
Yared what   NMLZ-buy-3MSG.SBJ ACC-Selam
ʔəgrim-u-ll-a?
surprise.GER-3MSG.SBJ-APPL-3FSG.OBJ
Lit. 'That Yared bought what surprised Selam?'

The fact that the above constructions are not sensitive to strong islands reinforces our initial analysis that this wh-interrogative strategy is indeed different from the other two.

With regard to the argument-adjunct asymmetry, Tigrinya shows a mixed picture.[11] The asymmetry holds with adjunct and (sentential) subject islands (compare (42) with (46), and (45) with (47)), but not with complex NP islands (compare (43) with (48)).[12]

From (46) and (47), we observe that extraction from adjunct islands and CP subjects is illicit, but no such restriction is observed with Complex NP islands, provided that there is no subject-verb inversion, as the examples in (48) illustrate.

(46) Adjunct island
   a. ?? selam [yared məʕas sɨləzɨ-rəxəb-o]   təħag*w*is-a?
      Selam Yared when with-find.PFV-3MSG.OBJ be.happy.GER-3FSG.SBJ
      Lit. 'Selam is happy because Yared met/found him when?'

   b. ?? selam [məʕas sɨləzɨ-rəxəb-o   yared]
      Selam why   when-find.PFV-3MSG.SBJ Yared
      təħag*w*is-a?
      be.happy.GER-3FSG.SBJ
      Lit. 'Selam is happy because Yared met/found him when?'

---

[11]A reviewer asks whether adjuncts like 'why' exhibit a different pattern in Tigrinya as it does in other languages. Recall Tigrinya has two 'why' forms: while the form *nɨmɨntay* can be tolerated in-situ, the form *sɨləmɨntay* cannot (see Irurtzun 2021 for a comprehensive typological overview).

[12]In many languages, islands including sentential islands, can be ameliorated by using resumptive pronouns –pronouns that function as variables bound by an operator in an A'-position (see Rouveret 2011 for a comprehensive overview). Some of the classical examples come from Semitic languages (see Borer 1984 on Hebrew, Aoun et al. 2010 on Arabic, a.o.). In Tigrinya, the obligatory use of OMs is generally dependent on the specificity/definiteness of the noun phrase (Gebregziabher 2019), and more importantly, the presence of OMs inside islands does not remedy the construction.

(47)   Sentential subject island

   a.   ?? [yared fərəs məʕas mɨ-gziʔ-u]   nɨ-selam
        Yared   horse when   NMLZ-buy-3MSG.SBJ ACC-Selam
        ʔəgrim-u-ll-a?
        surprise.GER-3MSG.SBJ-APPL-3FSG.OBJ
        Lit. 'That he bought a horse when surprised Selam?'

   b.   ?? [fərəs məʕas mɨ-gziʔ-u   yared] nɨ-selam
        horse when   NMLZ-buy-3MSG.SBJ Yared   ACC-Selam
        ʔəgrim-u-ll-a?
        surprise.GER-3MSG.SBJ-APPL-3FSG.OBJ
        Lit. 'That Yared bought a horse when surprised Selam?'

(48)   Complex NP island

   a.   selam [yared məʕas zɨ-s'əħaf-o   məs'ħaf]
        Selam Yared   when   REL-write.PFV-3MSG.OBJ book
        gəziʔ-a?
        buy.GER-3FSG.SBJ

   b.   ?? selam [məʕas zɨ-s'əħaf-o   məs'ħaf yared]
        Selam when   REL-write.PFV-3MSG.OBJ book   Yared
        gəziʔ-a?
        buy.GER-3FSG.SBJ
        Lit. 'Selam bought a book that Yared wrote when?'

   Thus, by the strong island account, there is a distinction among the three wh-strategies: the wh-movement strategy exhibits island effects across all strong islands, but the peripheral-wh does so only with (sentential) subject and complex NP islands. With the wh-in-situ strategy, however, no island effects are observed except an argument-adjunct asymmetry with adjunct and subject islands.[13]

---

[13]In addition, Tigrinya exhibits some variation between adverbial wh-phrases, such as 'when/where', and purpose/reason wh-phrases, namely 'how/why', in the context of a sentential subject island. (i) illustrates that the island sensitivity fares better with the latter than the former.

   (i)   a.   ?? [fərəs sɨləmɨntay/kəməy mɨ-gziʔ-u]   nɨ-selam
              horse why/how   NMLZ-buy-3MSG.SBJ ACC-Selam
              ʔəgrim-u-ll-a?
              surprise.GER-3MSG.SBJ-Appl-3FSG.OBJ
              (Lit. 'That he bought a horse how/why surprised Selam?')

2.5 Weak islands: Wh-islands

In contrast to strong islands, "weak" islands are considered to be "selective" when it comes to island violations. Wh-islands are considered to be weak because some wh-phrases can extract out of certain types of islands but others cannot (see a.o. Rizzi 1990, Lasnik & Saito 1994, Szabolcsi 2006 for discussion).

(49) a. ??What do you wonder why John bought?

    b. *Why/Where do you wonder what John bought?

The same contrast is observed in a wh-in-situ language, such as Chinese (Huang 1982). Examples from Bayer & Cheng (2017).

(50) Mandarin Chinese
Júdòu xiǎng-zhīdào shéi mǎi-le   shéme (?)
Judou want-know   who buy-ASP what
a. 'Judou wonders who bought what.'
b. 'For which y, y a thing, Judou wonders who bought y?'
c. 'For which x, x a person, Judou wonders what x bought?'

(51) Mandarin Chinese
Húfēi xiǎng-zhīdào shéi wèishéme shēngqì(?)
Hufei want-know   who why   get.angry
a. 'Hufei wonders who gets angry why.'
b. 'For which x, x a person, Hufei wonders why x gets angry?'
c. Intended but unavailable: 'What is the reason x, Hufei wonders who gets angry for x?'

Turning back to Tigrinya, we examine whether weak islands have some bearing on the distinction among the different wh-interrogatives. Here we use the example in (52) as a baseline to test the effects of wh-islands. Notice that the complement of the matrix verb 'wonder' is an interrogative clause with adjuncts

---

    b. ? [fərəs məʕas/ʔabəy mɨ-gzɨʔ-u]   ni-selam
        horse where/when NMLZ-buy-3MSG.SBJ ACC-Selam
        ʔəgrim-u-ll-a?
        surprise.GER-3MSG.SBJ-Appl-3FSG.OBJ
        Lit. 'That Yared bought a horse when/where surprised Selam?'

    Thus, there is a mild island effect with 'when/where' (compared to 'why/how') in Tigrinya. It appears, unlike in many other languages (see Irurtzun 2021 for a comprehensive list of examples), some interpretations of the in-situ 'why/how' are also not available in this context in Tigrinya.

ʔabʕɨdaga 'at market' and *nɨmǝmǝrǝk'i* 'for graduation' in their base-generated position.

(52) yared [selam wɨhbto ʔab-ʕɨdaga ni-mǝmǝrǝk'i   ʔɨntǝ-gǝziʔ-a]
      Yared Selam gift   at-market for-graduation COMP-buy.GER-3FSG.SBJ
      yǝ-sǝllasɨl
      3-wonder.IPFV
      'Yared wonders [whether Selam bought a gift for graduation at a market]'

The first observation is that, with the wh-movement strategy, a direct object wh-phrase, such as *ʔintay* 'what', cannot be moved out of a wh-island or an embedded interrogative as the ungrammaticality of the example in (53) illustrates.

(53) * ʔintay yǝ-sǝllasɨl   [selam ʔabǝy/mǝʕas ʔɨntǝ-gǝziʔ-a]?
        what   3-wonder.IPFV Selam where/when COMP-buy.GER-3FSG.SBJ
        'What does he wonder [where/when Selam bought]?'

Second, wh-adjuncts, such as *ʔabǝy* 'where' and *mǝʕas* 'when', compared to *sɨlǝmɨntay* 'why' and *kǝmǝy* 'how', give rise to a mild wh-island effect (comparable to the English examples in (49)).

(54) ?? ʔabǝy/mǝʕas yǝ-sǝllasɨl   [selam ʔintay ʔɨntǝ-gǝziʔ-a]?
        where/when 3-wonder.IGER Selam what   COMP-buy.GER-3FSG.SBJ
        'Where/when does he wonder [what Selam bought?]'

(55) * sɨlǝmɨntay/kǝmǝy yǝ-sǝllasɨl   [selam ʔintay
        why/how   3-wonder.IPFV Selam what
        ʔɨntǝ-gǝziʔ-a]?
        COMP-buy.GER-3FSG.SBJ
        'Why/how does he wonder [what Selam bought?]'

With the peripheral-wh strategy, movement of a wh-argument out of a wh-island generally has no effect (56), whereas dislocation of a wh-adjunct gives a mild weak island effect (this is especially true when we compare 'where' and 'when' (57) with 'why' (58)).

(56) ʔintay [selam ʔabǝy/mǝʕas ʔɨntǝ-gǝziʔ-a]   yǝ-sǝllasɨl?
      what   Selam where/when COMP-buy.GER-3FSG.SBJ 3-wonder.IPFV
      'What does he wonder [where/when Selam bought?]'

(57) ? ʔabəy/məʕas [selam ʔintay ʔintə-gəziʔ-a]   yə-səllasil?
      where/when Selam what   COMP-buy.GER-3FSG.SBJ 3-wonder.IPFV
      'What does he wonder [where/when Selam bought?]'

(58) ?? siləmintay/kəməy [selam ʔintay ʔintə-gəziʔ-a]
      why/how   Selam what   COMP-buy.GER-3FSG.SBJ
      yə-səllasil?
      3-wonder.IPFV
      'Why/how does he wonder [what Selam bought?]'

Finally, with the wh-in-situ strategy, while in-situ arguments do not give rise to weak island effects, in-situ wh-adjuncts show a mild effect, particularly with *siləmintay* 'why' or *kəməy* 'how' (59), but not with *məʕas* 'when' or *ʔabəy* 'where' (59).

(59) [selam ʔabəy/məʕas ʔintay ʔintə-gəziʔ-a] yə-səllasil?
      Selam where/when what   3-wonder.IPFV COMP-buy.IPFV-3FSG.SBJ
      'Where/when does he wonder [what Selam bought?]'

(60) ? [selam siləmintay/kəməy ʔintay ʔintə-gəziʔ-a]
      Selam why/how   what   COMP-buy.GER-3FSG.SBJ
      yə-səllasil?
      3-wonder.IPFV
      'Why/how does he wonder [what Selam bought?]'

Thus, by the count of weak wh-islands, it appears that in Tigrinya the three wh-interrogatives seem to contrast in exhibiting some minor argument-adjunct asymmetry, including a contrast between adjuncts like 'when/where' and 'how/why'.

2.6 Interim summary

Based on the above discussion, we can make the following observations: First, argument-adjunct asymmetries in both strong and weak islands are different from those observed in a typical wh-movement language, such as English, and in a wh-in-situ language, namely Chinese. For example, under the strong island constructions, wh-movement and peripheral-wh strategies do not show the asymmetry, but wh-in-situ does (save the CNP). On the other hand, under the weak islands, the adjunct-argument asymmetry appears to hold across the three strategies (often weaker or milder in some cases), but this is categorically different in

both English and Chinese (see Szabolcsi 2006 and references cited therein for an overview). In the literature, such asymmetries are often characterized in pragmatic or syntactic terms, such as nominals vs. adverbials (e.g., Huang 1982) or referential/individuals vs. non-referential/individuals (e.g., Cinque 1990), although many of the explanations still remain inadequate (see den Dikken 2018 for a recent overview).[14] Therefore, we suspect the nature of wh-expressions could be key in accounting for the observed contrasts.

Second, abstracting away from the argument-adjunct asymmetry issue, the construction we dubbed "wh-movement" shows the standard properties of typical wh-movement languages (namely, English). Similarly, the Tigrinya "wh-in-situ" construction shares most of its properties with bona-fide wh-in-situ languages (such as Chinese). One widely adopted approach is *unselective binding*, according to which in-situ adjunct wh-phrases that are bound by an operator in their scope position must raise at LF, thereby inducing island violations (a.o., Pesetsky 1987, 2000, Tsai 1994, 2008). The fact that the Tigrinya data exhibit a consistent pattern may lead one to claim that this approach is sufficient to license Tigrinya wh-in-situ arguments as well as wh-adjuncts discussed in this paper.

Our "peripheral-wh construction" contrasts with wh-movement in not systematically showing island effects. This suggests that even if a wh-dependency exists in these constructions, perhaps they may not be derived by A'-movement (presumably some pragmatic factors may be at play). Finally, as for the complex NP constraint (CNPC) it is not obvious how the pattern emerges. Recall, that peripheral-wh exhibits CNPC effects (but no adjunct or subject island effects), whereas wh-in-situ shows no CNPC effects (but does show adjunct and subject island effects). We suspect that this could be something to do with the syntax of relativization (cf. Cinque 2010 on some 'apparent' violations of the CNPC). We leave these issues open here until the CNPC facts in the contexts of NP + clausal complement is fully uncovered in Tigrinya. Table 1 summarizes the results obtained so far.[15] Note that with respect to Complex Noun phrases, extraction of the object is fine for peripheral-wh, but extraction of the subject is not.

---

[14]The fact that Tigrinya does not exhibit strong adjunct-argument asymmetry could be something to do with the nature of the wh-items themselves rather than their role (as the argument-adjunct) per se. In some wh-in-situ languages wh-items act like variables, whereas in wh-movement languages they generally act like pronominals (see Rizzi 1990, Szabolcsi & den Dikken 2003 for discussion). Thus, one way of interpreting the Tigrinya 'why/how' pattern is to say that they are reason/purpose wh-phrases, and that is what makes them less/unextractable from weak islands.

[15]In the table, "mild" refers to the judgement less grammatical signaled by ?/??.

Table 1: Diagnosing the syntactic properties of wh-interrogatives

		English	Chinese	Tigrinya		
		Wh-mvt	Wh-in-situ	Wh-mvt	Periph-wh	Wh-in-situ
Long-distance dependency		yes	yes	yes	yes	yes
Reconstruction effect		yes	NA	yes	yes	NA
WCO effect		yes	yes	yes	yes	(mild)
Island effect /argument	Adjunct	yes	no	yes	no	no
	CNP	yes	no	yes	no	no
	Subject	yes	no	yes	no	no
Island effect /adjunct	Adjunct	yes	yes	-	-	yes
	CNP	yes	yes	-	-	no
	Subject	yes	yes	-	-	yes
Wh-island effect /argument		no	no	yes	(mild)	no
Wh-island effect /adjuncts		yes	yes	(mild)	no	(no/ mild)

In what follows, we explore whether there are some semantic restrictions in the different syntactic wh-interrogatives we uncovered so far.

3 Interpretive properties

In our effort to characterize and delimit the patterns of wh-interrogative strategies in Tigrinya, we explore whether there are interpretive properties associated with the different word order configurations. In this regard, we follow previous literature (see É. Kiss 2010, Horvath 2013 on Hungarian, Bayer & Cheng 2017 on Chinese, Duguine & Irurtzun 2014 on Basque, Faure & Palasis 2021 on French, a.o.) and discuss interpretive properties, namely, presuppositionality and exhaustivity readings that have been often used to distinguish standard wh-in-situ from wh-movement (as well as focus movement) questions. See Keupdjio (2020) for a recent adaptation of the same idea on Medumba (Grassfields Bantu, spoken in Cameroon) to separate wh-in-situ, wh-movement and focus movement questions.

3.1 Exhaustivity

Exhaustivity entails the maximum set of entities or individuals the predicate satisfies in a given question. In the case of a wh-interrogative, exhaustivity often entails exclusivity or the exclusion of certain alternative answers in a given context (see Faure & Palasis 2021 for a recent discussion based on colloquial French). One of the ways in which exhaustivity is expressed is by using additive particles such as *else, other, also*, etc., and when such particles are associated with wh-phrases, as in (*who else, what else*, etc.), they presuppose a non-exhaustive or non-exclusive list of answers. In other words, such elements are incompatible in an exhaustive question (cf. É. Kiss 2010, Keupdjio 2020). For example, in Chinese, an answer with an additive particle *yě* 'also' cannot be felicitous for a contrastively focused wh-question (61a) that requires a unique answer because that can give rise to the exhaustivity reading that excludes a list answer (61b), whereas the same answer for the normal wh-in-situ question is felicitous because it does not result in exhaustivity reading that excludes a unique answer. See Pan (2019) for a recent discussion on Chinese.

(61) Mandarin Chinese (Cheung 2008: 54 cited in Pan 2014: 23)

  a. Q: (Shì) [shénme dōngxi]_{C-FOC} Mǎlì mǎi-le?
      be   what   thing   Mary buy-Perf
      'What thing(s) was it that Mary bought?'

  b. A: *Shì [màozi]_{C-FOC}, tā mǎi-le. Shì [wàitào]_{C-FOC}, tā yě
      be   hat   she buy-Perf be coat   she also
      mǎi-le.
      buy-Perf
        'It was a hat that she bought, and it was a coat that she also bought.'

In Tigrinya, there is an additive particle *kalɨʔ* 'else/other' that marks exhaustivity, and when wh-questions are associated with this additive particle, some differences among the three wh-interrogative strategies arise. First, notice that the additive particle can appear either to the right or left of the wh-phrase (62).

(62) a. kalɨʔ mən məsʔiʔ-u?
          else who come.GER-3MSG.SBJ

  b. mən kalɨʔ məsʔiʔ-u?
      who else   come.GER-3MSG.SBJ
      'Who else came?'

Second, when the additive particle precedes the wh-phrase, the wh-movement strategy appears somehow deviant for some speakers; however, no observable difference between the two other strategies is exhibited. Both appear felicitous (63b-c).

(63) a. Wh-movement
          ? kalɨʔ nɨmən sədid-a   selam dəbdabe?
            else who send.GER-3FSG.SBJ selam letter
      b. Peripheral-wh
          kalɨʔ nɨmən selam dəbdabe sədid-a?
          else who Selam letter send.GER-3FSG.SBJ
      c. Wh-in-situ
          selam dəbdabe kalɨʔ nɨmən sədid-a?
          Selam letter   else who send.GER-3FSG.SBJ
          'Who else did Selam send a letter to?'

On the other hand, when the additive particle appears to the right of the wh-phrase, while the wh-in-situ is felicitous, the peripheral-wh and wh-movement appear deviant (64). In this respect, the additive particle as an exhaustivity marker appears to differentiate wh-in-situ from peripheral-wh and wh-movement constructions.

(64) a. Wh-movement
          ?? nɨmən kalɨʔ sədid-a   selam dəbdabe?
             who   else send.GER-3FSG.SBJ Selam letter
      b. Peripheral-wh
          ? nɨmən kalɨʔ selam dəbdabe sədid-a?
            who   else Selam letter send.GER-3FSG.SBJ
      c. Wh-in-situ
          selam dəbdabe nɨmən kalɨʔ sədid-a?
          Selam letter   who   else send.GER-3FSG.SBJ
          'Who else did Selam send a letter?'

In many respects, the deviance of the above constructions with the peripheral-wh and wh-movement constructions recalls what É. Kiss (1998) labels exhaustive identification (see also Horvath 2010 on Hungarian), according to which additive markers are incompatible with contrastive focus due to their semantic requirement for exhaustive list answers. In Tigrinya, contrastive focus is often expressed with clefts, and clefts in Tigrinya are introduced by a copula ʔɨyy- preceded by the focalized element, as illustrated below:

(65) a. yared ʔɨyy-u   məs'iʔ-u
          Yared be-3MSG.SBJ give.GER-3MSG.SBJ
          'It was Yared who came.'

      b. məkina ʔɨyy-u   gəziʔ-u
          car   be-3MSG.SBJ buy.GER-3MSG.SBJ
          'It was a car that he bought.'

In the above cleft constructions, with the emphasis on *Yared* and *a car*, the sentences assert that Yared came and he bought a car, but they also express that the only person who came is Yared and the only thing that he bought is a car. Thus, the list of answers, namely, a car and Yared, exhaustively identify the relevant entities that have the property of the individuals who came and were bought, respectively.

According to É. Kiss (1998), some lexical items such as *else, even, also, again*, etc. are incompatible with bona-fide contrastive focus constructions, such as clefts (e.g., **It was even/also/else a hat that John picked out for himself*), due to their semantic properties.

In Tigrinya, identificational *it*-clefts, similar to the wh-movement (and peripheral-wh) constructions, are incompatible with additive markers such as 'else/other'.

(66) ?? nɨ-mən kalɨʔ ʔɨyy-a   selam dəbdabe sədid-a?
         ACC-who other be-3FSG.SBJ Selam letter   send.GER-3FSG.SBJ
         'Who else is it that Selam sent a letter to?'   Cleft construction

Thus, by the exhaustivity account, peripheral-wh and wh-movement constructions seem to be different from wh-in-situ strategies in Tigrinya, and they appear to exhibit semantic similarity with wh-clefts. This also appears consistent with what Keupdjio (2020) observed in Bamileke Medumba, where ex-situ wh-questions are exhaustive, but their in-situ counterparts are not.

3.2 Presuppositionality

In languages such as French, wh-clefts are associated with an existential presupposition. In turn, wh-movement or wh-in-situ constructions are not necessarily associated with such a presupposition. This is shown by the infelicity in answering 'nothing' to the cleft interrogative and its felicity to a wh-in-situ question, as illustrated below (examples from Shlonsky 2012; see also Duguine & Irurtzun 2014 for similar observation based on Basque).

(67) Wh-cleft (French)

    a. Q: C'est quoi que tu fais dans la vie?
           it's what that you do in   the life
           'What is it that you do in life?'

    b. A: #Rien. 'Nothing'

(68) Wh-in-situ (French)

    a. Q: Tu fais quoi dans la vie?
           you do what in   the life
           'What do you do in life?'

    b. A: Rien. 'Nothing'

In Tigrinya, there is some variability in the acceptability of denial responses to different types of wh-questions: while denial responses to wh-in-situ and wh-movement questions are well-formed (69-70), denial responses to the peripheral-wh questions are not (71).

(69) Wh-movement

    a. Q: nɨ-mən   sədid-u   yared nɨ-ʕdaga
           ACC-who send.GER-3MSG.SBJ Yared to-market
           'Who did Yared send to the market?'

    b. A: nɨwalaħadə/nɨmanɨm
           no-one/nobody
           'No one/nobody'

(70) Wh-in-situ

    a. Q: yared nɨ-mən   nɨ-ʕdaga sədid-u?
           Yared ACC-who to-market send.GER-3MSG.SBJ
           'Who did Yared send to the market?'

    b. A: nɨ-walaħadə/nɨ-manɨm
           ACC-no.one/ACC-nobody
           'Nobody/no one'

(71) Peripheral-wh

    a. Q: nɨ-mən yared nɨ-ʕdaga sədid-u?
           ACC-who Yared to-market send.GER-3MSG.SBJ
           'Who did Yared send to the market?'

b. A: #nɨ-walaħadə/nɨ-manɨm?
   ACC-no.one/ACC-nobody
   'Nobody/no one'

Again, there is no contrast between wh-movement and wh-in-situ constructions. However, peripheral-wh constructions appear different from the other two because they pattern together with Tigrinya (and French) wh-clefts as they are associated with an existential presupposition.

(72) Cleft-Construction

a. Q: mən ʔɨyy-u   kəyd-u   nɨ-ʕɨdaga?
      who be-3MSG.SBJ go.GER-3MSG.SBJ to-market
      'Who is it that went to the market?'

b. A: #walaħadə/manɨm
      no-one/-body
      'Nobody/no one'

Thus, it appears that, once again, the peripheral-wh is different from the wh-in-situ and wh-movement in being presuppositional (parallel to a cleft-construction).

3.3 Summary of interpretive properties

In this section we discussed the interpretive properties of wh-interrogatives in Tigrinya. We found that while some Tigrinya wh-interrogatives seem to exhibit some interpretive differences with respect to exhaustivity and presuppositionality, others do not. Consistent with our original assumption, although the in-situ and wh-movement strategies do not significantly differ semantically, the peripheral-wh pattern differs from the other two strategies in terms of these semantic properties.

Table 2 summarizes both the syntactic and semantic properties of Tigrinya wh-interrogatives.

Table 2: Diagnosing the interpretive properties of wh-interrogatives

	English	Chinese	Tigrinya		
	Wh-mvt	Wh-in-situ	Wh-mvt	Wh-peri	Wh-in-situ
Exhaustivity	no	no	yes	yes	no
Presuppositionality	no	no	no	yes	no

We tentatively suggest that the wh-fronting of the peripheral-wh strategy type (different from the standard wh-movement type) can be derived from a different type of structure (maybe driven by a feature other than [wh]) or a base-generated structure of a different question-formation type, which would account for the exhaustive and existential inferences the construction has (see Duguine & Irurtzun 2010 on Basque, Horvath 2013 on Hungarian, Keupdjio 2020 on Medumba, Faure & Palasis 2021 on Colloquial French, among others, for a similar approach). In particular, the Tigrinya peripheral-wh strategy could be treated parallel to what Duguine & Irurtzun (2014) called a "reinforced" wh-question in Labourdin (Northern dialects) Basque. They termed it 'reinforced' because it has a marked focus associated (with presuppositionality and contrastive focus) to it. Faure & Palasis (2021) also show that ex-situ interrogatives in French with no inversion (V-to-C movement), such as *Où elle va?* (as opposed to *Où va-t-elle?*) 'where does she go?', are exhaustive/exclusive and should not be driven by the standard wh-feature checking mechanism (see Cheng & Rooryck 2000, Mathieu 2016 for a recent prosody-based account). While these are suggestive potential accounts for Tigrinya, the exact analysis has to remain open for further investigation.

4 Conclusion

In this paper, we set out to describe and analyze wh-interrogatives in Tigrinya. Using a number of syntactic and semantic properties, we have shown that Tigrinya wh-interrogatives involve (at least) three independent strategies: (i) wh-movement, (ii) peripheral-wh, and (iii) wh-in-situ. While the first involves movement of the wh-phrase accompanied by V-to-C movement, exhibiting what appears like a V2-effect, the others do not.

In many respects, Tigrinya can hardly be considered either a strict wh-movement or an in-situ language; thus, standard syntactic theories including parametric approaches, such as (i) feature strength (Chomsky 1995), (ii) phase-based (Chomsky 2001), (iii) clause typing (cf. Cheng 1991, Cheng & Rooryck 2000) or (iv) externalization (as in Distributed Morphology, Richards 2010), cannot straightforwardly account for the data in Tigrinya, for the simple reason that Tigrinya wh-interrogatives exhibit mixed properties, not only with respect to violations of (strong) islands but also reconstruction and WCO effects. It is beyond the scope of this paper to delve into the pros and cons of these analyses but it will be a fruitful avenue to explore.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions:

A	answer	GER	gerund(ive)	Q	question

Acknowledgements

References

Abels, Klaus & Peter Muriungi. 2008. The focus marker in Kîîtharaka: Syntax and semantics. *Lingua* 118(5). 687–731.

Alexiadou, Artemis, Liliane Haegeman & Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin & New York: De Gruyter Mouton.

Amaechi, Mary & Doreen Georgi. 2020. On optional wh-/focus fronting in Igbo: A SYN-SEM-PHON interaction. *Zeitschrift für Sprachwissenschaft* 39(3). 299–327.

Aoun, Joseph, Elabbas Benmamoun & Lina Choueiri. 2010. *Arabic syntax*. Cambridge: Cambridge University Press.

Bayer, Josef & Lisa Lai-Shen Cheng. 2017. Wh-in-situ. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Wiley Blackwell companion to syntax. Second edition*, 1–44. Oxford: John Wiley & Sons.

Boeckx, Cedric. 2008. Islands. *Language and Linguistics Compass* 2(1). 151–167.

Borer, Hagit. 1984. *Parametric syntax: Case studies in Semitic and Romance languages*. Boston: De Gruyter Mouton.

Bošković, Željko. 2015. Wh-movement. In Antonio Fábregas, Jaume Mateu & Michael T. Putnam (eds.), *Contemporary linguistic parameters*, 251–279. London: Bloomsbury Publishing.

Bruening, Benjamin & Eman Al Khalaf. 2019. No argument-adjunct asymmetry in reconstruction for Binding Condition C. *Journal of Linguistics* 55(2). 247–276.

Cheng, Lisa Lai-Shen. 1991. *On the typology of wh-questions*. MIT. (Doctoral dissertation).

Cheng, Lisa Lai-Shen & Johan Rooryck. 2000. Licensing wh-in-situ. *Syntax* 3(1). 1–19.

Cheung, Candice Chi-Hang. 2008. *Wh-fronting in Chinese*. University of Southern California. (Doctoral dissertation).

Chomsky, Noam. 1977. On wh-movement. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on binding and government*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Barriers* (Linguistic Inquiry Monographs 13). Cambridge, MA: The MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge: MIT Press.

Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, MA: The MIT Press.

Cinque, Guglielmo. 2010. On a selective "violation" of the complex NP constraint. In C. Jan-Wouter Zwart & Mark de Vries (eds.), *Structure preserved: Studies in syntax for Jan Koster* (Linguistik Aktuell/Linguistics Today 164), 81–89. Amsterdam: John Benjamins.

den Dikken, Marcel. 2018. *Dependency and directionality*. Cambridge: Cambridge University Press.

Duguine, Maia & Aritz Irurtzun. 2010. Opérateurs d'éxclusivité dans les questions et réponses en basque. *Revue de Sémantique et Pragmatique* 27. 105–117.

Duguine, Maia & Aritz Irurtzun. 2014. From obligatory WH-movement to optional WH-in-situ in Labourdin Basque. *Language* 90(1). 1–30.

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74(2). 245–273.

É. Kiss, Katalin. 2010. Structural focus and exhaustivity. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological and experimental perspectives*, 64–88. Oxford: Oxford University Press.

Faure, Richard & Katerina Palasis. 2021. Exclusivity! Wh-fronting is not optional wh-movement in Colloquial French. *Natural Language & Linguistic Theory* 39(1). 57–95.

Fox, Danny & Jon Nissenbaum. 2004. Condition A and scope reconstruction. *Linguistic Inquiry* 35(3). 475–485.

Gavruseva, Elena. 2000. On the syntax of possessor extraction. *Lingua* 110(10). 743–772.

Gebregziabher, Keffyalew. 2012. The syntax of alienable and inalienable possession in Tigrinya. *Journal of Afroasiatic Languages* 5(1). 153–194.

Gebregziabher, Keffyalew. 2013. *Projecting possessors: A morphosyntactic investigation of nominal possession in Tigrinya*. University of Calgary. (Doctoral dissertation).

Gebregziabher, Keffyalew. 2019. Clitics or agreement markers: A view from Tigrinya clausal possession and modal necessity. In Akinbiyi, Akinlabi, Lee Bickmore, Michael Cahill, Michael Diercks, Laura Downing, James Essegbey, Katie Franich, Laura McPherson and and Sharon Rose (ed.), *Celebrating 50 years of ACAL: Selected papers from the 50th Annual Conference on African Linguistics*, 89–136. Berlin: Language Science Press.

Gebregziabher, Keffyalew. 2023. *Clausal structures and v2: A Tigrinya case study*. Unpublished Ms. IKER (Centre de Recherche sur la langue et les textes basques).

Green, Melanie & Philip J. Jaggar. 2003. Ex-situ and in-situ focus in Hausa. In Jacqueline Lecarme (ed.), *Research in Afroasiatic Grammar II: Selected papers from the Fifth Conference on Afroasiatic Languages, Paris, 2000* (Current Issues in Linguistic Theory), 187–214. Amsterdam: John Benjamins Publishing Company.

Hetzron, Robert. 1972. *Ethiopian Semitic: Studies in classification*. Manchester: Manchester University Press.

Horvath, Julia. 2010. Discourse-features, syntactic displacement and the status of contrast. *Lingua* 120(6). 1346–1369.

Horvath, Julia. 2013. Focus, exhaustivity and the syntax of wh-interrogatives. In Johan Brandtler, Valéria Molnár & Christer Platzack (eds.), *Approaches to Hungarian*, vol. 13, 97–132. Amsterdam: John Benjamins.

Huang, C-T. James. 1982. Move wh in a language without wh-movement. *The Linguistic Review* 1(4). 369–416.

Irurtzun, Aritz. 2021. Why questions break the residual V2 restriction (in Basque and beyond). In Gabriela Soare (ed.), *Why is 'why' unique?: Its syntactic and semantic properties*, 63–114. Boston: De Gruyter Mouton.

Keupdjio, Hermann. 2020. *The syntax of A'-dependencies in Bamileke Medumba*. University of British Colombia. (Doctoral dissertation).

Ko, Heejeong. 2005. Syntax of why-in-situ: Merge into [Spec, CP] in the overt syntax. *Natural Language & Linguistic Theory* 23(4). 867–916.

Kogan, Leonid E. 1997. Tigrinya. In Robert Hetzron (ed.), *The Semitic languages*, 424–445. London: Routledge.

Lasnik, Howard & Mamoru Saito. 1994. *Move alpha: Conditions on its application and output* (Current Studies in Linguistics 22). Cambridge, MA: The MIT Press.

Mathieu, Eric. 2016. The wh parameter and radical externalization. In Olga Fernández-Soriano Luis Eguren & Amaya Mendikoetxea (eds.), *Rethinking parameters*, 252–290. Oxford: Oxford University Press.

Pan, Victor. 2014. Wh-ex-situ in Mandarin Chinese: Mapping between information structure and split CP. *Linguistic Analysis* 39. 371–414.

Pan, Victor. 2019. *Architecture of the periphery in Chinese: Cartography and minimalism*. London: Routledge.

Pesetsky, David. 1987. Binding problems with experiencer verbs. *Linguistic Inquiry* 18(1). 126–140.

Pesetsky, David. 2000. *Phrasal movement and its kin* (Linguistic Inquiry Monographs 37). Cambridge, MA: MIT press.

Postal, Paul M. 1993. Remarks on weak crossover effects. *Linguistic Inquiry* 24(3). 539–556.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4). 657–720.

Richards, Norvin. 2010. *Uttering trees* (Linguistic Inquiry Monographs 56). Cambridge, MA: The MIT Press.

Rizzi, Luigi. 1990. *Relativized minimality* (Linguistic Inquiry Monographs 16). Cambridge, MA: The MIT Press.

Ross, John. 1967. *Constraints on variables in syntax*. Massachusetts Institute of Technology. (Doctoral dissertation).

Rouveret, Alain. 2011. Some issues in the theory of resumption: A presepective on early and recent research. In Alain Rouveret (ed.), *Resumptive pronouns at the interfaces* (Language faculty and beyond 5), 1–62. Amsterdam: John Benjamins.

Sabel, Joachim & Jochen Zeller. 2006. *wh*-question formation in Nguni. In John P. Hutchison John Mugane & Dee A. Worman (eds.), *Selected proceedings of the 35th Annual Conference on African Linguistics*, 271–283. Somerville, MA, USA: Cascadilla Proceedings Project.

Safir, Ken. 2004. *The syntax of anaphora*. Oxford: Oxford University Press.

Safir, Ken. 2017. Weak crossover. In Martin Everaert & Henk van Reimsdijk (eds.), *The Wiley Blackwell companion to syntax*, 2nd edn., 1–40. New York: Wiley-Blackwell.

Shlonsky, Ur. 2012. Notes on *wh* in situ in French. In L. Brugé., Anna Cardinaletti, G. Giustic, N. Munaro & C. Poletto (eds.), *Functional heads: The cartography of syntactic structures*, vol. 7, 242–252. Oxford: Oxford University Press.

Szabolcsi, Anna. 2006. Strong vs. weak islands. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. 4, 479–531. Oxford: Wiley-Blackwell.

Szabolcsi, Anna & Marcel den Dikken. 2003. Islands. In Lisa Lai-Shen Cheng & Rintje Sybesma (eds.), *The second glot international state-of-the-article Book*, 213–240. Berlin: De Gruyter.

Torrence, Harold & Jason Kandybowicz. 2015. Wh-question formation in Krachi. *Journal of African Languages and Linguistics* 36(2). 253–285.

Tsai, Wei-Tien Dylan. 1994. On nominal islands and LF extraction in Chinese. *Natural Language & Linguistic Theory* 12(1). 121–175.

Tsai, Wei-Tien Dylan. 2008. Left periphery and how-why alternations. *Journal of East Asian Linguistics* 17. 83–115.

Wasow, Thomas. 1979. *Anaphora in Generative Grammar*. Ghent: E. Story-Scientia.

Zentz, Jason. 2016. *Forming wh-questions in Shona: A comparative Bantu perspective*. Yale University. (Doctoral dissertation).

Chapter 12

On OV and VO at the Bantu/Bantoid borderlands

Elisabeth J. Kerr

Leiden University

While most Niger-Congo languages have SVO word order, a few outliers display S(Aux)OV(X) patterns, either in restricted TAM contexts or across the board. These OV languages include a handful of Guthrie zone A40 languages of Cameroon (Tunen, Nyokon), Bantoid languages (Tikar), and broader Niger-Congo (e.g. Gur and Kru languages). For the A40 languages, Nyokon ([nvo]/A45, Cameroon) has VO order in most tenses but OV order in the past tense, while its neighbour Tunen ([tvu]/A44, Cameroon) has OV consistently. This paper presents new data on Tunen and Nyokon, using controlled elicitation to test the accuracy of previous TAM and information structure (IS)-based accounts. I provide evidence that OV is the most pragmatically-neutral word order in Tunen, consistent across TAM contexts. For Nyokon, OV versus VO order is shown to be dependent on the TAM pattern and not directly conditioned by IS. I show that both languages otherwise consistently pattern as head-initial in their syntax. Based on these results, I reflect on the potential grammaticalisation source(s) of their OV syntax. Finally, I note that Tikar appears to pattern similarly to Nyokon in having a TAM-based OV/VO system.

1 Introduction

A basic feature of syntactic typology is the classification of a language's word order as having the verb preceding the object (VO) or the object preceding the verb (OV). Data from WALS feature 83A (Order of object and verb; Dryer 2013) show that the Niger-Congo languages in the WALS sample are overwhelmingly VO, with a few non-VO outliers, as reproduced in Figure 1 below. Note that there are
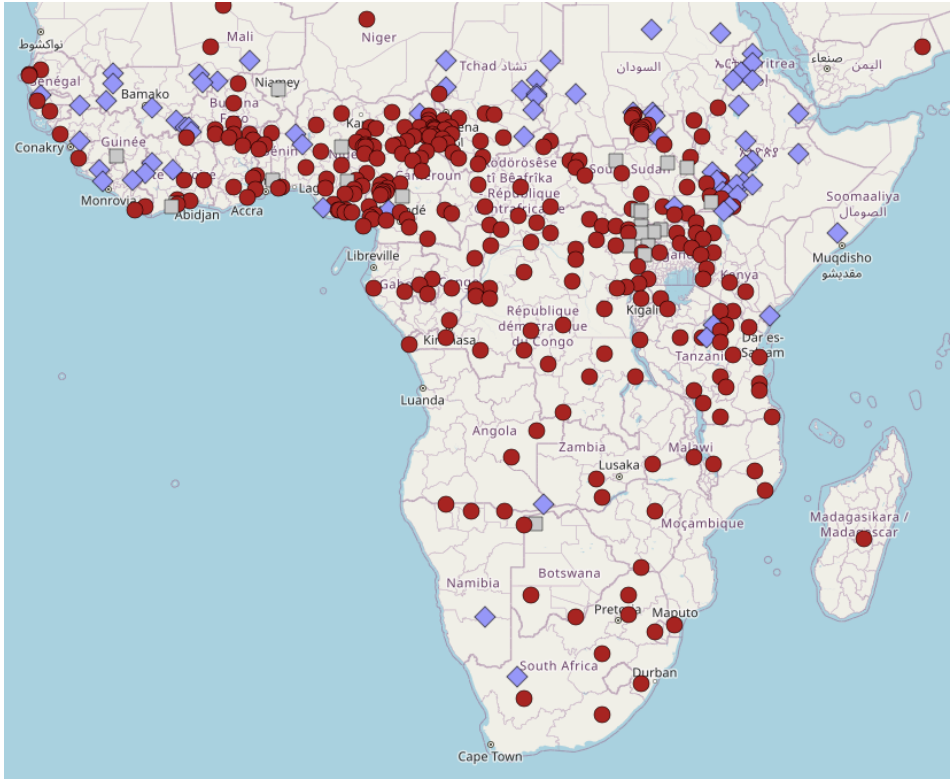
Figure 1: Map of WALS feature 83A (order of object and verb) data for sub-Saharan Africa (Dryer 2013). Key: red dot = VO word order, purple diamond = OV word order, grey square = no dominant word order.

no Niger-Congo languages on the map in the central, east, and southern Bantu-speaking areas that are coded as having OV; all the OV languages in those areas are from unrelated language families. In the Northwest of the Bantu region, Tunen (Bantu zone A, Cameroon) is one noticeable outlier in being coded as having OV as its base word order. Tikar (Bantoid, Cameroon) is another outlier in Cameroon, though is coded grey to indicate "no dominant order". All other Cameroonian Bantu and Bantoid languages of the WALS sample are VO. Turning to the west, outside of Narrow Bantu, a handful of (potentially) Niger-Congo OV languages are found in West Africa. These include Ịjọ (Ijoid, Nigeria) (Williamson 1965); Senufo languages (Gensler 1994); Kru languages, and Mande languages (Gensler 1994, Creissels 2005, Nikitina 2011, Sande et al. 2019).[1]

---

[1]The classification of several of these languages as Niger-Congo is doubtful, notably for the Ijoid and Mande languages (see e.g. Güldemann 2018). I do not take a stance on this question here, and simply note the presence of OV syntax in these languages.

In short, we see that Niger-Congo languages typically have VO syntax, but there are various outliers in Central and Western Africa. The syntax of many of these OV languages of West Africa has been discussed by Sande et al. (2019), who propose that they have different underlying syntactic derivations (related to earlier proposals of multiple causation, for which see e.g. Gensler & Güldemann 2003: 4, Güldemann 2007: 84). In this paper, I turn to the remaining Central African cases, considering the Bantu/Bantoid outliers in Cameroon: Tunen, Nyokon, and Tikar.

Tunen and Nyokon are classified as Narrow Bantu zone A40 (Guthrie no. A44 and A45 respectively), spoken in the Mbam-et-Inoubou area of the Centre region of Cameroon and therefore referred to as Mbam languages.[2] Previous work on Tunen provided evidence for OV as the basic word order (1; Dugast 1971, Mous 1997, 2003, 2005, 2014). While the editor of Dugast (1971: 6)'s grammar of Tunen writes in the introduction that Tunen OV order is "à ma connaissance absolument unique en bantu" [to my knowledge completely unique in Bantu], Mous (2005, 2022) reports that its neighbour Nyokon has a partial OV pattern, which he argues is determined by the TAM context (2):[3]

(1) a. bá-ndɔ *bɛ-kana* tála ɔ   yɔkɔ
        SM.2-PRS 8-basket put PREP 7.chair
        'They are putting baskets on the chair.'

     b. bá-ná   *bɛ-kana* tála ɔ   yɔkɔ
        SM.2-PST2 8-basket put PREP 7.chair
        'They put baskets on the chair.'   (Tunen; Mous 1997: 125, adapted)

(2) a. mù   nə̀ə́: yìl *wóó nìtān*
        SM.1SG COP take small stone
        'I take a small stone.'

     b. ù   *kìfá ús* yíl
        SM.1 stick short take
        'He took a short stick.'   (Nyokon, Mous 2005: 5)

---

[2]The Guthrie classification is a geographical classification of Narrow Bantu languages; see Maho (2003, 2009) for further details. Bantu is a sub-group of Bantoid, itself a subgroup of Benue-Congo and ultimately of Niger-Congo (see e.g. Marten 2020).

[3]See the Abbreviations section at the end of the paper for a list of glossing abbreviations. For clarity, the verb here is indicated in bold font, while the object is in italics. Tone marking is as follows: *á* = high tone; *ā* = mid tone; *à* or *a* = low tone; ' = downstep; [H] = floating high tone; [L] = floating low tone.

In their overview of S-Aux-O-V-Other word order patterns in Africa, Gensler & Güldemann (2003) and Güldemann (2007), based on data from Mous (1997), list Tunen as a language in which this word order is determined by information structure (IS), with S-Aux-O-V-Other treated as an exceptional pattern rather than the unmarked order. Güldemann (2007: 100) characterises Tunen under languages with unmarked VO order, with OV order synchronically determined by IS status of the object as less focal. In the absence of discourse context in Mous' data, it is hard to evaluate the extent to which IS conditions such word order patterns, i.e. whether S-Aux-O-V-Other in Tunen is an IS-conditioned word order variant or the pragmatically neutral order.[4] This paper will therefore investigate the word order of Tunen and Nyokon in closer detail by providing new data which controls the IS context versus the TAM context in order to test their respective influence on the use of OV versus VO order. While these zone A40 Mbam languages are the focus of the paper, it can also be argued based on secondary sources that the Tikar pattern coded in the WALS database as 'no dominant order' is a Nyokon-type system with a TAM-based alternation between OV and VO word order. I will thus return to Tikar at the end of the paper.

The rest of the paper is structured as follows. In section §2 I provide background on the Bantu/Bantoid borderlands region and its linguistic significance, provide background on proposals of the reconstruction of OV versus VO syntax in Niger-Congo, and formulate the research questions. Section §3 explains the methodology. Section §4 walks through the results in turn: §4.1 discusses the influence of IS versus TAM, §4.2 discusses the extent of head-finality within each language, and §4.3 considers the possible diachronic analysis. Finally, section §5 comments on Tikar, and section §6 concludes.

2 Background

2.1 The Bantu/Bantoid borderlands

In this paper I use the term "Bantu/Bantoid borderlands" to refer to the region where Narrow Bantu zone A borders (non-Bantu) Bantoid in central/West Cam-

---

[4]The full detail of the argument includes Mous (1997)'s proposal that there is an SVO strategy in Tunen used for contrast, where the object is preceded by a contrast marker *á*. In other work I argue that such constructions relate to biclausal clefts and ex-situ marking of contrastive focus, and are thus crucially different from the basic SVO constructions found in languages like Nyokon. See Kerr (to appear) for further detail. For space reasons, I will restrict the discussion in this paper to showing how S-Aux-O-V-Other in Tunen is, in contrast to the phrasing in Gensler & Güldemann (2003) and Güldemann (2007), synchronically the unmarked (i.e. pragmatically neutral) word order (i.e., not an IS-determined variant of an unmarked VO order).

eroon. To give an idea of the geography, on the road from Yaoundé to Bafoussam, the town Ndikiniméki is in the Tunen-speaking region. Ndikiniméki is followed by the town Makenene, where Nyokon (A45) is spoken, and then by the towns of Tonga and Bangangté, where Màdʉ̀mbà (Bantoid, East Bamileke) is spoken. Tunen and Nyokon are therefore the last Narrow Bantu languages before Bantoid languages are reached, as shown in Figure 2 below.[5]
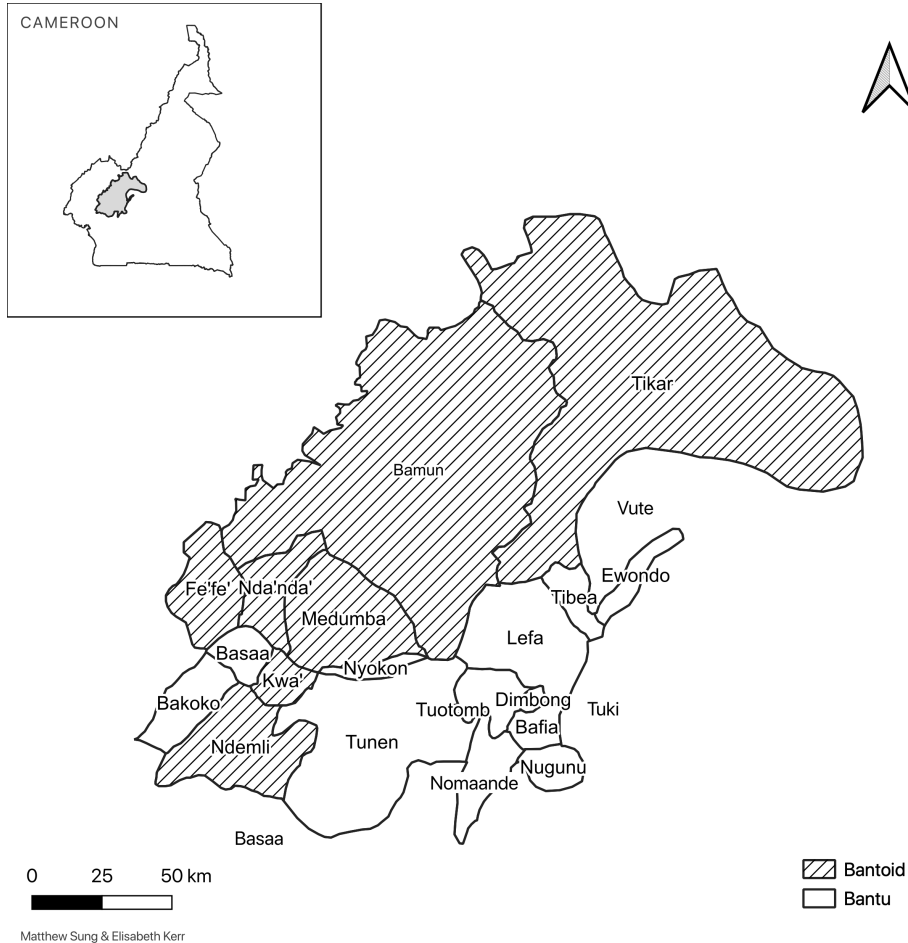
Figure 2: Map showing location of Tunen and Nyokon at the borderlands between Bantu and non-Bantu Bantoid languages of Cameroon, with Tikar a near neighbour to the Northeast (based on Ethnologue (Simons & Fennig 2018), ALCAM (Bikoi 2003), and Welaze Kongne.

[5]I thank Matthew Sung for help creating this map using QGIS.

While Tunen and Nyokon are typically considered to be Narrow Bantu languages, the proper classification of zone A40 as Narrow Bantu has been questioned (e.g. Nurse 2008, Blench 2022).[6] The A40 languages have been noted to be outliers for various linguistic properties, for example the loss of final vowels, the development of ATR harmony systems, and S-Aux-O-V-Other word order (Güldemann 2008, Nurse 2008, Downing & Hyman 2016, Rolle et al. 2020, a.o.). Note that this last feature is not found in all A40 languages – while I do not have original data on the other Mbam languages, other sources state that they are VO, not OV (Mous 2005, 2014). The presence of OV word order in Tunen and Nyokon is therefore particularly unusual and worth further investigation.

2.2 Reconstructions of word order

Having seen that certain Cameroonian languages are outliers in Niger-Congo in having OV word order, we can ask what this may tell us about the word order of the proto-language. There have been several different proposed reconstructions for the word order of Proto-Niger-Congo, as shown in Table 1 below.

Table 1: Hypotheses of Proto-Niger-Congo word order

Proposal	Source
*SOV	Givón (1975); Hyman (1975); Williamson (1986)
*SVO	Heine (1980); Heine & Reh (1984); Claudi (1993)
*S-Aux-O_{Pron}-V-Other/SVO	Gensler (1994); Gensler & Güldemann (2003)

Of these hypotheses, sole *SOV is the hardest to maintain, as it raises the question of why such a large number of present-day Niger-Congo languages have SVO word order (as argued by e.g. Mous 2005). It is therefore more generally preferred to take *SVO as the starting point (e.g. Heine & Claudi 2001, Nurse 2008: 58). There are however two points of complexity worth repeating here.

Firstly, Gensler (1994, 1997) points out that it is misleading to frame the choice of reconstruction as a dichotomy between OV and VO, because (i) there can also be the intermediate 'syntagm' S-Aux-O-V, and because (ii) a language may have

---

[6]Gensler & Güldemann (2003: 5) go so far as to exclude Tunen data from evaluation of the reconstruction of Proto-Bantu syntax, on the grounds of Tunen being "not Narrow Bantu". Relevant to the discussion below, if zone A40 languages are excluded from the definition of Narrow Bantu in this way, then the time depth of reconstruction necessarily has to be further back than Proto-Bantu in order to account for the origins of their syntax.

multiple orders at once to different extents. Therefore, instead of asking the question "OV or VO?", we should ask "what was the word order syntax of Proto-Niger-Congo actually like?" (Gensler 1997: 90). We have in fact already seen evidence in favour for such an approach, as (2) showed that Nyokon has both OV and VO patterns synchronically. In this paper I therefore formulate my discussion in terms of whether a language is 'canonically' OV or VO, not excluding that they may have other word orders available in particular discourse contexts.[7]

Secondly, there are different time depths of reconstruction to consider:

(3)   Proto-Niger-Congo > Proto-Benue-Congo > Proto-Bantoid[8] > Proto-Bantu

The larger the time depth, the more cycles of change are possible and likely to have occurred (Gensler 1997, Hyman 2011, a.o.). Proto-Niger-Congo is estimated as dating to 10000+ years ago (Blench 2006, cited in Nurse 2008: 226), while Proto-Bantu is put at 4000–5000 years ago (Nurse 2008: 228). This means that it is much harder to convincingly reconstruct Proto-Niger-Congo word order than Proto-Bantu or a close-by time depth that encapsulates zone A Bantu and (some of) the Bantoid varieties. This paper deals only with data from the Bantu/Bantoid borderlands, and I therefore focus solely on this lower-level reconstruction, leaving the question of Proto-Niger-Congo syntax aside.

2.2.1 SOV and reconstruction: The influence of IS

While the data that has been cited for Tunen and Nyokon to date is only at the level of the sentence,[9] more recent work has called for consideration of the possible impact of information structural context on word order. For example, Güldemann (2007) argues that information structure (IS) can be taken as a key factor conditioning the position of the object in Benue-Congo: OV word order

---

[7]See Kerr (to appear) for a more detailed description of the different word order patterns available in Tunen for various discourse contexts, e.g. clefting and the use of right- and left-dislocation. Such constructions can be seen as departures from the canonical word order which is the focus of this article.

[8]I use 'Proto-Bantoid' as a placeholder term for an intermediate stage further back than Proto-Bantu (capturing zone A Bantu and (some) non-Bantu Bantoid) but not as far removed as Proto-Benue-Congo or Proto-Niger-Congo, without committing to 'Proto-Bantoid' as a meaningful ontological stage.

[9]Some discourse-level data is available for Tunen in the selection of folktales (*contes*) transcribed in Dugast (1975). Mous refers to these *contes* in his work (e.g. Mous 2003), but the examples he provides in discussion of Tunen OV order are elicited examples without discourse context. I am not aware of any prior work on Nyokon beyond the level of the sentence.

is found when the object is "less focal or even extrafocal, non-asserted information" (Güldemann 2007: 83), with the proto-language having basic VO syntax (Güldemann 2007: 104). Linked to this is the form of the object as a pronoun versus a full lexical noun phrase, and the development of the agglutinative Narrow Bantu verbal template (which contains a pre-stem object marker; Meeussen 1967). Güldemann also follows Hyman & Watters (1984) in drawing a connection between certain TAM contexts, focus, and negation.

In terms of word order typology, Good (2010) points out that for Naki, a Bantoid language of Cameroon, the 'canonical' word order is SVO, but SOV, VSO, and other orders are also found. After considering the IS situation, he argues that Naki syntax is more accurately characterised as "Topic Field – Focus Field" than in terms of grammatical role ('subject'/'object'). This research raises the question as to whether Tunen and Nyokon word order is better captured in terms of IS notions, or whether the grammatical role notions of 'subject' and 'object' are the most appropriate. This is another motivation to test to what extent IS controls word order in Tunen and Nyokon. I will end up concluding that Tunen and Nyokon OV word order is not primarily determined by IS context, unlike what Good argues for Naki, and so I maintain the use of grammatical role-oriented terms in this paper (but see Kerr et al. (2023) for more detailed discussion for discourse role-oriented approaches to Tunen and Bantu more broadly).

2.2.2 Research questions

Having considered this background, we can develop three main research questions for the current paper, as listed below.

(4) **RQ1:** To what extent do TAM and IS influence OV versus VO word order in Tunen and Nyokon? Is it accurate to classify these borderland languages as canonically OV?

(5) **RQ2:** To what extent do the languages with OV orders differ syntactically from languages with VO orders?

(6) **RQ3:** Is OV at the borderlands historical or innovative? If innovative, how did it grammaticalise?

Before going through each of these research questions in turn, I will explain the methodology used for the study.

3 Methodology

The data in this paper, unless otherwise indicated, were collected as part of fieldwork on Tunen as part of the Bantu Syntax and Information Structure (BaSIS) project, for a period of approximately 3.5 months in 2019 and 3.5 months in 2021/2022, in Ndikiniméki and Yaoundé, Cameroon. The Nyokon data are from Mous (2005) and Lovestrand (2020) together with follow-up fieldwork I conducted with one Nyokon speaker in 2022 in Yaoundé, Cameroon.

Tunen data from my fieldwork are cited with the consultant's initials and unique form ID corresponding to the entry within the Tunen Dative database (to be archived open access at end of project, expected 2024). The Nyokon data were all elicited in Yaoundé with the consultant René Atchom (RA).[10] Some clarifications were also provided remotely in the preparation of this paper.

I used two field questionnaires for the data collection: the Bantu Syntax and Information Structure (BaSIS, Leiden University) project methodology on syntax & information structure, and (ii) a draft version of the Consequences of Head-Argument Order on Syntax (CHAOS/C08, Universität Potsdam) project questionnaire on OV/VO patterns.[11] The former questionnaire builds on the earlier Questionnaire on Information Structure (QUIS; Skopeteas et al. 2006) and investigates how information structure influences syntax. The latter questionnaire investigates which syntactic features correlate with VO versus OV order, testing to what extent the headedness of the verb phrase correlates with headedness in other syntactic domains. Together, and supplemented (for Tunen) with natural speech data, they address the research questions in §2.2.2 above.

4 Results

4.1 RQ1: TAM and IS

RQ1 asks to what extent TAM and IS influence OV versus VO word order, and how accurate it is to classify Tunen and Nyokon as canonically OV. In order to answer this research question, two steps are necessary: (i) investigate different

---

[10]One important caveat for the Nyokon data is that the transcriptions are initial versions based on the Nyokon community orthography developed by René, amongst others. This differs from the more IPA-based version used in Lovestrand (2011) in the use of replacements for certain IPA symbols, such as ⟨gh⟩ for the voiced velar fricative ɣ and an apostrophe ⟨'⟩ for the glottal stop ʔ. I also believe that there are some differences in the transcription of tone and vowel length, but have not investigated this in any detail. I also follow Mous (2022) in not glossing Nyokon noun classes, in absence of a detailed study of Nyokon's nominal classification system.

[11]The BaSIS methodology is freely available online at hdl.handle.net/1887/3608096. I thank the late Gisbert Fanselow for sharing and discussion of the C08 questionnaire.

TAM contexts, and (ii) investigate different IS contexts (e.g. topic, focus, contrast) using controlled elicitation and natural speech. The results are shown in Table 2.[12]

Table 2: Results for RQ1

Property	Tunen result	Nyokon result
Present tense	OV	VO
Past tenses	OV	OV
Future tense	OV	VO
Thetics	OV	VO/OV dependent on TAM
Object focus	OV	VO/OV dependent on TAM
Sub-NP focus	OV	VO/OV dependent on TAM
VP focus	OV	VO/OV dependent on TAM
Other PCF	OV	VO/OV dependent on TAM

In the interest of space, I will not walk through every single property for both languages, but rather give a selection of representative examples.[13]

The Tunen example in (7) below shows SOV for term focus on the direct object (with the recipient object preceding the theme object). The next examples show SOV in an out-of-the-blue discourse context (8), where there is no focus on the object, and SOV with VP focus (9).

(7) Context: 'What did the woman give to the other woman?'
    a   nɔ́   ɔsɔ́kɔ́ hɛtɛ́tɛ́   indi
    /a   nɔ́   ɔsɔ́kɔ́ hɛ-tɛ́tɛ́   índíɔ́/
    SM.1 PST1 other 19-gourd give
    'She gave [a gourd]_FOC to the other.'   (Tunen; PM 1541)

(8) Context: You enter the room and see a broken window. Someone announces...
    Biɔ́lɛ   a   ná   itúbɔ́   san.
    /Biɔ́lɛ   a   ná   ɛ-túbɔ́   sána/
    1.Pierre SM.1 PST2 7-window break
    'Pierre broke the window.'   (Tunen; EE+EB 1669)

---

[12]This table gives a simplified summary of the TAM systems. More detail will be provided when evaluating TAM-based grammaticalisation proposals in section 4.3.1 (see Table 5 and 6).

[13]A complete set of supporting evidence can be found in the Appendix to this paper and in the BaSIS archival deposit.

(9)   Context: What did Maria do?
       Malíá   a   ná   bilə́liə   fɔfɔ́kiə.
       /Malíá a   ná   bɛ-lə́liə   fɔfɔ́kíə́/
       1.Maria SM.1 PST2 8-varnish anoint.DUR
       'Maria [applied the varnish]_FOC.'   (Tunen; JO 2518)

Pronominal objects behave like full noun phrases, i.e. are canonically in a pre-verbal position (OV). This is exemplified in (10), where the class 1 pronoun *wɛ́ɛya* 'her' appears before the verb, just as the noun phrase objects did in the previous examples.

(10)   mɔndɔ ɔwá mɔ́ná ándɔ naa anɔ́ wɛ́ɛya ákánana ɔ ndɔ́kɛt
       /mɔ-ndɔ ɔwá mɔ-ná a   ᴴndɔ ná   a   nɔ́ wɛ́ɛya ákánana ɔ
       1-person REL.1 1-child SM.1 PRS be.sick SM.1 PST1 PRON.1 leave.APPL PREP
       ndɔ́kɛta/
       7.doctor
       'The man whose child is sick took her to the clinic.'   (Tunen; PM 2179)

From this, we see that OV word order is available both for focal and non-focal objects, and both for pronominal and non-pronominal objects, meaning that it should be considered the pragmatically neutral word order in Tunen.

In Nyokon, the results show a mixture of OV and VO patterns, directly dependent on TAM but not on IS context. For example, past tense examples are OV, consistent across different IS contexts. In (11) below, we see that OV can be used both for focus on the whole predicate and focus on the object in the past tense, while VO order is dispreferred.

(11)   a.   Context A: 'What did your father do?' (VP focus),
            Context B: 'What did your father kill?' (object focus)
            itə́   ángam ghó.
            father spider kill
            'My father [killed a spider]_FOC.'
            'My father killed [a spider]_FOC.'   (Nyokon; RA 216)
       b.   Context: 'What did your father kill?' (object focus)
            ? itə́   ghó ángam.
              father kill spider
            Intd.: 'My father killed [a spider]_FOC.'   (Nyokon; RA 224)

All present tense examples were VO regardless of IS context. This is exemplified below, where VO order is found both for narrow focus on the numeral

modifying the object (i.e. sub-NP term focus; (12)) and for predicate-centred focus (VP focus in (13) and truth focus in (14)).

(12) Context: 'How many chickens do you see?'
      táá   ndukŋ ikɔ'ɔ   itá.
      SM.1PL see   chickens three
      'We see [three]_{FOC} chickens.'   (Nyokon; RA 34)

(13) Context: 'What are you doing today?'
      taa   nə tsá'a   Rɛnɛ.
      SM.1PL PRS receive Rene
      'We are [hosting René]_{FOC}.'   (Nyokon; RA 233)

(14) Context: Kinyo is sick and unable to eat a lot. Someone asks 'Can Kinyo
      eat rice?'
      áa, u   fʉr anyɛ́ álif.
      yes SM.1 can eat rice
      'Yes, she [can]_{FOC} eat rice.'   (Nyokon; RA 151)

To sum up, we have seen that S(Aux)OV(Other) word order in Tunen is compatible with various different information structural contexts, including thetics, VP focus, and object focus. For Nyokon, OV or VO is dependent on the TAM context and consistent across different IS contexts. We can therefore conclude that TAM rather than IS is the primary synchronic conditioning factor for OV word order in Nyokon.[14]

4.2 RQ2: Head-finality versus head-initiality

Let us turn now to the second research question: "To what extent do the languages with OV orders differ syntactically from languages with VO orders?". To answer this question, we need to investigate the general syntactic profile of each language, considering whether we find head-finality in other domains. Table 3 gives the results for various diagnostics of head-finality versus head-initiality.

---

[14]As mentioned above, it can be argued that TAM is intrinsically linked to IS (Güldemann 2007), and therefore, a TAM-dependent system is also an IS-dependent system. For the purposes of this research question, I consider what is the *primary* conditioning factor, leaving this more indirect relation aside. I will return to discuss the possible diachronic relation between TAM and IS in section 4.3 below. Note that a convincing argument that TAM-based alternations reflect IS requires exposition that the TAM contexts for which Nyokon has OV versus VO map onto a coherent set of IS contexts. An alternative account for a TAM-based system is based on the presence or absence of segmental morphology in the TAM slot, as has been crucial for example in syntactic analyses of Kru languages, where the difference in OV and VO syntax is argued to depend on the ability for the verb to move to a higher tense node in the absence of a TAM marker (e.g. Koopman 1984, Sande et al. 2019); cf Mous (2005: 6, 11), Mous (2022: 29).

Table 3: Results for RQ2

Property	Tunen result	Nyokon result
Order of N and Mod	N-Mod	N-Mod
Adposition type	Prepositions	Prepositions
Order of N and Poss (pronoun)	Poss-N	Poss-N
Order of N and Gen	N-Gen	N-Gen
Canonical order of O and V	OV	OV/VO
Order of O and V in imperatives	VO (V-IO-DO)	VO (V-DO-IO)
Order of Aux and V	Aux-V	Aux-V
Order of C and Comp	C-Comp	C-Comp
Order of Cop and Compl	Cop-Compl	Cop-Compl
Order of V and manner adverb	V-Adv	V-Adv
Canonical adjunct position	SOV**X**	SOV**X**/SV**O**X
Low subjects (VS)?	✗	✗

These results present a more in-depth study of head-initiality in Tunen and Nyokon beyond the discussions in the literature that have focussed more narrowly on the verb phrase, and therefore allow us to evaluate the extent to which the languages pattern as head-final.[15] As we have already seen evidence for the order of the object and verb in §4.1 above, in this section I will present evidence for the other properties.[16]

Firstly, imperatives are interesting in being the only context where Tunen has VO syntax. Tunen affirmative imperatives have VO (V-IO-DO) order; Nyokon ones also have VO (V-DO-IO) order. Note that there is no TAM marker here.[17]

---

[15]As argued by Sande et al. (2019), having a deeper understanding of the syntax of the languages allows us to make more insightful comparisons between these languages and other languages with S-Aux-O-V-Other surface order, given that this surface order may in fact be derived in various different ways in different languages (both synchronically and in terms of the grammaticalisation processes), and taking general patterns of syntactic headedness as potentially diagnostic of underlying syntactic structure (as also part of the motivation behind the C08 questionnaire).

[16]The same note as for RQ1 about only providing a set of representative examples applies here.

[17]A reviewer points out the possible confound in that the Nyokon example has a pronominal where the Tunen has a full noun phrase. Note first of all that the theme object is a full noun phrase in both examples. Secondly, Tunen imperatives with pronominals are still VO. Finally, RA confirms remotely that the non-pronominal equivalent of the Nyokon example (15b) cannot be OV (while *nda manyí apɔ́ ghó!* ('give the child water!', V-DO-Prep-IO) is also grammatical).

(15) a. índíə́ mɔná imit!
         /índíə́ mɔ-ná ɛ-mítə́/
         give   1-child 7-calabash
         'Give the child a calabash!'   (Tunen; JO 1594)

      b. nda manyí ngɔ́m!
         give water PRON.1SG
         'Give me water!'   (Nyokon; RA 1)

The examples in (16) illustrate N-Mod order in the nominal domain (with D elements preceding NPs, as expected if the DP is head-initial), therefore providing further evidence for head-initial syntax.

(16) a. tɔ́ɔ́yɛ tɔbanána tɔté̍té tɔfítitiə tɔ́fandɛ
         /tɔ́ɔ́yɛ   tɔ-banána tɔ-té̍téá tɔ-fítitiə tɔ-ᴴfandɛ /
         13.DEM.PROX 13-banana 13-small 13-black 13-two
         'these two small black bananas'   (Tunen; JO 885)

      b. pí   pimbɔ́tɔ pífu
         DEM.PROX bananas two
         'these two bananas'   (Nyokon; RA 191)

      c. pimbɔ́tɔ kúpóm
         bananas ripe
         'ripe bananas'   (Nyokon; RA 188)

In terms of verbal and clausal syntax, we see Cop-Compl and C-Compl order, both head-initial properties.

(17) a. Context: 'Where are you?'
         mɛ lɛ ɔ nioní.
         /mɛ   léá ɔ   nɛ-oní/
         SM.1SG be PREP 5-market
         'I am at the market.'   (Tunen; PM 102)

      b. Context: 'Where are you?'
         mu   nə a   nyí.
         SM.1SG COP PREP market
         'I am at the market.'   (Nyokon; RA 174)

Finally, while a canonical head-final language would have SOV with the verb final (SOV#), in both Tunen and Nyokon, non-arguments typically appear postverbally (S-(Aux)-O-V-**Other**), with adverbs after the verb (V-Adv) (18).

(18)   a. Context: 'Where did he build his house?'

        a ka yayɛ́á miímə lúmə́kə́ ɔ iNdíki naánɛkɔl.

/a	ka	yayɛ́a	miímə lúm-aka	ɔ	iNdíki	naánɛkɔla/
SM.1	PST3	POSS.PRON.1.3	3.house build-DUR	PREP	Ndiki	yesterday

        'He built his house yesterday in Ndiki.'   (Tunen; JO 1121)

    b. Context: 'Where did you put the clothes?'

mu	piyɛ	ghá	avuə.
SM.1SG	clothes	put	outside

        'I put the clothes outside.'   (Nyokon; RA 72)

    c. Context: PM and EO are discussing how PD should be considered a true Munen [i.e. a local], despite being born somewhere else.

        a ka nyɔɔkɔ háaha ɔ uwəsú mɔŋɛŋ.

/a	ka	nyɔ-aka	háaha	ɔ	uwəsúə́	mɔŋɛ́ŋa/
SM.1	PST3	work-DUR	here	PREP	PRON.POSS.1PL	much

        'He worked a lot here in our region.'   (Tunen; EO 1043)

    d. Context: After visiting a friend's house, you announce:

punú	páa	liak	tsú.
POSS.2PL	children	cry	much

        'Your children cry a lot.'   (Nyokon; RA 7)

To summarise this section, we have seen that both Tunen and Nyokon show consistent evidence for head-initial syntax, with (partial) OV word order their only head-final property.[18] This makes them pattern more generally with VO languages than a canonical OV language. These results are significant when arguing for a syntactic analysis of the languages, as they provide evidence for an underlyingly head-initial syntax under a transformational model of syntax, as argued for West African languages by Sande et al. (2019). This therefore makes S-(Aux)-O-V-(Other) syntax in Tunen and Nyokon appear to be quite different from the S-(Aux)-O-V syntax discussed for consistently head-final languages such as Kru languages (e.g. Sande et al. 2019) and Ịjọ (Williamson 1965).

---

[18]Table 3 shows Poss-N in addition to N-Gen order, with Poss-N the order found with a possessive pronoun (e.g. *my house*). As will be discussed in section 4.3, there is evidence that N-Poss is the historic order. Note also that while there are sentence-final question particles, following Biberauer (2017) I do not take this to be evidence for head-finality, in that such particles are likely not syntactic heads.

4.3 RQ3: Diachrony

The third research question is: "Is OV at the borderlands historical or innovative?" As I stated in §2.2 above, the favoured null hypothesis is that OV is innovative for Bantu, given that positing OV as the sole historical order means that VO order has to be derived for the majority of present-day Niger-Congo languages.

Various grammaticalisation paths have been proposed in both Africanist and general typological literature for the development of OV word order, as listed in Table 4 below. Such grammaticalisations can arise independently in different languages, though they can also be influenced through contact. For a given language, multiple grammaticalisation pathways can interact, for example in the development of auxiliaries from verbs as part of earlier serial verb constructions (see e.g. Claudi 1993, Williamson 1986).

Table 4: Proposed grammaticalisation pathways for S-(Aux)-O-V-(Other) in Niger-Congo

Pathway	Example proponent(s)
V>Aux(>TAM)	Claudi (1993: 102)
Serial verbs	Givón (1975), Claudi (1993); Heine & Claudi (2001)
Periphrasis of nominal complement (Gen-N)	Claudi (1993); Heine & Claudi (2001)
IS status of object (extrafocal/focal, pronoun/NP)	Gensler & Güldemann (2003), Güldemann (2007)
Interaction between TAM, negation, and focus	Gensler & Güldemann (2003), Güldemann (2007) (building on Hyman & Watters 1984)
Nominalisation through infinitival constructions	Mous (2005)

In this section I consider whether we find evidence for one or more of these paths for Tunen and Nyokon, focussing on the V>Aux and infinitival paths. The other pathways are excluded for the reasons I give below.

Firstly, the serial verb constructions pathway is not considered because there is no evidence for serial verb constructions in the Mbam group of the kind found for the West African languages for which this path was proposed (Mous 2005: 11), and so this is not a plausible source for the Mbam OV patterns.

Secondly, for word order change related to Gen-N word order, it is significant that Tunen has N-Gen order, with the exception of possessives, which are Possessee-Possessor, i.e. Gen-N. Mous (1997: 124) provides evidence based on Dugast (1971, 1975) that indicates that the Poss-N order is a more recent development in Tunen, given that in certain frozen possessive expressions the order is N-Gen (e.g. *òŋwâm* 'my friend', from *wám òŋgwáyè*; Dugast 1971: 137). Gen-N syntax should therefore not be taken to be the trigger for the development of OV syntax, as N-Gen is the historic order in Tunen (Mous 2005: 11). The same N-Gen patterning applies to Nyokon, and so Gen-N is an unlikely grammaticalisation source of OV order in the Mbam languages.

Finally, it has been suggested that the development of OV versus VO order can be related to the differential position of given and non-given objects (e.g. Güldemann 2007 on Benue-Congo in general), which in turn relates to the difference between full noun phrases and pronouns. Of the pathways I do not discuss in detail, this is the most promising one and fairly likely to have played a role historically, as Güldemann argues. However, I have shown already in section §4.1 that the preverbal position is synchronically available both for focal and non-focal objects, with no difference in word order between pronominal forms and lexical noun phrases. The synchronic picture therefore does not provide convincing evidence for such a distinction.[19]

4.3.1 RQ3: V>Aux(>TAM)

The first grammaticalisation path to consider in detail is V>Aux(>TAM), meaning a situation where a main verb develops a function as an auxiliary verb (which can then develop further into a TAM marker). Such a development means that a secondary verb form (as part of a serial verb construction for example) is reanalysed as the main verb, as discussed for example in Claudi (1993) (see Mous 2005:

---

[19]One relevant note when assessing the likelihood of such a development is that Hyman (2011) writes that Tunen, Basaá, and Eton object/oblique pronouns are "clearly a secondary development", not retraceable to Proto-Bantu, with no evidence of object pronouns in a language like Tunen becoming procliticized onto the verb (as is taken to have been the case in the development of Narrow Bantu verbal object markers; see e.g. Nurse 2008). See also fn28 below on Ewondo.

11 for discussion related to the A40 languages). An example of such a grammaticalisation from Eastern Bantu is the development of the Swahili anterior TAM marker *-me* from the main verb *-mala* 'to finish' (Nurse 2008: 59-60).

In Tunen and Nyokon, the TAM marker appears together with the subject marker as a separate phonological word from the verb stem (as evidenced by the ability for other material, such as the object, to intervene) (19). Note that I consider the forms "TAM markers" rather than auxiliaries due to their invariant and reduced phonological forms and their inability to occur as main verbs, but the choice of terminology is not too significant for the current purposes.[20]

(19)   SM-TAM(-Deixis)(-O)-V(-Other)   (adapted from Mous 2003: 291)

The identification of verbal sources for the TAM markers found in Tunen and Nyokon would provide evidence for such a V>Aux pathway. I therefore provide the TAM paradigm for the two languages in Tables 5 and 6 below (with the Nyokon paradigm adapted from Mous 2022).[21]

Table 5: Tunen affirmative tense markers

TAM marker	Gloss	Time point
ŋɔ	FUT	future from tomorrow onwards
ᴴ*ndɔ*	PRS	present, immediate future
nɔ́	PST1	recent past, just an instant ago
ná	PST2	a few hours ago (hodiernal)
ka / ᴸ*ná*	PST3	yesterday and before (hestiernal)
lɛ	PST4	far past; many years ago, before birth

While V>Aux(>TAM) grammaticalisation is considered to be crosslinguistically common, I am not aware of previous work identifying verbal sources for

---

[20]The terminology becomes significant if one takes verbal origin to be a necessary criterion for an element to count as an auxiliary, as done by Dryer (1992). As this section will show that verbal origins are not always visible, I use TAM marker here as a more neutral term (although the term auxiliary could be maintained under a more generous definition, as for example employed by Gensler & Güldemann 2003 and Anderson 2007).

[21]I have simplified the Nyokon table to focus on the form of the TAM marker as opposed to the whole syntactic construction. "Unclear" means that both a preverbal and a postverbal object slot was indicated as possible, complicating the classification as OV or VO. The TAM markers and time points are given as in Mous (2022), with the imperative, subjunctive, negative tenses and focus tenses omitted. 'OV/VO' is written for the present tense to reflect the data presented.

Table 6: Nyokon affirmative tenses (adapted from Mous 2022: 4)

TAM marker	Time point	OV or VO
∅	present	OV/VO
∅	recent past	OV
∅	far past	OV
∅	remote past	OV
noó, nəkú	perfect	VO
nǎ	present continuous	VO
nɔ́	conditional	OV
mbɨ́ɔ́	past imperfective 1	unclear
mbɨ́ɔ́ ku	past imperfective 2	VO
mbɨ́ə	background	VO
mbɨɔ́	background (past remote)	VO
maa	background moment	unclear
mə	future	VO
pɨ	narrative	OV

tense markers in Tunen or Nyokon. Therefore, to consider the possible verbal origin of the present-day TAM markers, I turn to the verbs that are listed in Heine & Reh (1984) as common grammaticalisation sources for TAM markers (cf Heine & Kuteva 2004; Nurse 2008) (20). I provide the Tunen and Nyokon equivalents of these verbs in Table 7 below.

(20)   **Common verbal origins of TAM markers** (Heine & Reh 1984: 113-135)
       'to begin', 'to finish', 'to return', 'to come', 'to go, to leave', 'to enter', 'to exit', 'to see'

A first observation from Table 7 above is that there appear to be surprisingly few cognates between the Tunen verb forms and the Nyokon ones, which is notable considering that the languages are neighbours. Furthermore, when comparing these infinitival verb forms in Table 7 to the TAM markers for each language (Table 5/6), there are also no clear correspondences that lead to potential sources (i.e., there is no TAM marker which can be easily analysed as a contracted form of one of the verb forms from Table 7).

While this does not mean that the TAM forms could not have been derived from a verbal source, the lack of clear overlap is nevertheless interesting. This

Table 7: Tunen and Nyokon translations of verbs in (20)

Verb	Tunen	Nyokon
'to begin'	*ɔlúmɔ́*	*utɔ'ɔ*
'to finish'	*ɔhɔ́á*	*umán*
'to return	*utilə, ɔhíáná*	*uyám*
'to come'	*ɔsáá*	*ufɛ*
'to go, to leave'	*ɔwakáná*	*utʃɔp*
'to enter'	*ɔfínə*	*utʃó*
'to exit'	*ɔfámálána*	*utʃɔ́s*
'to see'	*ɔsinə*	*utiin, undukŋ*

is not to say that the TAM systems of the A40 languages are completely independent of each other, however. Mous (2005: 11) identifies various cognates between the A40-A60 languages, for example between Tunen and Nomaande (A46). He proposes one Tunen-Nyokon cognate of the Tunen PST2 *ná* and the Nyokon present tense continuous marker *nə* (in turn cognate with the certain future *ná* in Gunu (A62a)). Furthermore, the TAM systems share the common property of having a large number of tense gradations, which Nurse (2008: 126) (drawing on Watters 2003: 246-7) identifies as a common feature of zone A40 Bantu languages and Eastern Grassfields.

We are left with the question of where the TAM markers for Tunen and Nyokon came from, if not from one of the verbal sources from (20). Nurse (2008: 125-6) argues that the tense markers of the Northwestern region are likely local innovations that arose after loss of earlier TAM forms, with Tunen PST3 *ka* perhaps a rare example of retention of **k* in the Northwest, though possibly derived from another origin than the *-ka-* found in other Bantu tense systems. The identification of sources of the Mbam TAM markers in the area is therefore a question requiring further research.

In summary, we have seen that Tunen and Nyokon TAM markers appear together with an obligatory subject marker, with the SM-TAM cluster separate from the verb. While one common source for auxiliaries and TAM markers cross-linguistically is verbs like 'to come' and 'to go', the study of these forms for Tunen and Nyokon did not provide any clear correspondences.

4.3.2 The infinitival pathway

An alternative (or additional) hypothesis is the infinitival grammaticalisation path sketched briefly by Mous (2005) for Tunen and Nyokon. The basic idea is that Mbam languages had basic VO syntax, with an IS-conditioned variant as in (22), in which a given object could precede the verb (as in Güldemann 2007, and in relation to Mous 1997, 2005's analysis that the postverbal object position is more contrastive in Tunen). Mous suggests that the presence of an infinitival construction of form (21) led to generalisation of the OV syntax of (22) applying to all objects via analogy (together with V>Aux grammaticalisation of the initial verb in (21)); (Mous 2005, p.c.).

(21)   S V ɔ O (ɔ) V$_{INF}$

(22)   S TAM O$_{Pron}$ V

Mous writes that both OV and VO word order is found in Mbam infinitival constructions (23). He relates the different available positions of the object to a difference in interpretation between the object as the argument of the verb (e.g. OV 'at the field') or a circumstantial reading (e.g. VO 'work at the field').

(23)   S Aux/V [O V] ~ S Aux/V [V O]   (Mous 2005: 12, p.c.)

The data to support such a hypothesis are limited and Mous does not give many details on the exact proposal, making it difficult to evaluate. The origin of O-V$_{INF}$ order is not explained, despite the fact that it is surprising: while Tunen has N-Mod/N-Gen order in the nominal domain (i.e. head-initiality; cf Table 3), here the verbal head follows its complement (head-finality). The question of the origin of OV in general is therefore pushed back to the question of the origin of OV within infinitival constructions.

Mous notes that the limited data he had available for Nyokon did not show any OV word order in infinitives, unlike in Tunen. In this section I therefore provide some extra data from Tunen and Nyokon to illustrate the infinitival construction in question, highlighting some points of interest for further study.

In Tunen, the infinitive marker ɔ is homophonous with the general preposition ɔ.[22] For the current purposes, I gloss these forms as INF and PREP respec-

---

[22]The infinitive may surface as [u] as a predictable result of ATR harmony (for which see Boyd 2015 and references therein). While my consultants considered the infinitive and preposition to be homophonous, there is also potential dialectal variation in the presence of a vowel /ʊ/, argued to be found in the infinitival prefix but not the preposition (see Mous 2003, Boyd 2015).

tively. In Nyokon, the infinitive prefix is *u-*, while the general preposition is *a*.[23] Embedded clauses require the infinitive marker together with prepositions, as indicated below in boldface.[24]

(24) a. mɛ́ ndɔ siə **ɔ** mənífɔ́ **ɔ** ɔnyá.
          /mɛ ᴴndɔ siə ɔ   ma-nífɔ́ ɔ   ɔ-nyá/
          SM.1SG PRS   want PREP 6-water PREP INF-drink
          'I want to drink water.'   (Tunen; JO 609)

      b. miɔkɔ́ a lɛ́ ɔ́sɔ **ɔ** bɛŋgwɛtɛ (**ɔ**) ɔbáta.
          /miɔkɔ́ a lɛ óso ɔ   bɛ-ŋgwɛtɛ ɔ   ɔ-báta/
          9.chicken SM.1 NEG can PREP 8-potato PREP INF-collect
          'The chicken wasn't able to collect up her potatoes.' (Tunen; JO 1764)

Such contexts were also remarked upon by Dugast (1971), but, unlike in my data, her examples do not have an extra ɔ before the infinitival marker (25). The data from Mous (2005) pattern with Dugast's data in having no additional ɔ directly before an infinitive-marked verb form (26).[25]

(25) mɛ   nábékánɛ́n **o**   malex **o** wíndi
      SM.1SG PST2.try   PREP 6.advice INF give
      'J'ai essayé de donner des conseils.'
      ('I tried to give some advice.')   (Tunen; Dugast 1971: 309)

(26) a. a-ná   húánána **o** wâw **o** mʷəlúk **o**wíndi
          SM.1-PST2 must:H   PREP you PREP 6.wine INF.give
          'She/He was obliged to give you wine.'

      b. bá-ná lumín **o**   etɔpɔtɔpɔ **ɔ**-nyɔ
          2-child agree:H PREP 7.field   INF-work
          'They agreed to cultivate the field.'   (Tunen, Mous 2005: 10)

---

[23]A form *a* appears in other contexts in Nyokon, such as in sentence-final position after negation. I leave an analysis of whether these contexts are linked to the prepositional use as a topic for further research. Tunen does not have ɔ in the equivalent contexts.

[24]The parentheses around the second ɔ in (24b) is due to vowel elision making it inaudible in fast speech (this sentence was taken from a story). The consultant however said during transcription that there was an ɔ present.

[25]Note in terms of time depth that Mous' elicited data are based on fieldwork he conducted in 1984/1986. I adapt the glossing for consistency with my data, but leave the transcription line and translation unchanged. The English translation of the Dugast example is my own addition.

I suggest that this distinction is interesting because the prepositional ɔ in Tunen appears to be necessary to license postverbal nominals.[26] Supporting evidence for ɔ as a nominal licenser in Tunen comes also from the ability for such verb forms to appear as subjects (Dugast 1971), and its presence on topics in non-argument positions (Kerr to appear). The apparent systematic difference in number of ɔ forms between my own and Dugast/Mous' data therefore raises the question as to whether such a difference is linked to a change in progress with regards to nominal/verbal syntactic status. Mous (2005) suggests briefly that the homophony of the preposition and infinitive relates to the development of OV dominant order in Tunen, arguing that the infinitival verb in the OV construction was first marked by the infinitival ɔ and the preposition ɔ, and then, given homophony, the infinitival ɔ could be dropped. I suggest that the differences in our data could be understood as syntactic reanalysis of the scope of the nominal complement preceded by ɔ PREP, as sketched in (27) below.

(27) S V ɔ [O ɔ-V] → S V ɔ [O] ɔ [ɔ-V]

One issue with applying such an analysis to Nyokon is the lack of homophony between the infinitival prefix and the preposition in Nyokon, although the basic similarity in construction holds: infinitives in my Nyokon data have multiple marking, as in Tunen. However, the embedded non-finite verb appears before the object (VO; (28)). Unlike in Mous (2005)'s study, I found one example with OV syntax, although here the embedded clause was finite (29).

(28) Vivianə (nə) yár   a náám a kɔndáf
      Vivianne COP want A cook A pork
      'Vivianne wants to cook pork.'   (Nyokon; RA 51-2)

(29) Vivianə yár usáá   Roger kú a kɔndáf náám
      Vivianne want SM.1.say Roger TAM A pork   cook
      'Vivianne wants Roger to cook the pork.'   (Nyokon; RA 53)

The more accurate explanation for OV in (29) is likely the presence versus absence of a TAM marker, although it is also possible that the *kú* form in (29) is a verb rather than a TAM marker, therefore showing VO syntax with *a* licensing a postverbal object (S V [S V *a* [O V]]). This could in turn be related to the more general TAM-dependent alternation between OV and VO seen for Nyokon in

---

[26]An exception is a small number of inherently locative nouns, e.g. 'riverside', which do not require a preposition and can be taken as inherently licensed (see Mous 2003: 305).

§4.1 above. We therefore see that the infinitival pathway is harder to argue for as the source of OV syntax in Nyokon than it is for Tunen, at least with the limited data that are currently available.

4.4 Section summary

To sum up this research question, we have seen that there have been multiple grammaticalisation scenarios invoked in the literature for innovation of OV. While V>Aux(>TAM) grammaticalisation is crosslinguistically common, no clear examples were found for Tunen and Nyokon. Infinitival constructions were proposed by Mous (2005) as a possible grammaticalisation scenario for OV in Tunen and Nyokon, although Mous (2005) found no such OV examples in Nyokon. I showed that Nyokon retains VO in embedded non-finite clauses but could have OV in a finite example, while Tunen has OV consistently. I also highlighted a difference in the number of ɔ PREP/INF forms in my Tunen data and the 20th-century data of Dugast (1971); Mous (1997, 2003, 2005), which I suggested may relate to changes in nominal licensing (i.e. syntactic interpretation of an element as nominal versus verbal and thus requirement to be licensed by a preposition). In summary, the infinitival path still requires further development, but is an interesting possibility given the presence of the multiple *ɔ/a*-marking constructions in both languages. Such a grammaticalisation scenario could be better motivated if further relevant data are collected for Nyokon and neighbouring languages.

5 A note on Tikar

Before concluding, I will briefly comment on Tikar, given that it is the other Cameroonian language of the WALS sample which has been identified as having a (partial) OV system (Mous 2005, Dryer 2013), and is therefore of interest for understanding the development of OV syntax at the Bantu/Bantoid borderlands region. Recall from Figure 2 that Tikar is a close but not immediate neighbour of Tunen or Nyokon, being spoken approximately 50–150km to the North-East, and is classified as Bantoid (Northern Bantoid) rather than Narrow Bantu.

The classification of Tikar as having a mixed OV system is based primarily on the data from Stanley (1997), which are detailed but limited to sentence-level discussion. Stanley shows that Tikar has OV word order in imperfective tenses (30), while VO appears in the perfective (31) (Stanley 1997: 36).[27]

---

[27]The Tikar data are unchanged, with the exception of boldface/italics, translation of French lexical glosses into English, and the addition of an English free translation.

(30) à tǎ   *nye* **yìli**.
     he IPFP0 house sweep
     'Il balaie la maison.'
     ('He is sweeping the house.')   (Tikar; Stanley 1997: 103)

(31) mùn **kònnd-â** *kwìn*.
     I   add-PRF salt
     'J'ai ajouté du sel.'
     ('I have added salt.')   (Tikar; Stanley 1997: 139)

For Tikar, it is hard to fully evaluate the possible influence of IS due to the lack of discourse context in the sentences provided in the grammar. However, pronominal objects appear in the same position as lexical objects (e.g. (32)), which suggests that the givenness of the object does not affect its position relative to the verb, just as we saw above for RQ1 for Tunen and in contrast to the historical scenario discussed by Güldemann (2007).[28]

(32) à **ji-â**   *ɓon*.
     he eat-PRF them
     'Il les a mangés.' [sic.]
     ('He has eaten them.')   (Tikar; Stanley 1997: 248)

In terms of the other research questions RQ2 and RQ3, the data available in Stanley (1997) suggest that Tikar aligns with the results for Tunen and Nyokon. For example, it has VO (V-IO-DO) syntax in imperatives (33), N-Mod order, (34) and OV syntax of embedded non-finite verbs (35). Note however that there is no preposition or infinitival marking in (35), which is different from the Tunen and Nyokon construction (and thus significant for Mous 2005's infinitival grammaticalisation proposal).

(33) kònndi mùn nwɔ̀'.
     add   me meat
     'Donne-moi plus de viande.'
     ('Give me more meat.')   (Tikar; Stanley 1997: 139)

(34) kòn nywæb
     pot new
     'une marmite neuve'
     ('a new pot')   (Tikar; Stanley 1997: 273)

---

[28]Güldemann (2007) does mention data from Redden (1979) on Ewondo, a Bantu A72 language spoken near Tikar, which shows an IS-conditioned distinction in that only pronominal objects may be OV, and therefore would be a candidate for such a distinction holding synchronically.

(35) mùn yɛ̌   ɗunmi   ɓrìkì ywæli.
      I   IPFF1 AUX.begin bricks hit
      'Je vais commencer à fabriquer des briques.'
      ('I will start to make some bricks.')   (Tikar; Stanley 1997: 133)

In terms of grammaticalisation, Hyman (2011: 11-12) provides arguments based on Stanley (1997) that certain Aux components in Tikar are recent grammaticalisations from verbal sources, supporting a TAM-based V>Aux path from an earlier VO order. While a full analysis of Tikar requires further data on the language, these initial results suggest that it has a similar syntactic profile to Nyokon, differing from Tunen in having VO order in certain TAM contexts, and differing from both languages in its infinitival constructions.

6 Conclusion

This paper set out to investigate the rare OV word order patterns found in two languages at the Bantu/Bantoid borderlands region of Cameroon, testing the previous proposal of Mous (2003, 2005) in light of the possible influence of information structure as primary conditioning factor (Güldemann 2007, Good 2010). It was shown from new data with controlled discourse contexts that Tunen has a fully-established SOV order, while Nyokon has a partial SOV system with TAM as the primary synchronic conditioning factor. IS was shown to not be a directly significant factor conditioning order of the object and verb in either language. Both languages were shown to have a wide range of syntactic properties that fit the typical syntactic profile of a VO language, i.e. head-initial properties. Two grammaticalisation paths were then reflected upon. While V>Aux(>TAM) grammaticalisation is cross-linguistically common and likely, no evidence was found to support OV order having originated from this source in Tunen and Nyokon. Some extra data was then added to discussion of the possible grammaticalisation source via nominalisation in infinitival constructions sketched by Mous (2005). Finally, it was shown that Tikar patterns similarly to Nyokon in having a primarily TAM-based OV versus VO alternation, and differs from the Mbam languages in its infinitival constructions.

There are various questions for further work. One is empirical, as to whether other languages at the Bantu/Bantoid borderlands show OV patterns under certain TAM or IS contexts. This requires data collection on these languages, including discourse context (and ideally a corpus of natural speech). More detailed data on Nyokon is also important, especially for embedded non-finite clauses,

and considering the precise semantics of the TAM marker in relation to IS. Another question is what role contact has played. If OV is innovative, to what extent was it a shared innovation? This can be linked to the discussion of innovations in other domains in the proposed Macro-Sudan Belt region (Clements & Rialland 2008, Güldemann 2008, Hyman 2011). A more detailed comparison of the TAM systems of the Mbam languages would be of interest here. Finally, a syntactic analysis for each language would be desirable, considering other evidence for the structural height of the verb and object.

Abbreviations

Glossing conventions in this chapter follow the Leipzig Glossing Rules, with the following additions/differences.

Glossing conventions

1, 2, 3...	Bantu noun class marker	PST1	first-degree past tense (just now)
A	"Nyokon a form"		
ASSOC	associative marker	PST2	second-degree past tense (hodiernal)
IPFF1	imperfective "future premier degré" (from Stanley 1997)		
		PST3	third-degree past tense (hestiernal)
IPFP0	imperfective "non-passé" (from Stanley 1997)		
		PST4	fourth-degree past tense (ancient past)
IPFP1	imperfective "passé prémier degré" (from Stanley 1997)	REM	remote past
PREP	preposition	SM	subject marker
PRN	pronoun		

In-text abbreviations

ATR	Advanced Tongue Root	IS	Information structure
Aux	Auxiliary	Mod	Nominal modifier
C	Complementiser	NP	Noun phrase
Comp	Complement clause	O_{Pron}	Pronominal object
Compl	Complement	PCF	Predicate-centred focus
DP	Determiner Phrase	SOV	Subject-Object-Verb
\H	H grammatical tone (Mous 2022)	SVO	Subject-Verb-Object
Intd	Intended interpretation	TAM	tense/aspect/mood

Elisabeth J. Kerr

Acknowledgements

Appendix A RQ1 data sheet: TAM versus IS influence

This appendix provides supplementary supporting data for RQ1 of the paper "On OV and VO at the Bantu/ Bantoid borderlands".

A.1 TAM context
Present tense

Tunen OV, Nyokon VO

(36)   *Tunen present tense, focus on recipient object, OV:*
    Context: 'Who is the woman returning the calabash to?'

muəndú	á	ndɔ	imítə́	túmbiə	ɔ	mɔn.
/mɔ-əndú a		^Hndɔ	ɛ-mítə́	túmbiə	ɔ	mɔ-ná/
1-woman	SM.1 PRS		7-calebash	return	PREP	1-child

'The woman returns the calabash [to the child]_{FOC}.'   (Tunen; JO 1586)

(37)   *Nyokon present tense, VO/*OV:*

  a. | mu | nə | náám | alif | (ákitsia). |
     |---|---|---|---|---|
     | /mu | nə | náám | alif | ákitsia/ |
     | SM.1SG | PRS | cook | rice | today |

     'I am cooking rice today'   (Nyokon; RA 10)

b. *mu   nə alif náám (ákitsia)
/mu   nə alif náám ákitsia/
SM.1SG PRS rice cook   today
Intd.: I am cooking rice today.'   (Nyokon; RA 11)

Note that Mous (2022) distinguishes the Nyokon simple present from the present continuous, which differ in OV versus VO order.

Past tenses

Tunen OV, Nyokon OV
Tunen has four degrees of past tense, which in the affirmative have TAM markers *nɔ* PST1 (just now, a moment ago), *ná* PST2 (hodiernal), *ka* PST3 (hestiernal), and *lɛ* PST4 (ancient past). All have OV order.
Tunen, first/second/third-degree past tense, thetic, OV:

(38) a. Context: You are at the riverside outside the village and see an
elephant, which very rarely occurs, so run to tell the others.
mɛ   nɔ́ *misəku*   siəkin!
/mɛ   nɔ́ mi-səku   siəkinə/
SM.1SG PST1 3-elephant see.DUR
'I just saw an elephant!'   (Tunen; PM 316)

b. Context: Your friend asks what happened at church.
mɔtát   a   ná *imbə́nu yɛ*   *fəkin*   *nɛ́*   *Yɔ́səs   ɔ*
/mɔ-táta a   ná ɛ-mbə́nu yɛ   fəkinə   nɛ́   Yɔ́səsu ɔ
1-pastor SM.1 PST2 7-news   7.ASSOC 5.entrance 5.ASSOC Jesus   PREP
Yɛrúsalɛm nɔŋɔnak.
Yɛrúsalɛmɛ nɔŋɔn-aka/
Jerusalem   tell-DUR
'The pastor told the news of Jesus' entrance into Jerusalem.' (Tunen;
DM 166)

c. Context: 'What happened?'
yəmisə́   a   ka *wə́yíə́*   *mɔtɔ́á* lú.
/yamíá-isə́   a   ka wə́yíə́   mɔ-tɔ́á lúə/
POSS.PRON.1SG.9-9.father SM.1 PST3 POSS.PRON.1.3 3-car   sell
'My father sold his car.'   (Tunen; JO 2445)

Tunen, fourth-degree past, OV:

(39) mɛkɔ lɛ wamíá món ɔ́n.
      leopard PST4 POSS.PRON.1SG.1 1.child kill
      'Le léopard avait tué mon enfant.'
      ('The leopard killed my child.')   (Tunen; Dugast 1971: 182, adapted)

Nyokon past tense, OV, ?VO:

(40) mu álif náám akitsía
      /mu alif náám ákitsia/
      SM.1SG rice cook today
      'I cooked rice today.'   (Nyokon; RA 12)

(41) a. Context A: 'What did your father do?' (VP focus),
          Context B: 'What did your father kill?' (object focus)
          itɔ́ ángam ghó.
          father spider kill
          'My father [killed a spider]_FOC.'
          'My father killed [a spider]_FOC.'   (Nyokon; RA 216)
      b. Context: 'What did your father kill?' (object focus)
          ? itɔ́ ghó ángam.
            father kill spider
          Intd.: 'My father killed [a spider]_FOC.'   (Nyokon; RA 224)

More examples for Nyokon past tenses are available in Mous (2022). Mous (2005) also provides TAM paradigms for Tunen (for which see also Dugast 1971).

Future tenses

Tunen OV, Nyokon VO.
Tunen future tense, out-of-the-blue, OV:

(42) Samuɛ́lɛ á ŋɔ ɔlɛ́sa nɛ́ák.
      /Samuɛ́lɛ a ŋɔ ɔ-lɛ́sa nɛ́á-aka/
      1.Samuel SM.1 FUT 3-rice eat-DUR
      'Samuel will eat rice.' (tomorrow)   (Tunen; EE+EB 1656)

Tunen, future conditional, promoninal object, OV:

(43)   ábá bá   sɔ́   ŋɔ   bəsú   lúkə, bá   tuə́n.
         /ábá bá   sɔ́   ŋɔ   bəsúə́   lúkə bá   tuə́nə/
         if   SM.2 NEG FUT PRON.1PL feed SM.2 leave
         'If they don't feed us, then that's on them.'   (Tunen; EE 1701)

Nyokon, future tense, VO:

(44)   mə mu   náám álif (a   iŋgɛ)   aku'u.
         FUT SM.1SG cook   rice PREP toward evening
         'I will cook rice in the evening.'   (Nyokon; RA 13)

Nyokon, future conditional, VO:

(45)   nàá nə́   ngə́r mó.
         2PL COND.FUT have money
         'If you(pl) will have money.'   (Nyokon; Mous 2022: 19)

A.2 IS context
Thetics

Tunen OV, Nyokon VO/OV dependent on TAM.
Tunen, thetics, OV:

(46)   a.   Context: You enter the room and see a broken window. Someone
            announces...
            Biə́lɛ   a   ná   itúbə́   san.
            /Biə́lɛ   a   ná   ɛ-túbə́   sána/
            1.Pierre SM.1 PST2 7-window break
            'Pierre broke the window.'   (Tunen; EE+EB 1669)
         b.   Context: You are at the riverside outside the village and see an
            elephant, which very rarely occurs, so run to tell the others.
            mɛ   nɔ́   misəku   siəkin!
            /mɛ   nɔ́   mi-səku   siəkinə/
            SM.1SG PST1 3-elephant see.DUR
            'I just saw an elephant!'   (Tunen; PM 316)

c. Context: Your friend asks what happened at church.

    mɔtát   a   ná   imbə́nu   yɛ   fəkin   nɛ́   Yə́səs   ɔ
    /mɔ-táta a   ná   ɛ-mbə́nu   yɛ   fəkinə   nɛ́   Yə́səsu   ɔ
    1-pastor sm.1 pst2 7-news 7.assoc 5.entrance 5.assoc Jesus prep

    Yɛrúsalɛm   nɔŋɔnak.
    Yɛrúsalɛmɛ   nɔŋɔn-aka/
    Jerusalem   tell-dur

    'The pastor told the news of Jesus' entrance into Jerusalem.' (Tunen; DM 166)

Nyokon, thetic, present tense, VO:

(47)   mu   nə náám alif (ákitsia).
        /mu   nə náám alif ákitsia/
        sm.1sg prs cook rice today
        'I am cooking rice today.'   (Nyokon; RA 10)

Object focus

Tunen OV, Nyokon VO/OV dependent on TAM.
Tunen, term focus on theme object, OV:

(48)   a. Context: 'What did the woman give to the other woman?'
          a   nɔ́   ɔsɔ́kɔ́ hɛtɛ́tɛ́   indi
          /a   nɔ́   ɔsɔ́kɔ́ hɛ-tɛ́tɛ́   índíə́/
          sm.1 pst1 other 19-gourd give
          'She gave [a gourd]$_{FOC}$ to the other.'   (Tunen; PM 1541)

    b. Context: 'What is the man holding in his hand?'
          kalɔ́tɔ   á   ná itíə́ ɔ   mɔkat.
          /kalɔ́tɔ á-á   ná itíə́ ɔ   mɔ-kata/
          7.carrot cop-sm.1.rel pst2 hold prep 3-hand
          'He is holding [a carrot]$_{FOC}$ in his hand.'   (Tunen; JO 1630; Kerr to appear)

Nyokon, term focus on theme object, OV/?VO:

(49)   a. Context: 'What did your father kill?' (object focus)
          itə́   ángam ghó.
          father spider kill
          'My father killed [a spider]$_{FOC}$.'   (Nyokon; RA 216)

b. Context: 'What did your father kill?' (object focus)

? itə́   ghó ángam.
  father kill spider

Intd.: 'My father killed [a spider]_{FOC}.'   (Nyokon; RA 224)

Sub-NP focus

Tunen OV, Nyokon VO/OV dependent on TAM.
NB: In Tunen, there is a discontinuous noun phrase construction found for numeral, quantifier, and adjectival modifiers (S-O-V-Mod), which complicates this coding. However, sub-NP focus on the possessor shows that the word order follows the general OV pattern:
*Tunen, sub-NP focus on possessor, OV/*VO:*

(50)   Context: 'Whose book did Elisabeth buy?'

a. a   ka   híə́fúlə hέ   Jɔhána ɔ́nd.
   /a   ka   hɛ-əfulə hέ   Jɔhána ɔ́ndɔ/
   SM.1 PST3 19-book ASSOC.19 1.Jeanne buy

   'She bought Jeanne's book.'   (Tunen; EB+JO 2798)

b. *a   ka   híə́fúlə hέ   Jɔhána ɔ́nd.
   /a   ka   hɛ-əfulə hέ   Jɔhána ɔ́ndɔ/
   SM.1 PST3 buy   19-book ASSOC.19 1.Jeanne

   Intd.: 'She bought Jeanne's book.'   (Tunen; EB+JO 2801)

Nyokon, sub-NP focus on numeral, present tense, VO:

(51)   Context: 'How many chickens do you see?'
       táá   ndukŋ ikɔ'ɔ   itá.
       SM.1PL see   chickens three
       'We see [three]_{FOC} chickens.'   (Nyokon; RA 34)

VP focus

Tunen OV, Nyokon VO/OV dependent on TAM.
Tunen, VP focus, past tense, OV:

(52)   Context: What did Maria do?
       Malíá a ná bilə́liə fɔfɔ́kiə.
       /Malíá a   ná   bɛ-lə́liə   fɔfɔ́kíə́/
       1.Maria SM.1 PST2 8-varnish anoint.DUR
       'Maria [applied the varnish]_{FOC}.'   (Tunen; JO 2518)

Nyokon, VP focus, past tense, OV:

(53) Context A: 'What did your father do?' (VP focus),
      Context B: 'What did your father kill?' (object focus)
      itə́   ányam ghó.
      father spider kill
      'My father [killed a spider]꜀ꜰ꜀ꜱꜰ꜀.'
      'My father killed [a spider]꜀ꜰ꜀.'   (Nyokon; RA 216)

Nyokon, VP focus, present tense, VO:

(54) Context: 'What are you doing today?'
      taa   nə tsá'a   René.
      ꜱᴍ.1ᴘʟ ᴘʀꜱ receive René
      'We are [hosting René]꜀ꜰꜱ꜀.'   (Nyokon; RA 233)

Other PCF

Tunen OV, Nyokon VO/OV dependent on TAM.
Tunen, truth focus, present tense, OV:

(55) Context: 'Do you see the sheep?' (truth focus)
      mɛ́   nd(ɔ) ɛndɔ́mbá sin.
      /mɛ   ᴴndɔ ɛ-ndɔ́mbá sinə/
      ꜱᴍ.1ꜱɢ ᴘʀꜱ   7-sheep   see
      'I see the sheep.'   (Tunen; EO 695)

Tunen, state-of-affairs focus (verb focus), past tense, OV:

(56) Context: 'What did he do with the beans and the plantains?' (PCF)
      Context: 'What happened?' (thetic)
      a   ka   makɔnd͡ʒɛ nɛáka. a   ná   bilikó lu.
      /a   ka   ma-kɔnd͡ʒɛ néá-aka a   ná   bɛ-likó lu/
      ꜱᴍ.1 ᴘꜱᴛ3 6-plantain eat-ᴅᴜʀ ꜱᴍ.1 ᴘꜱᴛ2 8-bean sell
      'He [ate]꜀ꜰ꜀ the plantains. He [sold]꜀ꜰ꜀ the beans.' (JO 908; Kerr to
      appear)

Nyokon, truth focus, present tense, VO:

(57) Context: Kinyó is sick and unable to eat a lot. Someone asks 'Can Kinyó eat rice?'
áa, u   fʉr anyɛ́ álif.
yes sm.1 can eat   rice
'Yes, she [can]₍FOC₎ eat rice.'   (Nyokon; RA 151)

Other relevant data on the order of O and V

Pronominal objects

Pronominal objects pattern the same as lexical DPs in Tunen.
Tunen, pronominal object, OV:

(58) mɔndɔ ɔwá mɔ́ná á ndɔ naa a nɔ́ wɛ́ɛya ákánana ɔ ndɔ́kɛt
/mɔ-ndɔ ɔwá mɔ-ná a   ᴴndɔ ná   a   nɔ́ wɛ́ɛya ákánana   ɔ
1-person rel.1 1-child sm.1 prs   be.sick sm.1 pst1 pron.1 leave.appl prep
ndɔ́kɛta/
7.doctor
'The man whose child is sick took her to the clinic.'   (Tunen; PM 2179)

NB: In a paper published since this paper was submitted, Mous reports postverbal object pronouns for past tense in Nyokon:

(59) ò   lìá'   və̂.
ò   lə̀:K\H   və̂
3sg say:rem\H O3pl
'He had told them.'   (Nyokon; Mous 2022:15)

Negation

Order of O and V in negative clauses matches order of O and V in affirmative clauses in both Tunen and Nyokon.
Tunen, negation, pronominal and full noun phrase object, OV:

(60) a. mɛ   lɛ   aŋɔ́á   nimb.   (Neg-O-V, Tunen)
/mɛ   lɛ   aŋɔ́á   nimbə/
sm.1sg neg pron.2sg.emph deceive
'Je ne te trompe pas.'
'I'm not lying to you.'

b. mɛ   lɛ́   ndɔ tunəni ɔ́kɔ.   (Neg-O-V, Tunen)
/mɛ   lɛ   ᴴndɔ tu-nəní ɔ́kɔ/
SM.1SG NEG PRS   13-Nen understand
'I don't understand the Tunen language.'

Nyokon, negation, present tense, VO:

(61) Kinyó sɨ́   nyɛ́ alif a.
Kinyo NEG eat rice A
'Kinyo doesn't eat rice.'   (Nyokon; RA 126)

(62) ḿ sɨ́   swɔ́ ákɨ́ná.
1SG NEG wash calabash:F
'I don't wash the calabash (never).'   (Nyokon; Mous 2022: 25)

There is no difference in order of O and V between negation in matrix versus embedded clauses.

Embedded clauses

Order of O and V in embedded clauses matches order of O and V in matrix clauses in Tunen and Nyokon.
*Tunen, embedded clauses, OV/*VO:*

(63) a. mɛ́   ndɔ manya ɔwá Matɛ́ŋɛ a   ka   hiəfulə fanak.
/mɛ   ᴴndɔ manya ɔwá Matɛ́ŋɛ a   ka   hɛ-əfulə fana-aka/
SM.1SG PRS   know REL.1 1.Martin SM.1 PST3 19-book read-DUR
'I know that Martin has read the book.'   (Tunen; JO 905)

b. Malíá a   ná   láá ásɛa   Jɔhánɛs(ɛ) á   ndɔ bilíbílíbí
/Malíá a   ná   laa a-séá   Jɔhánɛsɛ a   ᴴndɔ bɛ-líbílíbí
1.Maria SM.1 PST2 say SM.1-say 1.Johannes SM.1 PRS   8-chili.pepper
nyɔ   ɔ   wayéá   ɔ́mbɛ́l.
nyɔ   ɔ   wayéá   ɔ-mbɛ́la/
cultivate PREP POSS.PRON.1.3 3-house
Maria said that Johannes grows chillies at home.'   (Tunen; JO 2450)

c. *Malíá a   ná   láá ásɛa   Jɔhánɛs á   ndɔ nyɔ
/Malíá a   ná   laa a-séá   Jɔhánɛsɛ a   ᴴndɔ nyɔ/
1.Maria SM.1 PST2 say SM.1-say 1.Johannes SM.1 PRS   cultivate

bilíbílíbí.
bɛ-líbílíbí/
8-chili.pepper
Maria said that Johannes grows chillies.'   (Tunen; JO 2451)

Nyokon, embedded clauses:

(64) Vivianə yár usáá   Roger kú a kɔndáf náám
      Vivianne want SM.1.say Roger TAM A pork   cook
      'Vivianne wants Roger to cook the pork.'   (Nyokon; RA 53)

Definiteness

Order of O and V with definite objects matches order of O and V with indefinite objects. Definiteness is generally unmarked, but can be seen in different discourse contexts and in Tunen through the use of the specific indefinite determiner *-mɔtɛ́* (see Kerr 2020) and modification by possessive pronouns (indicating definiteness).

Tunen, definite/non-specific indefinite/specific indefinite object, OV:

(65) mɛ́   ndɔ mɔ́ndɔ si.
      /mɛ   ᴴndɔ mɔ-ndɔ siə/
      SM.1SG PRS   1-person search
      'I'm looking for {someone/a person/the person}.'   (Tunen; JO 898; Kerr 2020: 246)

Tunen, indefinite object, OV:

(66) Context: You are looking for your friend Daniel.
      mɛ́   ndɔ wɔmɔté mɔ́ndɔ   si.   neayá   nínyə á Təniɛl.
      /mɛ   ᴴndɔ ɔ-ᴴmɔté mɔ-ndɔ siə   neayá   nɛ-nyə á Təniɛ́lɛ/
      SM.1SG PRS   1-one   1-person search 5.POSS.PRON.1 5-name COP 1.Daniel
      'I'm looking for someone. His name is Daniel.' (Tunen; JO, 891; Kerr 2020: 246)

*Tunen, definite object, OV/*VO:*

(67) Context: 'What happened?'
   a. yəmisə́   a   ka   wə́yíə́   mɔtɔ́á lú.
      /yamíá-isə́   a   ka   wə́yíə́   mɔ-tɔ́á lúə/
      POSS.PRON.1SG.9-9.father SM.1 PST3 POSS.PRON.1.3 3-car   sell
      'My father sold his car.'   (Tunen; JO 2445)

b. *yəmisə́   a   ka   lúə wə́yíə́   mɔtɔ́á.
   /yamíá-isə́   a   ka   lúə wə́yíə́   mɔ-tɔ́á/
   POSS.PRON.1SG.9-9.father SM.1 PST3 sell POSS.PRON.1.3 3-car
   Intd.: 'My father sold his car.'   (Tunen; JO 2446)

Appendix B RQ2 data sheet: Head-finality versus head-initiality

This appendix provides supplementary supporting data for RQ2 of the paper "On OV and VO at the Bantu Bantoid borderlands".

B.1 Order of N and Mod

Tunen N-Mod, D-NP order; Nyokon N-Mod, D-NP order.
Tunen, N-Mod, D-NP order:

(68) tɔ́ɔ́yɛ   tɔbanána tɔté꜖té   tɔfítitiə tɔ́fandɛ
      /tɔ́ɔ́yɛ   tɔ-banána tɔ-têtéá tɔ-fítitiə tɔ-ᴴfandɛ /
      13.DEM.PROX 13-banana 13-small 13-black 13-two
      'these two small black bananas'   (Tunen; JO 885)

Nyokon, N-Mod, D-NP order:

(69) a. pí   pimbɔ́tɔ pífu
         DEM.PROX bananas two
         'these two bananas'   (Nyokon; RA 191)

      b. pimbɔ́tɔ kúpóm
         bananas ripe
         'ripe bananas'   (Nyokon; RA 188)

Adposition type

Tunen prepositions, Nyokon prepositions:

(70) a. Context: Where are you?
         mɛ lɛ ɔ nioní.
         /mɛ   léá ɔ   nɛ-oní/
         SM.1SG be PREP 5-market
         'I am at the market.'   (Tunen; PM 102)

b. Context: 'Where are you?'

mu   nə   a   nyí.
SM.1SG COP PREP market

'I am at the market.'   (Nyokon; RA 94)

NB: There are some elements that appear like postpositions but are derived from nouns, e.g. *nuumə* 'inside'; see Dugast (1971) and Mous (2003:305).

B.2 Order of N and Poss (pronoun)

Tunen Poss$_{PRON}$-N; Nyokon Poss$_{PRON}$-N.
Tunen, Poss$_{PRON}$-N (for subjects, direct objects, indirect objects):

(71) a. wàmɛ   mòna lì ò   mǐm.
          POSS.PRON.1SG.1 1.child be PREP house
          'Mon enfant est dans la case.'

     b. níamìa   néhɔka nánìmin
         POSS.PRON.1SG.5 5-axe   PAST2-disappear
         'Ma hache a disparu.'

     c. mɛ̀kɔ   ná   yàm   ìmwiny ˈɛ́tà
         leopard PST2 PRON.POSS.1.7 7.goat   take
         'Le léopard a pris ma chèvre.'

     d. mɛ́   ndò wə́m   úmbienyi tìlin
         SM.1SG PRS POSS.PRON.1SG.3 sibling   write
         'J'écris à mon neveu utérin.'   (Tunen; Dugast 1971: 135)

Nyokon, possesive pronouns, Poss$_{PRON}$-N.

(72) Context: After visiting a friend's house, you announce:
      punú   páa   liak tsú.
      POSS.2PL children cry much
      'Your children cry a lot.'   (Nyokon; RA 7)

See §4.3 of the main paper re: N-Poss$_{PRON}$ order in certain frozen possessive expressions.

B.3 Order of N and Gen

Tunen N-Gen, Nyokon N-Gen.

Non-pronominal possessives are formed with the associative (aka connective, genitive) construction N$_1$-Assoc-N$_2$.

Tunen, N and Gen, N-Gen:

(73) a. *mùtɛká wà mònd* 'l'esclave de l'homme' ('the man's slave'), pl. *bàtɛká bá mònd* (1/2)

b. *mìɔkɔ́ yè mùənd* 'la poule de la femme' ('the woman's chicken'), pl. *mìɔkɔ́ yé múənd* (9/10) (Tunen; Dugast 1971: 133).

Nyokon, N and Gen, N-Gen:

(74) unyám yímambəŋ
man ASSOC.POSS.PRON.1SG.friend
'le mari de mon amie' (my friend's husband) (Nyokon; RA 179)

(75) pɨ́ nòó ŋgə́r ápín mápín mì twin.
3PL PRF must dance 6.dance 6:of twins
'One has to dance the twin-dance.' (Nyokon; Mous 2022: 18)

B.4 Canonical order of O and V
Tunen OV, Nyokon OV/VO dependent on TAM. The canonical order (aka 'basic word order') is based on thetics and VP focus (see section 1.2).

B.5 Order of O and V in imperatives
Tunen VO (V-IO-DO), Nyokon VO (V-DO-IO).
Tunen, Nyokon, imperatives, VO:

(76) a. índíə́ mɔná imit!
/índíə́ mɔ-ná ɛ-mítə́/
give 1-child 7-calabash
'Give the child a calabash!' (Tunen; JO 1594)

b. nda manyí ngɔ́m!
give water PRON.1SG
'Give me water!' (Nyokon; RA 1)

Note that Tunen imperatives with pronominals are still VO. Finally, RA confirms remotely that the non-pronominal equivalent of the Nyokon example (76b) cannot be OV (while *nda manyí apə́ ghó!* ('give the child water!', V-DO-Prep-IO) is grammatical).

B.6 Order of Aux and V

Tunen Aux-V, Nyokon Aux-V.

If TAM markers are considered as auxiliary elements (as in 'S-Aux-O-V-Other' word order), then both languages are Aux-V. For true auxiliaries, i.e. auxiliary verbs, the order Aux-V is found.

(77) tɔ   ná   tíkɔ́ sáá.
    /tɔ   ná   tíkɔ́ sá/
    SM.1PL PST2 stay come
    'We will follow.'   (Tunen; PM 1058)

B.7 Order of C and Comp

Tunen C-Comp, Nyokon C-Comp.
Tunen, C-Comp:

(78) a. mɛ́   ndɔ manya ɔwá Matéŋɛ a   ka   hiɔfulə fanak.
       /mɛ   ᴴndɔ manya ɔwá Matéŋɛ a   ka   hɛ-əfulə fana-aka/
       SM.1SG PRS   know   REL.1 1.Martin SM.1 PST3 19-book read-DUR
       'I know that Martin has read the book.'   (Tunen; JO 905)

   b. Malíá a   ná   láá ásɛa   Jɔhánɛs(ɛ) á   ndɔ bilíbílíbí
      /Malíá a   ná   laa a-séá   Jɔhánɛsɛ a   ᴴndɔ bɛ-líbílíbí
      1.Maria SM.1 PST2 say SM.1-say 1.Johannes SM.1 PRS   8-chili.pepper
      nyɔ   ɔ   wayɛ́á   ɔ́mbɛ́l.
      nyɔ   ɔ   wayɛ́á   ɔ-mbɛ́la/
      cultivate PREP POSS.PRON.1.3 3-house
      'Maria said that Johannes grows chillies at home.'   (Tunen; JO 2450)

Nyokon, C-Comp:

(79) a. pu   lə pusáá mu   fɛ́.
       /pu   lə pu-sáá mu   fɛ́/
       SM.3PL say 3PL-say SM.1SG come
       'They said that I should come.'   (Nyokon; RA 41)

   b. Vivianə yár usáá   Roger kú   a kɔndáf náám
      Vivianne want SM.1.say Roger TAM A pork   cook
      'Vivianne wants Roger to cook the pork.'   (Nyokon; RA 53)

B.8 Order of Cop and Compl

Tunen Cop-Compl, Nyokon Cop-Compl.
Tunen/Nyokon, locative complements, Cop-Compl:

(80)   a. Context: Where are you?
        mɛ lɛ ɔ nioní.
        /mɛ   lɛ́á ɔ   nɛ-oní/
        SM.1SG be PREP 5-market
        'I am at the market.'   (Tunen; PM 102)

       b. Context: 'Where are you?'
          mu   nə a   nyí.
          SM.1SG COP PREP market
          'I am at the market.'   (Nyokon; RA 174)

Tunen/Nyokon, predication, Cop-Compl:

(81)   a. Context: 'She was already set up.'
        á muəndú anyam.
        /á   mɔ-əndú a-nyama/
        COP 1-woman 1-brave
        'She's a brave/impressive woman.'   (Tunen; EO 1020)

       b. mu   nə mutsʉgʉ.
          SM.1SG COP teacher
          'I am a teacher.'   (Nyokon; RA 178)

NB: Cop-Compl also applies in negative copular clauses.

B.9 Order of V and manner adverb

Tunen, V-Adv (*Adv-V); Nyokon V-Adv (*Adv-V).
Tunen/Nyokon, manner adverbs, V-Adv:

(82)   a. Context: PM and EO are discussing how PD should be considered a
        true Munen [i.e. a local], despite being born somewhere else.
        aka nyɔɔkɔ háaha ɔ uwəsú mɔŋɛŋ.
        /a ka nyɔ-aka háaha ɔ   uwəsúə́   mɔŋɛ́ŋa/
        SM.1 PST3 work-DUR here   PREP PRON.POSS.1PL much
        'He worked a lot here in our region.'   (Tunen; EO 1043)

(83) a. Malíá a ná nyɔ́kɔ́ biabia.
/Malíá a ná nyɔ́-aka biabia/
1.Maria SM.1 PST2 work-DUR slowly
'Maria worked slowly.' (Tunen; JO 2560)

b. *ɛndánáná yɛ́ ˈná biabia yɔlaka.
/ɛ-ndánáná yɛ́ ˩ná biabia yɔ́la-aka/
7-ice SM.7 PST3.REL slowly melt-DUR
Intd.: 'The ice melted slowly.' (Tunen; JO 2558)

(84) a. Context: After visiting a friend's house, you announce:
punú páa liak tsú.
POSS.2PL children cry much
'Your children cry a lot.' (Nyokon; RA 7)

b. Context: As above.
*punú páa tsú liak.
POSS.2PL children much cry
Intd.: 'Your children cry a lot.' (Nyokon; RA 8)

B.10 Canonical adjunct position

Tunen: SOVX, Nyokon SOVX/SVOX dependent on TAM.
Tunen, adjuncts, SOVX:

(85) a. Context: 'Where did he build his house?'
a ka yayɛ́á miímə lúmə́kɔ́ ɔ iNdíki naánɛkɔl.
/a ka yayɛ́a miímə lúm-aka ɔ iNdíki naánɛkɔla/
SM.1 PST3 POSS.PRON.1.3 3.house build-DUR PREP Ndiki yesterday
'He built his house yesterday in Ndiki.' (Tunen; JO 1121)

b. Context: PM and EO are discussing how PD should be considered a
true Munen [i.e. a local], despite being born somewhere else.
a ka nyɔɔkɔ háaha ɔ uwəsú mɔŋɛŋ.
/a ka nyɔ-aka háaha ɔ uwəsúɔ́ mɔŋɛ́ŋa/
SM.1 PST3 work-DUR here PREP PRON.POSS.1PL much
'He worked a lot here in our region.' (Tunen; EO 1043)

Nyokon, adjuncts, SOVX/SVOX dependent on TAM:

(86) mu nə náám alif ákitsia.
/mu nə náám alif ákitsia/
SM.1SG PRS cook rice today
'I am cooking rice today.' (Nyokon; RA 10)

(87)   Context: 'Where did you put the clothes?'
mu   piyɛ   ghá avuə.
sm.1sg clothes put outside
'I put the clothes outside.'   (Nyokon; RA 72)

B.11 Low subjects (VS)?

Tunen: No, Nyokon: No.
   Tunen, low subjects, *VS:

(88)   a. *a   ka   nyɔkɔ   kíŋə   naánɛkɔl.
           /a   ka   nyɔ-aka   kíŋə   naánɛkɔla/
           sm.1 pst3 work-dur 7.chief yesterday
           Intd.: 'The chief worked yesterday.'   (Tunen; EO 2894)

       b. *a   ná   fámá mɔndɔ   həmə́tu.
           /a   ná   fámá mɔ-ndɔ   həmə́tuə/
           sm.1 pst2 arrive 1-person quickly
           Intd.: 'A man suddenly appeared.'   (Tunen; PM WA.70)

       c. *yɛ́ ná   yɔlaka   ɛndánáná biabia.
           /yɛ́ ná   yɔ́la-aka ɛ-ndánáná biabia/
           sm.7 pst2 melt-dur 7-ice   slowly
           Intd.: 'The ice melted slowly.'   (Tunen; PM WA.70)

   Nyokon, low subjects, *VS:[29]

(89)   a. *ghó itə́   anyam.
           kill   father spider
           Intd.: 'My father killed the spider.'   (Nyokon; RA 220)

       b. *tsə́s   vəs.
           appear somebody
           Intd.: 'Somebody came out.'   (Nyokon; RA 206)

       c. *ngərk alif.
           burn   rice
           Intd.: 'the rice burnt.'   (Nyokon; RA 209)

---

[29]Compare the grammatical SV versions: (a) *itə́ anyam ghó* 'My father killed the spider' (RA 216)
(good as A to question 'What did your father do?'); *vəs tsə́s váás apús* 'Someone appeared in
front of us' (RA 204); *alif ngərk* 'The rice is burnt' (RA 208); *Mari kiə a itsə́r* 'Marie walked
quickly' (RA 210).

d. *kɨə Mari a itsə́r.
walk Marie ᴀ quickly
Intd.: 'Marie walked quickly.'   (Nyokon; RA 213)

NB: The unavailability of VS order applies whenever the verb precedes the lexical subject (regardless of the position of the SM-TAM cluster). The pattern holds across verb types (transitive, verbs of appearance, unaccusatives, unergatives) and the sentences above are grammatical if the subject is in initial position (i.e., before the SM-TAM cluster).

References

Anderson, Gregory. 2007. *Auxiliary verb constructions.* Oxford: Oxford University Press.

Biberauer, Theresa. 2017. The Final-Over-Final Condition and particles. In Michelle Sheehan, Theresa Biberauer, Ian Roberts & Anders Holmberg (eds.), *The Final-Over-Final Condition: A syntactic universal*, 187–296. Cambridge: MIT Press.

Bikoi, Charles Binam (ed.). 2003. *Atlas linguistique du Cameroun: Tome 1: Inventaire des langues.* Yaoundé: Cerdotola.

Blench, Roger. 2006. *Language, archaeology and the African past.* Lanham: AltaMira Press.

Blench, Roger. 2022. *An overview of Bantoid: Recent research showing light on its relation to Bantu.* Presentation given at the 1st conference on Bantoid languages and linguistics (Banto1d). Hamburg: Universität Hamburg.

Boyd, Virginia L. 2015. *The phonological systems of the Mbam languages of Cameroon with a focus on vowels and vowel harmony.* Utrecht: LOT.

Claudi, Ulrike. 1993. *Die Stellung von Verb und Objekt in Niger-Kongo-Sprachen: Ein Beitrag zur Rekonstruktion historischer Syntax.* Köln: Universität zu Köln.

Clements, George N. & Annie Rialland. 2008. Africa as a phonological area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 36–85. Cambridge: Cambridge University Press.

Creissels, Denis. 2005. S-O-V-X constituent order and constituent order alternations in West African languages. In Yuni Kim & Rebecca Cover (eds.), *Proceedings of the Thirty-first Annual Meeting of the Berkeley Linguistics Society, Special session on languages of West Africa*, 37–51. Berkeley: Berkeley Linguistics Society.

Downing, Laura J. & Larry M. Hyman. 2016. Information structure in Bantu. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 790–813. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199642670.013.010.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68. 81–138.

Dryer, Matthew S. 2013. Order of object and verb. In Martin Dryer Matthew S. & Haspelmath (ed.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/83 (11 February, 2022).

Dugast, Idelette. 1971. *Grammaire du Tunen*, vol. 8. Paris: Klincksieck.

Dugast, Idelette. 1975. *Contes, proverbes et devinettes des Banen: Sud-Ouest du Cameroun*, vol. 12. Paris: Société d'études linguistiques et anthropologiques de France.

Gensler, Orin D. 1994. On reconstructing the syntagm S-Aux-O-V-Other to Proto-Niger-Congo. In David Peterson Kevin Moore & Comfort Wentum (eds.), *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: Special session on historical issues in African linguistics*, 1–20. Berkeley, CA: Berkeley Linguistics Society.

Gensler, Orin D. 1997. Grammaticalization, typology, and Niger-Congo word order: Progress on a still-unsolved problem. Review of *'Die Stellung von Verb und Objekt in Niger-Kongo-Sprachen'*, by Ulrike Claudi. *Journal of African Languages and Linguistics* 18. 57–93.

Gensler, Orin D. & Tom Güldemann. 2003. *S-Aux-O-V-Other in Africa: Typological and areal perspective*. Paper presented at the Workshop "Distributed Predicative Syntax (S P O V X)", 4th World Congress of African Linguistics (WOCAL), Rutgers University, 21 June 2003.

Givón, Talmy. 1975. Serial verbs and syntactic change: Niger-Congo. In Charles N. Li (ed.), *Word order and word order change*, 47–112. Austin: University of Texas Press.

Good, Jeff. 2010. Topic and focus fields in Naki. In Ines Fielder & Anne Schwarz (eds.), *The expression of information structure: A documentation of its diversity across Africa*, 35–67. Amsterdam: John Benjamins.

Güldemann, Tom. 2007. Preverbal objects and information structure in Benue-Congo. In Katherina Hartmann Aboh Enoch Oladé & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 83–112. Berlin: De Gruyter.

Güldemann, Tom. 2008. The Macro-Sudan belt: Towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151–185. Cambridge: Cambridge University Press.

Güldemann, Tom. 2018. Historical linguistics and genealogical language classification in Africa. In Tom Güldemann (ed.), *The languages and linguistics of Africa* (The World of Linguistics 11), 58–444. Berlin: Mouton de Gruyter.

Heine, Bernd. 1980. Language typology and linguistic reconstruction: The Niger-Congo case. *Journal of African Languages and Linguistics* 2:2. 95–112.

Heine, Bernd & Ulrike Claudi. 2001. On split word order: Explaining syntactic variation. *General Linguistics* 38. 41–74.

Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.

Heine, Bernd & Tania Kuteva. 2004. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Hyman, Larry M. 1975. On the change from SOV to SVO: Evidence from Niger-Congo. In Charles N Li (ed.), *Word order and word order change*, 113–147. Austin: University of Texas Press.

Hyman, Larry M. 2011. The Macro-Sudan Belt and Niger-Congo reconstruction. *Language dynamics and change* 1. 3–49.

Hyman, Larry M. & John R. Watters. 1984. Auxiliary focus. *Studies in African Linguistics* 15. 233–273.

Kerr, Elisabeth J. 2020. The grammaticalisation of a specific indefinite determiner: Prenominal -*mɔ̀té* in Tunen. In Jenneke van der Wal, Heleen J. Smits, Sara Petrollino, Victoria A. S. Nyst & Maarten G. Kossmann (eds.), *Essays on African languages and linguistics: In honour of Maarten Mous*, 241–266. Leiden: African Studies Centre Leiden Occasional Publications.

Kerr, Elisabeth J. to appear. The expression of information structure in Tunen (A44, Cameroon). In Jenneke van der Wal (ed.), *The expression of information structure in Bantu* (Contemporary African Linguistics). Berlin: Language Science Press.

Kerr, Elisabeth J., Allen Asiimwe, Patrick Kanampiu, Ernest Nshemezimana, Zhen Li & Jenneke van der Wal. 2023. Bantu word order between discourse and syntactic relations. *Linguistique et Langues Africaines* 9(1). DOI: 10.4000/lla.9496.

Koopman, Hilda J. 1984. *The syntax of verbs: From verb movement rules in the Kru languages to Universal Grammar*. Tilburg University. (Doctoral dissertation).

Lovestrand, Joseph. 2011. *Notes on Nyokon phonology (Bantu A.45, Cameroon)*. Manuscript. Yaoundé.

Lovestrand, Joseph. 2020. *Nyokon word list audio recordings.* DOI: 10.5281/zenodo. 3962412.

Maho, Jouni. 2003. A classification of the Bantu languages: An update of Guthrie's referential system. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages.* London: Routledge.

Maho, Jouni. 2009. *NUGL online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages.* (https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf).

Marten, Lutz. 2020. Bantu and Bantoid. In Rainer Vossen & Gerrit J. Dimmendaal (eds.), *The Oxford handbook of African languages*, 205–219. Oxford: Oxford University Press.

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana linguistica* 3(1). 79–121.

Mous, Maarten. 1997. The position of the object in Tunen. In Rose-Marie Déchaine & Victor Manfredi (eds.), *Object positions in Benue-Kwa*, 123–37. The Hague: Holland Academic Graphics.

Mous, Maarten. 2003. Nen (A44). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 283–306. London: Routledge.

Mous, Maarten. 2005. The innovative character of object-verb word order in Nen (Bantu A44, Cameroon). In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics with special focus on Bantu and Mande*, 411–24. Tervuren: Royal Museum for Central Africa.

Mous, Maarten. 2014. TAM-full object-verb order in the Mbam languages of Cameroon. *ZAS Papers in Linguistics* 57. 72–94.

Mous, Maarten. 2022. Tone and the verbal systems of Nyokon. *Journal of West African Languages (JWAL)* 49. 1–33.

Nikitina, Tatiana. 2011. Categorial reanalysis and the origin of the S-O-V-X word order in Mande. *Journal of African Languages and Linguistics (JALL)* 32. 251–273. DOI: 10.1515/JALL.2011.009.

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

Redden, James E. 1979. *A descriptive grammar of Ewondo.* Carbondale: Department of Linguistics, Southern Illinois University.

Rolle, Nicholas, Florian Lionnet & Matthew Faytak. 2020. Areal patterns in the vowel systems of the Macro-Sudan Belt. *Linguistic Typology* 24(1). 113–179. DOI: 10.1515/lingty-2019-0028.

Sande, Hannah, Nico Baier & Peter Jenks. 2019. The syntactic diversity of SAuxOV in West Africa. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and description in African linguistics: Selected papers from the 47th An-*

nual Conference on African Linguistics (Contemporary African Linguistics 4), 667–701. Berlin: Language Science Press. DOI: 10.5281/zenodo.3367195.

Simons, Gary F. & Charles D. Fennig (eds.). 2018. *Ethnologue: Languages of the world (21st ed.)* Dallas, Texas: SIL International.

Skopeteas, Stavros, Ines Fiedler, Samantha Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry & Manfred Krifka. 2006. *Questionnaire on information structure (QUIS): Reference manual.* Potsdam: Universitätsverlag Potsdam.

Stanley, Carol. 1997. *Description morpho-syntaxique de la langue Tikar (parlée au Cameroun).* Yaoundé: Société Internationale de Linguistique (SIL).

Watters, John R. 2003. Grassfields Bantu. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages.* London: Routledge.

Welaze Kongne, Jacquis. 2010. Morphologie verbale du Tunen. Diplôme d'études approfondies (DEA).

Williamson, Kay. 1965. *A grammar of the Kolokuma dialect of Ịjọ* (West African Mongraphs 2). Cambridge: University Press.

Williamson, Kay. 1986. Niger-Congo: SVO or SOV? *Journal of West African Languages* 16. 5–14.

Chapter 13

The position of Asá and Qwadza within Cushitic

Iris Kruijsdijk

Leiden University

This paper aims to re-examine lexical data on Asá and Qwadza in order to find out how they relate to each other and to other languages, in particular to the West Rift Southern Cushitic languages and their reconstruction. The data is analyzed on a phonological and morphological level. Some of the phones of the languages are compared in order to find regular sound correspondences, and certain morphemes are compared, providing more insight into common innovations, regarding both the nominal and verbal domain. This paper aims to answer the following research questions: "Are Asá and Qwadza Cushitic languages?" and "Are Asá and Qwadza Southern Cushitic languages forming East Rift Southern Cushitic?". The findings show that Asá and Qwadza are very likely Cushitic languages, but that there is no evidence for an East Rift Southern Cushitic group based on the evidence presented in this paper.

1 Introduction

The languages this paper is concerned with are Asá and Qwadza. These are two languages that were spoken in Northern Tanzania, but are now extinct. Asá and Qwadza are two languages that are classified together as East Rift Southern Cushitic, as can be seen in Figure 1. According to Blažek (2019), Asá and Qwadza can be classified within Southern Cushitic. Based on his lexicostatistics, Qwadza splits off first, then Asá (Blažek 2019: 44).

Data of these two languages were collected when there were only remember-ers of the languages, not active speakers. Speakers of Asá shifted to Maasai, a
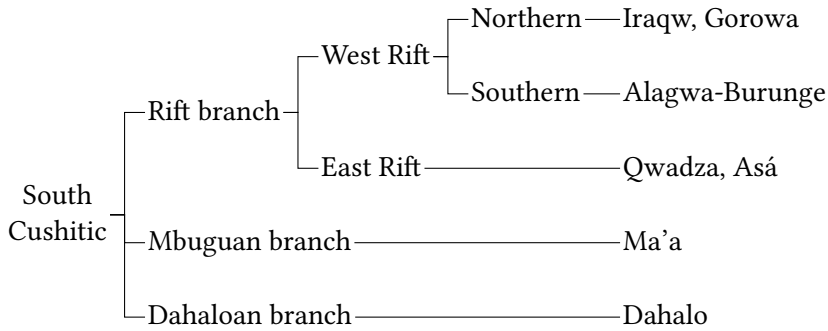
Figure 1: Southern Cushitic (Ehret 1980: 132, adapted)

Nilotic language, and speakers of Qwadza to Gogo, a Bantu language. It is important to keep this situation in mind when looking at the data. After a discussion of the previous research and the data used for this paper, the research questions will be answered. The first research question, "Are Asá and Qwadza Cushitic languages?", will be addressed in §2 and the second research question, "Are Asá and Qwadza Southern Cushitic languages forming East Rift Southern Cushitic?" in §3. These questions will be answered by looking at the phonological evidence to identify sound correspondences, and by looking at the morphology that is present on the lexemes to compare form, distribution and function. Finally, it is concluded that Asá and Qwadza are very likely Cushitic languages, but that there is no evidence for an East Rift Southern Cushitic group based on the evidence presented in this paper.

1.1 Previous research

Among the oldest sources on Asá is Merker (1904). Merker (1904) contains an ethnographic description and some lexical items. The linguistic data of this source is found in Fleming (1969). Winter (1979) collected much more extensive data from the last speakers of Asá, the language they remember from their youth. This is the data Ehret adapted and used (Ehret 1980). Petrollino & Mous (2010) were the last to collect data on Asá.

The earliest data for Qwadza were collected in 1908 by colonial administrators who, most likely, worked mainly with rememberers, who were shifting to Gogo (Bantu). The two sources, Claus (1910) and Kohl-Larsen (1943) both contain linguistic data. Pearce (1954) contains an unpublished wordlist, which is referred to in Whiteley (1958). Ehret collected data from 1973-74. Lastly, Kießling reports a word list of Qwadza (Kießling 1999). The linguistic data on Qwadza are probably

more reliable in the two earliest collections even though these were collected by non-linguists. The largest collection is Ehret (1974), who was possibly too late to collect reliable data. The same goes for Asá; the data were collected when speakers were shifting languages, and when it was possibly too late for reliable data. This means that the data have to be taken with care.

The Gogo name for the Qwadza people and their clan among the Gogo is Ng'omvia. This is also the name used in the literature, until Ehret uses the name the people use themselves, which is Kw'adza. Variants of that name are currently most used. I will be using Qwadza.

1.2 Data

As mentioned above, it is important to point out that all of the data collection on these languages happened during the shift to another language, or when it was only remembered. In addition, the time period in which data could be collected was fairly short, and data collectors mainly focused on lexical information. Also, it was conducted partly by non-linguists. For Asá, the data consist of 334 lexemes, including derivations. The data are all taken from Ehret (1980). The lexemes in this dataset can be subdivided into several domains: There are around 60 terms for animals, consisting of terms for non-domestic animals and general animal terms, like the word for *du'umok* 'leopard' and terms relating to domestic animals, like *ŋeris* 'to herd, or feed'. There are around 38 words referring to movement, often involving the body. There are 33 nouns denoting body parts and 28 nouns relating to the environment. There is also a category of abstract lexemes, a group of 28 words, including both nouns, like the word *lawala* 'truth', and verbs, like the word *šah-* 'to know'. There are 25 lexemes that refer to cultivating and cooking. There are some 20 lexemes that have to do with hunting, including words relating to honey. There are also some adjectives, quantifiers, numerals, and possessive pronouns.

The Qwadza data consist of 973 lexemes, including derivations. The data are taken from Ehret (1980), Kießling (1999) (unpublished word list) and Kohl-Larsen (1943). There is a large group of general terms (156) relating to animals, and within that group, terms for wild animals, like *tsaayiko* 'weaver bird', but also a group of 38 words relating to domestic animals. There are 100 words for body parts. Just like in Asá, some of the semantic categories concern terms relating to environment (92), cultivating (92) and hunting (25). Qwadza has relatively more words to refer to people or their status (56). Something that is not present in the Asá data, but is present in the Qwadza data, are terms for what people wear, like *daliko* 'women's upper garment'.

2 Are Asá and Qwadza Cushitic languages?

This section examines the evidence there is to classify Asá and Qwadza as Cushitic languages. It has been shown that Asá is a Cushitic language by Fleming (1969). He bases this on the types of lexical items that occur in the languages, showing that lexical items that are not easily borrowed are cognate to other Cushitic languages and showing that such cognates do not exist with Nilotic and Bantu. In addition, Ehret claims Asá and Qwadza are Cushitic languages based on the gender distinctions on nouns. The suffixes are /-k, -g/ for masculine and /-t, -to/ for feminine (Ehret 1980). The fact that these suffixes occur on nouns supports that these are Cushitic languages, since these endings occur in the whole of Cushitic. This argument is explained further in §3.2 in relation to the hypothesis of an East Rift branch. A third argument for considering Asá and Qwadza Cushitic languages is the existence of certain verbal extensions, which is discussed in the following.

Some verbs in Asá and Qwadza seem to have verbal extensions. In particular, there are three formatives that can be identified in both Asá and Qwadza: *-m*, *-t*, and *-s*. They resemble verbal extensions in other Cushitic languages and they are separable from the root when contrasted to similar roots, changing the meaning of the root slightly. The formatives are shown in Table 1. It is not clear if the vowel before the suffixes is actually part of it or not, and no function can be established for these formatives, since only lexical data is available.

Table 1: Verbal formatives

	Asá	Qwadza
-m	*'adam* 'to see'	*sa'im* 'burn (intr.)'
-t	*ji'it* 'to jump, to fly'	*tsalahet* 'curse'
-s	*wa'alis* 'to exchange'	*tulas* 'split in two'

In some instances, the vowel before the formative fully assimilates to the vowel of the stem, suggesting that the vowel before the *-m, -t* or *-s* is part of the formative. This assimilation can occur with all vowels and all three formatives, as can be seen in (1)–(2), but it remains unclear what the conditioning of the assimilation is. Compare the examples in Table 1, where the assimilation does not always occur. An explanation for some of the examples would be progressive trans-guttural assimilation, as proposed for Cushitic by Hayward & Hayward (1989). This makes sense in (1)–(2), since the vowel of the formative fully assimilates to the root vowel when a guttural phone is in between. However, it would

not explain why this assimilation does *not* occur in the lexeme *sa'im* 'burn (intr.)' in Table 1.

(1) Asá
 a. *duhum*
    'to marry'
 b. *hla'at*
    'to love, to like'
 c. *hidis*
    'to try, to try out, to prove'

(2) Qwadza
 a. *fo'om*
    'bathe'
 b. *fe'em*
    'measure'

On the basis of the forms of these verbal extensions, they are posed as cognates to the m-suffix, t-suffix, and s-suffix in Cushitic. The m-suffix is cognate to the passive or durative in Cushitic: "The form of the passive is *-am* (Somali, Konso, Dullay, Afar) or *-s(t)* (Khamtanga, Bilin)" (Mous 2020). This suffix is known to have shifted to the meaning of durative in West Rift (Kießling 2000: 14). Although a passive meaning seems to be missing, the presence of this suffix suggests Cushitic membership. The t-suffix is cognate to the middle in Cushitic: "The Cushitic languages with middle derivation in t (or related sounds) include all of East Cushitic (...) and all of the Southern Cushitic languages where there is no productive autobenefactive meaning" (Mous 2014: 77). The s-suffix is cognate to the causative in Cushitic: "the causative is a suffix *-s* or *-sh* preceded by a vowel *i* which is sometimes analysed as epenthetic" (Mous 2020: 3).

3 Do Asá and Qwadza form East Rift Southern Cushitic?

Considering Cushitic membership is proven, the next question is what the subclassification of Asá and Qwadza is. As has been previously shown, Ehret classifies Asá and Qwadza together in the East Rift branch of Southern Cushitic, based on the gender distinctions on nouns, among others. The suffixes are /-k, -g/ and /-t, -to/ (Ehret 1980). Although this supports the hypothesis that Asá and Qwadza are Cushitic languages, for these endings occur in the whole of Cushitic, it is not an argument for grouping Asá and Qwadza together. This section aims to

compare phonological and morphological evidence of Asá and Qwadza, and will show show that Asá stands further from the West Rift branch of Southern Cushitic than Qwadza does, implying that grouping these languages together might not be the right analysis. §3.1 discusses phonological evidence, §3.2 discusses gender suffixes, and §3.3 discusses verbal suffixes.

3.1 Phonological evidence

Phonological differences and similarities can provide insights into the relatedness of these languages. It is important to keep in mind that the data do not show underlying forms, but merely a transcription of the word. There is far too little morphology on these lexemes to identify phonological alternations. Therefore, the transcribed form is taken as underlying form, and all the phones that are found in the language are included in the phonological inventories.

The cognates in Table 2 show that Asá has a *d* where West Rift and Qwadza have a *tl.* There are not many cognates containing this sound that occur in Asá, Qwadza and Proto-West Rift, but for the ones that were found, this is a consistent sound correspondence, which places Asá further from West Rift than Qwadza. The reconstructed phone for Proto-East Cushitic is **d'* (Kießling & Mous 2003:36), while the reconstructed phone for Proto-Cushitic is **tl* (Mous 2021, personal communication), linking Asá to East Cushitic. This leads to two possible hypotheses: 1. Asá is part of East Cushitic and inherited this sound from Proto-East Cushitic. Or 2. Asá is part of an East Rift branch and underwent the sound change **tl >d.* The first hypothesis involves fewer changes and is therefore more efficient.[1]

The same applies to the phones *š* (presumably [ʃ]) and *ts*. Asá has *š* where Qwadza and Proto-West Rift have *ts*, as can be seen in Table 3. This places Asá further from Proto-West Rift and Qwadza. Qwadza and Proto-West Rift, again, have the exact same phone, if the orthographies used are representative of the sounds that were spoken.

There are a few exceptions to this correspondence in the data. These are shown in Table 4. The first row shows that Asá *š* corresponds to West Rift and Qwadza *t,* and the second row shows that Asá *š* corresponds to West Rift *s.* However, even when there is a correspondence between West Rift and Asá, Qwadza can either have a corresponding *c* or *s*, as is shown in the third and fourth row.

Although the following is a small set of cognates, the sound correspondence in Table 5 is completely regular. For these cognates, Qwadza and Proto-West Rift can be grouped together. These languages both have a labialized velar, while Asá

---

[1]Slashes are used to indicate a voiced pharyngeal fricative [ʕ], as in e.g. **tlaa/a* 'rock' in Table 2.

Table 2: d ~ tl

d ~ tl	Asá	Qwadza	West Rift
root-initial	*de'ok* 'stone'	*tlayiko* 'stone'	PWR **tlaa/a* 'rock'
	di' 'to stink'	-	PWR **tli/* 'to stink'
root-final	*hud-* 'to twist fibers into string or rope'	-	PWR **ʕutl* 'plait, braid, weave, knit, twist'
	-	*qutl* 'cut up'	PWR **quutl* 'cut up, cut into pieces'

Table 3: š ~ ts

š ~ ts	Asá	Qwadza	West Rift
root-initial	*šo'ok* 'wind'	*tsa'u* 'cold' ~ *tsa'uko* 'wind'	PWR **tsaaqwa* 'cold wind, coldness'
	-	*tsaluko* 'oryx'	PWR **tsawadu* 'water-bucks'
root-internal	*hašok* 'stick'	-	PWR **xwaytsaa* 'switch, flexible stick'
root-final	*haš-* 'to be full'	*hats-umo* 'much'	PWR **hats* 'to be full'

Table 4: Exceptions

Asá	Qwadza	West Rift
šike'e 'fly'		ALBU **tika/iya* 'tsetse fly, tsetse flies'
šada'ak 'buffalo'	-	PWR **sadee/a* 'buffalo'
šat 'to cut (down)'	*cat* 'to cut down'	PWR **tsat* 'cut'
šira'a 'bird'	*sil* 'to pluck a bird'	PWR **tsira/a* 'bird'

has a plain velar. Considering Proto Cushitic has **kw,* and Proto East Cushitic has **k,* Asá is more similar to Proto East Cushitic.

Table 5: k/g ~ kw/gw

k/g ~ kw/gw	Asá	Qwadza	West Rift
root-internal	*daka* 'bad, etc; rotten'	*tlakwa* 'bad'	PWR **tlakw* 'bad, wicked'
root-internal	*yaga* 'cattle'	*yaagwa ~ yaagway* 'cattle'	PWR **yakwaa ~ *hikwaa* 'cattle'

West Rift *x* can correspond to *x* and *h* in Asá and Qwadza. The examples in Table 6 show that Qwadza joins West Rift in some cognates and joins Asá in others. Asá differs from West Rift in all occurrences of this correspondence, and therefore seems the least similar to West Rift, as opposed to Qwadza. Proto-sound **ʕ* likely became *x* in Qwadza, after which it changed to *h* in some instances, while the **ʕ* likely became *h* in Asá, implying a separate development of these sounds. However, in both instances, a marked pharyngeal sound is replaced by a less marked sound. As such, these changes cannot be used for subclassification, but the lexical cognates of Qwadza and West-Rift show that these are related.

Based on these sound correspondences, Asá seems consistently more different from West Rift than Qwadza.

3.2 Gender suffixes

Old Afro-Asiatic gender morphemes are **kw* for masculine and **t* for feminine (Kießling 2000: 7), and the current suffixes on nouns correspond to this. Asá has two allomorphs of the masculine suffix, *-ok* and *-k*, and there are two allomorphs of the feminine suffix, *-et*/*-t*. The proto-suffix **-ku* developed to *-(o)k* and the proto-suffix **-ta* developed to *-(e)t*. The allomorphs are exemplified below in (3). Note that (3a) and (3d) are affected by vowel harmony; where the root of 'eland' in (3b) is *dam* and is followed by a suffix containing [a], the root is *dom* in (3a) and is followed by a suffix containing [o]. The root of 'rhinoceros' is *dofu* in (3c), but it has become *def* in (3d) under influence of the *e* in the suffix *-et*. Also note that the root seems to be shortened after the addition of the *-et* suffix.

Table 6: h ~ x

h ~ x	Asá	Qwadza	West Rift
h ~ h ~ x root-initial	*ho'orus* 'to snore, to growl'	*hil* 'to snore'	PWR **xuruutl* 'to snore'
- ~ x ~ x root-initial	-	*xab* 'to marry (of man)'	PWR **xab* 'to marry'
- ~ x ~ x root-internal	-	*tsoxoondo* 'soda'	PWR **tsaxasaa* 'salt'
h ~ x ~ x root-initial	*hadonk* 'horn'	*xalinko* 'horn'	PWR **xadaangw* ~ **xadamu* 'horn, tusk; something protruding'

(3)   Asá

    a.   *-ok* (m)
        dom-ok
        eland-MASC
        'male eland'

    b.   *-t* (f)
        dam-ayi-t
        eland-SGV-FEM
        'female eland'

    c.   *-k* (m)
        dofu-k
        rhinoceros-MASC
        'male rhinoceros'

    d.   *-et* (f)
        def-et
        rhinoceros-FEM
        'female rhinoceros'

Qwadza seems to have one allomorph of the masculine suffix, *-ko*, and two allomorphs of the feminine suffix, *-to/-o.* In addition, there is the neuter suffix -

wa. The *-ku developed to *-ko* and the *-ta developed to *-(t)o*. The allomorphs are shown in (4). Examples (4a-ii), (4b-ii), (4c-ii), and (4d-ii) serve as evidence that these are suffixes. Note the vowel harmony in (4c). Either the *a* in the suffix *-wa* causes full assimilation of the root vowel, resulting in the surface form *yawa*, or the feminine suffix *-o* causes the root vowel to change to *o*, resulting in *yo'o*. As will be shown later on, the Qwadza neuter gender mostly corresponds to West Rift (p) gender, a value of gender that causes the same agreement as a third person plural subject.

(4) Qwadza

    a. *-ko* (m)

        i. sag-i-ko
           head-SGV-MASC
           'head'

        ii. sag-ela
           head-PL
           'heads'

    b. *-to* (f)

        i. 'ag
           eat
           'eat'

        ii. 'ag-o-to
           eat-NMLZ-FEM
           'food'

    c. *-o* (f)

        i. ya-wa
           child-NEUT
           'children'

        ii. yo'-o
           child-FEM
           'child, children'

    d. *-wa* (n)

        i. tlunga-to
           ?-FEM
           'sky'

> ii. tlunga-wa
> ?-NEUT
> 'cloud'

The feminine suffixes -(*e*)*t* of Asá and -*to* and -*o* of Qwadza correspond to feminine suffixes found across Cushitic. According to Castellino (1970), there are two types of suffixes: "A, a first type having as principal feature a vowel or a vowel plus consonant other than t; B, a second type in which (feminine) gender is marked by means of morphemes characterized by the presence of a consonant which, in most cases, is -*t*." The latter type is said to be more common (Castellino 1970: 349-350). The -*to* suffix matches the description of Castellino's type B, and -*o* matches his type A. In the discussion of this article, it is added that many Cushitic languages have a k/t alternation for masculine and feminine, respectively (Castellino 1970: 358-359). This reflects the feminine suffixes as exemplified above, but also the masculine -(*o*)*k* suffix in Asá and the masculine -*ko* suffix in Qwadza.

There are 115 nouns in the database, of which 52 have cognates in Asá, Qwadza and (Proto-)West Rift. From these 52, there are 19 instances of a cognate with the same gender in all three languages (shown in Table 7 and Table 8). This mostly consists of basic lexicon, like words for body parts, environment, and some animals.[2]

In addition, of the 52 nouns that have cognates in all three languages, there are 15 nouns that show the same gender suffix in Qwadza and Proto-West Rift, but a different one in Asá. The gender in Asá is in these cases unrecognizable, as can be seen in Table 9, example 24 for masculine, 25 for feminine, and 26 for neuter. The lack of a suffix in Asá may point to truncation of the suffix, or indicate an unidentified gender in -*a*. The last possibility seems likely when looking at example 27, since the Qwadza root *yaag* is followed by the neuter suffix -*wa*, which would correspond to an Asá root *yag* followed by a possible neuter suffix -*a*.

Of the 52 cognate nouns, there are only 7 nouns that have the same gender suffix in Asá and Qwadza, and a different one in West Rift. This is shown in Table 10: example 28 for the masculine in Asá and Qwadza, and example 29 for the feminine in Asá and Qwadza.

---

[2]Examples in brackets indicate that the cognacy between the lexemes is not convincing. However, these lexemes are shown for reasons of transparency and potential cognacy.

Table 7: Cognates with the same gender (masculine) in Asá, Qwadza, and West Rift

	Asá	Qwadza	West Rift
5	*mono-k* 'heart'	*muna-ku* 'heart'	PWR **muuná* (sg.m) 'heart'
6	*tibiš-ok* 'gravy, broth'	*dabadzi-ko* 'broth, gravy'	PWR **dabatsuu'a* (col.m) 'gravy, broth'
7	*lupa'-ok* 'upper arm'	*lupa'i-ko* 'hand'	IRGO **tluba/a* (sg.m) 'upper arm'
8	*lehe-k* 'moon'	*slahayi-ko* 'moon'	PWR **slaʕaangw* (sg.m) 'moon, month'
9	*hadon-k* 'horn'	*xalin-ko* 'horn'	PWR **xadaangw* ~ **xadamu* (sg.m) 'horn, tusk; something protuding'
10	*'af-ok* 'mouth'	*'afu-ko* 'mouth'	PWR **'afa* (sg.m) 'mouth'
11	*hade-k* 'sweat'	(*dulutu-ko* 'sweat')	PWR **haru'u* (sg.m) 'dew'
12	*mor-ok* 'house'	*mali-ko* 'house'	PWR **mara* (sg.m) 'house'
13	*sog-ok* 'head'	*sagi-ko* 'head'	PWR **saga* (sg.m) 'head'
14	*šugum-ok* 'bushbuck'	*tsa'u-ko* 'kid, lamb'	PWR **tsakanáy* (pl.m) 'klipspringers'
15	*dem-ok* 'morning'	*letlemu-ko* 'God'	PWR **tleheema* (sg.tant.m) 'sunshine'
16	*kunduf-ok* 'jigger'	(*holofe'idu-ko* 'beetle')	PWR **kuundurufu* (n.col.m) 'spirilum ticks'
17	*gwaran-k* 'rat'	*gilati-ko* 'fat-mouse'	PIRQ **gwaraangw* (sg.m) 'rat'
18	*'id-ok* 'person'	*hila-ku* 'person, goat'	PWR **heedi* (sg.m) 'person'

Table 8: Cognates with the same gender (feminine) in Asá, Qwadza, and West Rift

	Asá	Qwadza	West Rift
19	*'efer-et* 'goat'	*'afula-tu* 'he-goat'[a]	PWR **'afaquraa* ~ **'afuraa* (pl.f) 'warthogs'
20	*isa-t* 'neck'	*isi-to* 'neck'	PWR **'isa* (sg.f) 'neck'
21	*ila-t* 'eye'	*'ili-to* 'eye'	PWR **'ila* (sg.f) 'eye; spring'
22	*fari-t* 'bone'	*fala'a-to* 'bone'	PWR **fara* (sg.f) 'bone'
23	*ya'a-ta* 'sandal'	*ya'o-o* 'foot, leg'	PWR **ya/ata* (sg.f) 'sandal'

[a]This suffix deviates from the regular *-to* suffix.

Table 9: Cognates with corresponding gender in Qwadza and PWR

	Asá	Qwadza	West Rift
24	*waya* 'intestines'	*wa'itu-ko* 'belly'	PWR **wa'ay* (sg.m) 'inside'
25	*iringa* 'nose'	*ningwa-to* 'nose'	PWR **'uruunga* (sg.f) 'nostril, pair of nostrils, nose'
26	*hesa* 'tail'	*hayiso-wa* 'tail'	PWR **ʕaysoo* (sg.n) 'tail, hair of tail'
27	*yaga* 'cattle'	*yaag-wa* ~ *yaagway* 'cattle'	PWR **yakwaa* ~ **hikwaa* (pl.n) 'cattle'

Table 10: Same gender suffix in Asá and Qwadza, different in West Rift

	Asá	Qwadza	West Rift
28	*'ofolon-k* 'navel'	*'ofu'u-ko* 'navel'	(PIRQ **'afeetloo* (sg.n) 'hip, waist, loin')
29	*ŋulu-et* 'knee'	*'ugun-o* 'knee'	PWR **guruguunda* (sg.m) 'knee'

Lastly, there are only 5 of 52 nouns that have the same gender suffix in Asá and Proto-West Rift, and a different one in Qwadza. This is shown in Table 11: example 30 for the masculine in Asá and West Rift, and example 31 for the feminine in Asá and West Rift (where the corresponding word *'aama* 'mother' in Qwadza lacks a gender suffix).

Table 11: Same gender suffix in Asá and PWR, different in Qwadza

	Asá	Qwadza	West Rift
30	*mong-ok* 'arm'	*munga'ay-o* 'bracelet'	PWR **maangaa'u* (sg.m) 'front leg'
31	*'ama'e-to* 'older girl'	*'aama* 'mother' (term of address)	PWR **'aama ~ *'aa-maa* (sg.f) 'mother, old woman'

Regarding gender suffix correspondences, most nouns of which three cognates are present, have the same gender. This, again, confirms the relatedness of the languages. However, outside of this core group of nouns, Asá is the least similar to West Rift, and also not convincingly similar to Qwadza. The corresponding suffixes show that Qwadza and West Rift have a lot of gender correspondences, and are thus more similar. This suggests that Asá is the least similar to both Qwadza and Proto-West Rift.

As was explained, (p) gender is a value of gender that causes the same agreement as a third person plural subject. Of the cognates Asá and Proto-West Rift have, the instances of (p) gender in Proto-West Rift (a total of 10 nouns) correspond to either a zero suffix (example 32) or a masculine suffix in Asá (example 33 and 34). In addition, all of the zero gendered nouns in Asá end in *a,* except example 35 in Table 12, which could be assimilation to the root vowel *e.*

Table 12: West Rift (p) gender vs. Asá cognates

	Asá	West Rift
32	*liba* 'milk'	PWR **'ilibaa* (pl.tant.n) 'milk'
33	(*'agun-k* 'honey')	(PNWR **'aanxari* (n.acti.n) 'phlegm')
34	(*'ofolon-k* 'navel')	(PIRQ **'afeetloo* (sg.n) 'hip, waist, loin')
35	*gide'e* 'ribs'	PIRQ **gwe'eedoo* (sg.n) 'pair of buttocks, bottom'

In Qwadza, cognates of the nouns can have masculine (example 36), feminine (example 37) or neuter gender (example 38), shown in Table 13.

Table 13: West Rift (p) gender vs. Qwadza cognates

	Qwadza	West Rift
36	*'angali-ko* 'honey'	PNWR **'aanxari* (n.acti.n) 'phlegm'
37	*ge'eli-to* 'loins'	PIRQ **gwe'eedoo* (sg.n) 'pair of buttocks, bottom'
38	*hayiso-wa* 'tail'	PWR **ʕaysoo* (sg.n) 'tail, hair of tail'

In fact, underived neuter nouns in Qwadza usually have a Proto-West Rift cognate in (p) gender, as is shown in example 39. The only exceptions that were found in the database, are two nouns that do not have a cognate in Proto-West Rift, and therefore it cannot be confirmed what the gender in Proto-West Rift would be. However, the Asá cognates have unmarked / *-a* gender. The two nouns are shown in example 40 and 41. Other instances of Qwadza neuter nouns are deverbal derivations, based on cognates in Asá and Proto-West Rift. This is shown in example 42, all in Table 14.

In sum, Asá and Qwadza nouns correspond less to each other in terms of gender suffixes than Qwadza and Proto-West Rift do, and Asá and Proto-West Rift correspond less than Qwadza and Proto-West Rift. So, again, we can conclude that Asá stands further from Proto-West Rift than Qwadza does.

Table 14: Asá and Qwadza 'neuter' nouns vs. West Rift cognates

	Asá	Qwadza	West Rift
39	*yaga* 'cattle'	*yaag-wa ~ yaag-way* 'cattle'	PWR **yakwaa ~ *hik-waa* (pl.n) 'cattle'
40	*'ajota* 'day'	*'atso-wa* 'day'	-
41	*liga* 'tooth'	*'islik-wa* 'teeth'	-
42	*ma'-* 'to avoid, to leave alone'	*mayik-wa* 'avoidance object'	PWR **maw* (v) 'let, leave, avoid'

3.3 Verbal extensions

As has been shown in §2, Asá and Qwadza both have formative suffixes that correspond to the formative suffixes across Cushitic. However, the same suffixes are not always found on the same words in Asá and Qwadza. This is also not the case for Qwadza and Proto-West Rift, but even less so for Asá and Qwadza. See Table 15.

Table 15: Non-corresponding formative suffixes

	Asá	Qwadza	West Rift
43	*hla'-at* 'to love, to like'	*sla'-as* 'love, like' *sla'* 'purify'	PWR **sla'* (v) 'love, like, want'
44	*ra'* 'to stay, to remain'	*da'am* 'to watch'	PIRQ **daam* 'to wait, to expect'

However, the limitations of the data come into play here. There are only few instances of recognizable verbal morphology and when one verbal extension is found, this does not mean another verbal extension does not exist.

4 Conclusion

The goal of this study has been to answer two questions: "1. Are Asá and Qwadza Cushitic languages?" and "2. Are Asá and Qwadza Southern Cushitic languages forming East Rift Southern Cushitic?". Answering these questions have led to newly gained analyses and insights about the subclassification of the Cushitic language family.

Before concluding this study, it should be noted that there is a large lack of data, and it is not possible to collect more. This affects the quality of any analysis formed about the classification of Asá and Qwadza. In addition, it is difficult to fully rely on the data, since especially the earlier sources were not written by linguists. Also, since gender suffixes on nouns are a retention of Cushitic, it is difficult to use them for subclassification, and because gender and number are encoded together, one cannot be sure that the gender on a noun in these languages is underived.

Despite these challenges, it has been shown that Asá and Qwadza have convincing cognates with West Rift. Gender and derivational morphology are present on the lexemes, suggesting membership in the Cushitic language family and relatively close relatedness to the West Rift language branch of Cushitic. Phonological evidence places Asá as the more distant from West Rift. Qwadza is closer to West Rift, as it has retained (or innovated) the same phones in cognates. A classification of Asá under East Cushitic is implied by one of the sound correspondences, as this would require the least sound changes. However, this is highly speculative and more research is needed into this. The gender suffixes also place Asá as most distant from West Rift. For both Asá and Qwadza, different derivational morphology is present on the same cognates. This suggests that this is not a common innovation. Since Asá developed many features separately, it seems to have split off from Rift first (or it belongs to East Cushitic). Qwadza has some common developments with West Rift, so it seems to have split off second. At this point, there is not much proof for Asá and Qwadza forming the East Rift group together. However, more research is needed to find out how exactly these languages fit into the Cushitic language family. In order to argue for a more accurate subclassification of Asá and Qwadza, it is necessary to find how Southern Cushitic fits in Cushitic and if and how Asá would fit into East Cushitic. Other topics for further research are the lack of gender suffixes on some Asá and Qwadza nouns, the variation of forms of suffixes, and the Asá connection with East Cushitic.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

ALBU	Alagwa-Burunge	n.acti	nomen acti
.f/FEM	feminine gender	PIRQ	Proto-Iraqw
IRGO	Iraqw-Gorwaa	pl.tant	plurale tantum
.m/MASC	masculine gender	PWR	Proto-West Rift
.n/NEUT	neuter gender	SGV	singulative
n.	noun	v.	verb

Acknowledgements

References

Blažek, Václav. 2019. South Cushitic classification in lexicostatistic perspective. *Folia Orientalia* LVI. 41–94.

Castellino, Giorgio R. 1970. Gender in Cushitic. Hamito-Semitica. In James Bynon (ed.), *Proceedings of a colloquium held by the Historical section of the Linguistics Association (Great Britain) at the School of Oriental and African Studies*, 333–360. Berlin, Boston: De Gruyter Mouton.

Claus, Heinrich. 1910. Die Wangómvia mit Bemerkungen zu den Sprachproben von K. Meinhof. *Zeitschrift für Ethnologie* 42. 489–497.

Ehret, Christopher. 1974. *Kw'adza word list*. Unpublished manuscript.

Ehret, Christopher. 1980. *The historical reconstruction of Southern Cushitic phonology and vocabulary* (Kölner Beiträge zur Afrikanistik 5). Berlin: Dietrich Reimer.

Fleming, Harold. 1969. Asa and Aramanik: Cushitic hunters in Masai-land. *Ethnology* 8. 1–36.

Hayward, Katrina M. & Richard J. Hayward. 1989. "Guttural": Arguments for a new distinctive feature. *Transactions of the Philological Society* 87(2). 179–193.

Kießling, Roland. 1999. *Notes on the death of Qwadza*. Unpublished manuscript.

Kießling, Roland. 2000. Some salient features of Southern Cushitic (Common West Rift). *Lingua Posnaniensis* 42. 69–89.

Kießling, Roland & Maarten Mous. 2003. *The lexical reconstruction of West-Rift Southern Cushitic*. Cologne: Rüdiger Köppe Verlag.

Kohl-Larsen, Ludwig. 1943. *Auf den Spuren des Vormenschen: Forschungen, Fahrten und Erlebnisse in Deutsch-Ostafrika.* 2 vols. Stuggart: Strecker und Schröder, Stuttgart.

Merker, Moritz. 1904. *Die Masai: Ethnographische Monographie eines ostafrikanischen Semitenvolkes.* Berlin: Dietrich Reimer.

Mous, Maarten. 2014. The Middle in Cushitic languages. In Andrew Simpson (ed.), *Proceedings of the Twenty-Seventh Annual Meeting of the Berkeley Linguistics Society, March 22-25, 2001: Special session on Afroasiatic languages*, 75–86. Berkeley, CA: Berkeley Linguistic Society.

Mous, Maarten. 2020. *Cushitic verbal derivation.* Unpublished manuscript.

Pearce, John. 1954. *Ng'omvia wordlist.* Unpublished manuscript.

Petrollino, Sara & Maarten Mous. 2010. Recollecting words and expressions in Aasá, a dead language in Tanzania. *Anthropological Linguistics* 52(2). 206–216.

Whiteley, William H. 1958. *A short description of item categories in Iraqw (with material on Gorowa, Alagwa and Burunge).* Kampala: East African Institute of Social Research.

Winter, Christopher. 1979. Language shift among the Aasáx, a hunter-gatherer tribe in Tanzania. *Sprache und Geschichte in Afrika* 1. 175–204.

Chapter 14

Emphatic properties of object marking in Ikalanga

Rose Letsholo

University of Botswana

Madelyn Colantes

Pomona College

Michael Diercks

Pomona College

This paper builds on Letsholo (2013), documenting additional properties of object marking in Ikalanga (Bantu, Botswana and Zimbabwe). We show that the ability of object markers to co-occur with overt objects is highly constrained by pragmatic context. These co-occurrence constructions are linked with corrective focus readings and mirative focus readings, along with verum focus readings (as first noticed by Letsholo 2013). We also detail some persistent analytical puzzles regarding the focus interpretations associated with OM-Obj co-occurrence in Ikalanga.

1 Introduction

Object marking is a linguistic strategy to refer to discourse-familiar objects, similar to English pronominalization. Across Bantu languages, the object marker appears as a prefix within the verb form, typically affixing before the verb stem and after the tense morpheme. (1b) below illustrates object marking in Ikalanga, a

Rose Letsholo, Madelyn Colantes & Michael Diercks

Bantu language spoken in Botswana and Zimbabwe: the object marker *-í-* affixes to the verb and replaces the noun *búka* 'book'.[1]

(1) Ikalanga (Botswana)

   a. Joni   w-áká-bál-á   búka.
       1.John 1SM-PST-read-FV 9.book
       'John read the book.'

   b. Joni   w-áká-í- bál-a.
       1.John 1SM-PST-9OM-read-FV
       'John read it.'

There is a broad range of previous work on object marking (OMing) in Bantu languages (Bresnan & Mchombo 1987, Riedel 2009, Zeller 2012, 2015, van der Wal 2015, 2020, 2022, Sikuku & Diercks 2021a). A central area of investigation is whether the object marker may co-occur with the *in situ* lexical object that it refers to, a mechanism generally referred to as OM-doubling. The acceptability of OM-doubling in Bantu languages is heavily dependent on discourse contexts. Recent work on Lubukusu (Sikuku et al. 2018, Sikuku & Diercks 2021a), Tiriki (Liu 2022), Wanga (Kuzmik 2022), and Cinyungwe (Langa da Câmara et al. 2023, Lippard et al. forthcoming) has investigated this interface of syntax and pragmatics, delving into the precise discourse conditions that license doubling. In this paper, we expand on earlier work on Ikalanga object marking (Letsholo 2013) to further describe a subset of the interpretive effects of Ikalanga object marking.

Specifically, we investigate how Ikalanga co-occurrence of OMs and their associated objects interacts with focus, and what emphatic interpretations arise from an OM-Obj co-occurrence (OOC) sentence. We summarize basic Ikalanga morphology and parameters of object marking in §2, and key generalizations of Ikalanga OM-Obj co-occurrence in §3 as background. §4 discusses object marking in verum contexts, and §5 explores the interaction between object marking and focus within *v*P. In §6 we show that the co-occurrence of OMs and their associated objects is also associated with a mirative reading. §7 outlines areas for future research, and §8 concludes.

We find that Ikalanga shares a number of properties that have been recently documented for related Bantu languages like Lubukusu, Cinyungwe, and Zulu. Notably, it shares properties with all of them, but in a different configuration

---

[1]All uncited data points are provided by the first author. We annotate surface tone, though more research is necessary to understand the tonology of Ikalanga more generally and of these constructions specifically.

than has been previously documented. As in Zulu, object marking in Ikalanga is correlated with right-dislocation of objects (which is unlike Cinyungwe and Lubukusu). But Ikalanga is similar to Cinyungwe and Lubukusu in that the co-occurrence of OMs and overt objects is necessarily linked with emphatic readings of the clause (which is unlike Zulu). The result leaves a lot of unsettled questions for Ikalanga, but continues to contribute to understanding the conditions on OM-ing in Bantu languages, and the typological patterns that recur across Bantu languages.

2 OMing basics in Ikalanga

2.1 Morphology of OMs

Like most Bantu languages, Ikalanga has a rich noun class system. Nouns are organized into one of 17 classes, each distinguished by its own prefix. Ikalanga's noun class system also dictates the morphological forms of object markers: each OM agrees in noun class with its co-referent. Table 1 below shows the prefixes for each noun class and pronominal, as well as the corresponding object marker in Ikalanga.

2.2 Basic parameters of Ikalanga OMs

Typological descriptions of Bantu OMs have often been categorized along the following parameters (as delineated in van der Wal 2020):[2]

(2)   Properties of Bantu object markers (van der Wal 2020):
- Position of object markers – Is it a pre-stem affix or an enclitic?
- Types of objects marked – Does animacy, definiteness, or other factors influence whether an object can be OMed?
- Number of object markers – How many OMs can a verb stem take?
- Behavior in ditransitives – Can either object in a ditranstive be OMed?
- Nature of object markers – Are they a form of agreement morphology (and thus allow doubling), or are they pronominal enclitics (and thus cannot double)?

Regarding the first two parameters, Ikalanga OMs are pre-stem affixes, and to our knowledge animacy does not affect object marking (in any obvious/direct

---

[2]Also see Marten & Kula (2012), Marlo (2014, 2015), and Riedel (2009) for relevant overviews.

Table 1: Ikalanga noun class and object marker morphology (adapted from Letsholo 2002)

Class	Prefix	Noun example	OM	OM example	Gloss
1st sg	n-	-	ndi-	bánóndítóla	'They'll take me.'
1st pl	ba-	-	ti-	bánóítóla	'They'll take us.'
2nd sg	n-	-	ku-	bánókútóla	'They'll take you.'
2nd pl	ba-	-	mu-	bánómútóla	'They'll take you (pl).'
3rd sg/1	n-	nthu 'person'	n-	bánóntóla	'They'll take him/her.'
1a	-	mme 'mother'	ba-	bánóbatóla	'They'll take him/her.'
2	ba-	bathu 'people'	ba-	bánóbatóla	'They'll take them.'
3	n-	nti 'tree'	u-	bánóútóla	'They'll take it.'
4	mi-	miti 'tree'	i-	bánóítóla	'They'll take them.'
5	-	zhani 'leaf'	li-	bánólítóla	'They'll take it.'
6	ma-	mazhani 'leaves'	a-	bánóátóla	'They'll take them.'
7	chi-/i-	chibululu 'lizard'	chi-	bánóchítóla	'They'll take it.'
8	zwi-	zwibululu 'lizards'	zwi-	bánózwítóla	'They'll take them.'
9	N-	mbga 'dog'	i-	bánóítóla	'They'll take it.'
10	N-/dzi-	mbga 'dogs'	dzi-	bánódzítóla	'They'll take them.'
11	li-	likuni 'log'	gu-	bánógútóla	'They'll take it.'
14	bu-	bushwa 'grass'	gu-	bánógútóla	'They'll take it.'
15	ku-	ku izela 'to sleep'	k-	únókuda	'He likes it.'
16	pa-	pa ngwina 'by the hole'	-	-	-
17	ku-	ku nzi 'at home'	-	-	-
18	mu-	mu ngumba 'in the house'	-	-	-
21	zhi-	zhingwana 'enormous child'	li-	bánólítóla	'They'll take it.'

manner). Letsholo (2013) does report that co-occurence of object markers and overt objects yields a definite reading of the object noun phrase in question. Ikalanga generally allows just one object marker on a verbal stem. Attempts to add two OMs on either a lexical ditransitive or benefactive applicative verb both result in ungrammaticality:

(3) *Nchídzí w-áká-[bá-] [i-] pa.
    1.Nchidzi 1SM-PST-2OM-9OM-give
    Intd. 'Nchidzi gave it (the toy) to them (the children).'

(4) *Ludó w-áká-[zwí-] [m-] bík-il-a.
    1.Ludo 1SM-PST-8OM-1OM-cook-APPL-FV
    Intd. 'Ludo cooked it (food) for him.'

Causatives seem to be a possible exception to this generalization.

(5) Nchídzi w-áka-[bá-] [dzi-] séng-és-a.
    1.Nchidzi 1SM-PST-2OM-10OM-carry-CAUS-FV
    'Nchidzi made them (the children) carry it (the firewood).'

Not all instances of causative constructions, however, naturally take multiple OMs; further research is needed to establish when this is possible.

Regarding the fourth parameter, in double object constructions, either object can be object marked in Ikalanga – that is, it displays object marking symmetry.[3]

(6) a. Ludó w-áká-[m-] bík-íl-á nyama.
       1.Ludo 1SM-PST-1OM-cook-APPL-FV 9.meat
       'Ludo cooked meat for him.'

    b. Ludó w-áká-[í-] bík-íl-a Nchídzi.
       1.Ludo 1SM-PST-9OM-cook-APPL-FV 1.Nchidzi
       'Ludo cooked it for Nchidzi.'

The final typological parameter seeks to classify the OM as either agreement morphology or a pronominal enclitic. One key diagnostic for this classification is whether the OM is allowed to co-occur with its associated lexical object DP when that object is *in situ* inside the verb phrase.

Letsholo (2013) argues for a pronominal incorporation analysis of Ikalanga OMs. Namely, while an OM can co-occur with a lexical object, the object cannot

---

[3]For background on the properties of (a)symmetrical object properties in Bantu languages, see Bresnan & Moshi (1990), Jerro (2015), Riedel (2009), and Zeller (2015), among others.

remain within the verb phrase. In (7) below, the OMed object is obligatorily offset by a prosodic break, suggesting it has been dislocated to a different syntactic position. (§3.2 will discuss in more detail the evidence for dislocation of the OMed object.)

(7) Nchídzí   w-á-gu- ngw-á   *(,) búsúkwa .
     1.Nchidzi 1SM-PST-14OM-drink-FV   14.beer
     'Nchidzi DID drink the beer.'

(7) also provides evidence against an analysis of the OM as agreement morphology, per Letsholo (2013); agreement morphemes are generally assumed to be grammatically obligatory and should not alter the sentence's meaning in any way. As the translation in (7) suggests, OOC in this sentence creates an emphatic reading.[4] Removing the OM is not only grammatical, but also removes this emphatic effect:

(8) Nchídzí   w-á-ngw-á   búsúkwa.
     1.Nchidzi 1SM-PST-drink-FV 14.beer
     'Nchidzi drank the beer.'

Letsholo (2013) thus argues that because Ikalanga OMing results in dislocation of the associated object, does not require the overt object, and introduces an emphatic reading, Ikalanga OMs are best analyzed as pronominal clitics rather than agreement morphemes. This paper doesn't specifically comment on this analysis, though these new observations add complexity that any analysis of OMing will need to account for.

3 Basics of OM-Obj co-occurrence in Ikalanga

This section presents two core generalizations about the co-occurrence of object markers and lexical DP objects in Ikalanga: OOC is unnatural in discourse-neutral contexts, and OOC and object dislocation are obligatorily linked. The co-occurrence of OMs and objects has generally been discussed under the term *OM-doubling* (based on the familiar term *clitic-doubling*): this is generally used to refer to the co-occurrence of a clitic or object marker with an *in situ* lexical object. As we will show, languages that unambiguously display OM-doubling (e.g. Lubukusu and Cinyungwe) share a broad range of patterns with Ikalanga. Nonetheless, Ikalanga behaves relatively Zulu-like in that it appears that an OM

---

[4]The specific emphatic interpretation generated here will be further discussed in §4.

on the verb is incompatible with an *in situ* lexical object. As such, we avoid the term *doubling* for the most part, using the pre-analytic term *OM-Object co-occurrence* instead.

3.1 OOC is unnatural in all-new contexts

In all-new contexts, such as out-of-the-blue situations where no prior knowledge is assumed between interlocutors, OOC is infelicitous:

(9) *Out of the blue:*
     #Nchídzí w-á-⟨gu-⟩ ngw-á   , ⟨búsúkwa⟩ .
     1.Nchidzi 1SM-PST-14OM-drink-FV   14.beer
     Intd. 'Nchidzi DID drink the beer.'

As noted in the translation, OOC in (9) creates an emphatic, insistent reading that we identify as verum (see §4). In an out-of-the-blue context, however, there is no one doubting the speaker; the emphatic reading introduced by OOC therefore sounds unnatural. Similar verum readings have been identified in OM-doubling in Lubukusu (Sikuku et al. 2018, Sikuku & Diercks 2021a), Tiriki (Liu 2022), Wanga (Kuzmik 2022), Cinyungwe (Langa da Câmara et al. 2023, Lippard et al. forthcoming), and Rukiga (Allen Asiimwe, pc).

3.2 OM-Obj co-occurrence requires dislocation

3.2.1 Relevant case study: Zulu OM-Obj co-occurrence

Zeller (2015) shows that in the Southern Bantu language Zulu, OMs and object dislocation are obligatorily linked. (10) below shows the canonical word order, with the manner adverb right-adjoined to *v*P and thus following the object. (This is consistent with the generalization across languages that 'low' adverbs mark the edge of *v*P).

(10) Zulu
     Si-bon-a   i-n-kosi   kahle.
     1PL-see-FV AUG-9-chief well
     'We are seeing the chief well.' (Zeller 2015: 20)

When an object co-occurs with its OM, it obligatorily moves to the right, past the adverb and outside of *v*P (11a). Leaving the object *in situ* as in (11b) is ungrammatical, as is moving the object without an OM appearing on the verb (11c).

(11) Zulu

  a. Si-[yi-] bon-a   kahle [i-n-kosi] .
     1PL-9OM-see-FV well   AUG-9-chief
     'We are seeing him well, the chief.'

  b. *Si-[yi-] bon-a   [i-n-kosi]   kahle.
     1PL-9OM-see-FV AUG-9-chief well
     Intd. 'We are seeing him well, the chief.'

  c. *Si-bon-a   kahle i-n-kosi.
     1PL-see-FV well   AUG-9-chief
     Intd. 'We are seeing the chief well.'
     (Zeller 2015: 20)

In OOC double object constructions in Zulu where the verb is in its conjoint form (signaling that the *v*P is non-empty), the OM-doubled benefactive object is obligatorily dislocated to the right of the theme object, as we might expect (12b). Leaving the object in its canonical position is ungrammatical (12c).

(12) Zulu OMing: conjoint verb forms

  a. Ngi-theng-el-a   u-Sipho   u-bisi.
     1SM-buy-APPL-FV AUG-1a.Sipho AUG-11.milk
     'I'm buying milk for Sipho.'

  b. Ngi-[m-] theng-el-a   u-bisi   [u-Sipho] .
     1SM-1OM-buy-APPL-FV AUG-11.milk AUG-1a.Sipho
     'I'm buying him milk, Sipho.'

  c. *?Ngi-[m-] theng-el-a [u-Sipho]   u-bisi.
     1SM-1OM-buy-APPL-FV AUG-1a.Sipho AUG-11.milk
     Intd. 'I'm buying him milk, Sipho.'
     (Zeller 2015: 21)

The word order rigidity of objects disappears, however, when the verb takes on its disjoint form, indicated by the *-ile* affix below. That is, while the two objects were restricted in their order in (12), using the verb's disjoint form allows the two objects to occur in *either* order:

(13) Zulu OMing: disjoint verb forms

  a. U-John   u-[ba-] nik-ile   [a-ba-ntwana]   i-mali.
     AUG-1a.John 1SM-2OM-give-PST.DJ AUG-2-child   AUG-9.money
     'John *did* give the children the money.'

b. U-John   u-ba- nik-ile   i-mali   a-ba-ntwana .
AUG-1a.John 1SM-2OM-give-PST.DJ AUG-9.money AUG-2-child
'John *did* give the children the money.'
(Zeller 2015: 23)

Zeller (2015) analyzes these Zulu constructions where both objects can follow the disjoint verb form in either order as *double right dislocation* constructions – that is, both object DPs are dislocated out of the *v*P. Per Zeller, the two dislocated objects become right-adjoined to a higher maximal projection as adjuncts, and it is this adjunct classification that accounts for the flexible word order.

Zulu OMing patterns are characterized by an obligatory association between OMs and dislocation, and by double right dislocation constructions: while word order of objects is asymmetrical when the verb is in its conjoint form, a verb in its disjoint form requires both objects to be right-dislocated as adjuncts, allowing for word order flexibility between the objects.

3.2.2 Ikalanga: Similarities to Zulu

Just as in Zulu, Ikalanga requires that objects co-occurring with OMs be dislocated outside of *v*P, and disallows OOC and dislocation to occur independently from each other. (14) below establishes a context that facilitates OOC (corrective focus on an adverb, discussed in §5.1), and offers four attempted responses:[5]

(14)   a. W-áká-bón-a   baná   íbábájé bé   ikwélé tshípí yáká píndá?
1SM-PST-see-FV 2.child 2.DEM   ASSOC 7.school week last
'Did you see those students last week?'

b. Á, nd-aká-bon-a   baná   íbábájé **mádékwe**.
No 1SG.SM-PST-see-FV 2.child 2.DEM   yesterday
'No, I saw those children yesterday.'

c. *Á, nd-aká-bon-a   **madékwé** , baná   íbábáje.
No 1SG.SM-PST-see-FV yesterday   2.child 2.DEM
Intd. 'No, I saw those children yesterday.'

d. *Á, nd-aká-bá- bón-á   báná íbábájé **mádékwe**.
No 1SG.SM-PST-2OM-see-FV 2.child 2.DEM   yesterday
Intd. 'No, I saw those children yesterday.'

---

[5]Note that the two words *yáká píndá* in (14a) together mean 'last'.

e. Á, nd-aká-⟨bá-⟩ bón-á   **mádékwe** , ⟨baná íbábáje⟩ .
    No 1SG.SM-PST-2OM-see-FV yesterday   2.child 2.DEM
    'No, I saw those children yesterday.'

(14b) shows the canonical word order in a monotransitive sentence, with the object to the left of the temporal adverb. Dislocating the object outside of *v*P – indicated by the prosodic break preceding the object, and its position to the right of the adverb – is ungrammatical without OOC (14c), as is OOC without dislocation (14d). (14e), which has both OOC and dislocation, is a grammatical and natural response.[6] We see then that, similar to the Zulu patterns, Ikalanga OM-Obj co-occurrence requires object dislocation, and object dislocation requires an OM on the verb.

We see further similarities to Zulu in Ikalanga's flexible word order when both the adverb and object are dislocated. Above, when just the object DP was dislocated, the word order was obligatorily Adv OMed-Obj (14e). However, just as dislocating both objects in Zulu allows for word order symmetry, dislocating both the adverb and OMed object in Ikalanga allows for the otherwise unacceptable order of OMed-Obj Adv. In (15), both the adverb *madekwe* 'yesterday' and OMed object DP *bana ibabaje* 'those children' are offset by a prosodic break (indicated by a comma), suggesting both elements have been dislocated outside of *v*P. In this case of double dislocation with an OM on the verb, the word order can either be OMed-Obj Adv *or* Adv OMed-Obj.

(15) a. Did you really see those children yesterday?[7]

 b. Nd-aká-⟨bá-⟩ bón-á   , ⟨báná íbábáje⟩ mádékwe!
     1SM.SG-PST-2OM-see-FV   2.child 2.DEM yesterday

     'I did see those children yesterday!'[8]

 c. Nd-aká-⟨bá-⟩ bón-á   , madékwé ⟨baná íbábájé⟩ !
     1SM.SG-PST-2OM-see-FV   yesterday 2.child 2.DEM

     'I did see those children yesterday!'

This pattern of flexibility (when both elements are dislocated) is consistent with the Zulu patterns of double right dislocation constructions; it is also consistent with Letsholo's (2013) analysis of objects co-occurring with OMs as adjuncts

---

[6]Though both (14b) and (14e) are acceptable responses to (14a), there is a slight interpretive difference between the two, with the OOC in (14e) implying there is something noteworthy and important about the statement. §6 will discuss this interpretation further.

[7]The context provided here that facilitates OOC in the response will be explored in detail in §5.1.

[8]Though this is still an acceptable response to the given question, it is slightly less natural compared to (15c).

outside of *v*P: if the dislocated object is assumed to be an adjunct along with the temporal adverb, it follows logically that the two adjuncts would have flexible word order relative to each other.

With these two key generalizations about OOC – its unacceptability in discourse-neutral contexts, and its obligatory association with object dislocation – we are now equipped to delve into the particular pragmatic contexts that allow for OOC.

4 Verum facilitates OOC

As noted above, Letsholo (2013) first observed that OOC has an emphatic bearing on a sentence. Here, we expand on her observation, identifying this emphatic effect as a verum reading. Verum readings are licit when the proposition at hand is being doubted, and the speaker is attempting to assert their confidence and settle the issue, similar to the use of English emphatic *do* (Gutzmann & Castroviejo Miró 2011, Gutzmann et al. 2020, Güldemann 2003, 2016, Kerr & van der Wal 2023). Natural contexts for verum interpretations include addressing listener denial or doubt, affirmation of a preceding assertion, and opposite polarity contexts. As noted above, verum readings are becoming a well-documented property of object marking constructions across Bantu languages.

Another key property of verum constructions that we can apply to Ikalanga is their non-deniability; that is, the speaker-certainty portion of the meaning ought to be considered non-propositional. Gutzmann & Castroviejo Miró (2011) analyze verum as a multidimensional conversational operator, whose propositional content is separable from its verum dimension. Building on a diagnostic from Gutzmann & Castroviejo Miró (2011), we see that this is true for the verum interpretation created by OOC in Ikalanga: denying the propositional content of an utterance does not simultaneously deny the speaker's certainty.

(16)   A: Nchídzí   w-á-gu- ngw-á   , busukwá .
       1.Nchidzi 1SM-PST-14OM-drink-FV   14.beer
       'Nchidzi DID drink the beer.' *(in an appropriate context)*

       B: Até málébeswa!
       NEG truth
       'That's not true!'
         • OK: It's not true that Nchidzi drank the beer.
         • #It's not true that you are certain of that.

We see, then, that Ikalanga OOC displays patterns that are hallmarks of verum interpretations: namely, OOC introduces a sense of speaker confidence that is separable from the sentence's propositional content, and is thus appropriate in classic verum contexts such as addressing listener doubt and denial. As far as we know, non-OOC sentences are acceptable in verum contexts (as in the other contexts noted below) but do not carry the emphatic import of the OOC sentences.

5 Focus on *v*P-internal constituents facilitates OOC

Apart from verum contexts, OOC in Ikalanga can also be facilitated by certain types of focus: specifically, OOC is natural when corrective or exhaustive focus falls on a structurally low adverb or on a distinct object. New information focus, however, has no such effect.[9] Focus on a *v*P-external constituent also fails to facilitate OOC. Sikuku & Diercks (2021a) present robust evidence for the effect that focus on *v*P-internal elements has on OM-doubling – they show that new information focus, 'only' (exhaustive) focus, and contrastive focus on a distinct *v*P-internal constituent facilitate doubling, a pattern that persists across monotransitives and ditransitives alike. In all instances, though, doubling constructions have a strong emphatic sense (e.g. a mirative reading, or other intensity-type readings). (17) illustrates a Lubukusu example where OM-doubling creates an emphatic interpretation on a low temporal adverbial.

(17)   Lubukusu (Sikuku & Diercks 2021a)

Q: Ba-ba-ana   ba-a-kes-a   ka-ma-indi **liina**?
   2-2-children 2SM-PST-harvest-FV 6-6-maize   when
   'When did the children harvest the maize?'

A: Ba-ba-ana   ba-(ka)-kes-ile   ka-ma-indi **likolooba**.
   2-2-children 2SM-6OM-harvest-PFV 6-6-maize   yesterday
   'The children harvested the maize YESTERDAY.'
   (It is particularly notable, for some reason, that it was yesterday that this happened.)

5.1 Corrective focus facilitates OOC

A number of the patterns seen in Lubukusu are similar to those in Ikalanga – for one, we see that corrective focus[10] on a *v*P-internal element facilitates OOC in

---

[9]These patterns align with the hierarchy of degrees of contrast proposed by Cruschina (2021).

[10]For the purposes of this paper, we assume corrective and contrastive focus to be equivalent.

Ikalanga. (18) establishes a context where corrective focus falls on the temporal adverb *mádékwe* 'yesterday'; in turn, the object *baná íbábáje* 'those children' naturally co-occurs with an OM. Notably, a verum reading is not necessary here.[11]

(18) Q: W-aká-bón-á   baná   íbábájé bé   ikwélé tshípí yáká píndá?
      1SM-PST-see-FV 2.child 2.DEM   ASSOC 7.school week last
      'Did you see those students last week?'

      A: Á, nd-aká-⸤bá-⸥ bón-á   **mádékwe** , ⸤baná íbábáje⸥ .
      No 1SG.SM-PST-2OM-see-FV yesterday   2.child 2.DEM
      'No, I saw those children YESTERDAY.'

Ditransitive constructions follow a similar pattern: in a lexical ditransitive, corrective focus on the recipient object facilitates OOC for the theme object, as in (19)A1. Consistent with our previous observations of OOC and dislocation, we see the OMed object *shangu* 'shoes' right-dislocated outside of the verb phrase. The focused recipient object *Lúdo* remains within *v*P. Notably, (19)A2 shows that the object cannot co-occur with an OM and be correctively focused; that is, corrective focus and OOC must be done on *different* objects, and dislocating the focused object is infelicitous.

(19) Q: A Nkádzí w-áka-p-á   Shátho shangú?
      Q 1.Nkadzi 1SM-PST-give-FV 1.Shatho 10.shoes
      'Did Nkadzi give Shatho shoes?'

      A1: Á, Nkádzí   w-áka-⸤dzí-⸥ p-á   **Lúdo** , ⸤shangu⸥ .
       No 1.Nkadzi 1SM-PST-10OM-give-FV 1.Ludo 10.shoes
      'No, Nkadzi gave LUDO shoes.'[12]

      A2: #Á, Nkádzí   w-áka-⸤m-⸥ p-á   shángú , ⸤**Lúdo**⸥ .
       No 1.Nkadzi 1SM-PST-1OM-give-FV 10.shoes 1.Ludo
      Intd. 'No, Nkadzi gave LUDO shoes.'

These facts are strongly reminiscent of Zulu, where it has been analyzed that *v*P is a focal domain (Cheng & Downing 2012 and Zeller 2015, among others);

---

[11]Here we adopt the common assumption that temporal adverbs like *mádékwe* 'yesterday' are low adverbs, adjoined to *v*P (see Henderson 2006, Sikuku & Diercks 2021a for similar assumptions and similar behaviors of low temporal adverbials).

[12]There seem to be multiple interpretive layers to this sentence beyond the corrective focus aspect. This sentence is emphasizing the fact that without a doubt, Nkadzi gave the shoes, and not anything else, to Ludo. It seems that the OMed object *shangu* 'shoes' is perhaps receiving some exhaustive emphasis, but a verum reading of certainty is also seemingly intertwined here.

focused material remains within *v*P, while non-focused material moves outside of *v*P (Buell 2006).

Corrective focus can also fall on the theme object, in which case it facilitates OOC for the recipient object:

(20) Q: A Nchídzí w-áka-p-á   Lúdó búrúkhwí?
         ǫ 1.Nchidzi 1sm-pst-give-fv 1.Ludo trousers
         'Did Nchidzi give Ludo trousers?'

      A: Á, Nchídzí w-áka-(m-) p-á   **shángú** , (Lúdo) .
         No 1.Nchidzi 1sm-pst-1om-give-fv 10.shoes   1.Ludo
         'No, Nchidzi gave Ludo sнoes.'

Benefactive applicatives follow the same pattern: in (21), corrective focus on the benefactive object facilitates OOC for the theme object. And again, we see that OOC for the same object that is correctively focused is unnatural.

(21) Q: A Ludó w-áka-bík-íl-á   Mpaphi nyama?
         ǫ 1.Ludo 1sm-pst-cook-appl-fv 1.Mpaphi 9.meat
         'Did Ludo cook meat for Mpaphi?'

      A1: Á, Ludó   w-áka-(i-) bík-íl-á   **Nchídzi** , (nyáma) .
          No 1.Ludo 1sm-pst-9om-cook-appl-fv 1.Nchidzi   9.meat
          'No, Ludo cooked meat for Nchidzi.'[13]

      A2: #Á, Ludó   w-áka-(m-) bík-íl-á   nyáma , (**Nchídzi**) .
          No 1.Ludo 1sm-pst-1om-cook-appl-fv 9.meat   1.Nchidzi
          Intd. 'No, Ludo cooked meat for Nchidzi.'

5.2 Exhaustive focus facilitates OOC

Exhaustive focus shows the same general effect as the patterns outlined above: when a *v*P-internal constituent is exhaustively focused with a focus particle, OOC is natural. (22) demonstrates this with a temporal adverb:

(22) Nd-aká-(bá-) bón-á   **mádékwé kóga** , (baná íbábáje) .
      1sm.sg-pst-2om-see-fv yesterday only   2.child 2.dem
      'I saw those children only yesterday.' *(i.e. not any other day)*

---

[13]Again, here we get an interesting multi-layered interpretation: per the first author, this sentence is emphasizing the fact that without a doubt, Ludo cooked the meat, and not anything else, for Nchidzi.

Additionally, in ditransitive constructions, exhaustively focusing one object facilitates OOC for the other:

(23) a. Nchídzí w-áka-[m-]p-á   **shángú kóga** , [Ludó] .
          1.Nchidzi 1SM-PST-1OM-give-FV 10.shoes only   1.Ludo

      'Nchidzi gave ONLY SHOES to Ludo.'

      b. Nchídzí w-áka-[dzí-]p-á   **Lúdó kóga** , [shángu] .
          1.Nchidzi 1SM-PST-10OM-give-FV 1.Ludo only   10.shoes

      'Nchidzi gave ONLY LUDO shoes.'[14]

Unlike the patterns shown for contrastive focus, however, an exhaustively focused object can also be associated with a co-occurring OM, but only when both the adverb and object have been dislocated. Further research is needed to discern the reason for this difference.

(24) a. Nd-áka-[bá-]bón-á   , [**báná íbábájé kóga**] madekwé.
          1SM.SG-PST-2OM-see-FV   2.child 2.DEM only   yesterday

      'I saw ONLY THOSE CHILDREN yesterday.' *(i.e. not any other children)*

      b. #Nd-áka-[bá-]bón-á   madekwe , [**baná íbábájé kóga**] .
          1SM.SG-PST-2OM-see-FV yesterday   2.child 2.DEM only

      Intd. 'I saw ONLY THOSE CHILDREN yesterday.' *(i.e. not any other children)*

5.3 New information focus does not facilitate OOC

While corrective and exhaustive focus on a *v*P-internal constituent allow for OOC, new information focus fails to do so on its own, regardless of whether it falls on an adverb (25), direct object (26), or indirect object (27):[15]

(25) Q: How did the children eat the okra?

      A: #Baná b-áka-[lí-]j-á   **ngébúnya** , [delele] .
          2.child 2SM-PST-5OM-eat-FV slowly   5.okra

      Intd. 'The children ate the okra SLOWLY.'

---

[14]For both these examples, they seem to require a context of doubt about what was given to be entirely natural, though further investigation into this is required.

[15](26) and (27) are unnatural responses to the question because they seem to overemphasize the object associated with the OM. For example, (26) sounds unnatural because though the question seeks emphasis on *shángú* 'shoes', OOC for *Lúdo* has the effect of overemphasizing it over *shángú*.

(26) Q: What did Nchidzi give Ludo?

   A: #Nchídzí w-áka-[m-] p-á   **shángú** , [Lúdo] .
   1.Nchidzi 1SM-PST-1OM-give-FV 10.shoes   1.Ludo
   Intd. 'Nchidzi gave Ludo SHOES.'

(27) Q: Who did Nchidzi give shoes to?

   A: #Nchídzí w-áka-[dzí-] p-á   **Lúdo** , [shangú] .
   1.Nchidzi 1SM-PST-1OM-give-FV 1.Ludo   10.shoes
   Intd. 'Nchidzi gave LUDO shoes.'

The answers in examples (25)–(27) are all grammatical sentences, but new information focus on its own is insufficient to license OOC; all of these examples would require additional emphatic readings (with appropriate licensing context) in order to be acceptable. In this way Ikalanga is similar to Lubukusu (which similarly requires emphasis in this way) and unlike Zulu, which appears to lack these emphatic readings with OMing constructions.

5.4 *v*P-external focus does not facilitate OOC

We have been careful to note thus far that the focus types discussed facilitate OOC when applied to an element *within the verb phrase* (namely, low adverbs and objects). This is an important specification to make, as focus on a *v*P-external constituent generally fails to facilitate OOC. For example, (28) below shows that corrective focus on a preverbal subject does not allow OOC.

(28) Q: Á Ludó w-áka-bón-a   báná íbábájé bé   íkwélé madékwé?
   Q 1.Ludo 1SM-PST-see-FV 2.child 2.DEM ASSOC 7.school yesterday
   'Did Ludo see those children yesterday?'

   A: #Á, **Nchidzí** w-áka-[bá-] bón-á   mádekwe , [baná íbabájé] .
   No 1.Nchidzi 1SM-PST-2OM-see-FV yesterday   2.child 2.DEM
   Intd. 'No, NCHIDZI saw those children yesterday.'

The overall generalization from this section remains clear: focus on *v*P-internal constituents serves as another licensing condition for OOC. Specifically, we have seen that corrective and exhaustive focus on low adverbs and distinct objects facilitate OOC, while new information focus on the same elements does not, nor does focus on a *v*P-external subject. This connection to *v*P-internal focus enabling OMing is familiar from Lubukusu (Sikuku & Diercks 2021b,a), Tiriki (Liu 2022), Wanga (Kuzmik 2022), Cinyungwe (Langa da Câmara et al. 2023), Zulu (Zeller 2015), and others.

6 Mirative focus facilitates OOC

The previous section explored how focus on a particular constituent can facilitate OOC; however, it is also possible for an emphatic interpretation to arise on the entire sentence as a whole, generating an expressive reading. Specifically, OOC in Ikalanga is licit if there is something unexpected and/or surprising within the construction, an interpretation that we analyze as mirative focus constructions (following Sikuku & Diercks 2021a and Lippard et al. forthcoming).

Recent research by Sikuku & Diercks (2021a), Langa da Câmara et al. (2023), and Lippard et al. (forthcoming) has analyzed mirativity as one of the emphatic interpretations generated by OOC across various Bantu languages. Similarly, past research has identified mirativity as one of a similar range of interpretations for predicate clefting (Lusekelo et al. 2023, Jerro & van der Wal 2022) and for a nominal-modifying particle (Asiimwe & van der Wal 2021). The most recognizable mirative context, perhaps, is one that makes all or part of a sentence surprising, unexpected, or shocking. Consequently, mirative interpretations are highly context-dependent. Mirative contexts are one of multiple kinds of emphatic contexts that naturally license OM-doubling in Lubukusu and Cinyungwe. Also noted by Sikuku & Diercks (2021a) regarding OM-doubling and mirativity is that – just as discourse context can create the licensing conditions for OM-doubling – context can just as well 'undo' them. That is, if an utterance that was once surprising and thus naturally OM-doubled becomes expected and unsurprising in a different context, OM-doubling is no longer natural (Lippard et al. forthcoming note the same for Cinyungwe).

In Ikalanga – just as in Lubukusu and Cinyungwe – OOC is readily associated with a mirative interpretation, and is facilitated by mirative contexts. OOC is licit when an utterance is especially surprising or unexpected, as in (29a):

(29)   *Context: The children love beans, and every time they are served beans, they eat them incredibly quickly because they like them so much. This time, however, when they sit down to eat, they eat the beans very slowly, which is quite unusual for them. In response, someone could say:*

   a. Baná   b-áka-⌊dzí-⌋ j-á   ngébúnya , ⌊nyemba⌋ .
       2.child 2SM-PST-10OM-eat-FV slowly   10.beans

       'The children ate the beans slowly!'[16]

---

[16]A double-dislocation construction is also possible here (notably, with two prosodic breaks): Baná b-áka-dzí-j-a, nyémba, ngebúnya.

    b. #Baná b-áka-j-á   nyémbá ngebúnya.
        2.child 2sm-pst-eat-fv 10.beans slowly
        Intd. 'The children ate the beans slowly.'

Notably, the canonical non-OMed sentence in (29b) is unnatural in the given context: the context makes the utterance unexpected, and without OOC, there is no surprise conveyed – a non-OMed sentence is thus less natural.[17]

The mirative emphasis can also fall on the entire sentence, indicating the whole event is shocking:

(30)   *Context: Nchidzi is known to not get along with Ludo; he is always very rude to her. But one day, he presents Ludo with a gift, and everyone is shocked.*
     Nchídzí   w-áka-m- p-á   shángú , Lúdo .
     1.Nchidzi 1sm-pst-1om-give-fv 10.shoes 1.Ludo
     'Nchidzi gave Ludo shoes.'

In addition to being licit in mirative contexts where a sentence is surprising and unexpected, OOC in Ikalanga is also licit when an utterance is highly informative (similar to Lubukusu; Sikuku & Diercks 2021a):

(31)   *Context: You arrive home to find that Nchidzi is acting very strangely and stumbling around. You ask someone what has happened to him. They respond:*
     Nchídzí   w-á-gu- ngw-á   , busukwá .
     1.Nchidzi 1sm-pst-14om-drink-fv   14.beer
     'Nchidzi drank beer.'

Here, the statement is particularly informative in that it offers an explanation to Nchidzi's bizarre behavior; OOC is thus acceptable.

---

[17]An anonymous reviewer points out that contexts like these suggest that there is focus on the adverb. "If Ikalanga, like many other Bantu languages, has the *v*P as a focus domain, or even has an immediate after verb (IAV) focus position, then what happens here is the non-focal object evacuating the focus domain (see Buell 2006)." The reviewer suggests that OOC here is not related to a mirative effect. If the analysis set forward by Sikuku & Diercks (2021a) and Lippard et al. (forthcoming) is on the right track for a variety of other Bantu languages, there is in fact a direct correlation between the emphatic reading and the focus semantics. Specifically, Sikuku & Diercks (2021a) argue that there is *both* a focal effect and an additional emphatic interpretation in Lubukusu OM-doubling, following the analysis of emphatic focus fronting by Bianchi et al. (2015) and Cruschina (2021). That is to say, the presence of a focal effect does not rule out the relevance of the emphatic readings, but rather the emphatic effects appear to themselves be focus-associated, requiring focus but adding a layer of interpretation in addition to focus.

Sentences can also be informative because of the sense of importance or gravity that they carry. In the example below, the speaker's use of OOC conveys the seriousness and newsworthiness of the situation; that is, it indicates there is something notable and important about having seen the children:[18]

(32) *Context: Last week, a group of children were being naughty, playing with a*
      water tap and wasting water, which is a rare commodity in their town.

      W-aká-ba- bón-á   baná íbábájé bé   ikwélé tshípí yáká píndá?
      1SM-PST-2OM-see-FV 2.child 2.DEM ASSOC 7.school week last

      'Did you see those students last week?'

Again consistent with Lubukusu (Sikuku & Diercks 2021a), we can see that if a mirative interpretation is 'undone,' OOC is no longer licit. The sentence in (33a) is surprising in the given context, and OOC is therefore natural (and preferred to a non-OMed sentence). When the same utterance becomes expected and no longer surprising in (34), however, OOC ceases to be acceptable, and the most appropriate response is a non-OMed sentence.

(33) *Context: The children hate eating beans, and never want to eat them. But*
      one day, they do eat the beans. Someone says, very surprised:

      a. Baná   b-áka-dzí- j-a   , nyémba !
          2.child 2SM-PST-10OM-eat-FV 10.beans

          'The children ate the beans!'

      b. Baná   b-áka-j-á   nyémba!
          2.child 2SM-PST-eat-FV 10.beans

          'The children ate the beans!' (*Less felicitous than* (33a))

(34) *Context: After that one miraculous day, the children realize that they*
      actually do like beans, and so now they eat them every day. One evening,
      someone asks what happened at dinnertime. You say:

      a. #Baná b-áka-dzí- j-a   , nyemba .
          2.child 2SM-PST-10OM-eat-FV 10.beans

          Intd. 'The children ate the beans.'

---

[18]Example (32) is an apparent exception to the requirement that objects co-occurring with OMs require dislocation of the object. Notably, this context (where a whole situation is surprising/upsetting) is precisely the same context that Sikuku & Diercks (2021a) identify as exceptional in Lubukusu as well (in that case, OM-doubling is exceptionally possible absent the normal requirement of OM-doubled objects to be discourse-given).

b. Baná   b-áka-j-á   nyémba.
  2.child 2SM-PST-eat-FV 10.beans
  'The children ate the beans.'

Following analyses of Romance focus fronting constructions (Bianchi et al. 2016, Cruschina 2021, 2019), Sikuku & Diercks (2021a) and Lippard et al. (forthcoming) analyze the emphatic readings in OM-doubling constructions as conventional implicatures, situated in a distinct tier of meaning from at-issue truth-conditional meaning. We can see that Ikalanga OOC in mirative contexts displays the expected characteristics of a conventional implicature. Specifically, the mirative 'surprise' aspect of an utterance is separable from the propositional content:

(35)   A: Baná   b-áka-dzí- j-á   ngébúnya , nyemba !
          2.child 2SM-PST-10OM-eat-FV slowly   10.beans
          'The children ate the beans slowly!' *(In an appropriate surprising context)*

       B: A-kútó-chénámísa!
          NEG-PRS-surprise
          'It's not surprising!'

Here, Speaker B denies that the event is surprising, but does so without denying the fact that the children ate the beans slowly. It is therefore evident that the not-at-issue (mirative) content and at-issue (propositional) content occupy different dimensions and are independent from the other, as is typical of a conventional implicature.

A second diagnostic for conventional implicatures, however, behaves unexpectedly in Ikalanga. Because conventional implicatures are non-cancellable, we would expect that a speaker who says a sentence with mirative focus *cannot* then deny that the sentence is surprising. That is, they must remain committed to the implicature that the utterance is surprising. However, this does not seem to be the case in Ikalanga; it seems to be possible for the speaker to deny their own surprise. This may be related to the fact that OOC in Ikalanga can convey a variety of interpretive effects beyond just mirativity (see Lippard et al. forthcoming).

(36)   a.   Baná   b-áka-dzí- j-á   ngébúnya , nyemba ...
          2.child 2SM-PST-10OM-eat-FV slowly   10.bean
          'The children ate the beans slowly!' *(In an appropriate surprising context)*

b. ... Ngóno akúna chinó chénámísá ípápo.
    but   NEG that surprise   there
    '...but there is nothing surprising there.'

Further research is needed to clarify this data point.

7 Unsolved puzzles: Areas for future research

7.1 Focused objects with object markers

The most notable area for future research – hinted at in footnotes throughout this paper – is the pattern of OMed objects sometimes seemingly being focused, and sometimes not. §5 presented numerous pieces of data that substantiated the generalization that OOC is possible if *some other* vP-internal constituent was focused. This observation is consistent with other Bantu languages as well (cf. Sikuku & Diercks (2021a) for Lubukusu, Langa da Câmara et al. (2023) for Cinyungwe, and Zeller (2015) for Zulu). Though there is robust evidence for this pattern, there also remain a number of data points that complicate, if not contradict, it – that is, at times, the object that co-occurs with an OM seems to be the element receiving emphasis or focus, rather than some distinct element remaining in vP. As an initial foray into this puzzle, consider this example:

(37) Q: Baná   b-áka-j-á   déléle chiní?
         2.child 2SM-PST-eat-FV 5.okra how
         'How did the children eat the okra?'

     A: #Baná b-áka-[lí-] j-á   ngébúnya , [déléle] .
        2.child 2SM-PST-5OM-eat-FV slowly   5.okra
        Intd. 'The children ate the okra slowly.'

Cruschina (2021) argues that different types of emphatic focus may be distinguished by the degrees of contrastive interpretation that they create. The low degree of contrast that new information focus carries can potentially explain why OOC is unacceptable in (37). However, there seem to be reasons beyond just this – the response in (37) is infelicitous because though the question searches for focus on the adverb *ngébúnya* 'slowly', the OMed object *déléle* 'okra' seems to be unnecessarily prominent; the sentence is emphasizing that it is the *déléle* that they ate, to the first author's ear.

This appears to contradict the interpretation of the same sentence in a corrective focal context:

(38) Q: Did the children eat the okra quickly?

A: A,   baná   b-áka-⎡lí-⎤j-á   ngébúnya , ⎡déléle⎤.
No 2.child 2sm-pst-5om-eat-fv slowly   5.okra
'No, the children ate the okra slowly.'

In (38), the OMing sentence sounds natural in response to the given question, and focus falls on the adverb as intended, rather than the doubled object. Yet the unexpected emphasis on the OMed object appears in a different contrastive focus context:

(39) Q: A Nchídzi   w-áka-p-á   Lúdó   búrúkhwí?
Q 1.Nchidzi 1sm-pst-give-fv 1.Ludo trousers
'Did Nchidzi give Ludo trousers?'

A: Á, Nchídzi   w-áka-⎡m-⎤p-á   shángú , ⎡Lúdo⎤.
No 1.Nchidzi 1sm-pst-1om-give-fv 10.shoes   1.Ludo
'No, Nchidzi gave Ludo shoes.'

The response in (39) is appropriate, and contrastive focus falls on *shángú* 'shoes' as expected, but the OMed object *Lúdo* does indeed feel emphasized here. Because OOC is not obligatory (leaving *Lúdo* out of the response would be a felicitous answer), it seems as if the speaker is including *Lúdo* for some significant reason. In the intuitions of the first author, whatever co-occurs with an OM feels like information that can be left out, and so by deliberately including the object, the speaker is emphasizing *Lúdo* (though we have yet to clarify the nature of the emphasis).

Exhaustive focus contexts also illustrate this puzzle – an object can co-occur with an OM and be exhaustively focused, which diverges from the patterns seen with corrective focus in (19) and (21) from §5.1:

(40) a. Nchídzi   w-áka-⎡m-⎤p-á   shángú , ⎡Lúdó kóga⎤.
1.Nchidzi 1sm-pst-1om-give-fv 10.shoes   1.Ludo only
'Nchidzi gave only Ludo shoes (no one else).'

b. Nchídzi   w-áka-⎡dzí-⎤p-á   Lúdó , ⎡shángú kóga⎤.
1.Nchidzi 1sm-pst-10om-give-fv 1.Ludo   10.shoes only
'Nchidzi gave Ludo only shoes (nothing else).'

Clearly, in both examples above, the object that co-occurs with an OM is also focused – yet, if we are to assume again here that OMed objects are 'extra' information that can optionally be left out, it is unclear how an OMed object could

also be focused and carry key information. It is also unclear why the non-doubled objects remaining in *v*P (above, *shángú* and *Lúdó*, respectively) do not appear to be receiving any emphasis or focus. This is a clear divergence from the generalization that focusing one object facilitates OOC of the other.

Although exhaustive focus fell on the co-occuring objects above, it is still possible to focus and OM distinct objects. (41) below follows the expected pattern:

(41) Nchídzi w-áka-[m-] p-á   shángú kóga , [Ludó] .
      1.Nchidzi 1sm-pst-1om-give-fv 10.shoes only   1.Ludo
      'Nchidzi gave Ludo only shoes (nothing else).'

We are thus presented with a puzzle in which an object co-occurring with an OM is sometimes simultaneously focused, and sometimes not; still other times, the OMed object seems to receive some type of emphasis by virtue of being included in the utterance. Future research is needed to, first, more precisely delineate the empirical facts of these patterns, and ultimately analyze it from a theoretical standpoint.

7.2 Overlapping interpretive readings

Though in this paper we present verum, *v*P-internal focus, and mirative interpretations as distinct phenomena, there are multiple instances in which various readings seem to be intertwined. Lippard et al. (forthcoming) broach the idea that various emphatic interpretations (such as mirativity, reprimand readings, verum, and exhaustivity) may be related to others. Indeed, in Ikalanga, it is difficult at times to disentangle different readings.

As an example of possibly overlapping interpretations, consider (42), replicated from (21) in §5.1:

(42) Q: A Ludó   w-áka-bík-íl-á   Mpaphi nyamá?
          q 1.Ludo 1sm-pst-cook-appl-fv 1.Mpaphi 9.meat
          'Did Ludo cook meat for Mpaphi?'
      A: Á, Ludó   w-áka-[í-] bík-íl-á   Nchídzi , [nyáma] .
          No 1.Ludo 1sm-pst-9om-cook-appl-fv 1.Nchidzi   9.meat
          'No, Ludo cooked meat for Nᴄʜɪᴅᴢɪ.'

The response in (42) is emphasizing the fact that, without a doubt, Ludo cooked the meat and not anything else for Nchidzi. There are multiple layers to this comment. First, because the response is licit in the established corrective focus context, we know the recipient object *Nchídzi* is being correctively focused. Yet the

meaning extends beyond this focus: there also seems to be a verum-like reading, since the sentence is uttered "without a doubt". Furthermore, the interpretation that Ludo cooked meat "and not anything else" suggests an exhaustive reading is also present. The question therefore arises of whether some of these various readings are related, or if they are all indeed distinct phenomena. If the latter is true, it must then also be determined in which situations they all arise.

8 Conclusions

From the empirical facts presented in this paper, it is abundantly clear that the focal and emphatic effects that OOC has in Ikalanga are wide-ranging. Though OOC is infelicitous in out-of-the-blue situations, specific pragmatic contexts can make OOC entirely natural – specifically, OOC can create a verum reading, licit in verum contexts where the speaker intends to assert their confidence and address listener denial or doubt, similar to English emphatic *do*. OOC is also associated with a mirative reading of surprise: when context makes all or part of an utterance shocking, newsworthy, or highly informative, OOC is acceptable. We have also seen that OOC is facilitated by certain focus environments. When a vP-internal constituent is focused with corrective or exhaustive focus, OOC is licit; new information focus on a vP-internal constituent or focus on a vP-external element, however, are insufficient licensing conditions. In all cases of OOC, we saw that the OMed object is obligatorily dislocated, and that OOC and dislocation cannot occur independent of each other.

The Ikalanga facts here are significant in how they corroborate and overlap with ongoing work in other related Bantu languages, yet also offer new and distinct patterns that further enrich the current research on Bantu OOC. There is a lot of similarity between Lubukusu and Ikalanga in that there are clear emphatic interpretations associated with OMing constructions that appear to be focus-related; that said, Lubukusu OM-doubling leaves all relevant objects and emphasized elements inside the vP. Ikalanga is in this way more Zulu-like, with necessary links between OMing and dislocation of the associated object. But Zulu OMing appears to be strictly a backgrounding operation, whereas a broader range of interpretive effects are associated with Ikalanga OMing. In this way, Ikalanga appears to be adding a new constellation of facts regarding object marking, but more work is necessary to fully understand (and analyze) the relevant constructions.

Acknowledgments

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

ASSOC	associative marker	FV	final vowel
AUG	augment	OM	object marker
DJ	disjoint	SM	subject marker

References

Asiimwe, Allen & Jenneke van der Wal. 2021. The multifunctionality of -o in Rukiga: Pronoun, contrastive topic, and mirative marker. *Nordic Journal of African Studies* 30(1). 1–26.

Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2015. Focus fronting and its implicatures. In Enoch Oladé Aboh, Jeanette Schaeffer & Petra Sleeman (eds.), *Romance languages and linguistic theory 2013: Selected papers from 'Going Romance' Amsterdam 2013*, 1–20. Amsterdam: John Benjamins.

Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2016. Focus fronting, unexpectedness, and evaluative implicatures. *Semantics and Pragmatics* 9. Article 3: 1–54. DOI: DOI:10.3765/sp.9.3.

Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63. 741–782.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 22(2). 147–186.

Buell, Leston. 2006. The Zulu conjoint/disjoint verb alternation: Focus or constituency? *ZAS Papers in Linguistics* 43. 9–30.

Cheng, Lisa Lai-Shen & Laura J. Downing. 2012. Against FocusP: Arguments from Zulu. In Ivona Kučerová & Ad Neeleman (eds.), *Contrasts and positions in information structure*, 246–266. Cambridge: Cambridge University Press.

Cruschina, Silvio. 2019. Focus fronting in Spanish: Mirative implicature and information structure. *Probus* 29(1). 119–146. DOI: DOI:10.1515/probus-2018-0008.

Cruschina, Silvio. 2021. The greater the contrast, the greater the potential: On the effects of focus in syntax. *Glossa: A journal of general linguistics* 6(1). 3. 1–30. DOI: DOI:10.5334/gjgl.1100.

Güldemann, Tom. 2003. Present progressive vis-à-vis predication focus in Bantu: A verbal category between semantics and pragmatics. *Studies in Language* 27(2). 323–360. DOI: DOI:10.1075/sl.27.s.05gul.

Güldemann, Tom. 2016. Maximal backgrounding = focus without (necessary) focus encoding. *Studies in Language* 40(3). 551–590. DOI: 10.1075/sl.40.3.03gul.

Gutzmann, Daniel & Elena Castroviejo Miró. 2011. The dimensions of verum. In Oliver Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics 8*, 143–165.

Gutzmann, Daniel, Katharina Hartmann & Lisa Matthewson. 2020. Verum focus is verum, not focus: Cross-linguistic evidence. *Glossa: A Journal of General Linguistics* 5(1). 51. 1–48.

Henderson, Brent. 2006. *The syntax and typology of Bantu relative clauses.* Urbana, IL: University of Illinois at Urbana-Champaign. (Doctoral dissertation).

Jerro, Kyle. 2015. Revisiting object symmetry in Bantu. In Ruth Kramer, Elizabeth Zsiga & One Tlale Boyer (eds.), *The Selected Proceedings of the 44th Annual Conference on African Linguistics*, 130–145. Somerville, MA: Cascadilla Proceedings Project.

Jerro, Kyle & Jenneke van der Wal. 2022. Contrastive topic and predicate doubling in Bantu. Poster presented at Sinn und Bedeutung 27.

Kerr, Elisabeth J. & Jenneke van der Wal. 2023. Indirect truth marking via backgrounding: Evidence from Bantu. Ms, Leiden University.

Kuzmik, Jackson. 2022. *Object marking in Wanga.* Cambridge University. (MA thesis).

Langa da Câmara, Crisófia, Michael Diercks, Madelyn Colantes, Brendan Ly, Jackson Kuzmik & Hannah Lippard. 2023. An initial look at object marking in Cinyungwe. To appear in the Proceedings of ACAL 51/52.

Letsholo, Rose. 2002. *Syntactic domains in Ikalanga.* University of Michigan. (Doctoral dissertation).

Letsholo, Rose. 2013. Object markers in Ikalanga. *Linguistic Discovery* 11(1). 105–128.

Lippard, Hannah, Justine Sikuku, Crisófia Lánga da Câmara, Madelyn Colantes, Kang (Franco) Liu & Michael Diercks. Forthcoming. Emphatic interpretations of object marking in Bantu languages. *Studies in African Linguistics.*

Liu, Kang. 2022. *Object marking in Tiriki.* Cambridge University. (MA thesis).

Lusekelo, Amani, Allen Asiimwe, Patrick Kanampiu, Nelsa Nhantumbo, Ernest Nshemezimana & Jenneke van der Wal. 2023. Predicate doubling and information structure in eastern Bantu languages. Presentation at the BaSIS conference at Leiden University, June 8–9, 2023.

Marlo, Michael R. 2014. Exceptional patterns of object marking in Bantu. *Studies in African Linguistics* 43. 85–123.

Marlo, Michael R. 2015. On the number of object markers in Bantu languages. *Journal of African Languages and Linguistics* 36. 1–65.

Marten, Lutz & Nancy Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30(2). 237–253.

Riedel, Kristina. 2009. *The syntax of object marking in Sambaa: A comparative Bantu perspective.* Leiden: Universiteit Leiden. (Doctoral dissertation).

Sikuku, Justine & Michael Diercks. 2021a. *Object marking in Bukusu: At the interface of pragmatics and syntax.* Ms, Moi University and Pomona College. https://research.pomona.edu/mjkd/papers/.

Sikuku, Justine & Michael Diercks. 2021b. Object marking in Lubukusu: Information structure in the verb phrase. In Akinbiyi Akinlabi, Lee Bickmore, Michael Cahill, Michael Diercks, Laura J. Downing, James Essegbey, Katie Franich, Laura McPherson & Sharon Rose (eds.), *Celebrating 50 years of ACAL: Selected papers from the 50th Annual Conference on African Linguistics*, 305–330. Berlin: Language Science Press.

Sikuku, Justine, Michael Diercks & Michael R. Marlo. 2018. Pragmatic effects of clitic doubling: Two kinds of object markers in Lubukusu. *Linguistic Variation* 18(2). 359–429.

van der Wal, Jenneke. 2015. Bantu object clitics as defective goals. *Revue Roumaine de Linguistique* LX(2-3). 277–296.

van der Wal, Jenneke. 2020. The AWSOM correlation in comparative Bantu object marking. In Peter W. Smith, Johannes Mursell & Katharina Hartmann (eds.), *Agree to agree: Agreement in the Minimalist Programme*, 199–234. Berlin: Language Science Press. DOI: 10.5281/zenodo.3541755.

van der Wal, Jenneke. 2022. *A featural typology of Bantu agreement.* Oxford: Oxford University Press.

Zeller, Jochen. 2012. Object marking in isiZulu. *Southern African Linguistics and Applied Language Studies* 30. 219–325. DOI: 10.2989/16073614.2012.737600.

Zeller, Jochen. 2015. Argument prominence and agreement: Explaining an unexpected object asymmetry in Zulu. *Lingua* 156. 17–39. DOI: 10.1016/j.lingua.2014.11.009.

Chapter 15

(A)symmetry in double object constructions in Tiriki

Kang Franco Liu

University of California, Berkeley

In this paper, I describe and analyze a range of object properties exhibited in double object constructions (DOCs) in Tiriki (Bantu, Luyia; JE413). The preliminary investigation makes clear that Tiriki is symmetrical: both objects of Tiriki DOCs display primary object properties in object-marking, passivization, *wh*-clefting, and relativization. Asymmetry, however, surfaces when one object is passivized and the other undergoes Ā-movements (cf. 'Double Object Movement Asymmetry' or DOMA in Holmberg et al. 2019). While the aforementioned symmetry and emergent asymmetry are fully captured by a high applicative structure that allows for flexible licensing of internal arguments and exhibits phasal properties, Tiriki instantiates two additional DOMA-triggering contexts: passivization combined with left dislocation and *wh*-in-situ. Based on various diagnostics targeting Ā-dependencies, I show that neither left dislocation nor *wh*-in-situ requires a movement analysis, thereby articulating the analytical challenges that these configurations, especially left dislocation, pose to the current movement-based DOMA analysis.

1 Introduction

This paper explores aspects of the morphosyntax of Tiriki (Bantu, Luyia; JE413), an under-documented language of Western Kenya and Eastern Uganda (Lewis et al. 2016). Specifically, I investigate issues related to object properties in Tiriki double object constructions (DOCs), with an emphasis on describing and deriving their (a)symmetry under a range of syntactic operations.

As is common in most Bantu languages, Tiriki DOCs can be either introduced with lexical ditransitive predicates or derived with valency-increasing suffixes. See below for examples of Tiriki DOCs:

(1) a. Lexical ditransitive *-manyinya* 'show'
Hardley a-manyiny-e   va-somi   vi-tapu
Hardley 1SM-show-FV.PST 2-student 8-book
'Hardley showed students books.'

b. Benefactive applicative *-tekhela* 'cook for'
Ebby a-tekh-el-e   va-cheni vy-apati
Ebby 1SM-cook-APPL-FV.PST 2-guest 8-chapati
'Ebby cooked chapatis for the guests.'

c. Causative *-ng'wekhitsa* 'feed (make … eat)'
Ebby a-ng'wekh-its-e   shi-paka ma-vele
Ebby 1SM-drink-CAUS-FV.PST 7-cat   6-milk
'Ebby fed the cat milk.'

As proposed in previous work (cf. Kimenyi 1980, Baker 1988, Bresnan & Moshi 1990), Bantu languages vary with regard to whether one or both internal arguments of DOCs show primary object properties. Tiriki is by and large a symmetrical language, meaning that "different [internal] arguments can simultaneously have primary object properties" (Bresnan & Moshi 1990: 153). As shown below, both objects can be promoted to be the grammatical subject in passives or the head noun in relative clauses:

(2) a. Passivization
i.   [va-somi] va-manyiny-w-e   vi-tapu
2-student 2SM-show-PASS-FV.PST 8-book
'Students were shown books.'

ii. [vi-tapu] vi-manyiny-w-e   va-somi
8-book   8SM-show-PASS-FV.PST 2-student
'Books were shown to the students.'

b. Object relative clauses (RCs)
i.   [va-somi] v-a   Hardley a-manyiny-e   vi-tapu
2-student 2-REL Hardley 1SM-show-FV.PST 8-book
'the students that Hardley showed the books to'

ii. [vi-tapu] vy-a Hardley a-manyiny-e   va-somi
8-book   8-REL Hardley 1SM-show-FV.PST 2-student
'the books that Hardley showed the students'

Asymmetry, however, emerges when the internal arguments undergo a combination of syntactic operations. When the theme object is passivized, the indirect/applied object is unable to undergo Ā-movements like relativization whereas the reverse is grammatical.

(3) Object RCs with passivization

    a. *va-somi v-a   vi-tapu vi-manyiny-w-e
        2-student 2-REL 8-book 8SM-show-PASS-FV.PST
        Intended: 'the students that the books were shown to'

    b. vi-tapu vy-a va-somi   va-manyiny-w-e
        8-book 8-REL 2-student 2SM-show-PASS-FV.PST
        'the books that the students were shown'

Similar emergent object asymmetries have been attested in other symmetrical Bantu languages like Zulu and Lubukusu (see Holmberg et al. 2019 and references therein). What Tiriki patterns contribute to the existing body of literature is 1) another case study where the currently available analysis and relevant predictions can be tested, and 2) novel asymmetry patterns that pose challenges to the current analysis. For example, *wh*-in-situ, where no overt *wh*-movement is observed, also incurs asymmetry when combined with passivization in Tiriki:

(4) Object *wh*-in-situ with passivization

    a. *vi-tapu vi-manyiny-w-e   w-ina
        8-book 8SM-show-PASS-FV.PST 1-who
        Intended: 'The books were shown who?' (*non-echo question*)

    b. va-somi va-manyiny-w-e   sh-ina
        2-student 2SM-show-PASS-FV.PST 7-what
        'The students were shown what?' (*non-echo question*)

In this paper, I focus on only lexical and derived ditransitive predicates whose internal arguments bear thematic roles of beneficiary (BEN), recipient (REC), and Theme (THE).[1] I use 'theme objects' and 'direct objects' (DOs) interchangeably

---

[1]As for applied objects with other thematic roles, causee objects show generally the same patterns as recipient/beneficiary objects and are therefore not elaborated on. I also exclude instrumental and locative applied objects from the discussion. It is still disputed whether objects with these thematic roles are base-generated in the same underlying structure as that of other DOCs (Baker 1988, Nakamura 1997, Jerro 2016), and clarifying the answer to this dispute is outside the scope of this paper.

throughout this paper and refer to non-theme objects collectively as indirect objects (IOs) for ease of exposition.

The rest of the paper is structured as follows: In §2, I introduce basic object symmetry patterns in Tiriki across a range of syntactic contexts (e.g. object-marking, passivization, *wh*-clefts, relative clauses) and analyze them by applying the flexible licensing approach (van der Wal 2017, 2022) to a high applicative DOC structure (Pylkkänen 2008). §3 details how object symmetry is lost when one object is passivized and the other attempts to be Ā-extracted. Besides confirming the general compatibility between the Tiriki facts and the extant account of emergent object asymmetry (Holmberg et al. 2019), I present data on two additional configurations that result in asymmetry, namely left dislocation and *wh*-in-situ in passives. Multiple diagnostics are also performed to ascertain the nature of the syntactic dependencies involved and the implications for a movement-centered account of asymmetry. §4 summarizes the findings and concludes the paper.

2 Deriving object symmetry from structural asymmetry

In this section, I illustrate a range of object properties displayed in Tiriki DOCs with new empirical data. Then, I draw on insights from previous work on Bantu DOCs (chiefly van der Wal 2022, but also see among others Bresnan & Moshi 1990, Jerro 2016, van der Wal 2017) and develop a working analysis of Tiriki object symmetry.

2.1 Object properties in Tiriki DOCs

At first blush, Tiriki exhibits largely symmetrical object properties. Both objects can be object-marked (in (5a)), be promoted to the subject position (i.e. passivization in (5b)), and be involved in Ā-movement dependencies (e.g. *wh*-clefts in (5c) and relative clauses in (5d)). In summary, both internal arguments of Tiriki DOCs are capable of undergoing movement and non-movement operations alike.

(5) a. Object-marking
     i. Hardley a-va -manyiny-e   vi-tapu
         Hardley 1sm-2om-show-fv.pst 8-book
         'Hardley showed them (i.e. students) books.'
     ii. Hardley a-vi -manyiny-e   va-somi
         Hardley 1sm-8om-show-fv.pst 2-student
         'Hardley showed them (i.e. books) to the students.'

b. Passivization

i. [va-somi] va-manyiny-w-e vi-tapu
2-student 2SM-show-PASS-FV.PST 8-book
'Students were shown books.'

ii. [vi-tapu] vi-manyiny-w-e va-somi
8-book 8SM-show-PASS-FV.PST 2-student
'Books were shown to the students.'

c. Object *wh*-clefts

i. [wina] w-a Hardley a-manyiny-e vi-tapu
1-who 1-REL Hardley 1SM-show-FV.PST 8-book
'Who did Hardley show the books to?'

ii. [shina] sh-a Hardley a-manyiny-e va-somi
7-what 7-REL Hardley 1SM-show-FV.PST 2-student
'What did Hardley show to the students?'

d. Object relative clauses (RCs)

i. [va-somi] v-a Hardley a-manyiny-e vi-tapu
2-student 2-REL Hardley 1SM-show-FV.PST 8-book
'the students that Hardley showed the books to'

ii. [vi-tapu] vy-a Hardley a-manyiny-e va-somi
8-book 8-REL Hardley 1SM-show-FV.PST 2-student
'the books that Hardley showed the students'

Notably, the only object asymmetry in Tiriki is manifested in post-verbal word order. In neutral discourse contexts, Tiriki speakers always default to the REC/BEN > THE word order and judge the inverse infelicitous:

(6) Canonical post-verbal word order in Tiriki DOCs

a. Hardley a-manyiny-e va-somi vi-tapu
Hardley 1SM-show-FV.PST 2-student 8-book

b. #Hardley a-manyiny-e vi-tapu va-somi
Hardley 1SM-show-FV.PST 8-book 2-student
'Hardley showed students books.' (*in neutral discourse contexts*)

In fact, the inverse word order is perfectly natural in the appropriate contexts. It has been widely documented across Bantu languages that information structure affects post-verbal word orders (see van der Wal 2006, 2009 on Makhuwa; Buell 2009, Cheng & Downing 2012, Zeller 2014 on Zulu; Selvanathan 2019, Sikuku &

Diercks 2021 on Lubukusu, among others). For example, Aghem displays focus-related word order variation, where post-verbal focused elements can be scrambled to the immediately-after-the-verb (IAV) position (Watters 1979). In Tiriki, a topicality-related IAV effect can be observed. For all DOCs exemplified above, an aboutness-topic prompt allows the lower object to surface in the IAV position:

(7)   Topicality-driven IAV effect in Tiriki

A: m-bol-el-a   shi-ndu khu   vy-apati
    1SG.OM-say-APPL-FV 7-thing about 8-chapati
    'Tell me something about the chapatis.'

B: Ebby a-tekh-el-e   **vy-apati** va-cheni
    Ebby 1SM-cook-APPL-FV.PST 8-chapati 2-guest
    'Ebby cooked chapatis for the guests.'

I will not further elaborate on this pragmatically regulated word-order variation as it is not directly pertinent to the inquiry at hand. This brief description of IAV effect is intended to clarify what I refer to as 'canonical word order,' the use of # diacritics, and that this word-order asymmetry is due to information structure.

2.2 Analysis of Tiriki object symmetry

In terms of the underlying structure of Tiriki DOCs, I take it that the canonical post-verbal word order straightforwardly encodes the c-command asymmetry between the two internal arguments' base-generated positions (given the robust relationship between c-command asymmetry and linear order especially in head-initial languages, cf. Kayne 1994). The default word order of IO > DO thus suggests that the IO asymmetrically c-commands the DO in Tiriki DOCs, which is common in most DOCs cross-linguistically (Barss & Lasnik 1986, Marantz 1993). I adopt Pylkkänen's (2008) high applicative structure for Tiriki DOCs, in which the applicative head (Appl) introduces an external-argument-like IO and relates said object and the event described by the predicate. The high applicative structure is schematized in Figure 1.[2]

This syntactic structure for Tiriki DOCs is supported by the fact that IOs in Tiriki ('student' in the example below) can be modified by the depictive secondary predicate ('tired' in this case), a property unique to high applicative structures (Pylkkänen 2008).

---

[2]See Pylkkänen 2008 for discussion on *v*P versus VoiceP (Kratzer 1996). In this paper, I conflate the two and use *v*P throughout.
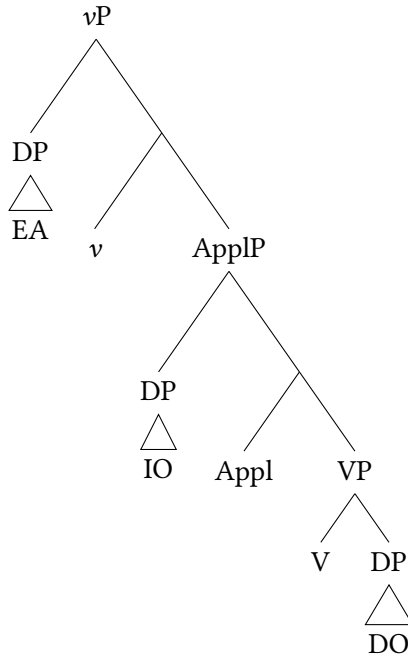
Figure 1: Structure of Tiriki DOCs

(8) High applicative DOCs: Depictives
Hardley a-manyiny-e   mu-somi vi-tapu nachoti
Hardley 1sm-show-fv.pst 1-student 8-book 1tired
'Hardley showed a student books (while the student is) tired.'

Also, Tiriki patterns with other high applicative Bantu languages like Kichaga (Bresnan & Moshi 1990; also see Pylkkänen 2008 for the high applicative analysis thereof) in that only discourse-familiar DOs but *not* IOs can undergo null object drop (or unspecified object deletion). Henderson (2018) explains this generalization by claiming that the absence of discourse-familiar DOs simply results from externally merging the applied object above an intransitive VP.

(9) High applicative DOCs: Null object drop
   a. Ebby a-tekh-el-e   va-cheni ma-kaanda
      Ebby 1sm-cook-appl-fv.pst 2-guest   6-bean
      'Ebby cooked beans for the guests.'

    b. *Ebby a-tekh-el-e   ∅ ma-kaanda
       Ebby 1SM-cook-APPL-FV.PST   6-bean
       'Ebby cooked beans (for them).'

    c. Ebby a-tekh-el-e   va-cheni ∅
       Ebby 1SM-cook-APPL-FV.PST 2-guest
       'Ebby cooked (them) for the guests.' (Michael Diercks, personal communication)

Now, I derive Tiriki object symmetry from the high applicative structure using largely the same theoretical machinery presented in Holmberg et al. (2019). In the case of object-marking, I assume the following: First, Tiriki object markers (OMs) arise in the lower domain, as opposed to an independent functional head base-generated atop the verb's landing site like other Bantu prefixes (Julien 2002)[3]; specifically, they are the spellout of ϕ-features on v (contra Sikuku et al. 2018, Sikuku & Diercks 2021). Second, I take Tiriki OMs to reflect the ϕ-agreement between a Probe on v and a defective DP Goal in the thematic object position, following Roberts (2010) and relatedly van der Wal (2015, 2022). I adopt van der Wal's (2022) definition of goal defectivity (modified based on the original conception in Roberts 2010: 62), which states that a defective Goal contains only a subset of the Probe's features. Concretely for ϕ-agreement, a defective Goal would be maximally a ϕP (Déchaine & Wiltschko 2002), which has no D feature. In this Probe-Goal relation, ϕ-features are shared between the Probe and the Goal upon Agree, forming a chain of copies, and only the highest copy gets spelled out and realized as an OM on v due to chain reduction (cf. Nunes 2004). In the case where a non-defective Goal (i.e. a full object DP) is probed, no such chain reduction takes place as the features on the Goal DP do not constitute a subset of those on the Probe. Upon Agree, the ϕ-Probe on v copies back ϕ-features from the Goal DP, which remain unpronounced, and only the full Goal DP spells out. These two configurations are illustrated below:

---

[3]This view on the status of Bantu OMs is informed by work on Bantu morphophonology (e.g. Hyman 2003, Marlo 2013, 2015). They show that Bantu OMs, despite their prefixal position, behave as though they form one unit with their verb stems, i.e. a macrostem. Together, they are subject to tonal processes, such as inflectional tone assignments and repairs to potential violations of the Obligatory Contour Principle.

(10) Object agreement in Tiriki

a.

b.
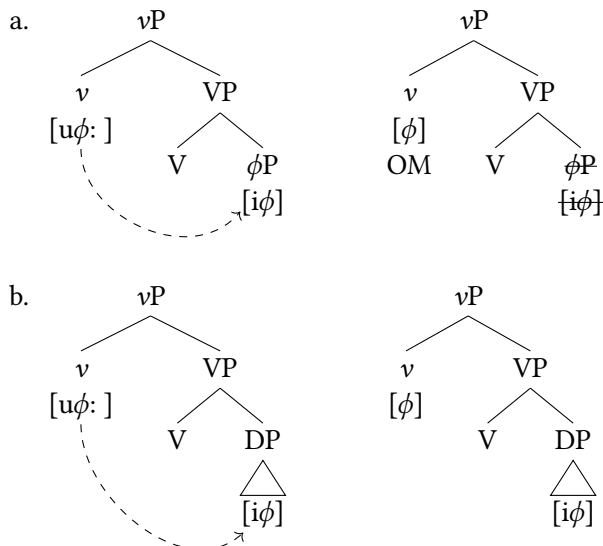
This analysis correctly captures the fact that Tiriki OMs cannot co-occur with their co-referential *in-situ* object DPs in neutral discourse contexts, thereby accomplishing an incorporation effect of object-marking using a purely Agree-based approach:

(11) Illicit OM doubling in Tiriki

#a-[mu] -lol-i   [∅-raisi]
1SM-1OM-see-FV.PST 1-president

'He saw him, the president.' (*neutral discourse context*)[4]

On a high applicative analysis of Tiriki DOCs, object-marking a DO in the presence of an IO is impossible without violating principles of locality and minimality (Rizzi 1990, Chomsky 1995), for the IO always intervenes as the closer Goal. van der Wal (2022) points out the same problem in symmetrical Bantu languages and identifies several possible solutions which either make the DO an equally close,

---

[4]It has been well-documented for Bantu languages that while OMs behave like incorporated pronouns in neutral discourse contexts, they can also co-occur with their co-referent object DPs (i.e. OM/clitic doubling) in certain pragmatic contexts. I do not discuss this further in this paper, but I direct the interested reader to Liu (2022) on Tiriki, Sikuku et al. (2018) and Sikuku & Diercks (2021) on Lubukusu, Bax & Diercks (2012) on Manyika, and Lippard et al. (forthcoming) for a comparative overview.

if not closer, Goal, or render the closer IO invisible to the Probe and thus inert for Agree. She eventually advocates for the approach of flexible licensing, following van der Wal (2017) and Haddican & Holmberg (2019).

The details are as follows: In Bantu DOCs, both internal arguments require Case-licensing, a mechanism independent of the realization of ϕ-features (following Carstens 2005). Rather than making v the only Case-licensor of the clause, this approach allows the lower functional head Appl to flexibly Case-license either the structurally higher IO or the lower DO (Figure 2).

The object-marking symmetry readily follows from this flexibility in Case-licensing. When Appl assigns Case (and theta role) upward to the IO, the IO becomes deactivated and thus invisible to v (Chomsky 2001). v can then probe downward to a defective DO, Case-license it, and realize it as an OM via Agree. When the DO gets Case-licensed first by Appl, v then proceeds to assign Case and Agree with the IO.[5] I will adopt this analysis for Tiriki object-marking symmetry in DOCs.

In the same vein, the object symmetry in passivization, *wh*-clefts, and relativization can be achieved. Concretely for passivization, the higher Probe becomes T as v does not assign Case in passive voice. DO passives, namely the curious case where the higher Probe T successfully assigns nominative Case to the lower Goal and attracts it to Spec,TP, are made possible by a version of Phase Impenetrability Condition (Chomsky 2001) and a contextually based definition of the lower phase (Bošković 2014, 2015):

(12)   Phase Impenetrability Condition (see Holmberg et al. 2019 and references therein)
Given a structure [$_{ZP}$ Z ...[$_{XP}$ X [HP α [H YP]]]] where H and Z are phase heads, the domain of H is not accessible to operations at ZP; only H and its edge (i.e. the outermost specifier of HP) are accessible to such operations.

(13)   Bošković's (2015) definition of phase
α is the head of a phase Ph making up a thematic domain if and only if α is the highest head introducing an argument in Ph.

In the case of passive DOCs, the lower phase is ApplP. Appl first licenses the IO, and upon completion of the lower phase, the DO undergoes movement to outer Spec,ApplP with its [uCase] feature (McGinnis 2001, Aldridge 2004, Bošković

---

[5]Note that this is the *only* available Case-licensing mechanism in asymmetrical single-object-marking Bantu languages (e.g. Swahili). See van der Wal (2020) for more discussion on this.
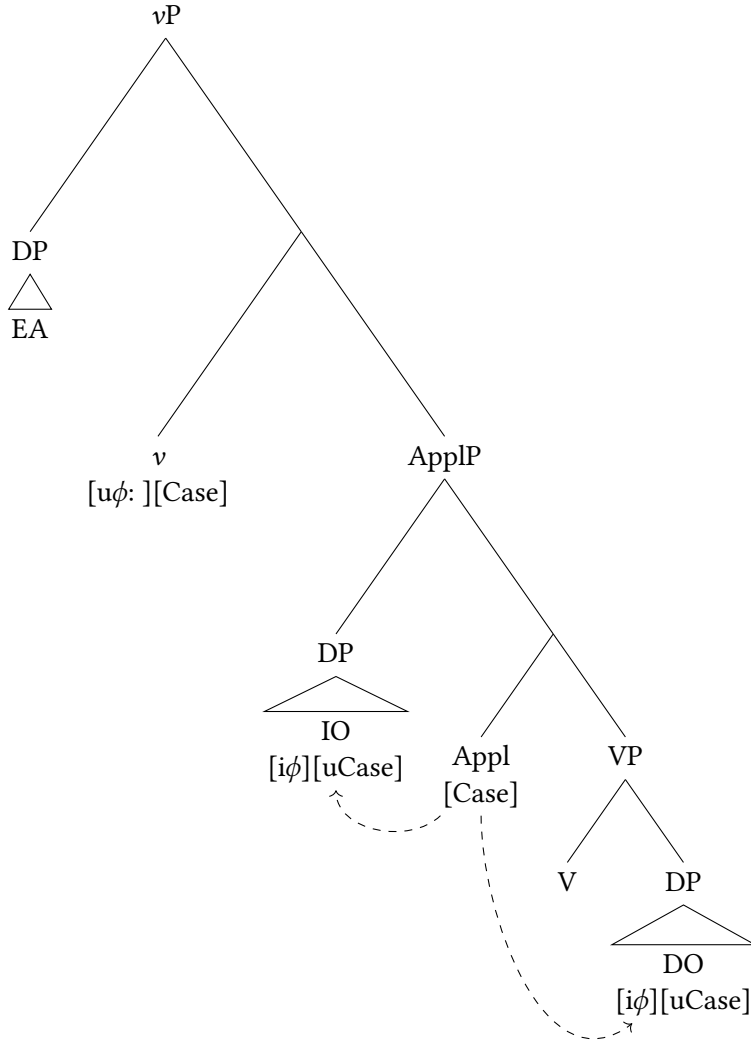
Figure 2: Flexible licensing via Appl (adapted from Holmberg et al. 2019: 687)

2016). This movement follows from Bošković's (2007) account of greed-driven movement, which stipulates that XPs whose uninterpretable features cannot be valued phase-internally must escape through the phase edge to avoid derivational crash. Following this movement to the lower phase edge, the DO gets Case-licensed by T. The opposite mechanism, where Appl licenses the DO and T probes the IO, derives IO passives. In the contexts of *wh*-questions and relativization (in active voice), the lower phase is vP. IO and DO receive Case from v and Appl, and the object that still bears an unvalued [u*wh*] feature moves to the outermost layer of Spec,vP and is probed for subsequent Ā-movement when C is merged and the complement of vP is transferred.

All object symmetry facts in Tiriki DOCs have been accounted for thus far with the help of flexible licensing and phasal properties. In the next section, I shed light on the unexpected object asymmetry that emerges from combinations of A-movement, Ā-movement, and non-movement operations in Tiriki—the latter of which is not yet discussed in previous Bantu literature.

3 Deriving (movement) asymmetry from object symmetry

Interestingly, though Tiriki DOCs show symmetry in object-marking, passivization, *wh*-questions, and relativization, asymmetry surfaces under combinations of the aforementioned syntactic operations. Specifically, although both the IO and the DO can be A- and Ā-moved on their own in these languages, a DO can be Ā-extracted from an IO passive, but not vice versa. Such unexpected asymmetries in generally symmetrical DOCs have been widely reported in previous Africanist literature (e.g. Visser 1986 on Xhosa; De Guzman 1987 on Swati; Adams 2010 and Zeller 2012 on Zulu, among others). Holmberg et al. (2019), one of the most recent analytical accounts, provide a cross-linguistic survey of relevant patterns in Norwegian, Northwest British English, Zulu, and Lubukusu. I fully adopt their theoretical assumptions and final proposal for my analysis of similar asymmetries in Tiriki DOCs.

In the following subsections, I review Holmberg et al.'s (2019) account of object movement asymmetry and demonstrate that not only does this asymmetry hold true in Tiriki, but it also arises in different guises that don't immediately lend themselves to their account.

3.1 Double object movement asymmetry (DOMA) in Tiriki DOCs

Based on data from Zulu and Lubukusu, Holmberg et al. (2019) articulate a movement restriction on these otherwise fully symmetrical languages, where the IO

cannot be relativized when the DO is passivized despite the reverse being grammatical. They summarize this restriction as the 'Double Object Movement Asymmetry' (DOMA). The same DOMA effect is borne out in Tiriki; when the DO is promoted to subject position, the IO is prohibited from being Ā-extracted:

(14)   a.   Object *wh*-clefts with passivization

    i.   *wina wa   vi-tapu vi-manyiny-w-e
        1-who 1-REL 8-book 8SM-show-PASS-FV.PST
        Intended: 'Who were the books shown to?'

    ii.   shina   sha   va-somi va-manyiny-w-e
        7-what 7-REL 2-student 2SM-show-PASS-FV.PST
        'What were the students shown?'

  b.   Object RCs with passivization

    i.   *va-somi va   vi-tapu vi-manyiny-w-e
        2-student 2-REL 8-book 8SM-show-PASS-FV.PST
        Intended: 'the students that the books were shown to'

    ii.   vi-tapu vya   va-somi va-manyiny-w-e
        8-book 8-REL 2-student 2SM-show-PASS-FV.PST
        'the books that the students were shown'

In the presence of a passivized DO, an IO also cannot be object-marked:

(15)   Object-marking with passivization

  a.   Alulu a-kumil-il-e   mu-saakhulu tsi-haywa
      Alulu 1SM-touch-APPL-FV.PST 1-grandfather 10-axe
      'Alulu held the axes for grandfather.'

  b.   *tsi-haywa tsi-[mu] -kumil-il-w-e   (mu-saakhulu)
      10-axe   10SM-1OM-touch-APPL-PASS-FV.PST 1-grandfather
      Intended: 'Axes were held for grandfather.'

  c.   mu-saakhulu a-[tsi] -kumil-il-w-e   (tsi-haywa)
      1-grandfather 1SM-10OM-touch-APPL-PASS-FV.PST 10-axe
      'Grandfather was held axes for.' (Michael Diercks, personal communication)

This inevitable derivational crash incurred by the Ā-movement of IO out of a DO passive is in fact predicted by the analysis presented in the previous section (à la Holmberg et al. 2019). Recall that ApplP becomes the lower phase in passive DOCs. In the grammatical case, e.g. a DO *wh*-question with a passivized IO,

the derivation can be modeled as in Figure 3: First, Appl Agrees with and Case-licenses the DO (arrow (1)). T Agrees with and Case-licenses the IO (arrow (2)), which attracts it to A-move to Spec,TP and fulfill EPP as the grammatical subject. Due to the unvalued [u*wh*], the DO moves to the edge of the lower phase at Spec,ApplP. Upon the spellout of the lower phase, the DO survives and proceeds to the C domain, driven by the [*wh*] on C (arrow (3)).

When it is the other way around, however, the IO gets inevitably stranded within the lower phase, and its unvalued [u*wh*] causes ungrammaticality. This failed derivation is shown in Figure 4.

In order to derive a DO passive, the DO cannot be Case-licensed by Appl this time. Instead, Appl probes upward to license the IO (arrow (1)), leaving [uCase] on the DO unvalued. The DO then proceeds to move over the IO to the phase edge (outermost Spec,ApplP) and get Case-licensed by T (arrow (2)). In this scenario, the IO will not be able to escape the lower phase, for it is stranded in the innermost Spec,ApplP and cannot move to a higher edge position due to anti-locality (Abels 2003, Grohmann 2003). When C is merged, the [u*wh*]-bearing IO will be transferred to PF, causing the derivation to crash.

For object-marking under passivization, ApplP still demarcates the lower phase. Here, I revise the previously introduced assumption that *v* is the locus of ϕ-agreement for object-marking (cf. (10)) and restrict it to only active contexts. Following Holmberg et al. (2019), I assume that Appl doubles as the lower Case-licensor *and* the ϕ-Probe in passives. In the ungrammatical case of object-marking the IO in a DO passive, spelling out the ϕ-features on Appl becomes impossible as the Goal sits right above the Agreeing head associated with object-marking. After Appl licenses and Agrees with the IO upward (and copies back the ϕ-features), the IO itself becomes the highest copy in the chain, causing the lower copy to be deleted at PF upon transfer of the lower phase. As a result, an IO OM can never be pronounced.

In summary, the emergent asymmetries come down to purely structural constraints, namely the c-command asymmetry between two internal arguments and the cyclic nature of syntactic derivations.

3.2 Other emergent asymmetries in Tiriki DOCs

In this section, I explore the object properties exhibited in the cases where passivization is combined with left dislocation and *wh*-in-situ in Tiriki DOCs. The goal for this section is not to sketch out a complete analysis that accounts for the additional asymmetries; rather, I draw on relevant diagnostics for (covert) movement to develop preliminary arguments on how these asymmetry patterns
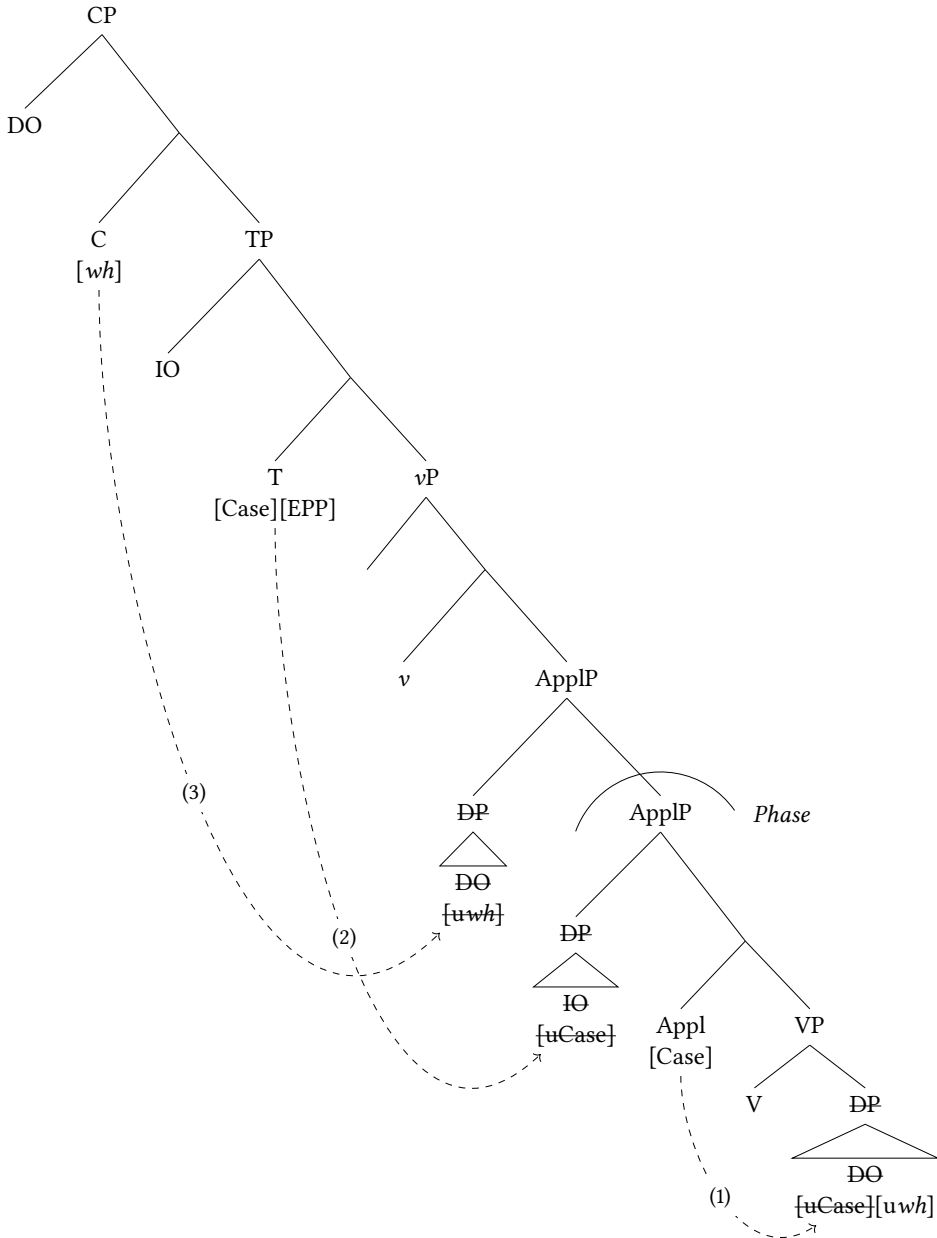
Figure 3: Derivation: Grammatical DO *wh*-clefts with IO passive
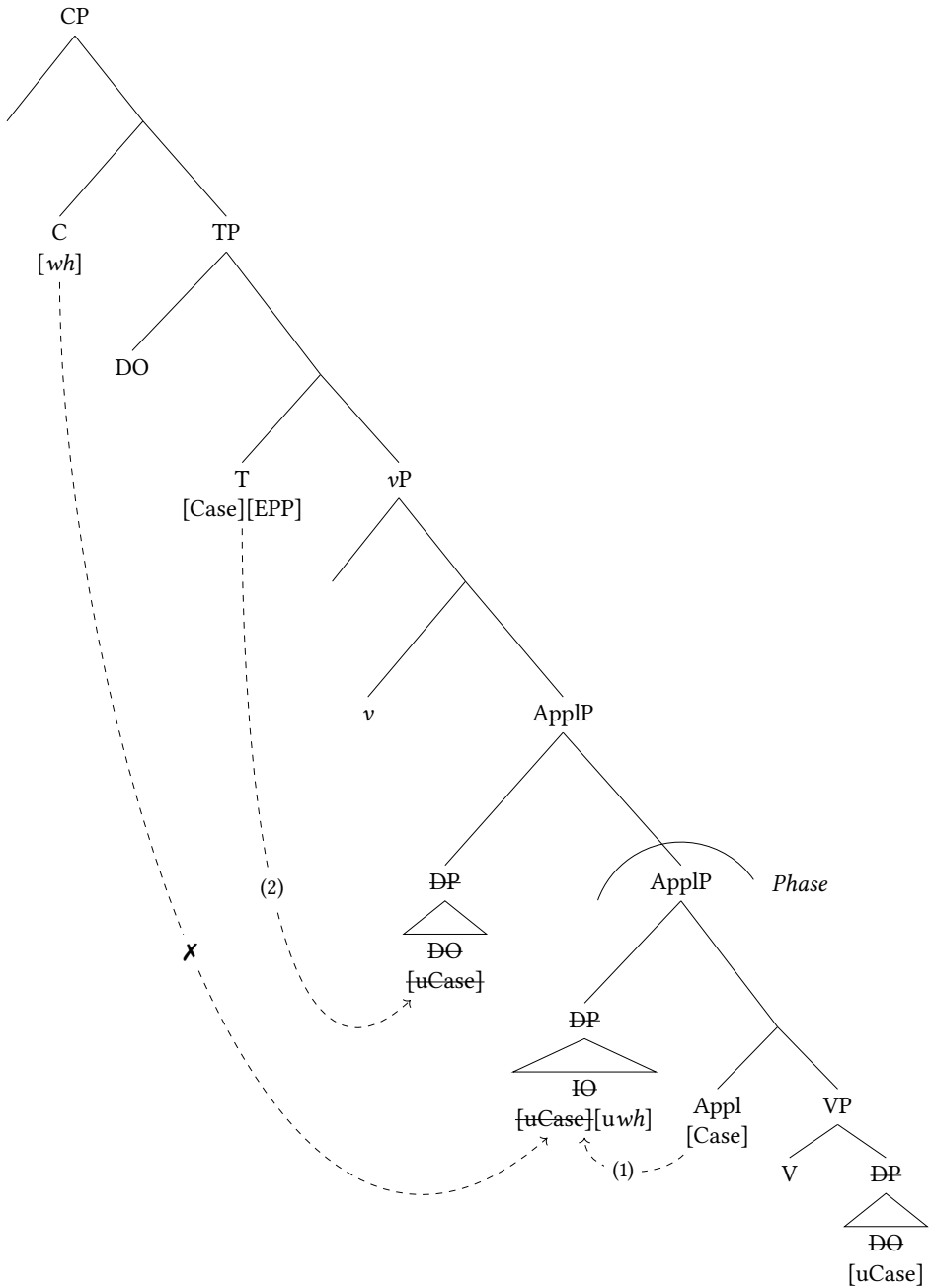
Figure 4: Derivation: Ungrammatical IO *wh*-clefts with DO passive

would fare on Holmberg et al.'s (2019) analysis. I leave the investigation into the syntactic nature of left dislocation and *wh*-in-situ in Tiriki for future research.

3.2.1 Left dislocation in passives

In Tiriki, the object has the option to surface clause-initially (i.e. be left-dislocated) when it bears an aboutness topic reading, a construction reminiscent of English topicalization:

(16) Object left dislocation

    a. va-somi , Hardley a-*(va)-manyiny-e   vi-tapu
        2-student Hardley 1SM-2OM-show-FV.PST 8-book
        'The students, Hardley showed them books.'

    b. vi-tapu , Hardley a-(*vi)-manyiny-e   va-somi   'The books,
        8-book Hardley 1SM-8OM-show-FV.PST 2-student   Hardley showed them to the students.'

As shown above, left dislocation (LD) is symmetrical in Tiriki DOCs, where either internal argument can be fronted.[6] However, a DOMA-like asymmetry is observed when left dislocation is combined with passivization:

(17) Object left dislocation with passivization

    a. *va-somi , vi-tapu vi-(va)-manyiny-w-e
        2-student 8-book 8SM-2OM-show-PASS-FV.PST
        Intended: 'The students, the books were shown to.'

    b. vi-tapu , va-somi va-(*vi)-amanyiny-w-e
        8-book 2-student 2SM-8OM-show-PASS-FV.PST
        'The books, the students were shown.'

While these DOMA-like facts may seem unsurprising on an Ā-movement analysis of LD (along the lines of Zeller 2009), further exploration into the properties of Tiriki LD suggests otherwise. As it turns out, LD in Tiriki patterns

---

[6]Both anonymous reviewers noted that object-marking interacts with left dislocation in interesting ways, given its obligatoriness in (16a) and ungrammaticality in (16b). I do not elaborate on this object-marking asymmetry in this paper, but the findings reported in Liu (2022) suggest that it is due to a combination of animacy effects and asymmetrical object properties under topicalization. Essentially, in monotransitives, it is obligatory to object-mark animate internal arguments, but optional for inanimates; in DOCs, object-marking left-dislocated animate IOs is mandatory, but left-dislocated (in)animate DOs and inanimate IOs cannot be realized as OMs. I leave the syntactic nature of this curious interaction between object-marking and left dislocation, as well as its implications for DOMA, for future research.

more closely with hanging topic left dislocation (HTLD, following Cinque 1977) cross-linguistically, which is generally analyzed as base-generation. In the paragraphs to come, I discuss the basic properties of HTLD and employ several syntactic diagnostics to show that left-dislocated objects in Tiriki are base-generated, rather than moved, clause-initially (see Alexiadou 2017 for an overview, as well as Ranero 2019 for a similar approach to object left dislocation in Luganda).

First, a discernible prosodic break between the dislocated element and the rest of the clause is observed, if not required, in HTLD. Tiriki exhibits the same requirement, as evidenced by the commas in all Tiriki left-dislocation examples. Second, the ban on left dislocation in embedded contexts (such as the adjunct relative clause below) in Tiriki is consistent with the generalization that HTLD is typically only allowed in root contexts.

(18) Illicit object left dislocation inside RC
*d-ukh-i   [_{DP} ha-vundu h-a   tsi-khuyi   , Anangwe
1SG.SM-arrive-FV.PST   16-place 16-REL 10-firewood Anangwe
a-tsi-isiyak-il-e]
1SM-10OM-chop-PFV-FV
Intended: 'I arrived at the place where the firewood, Anangwe chopped it.'

Another hallmark property of HTLD in contrast with other similar constructions (e.g. clitic left dislocation) is the lack of connectivity between the dislocated element and its corresponding thematic position. Such absence of connectivity effects can be shown with idiom chunks. In Tiriki, idioms appear to be not just a surface phenomenon in the sense that the idiomatic interpretation of a full-clause idiom can still be retained when its subparts do not immediately follow each other linearly. For example, the idiomatic reading persists even when the idiom is separated by a (hyper-)raising predicate *-lolekha* 'seem':

(19) Idiom connectivity with raising predicate *-lolekha* 'seem'
i-mbisi_i i-lolekh-a   khuli ____i i-hulil-e   mu-lilo
9-hyena 9SM-seem-FV that   9SM-feel-FV.PFV 3-fire
✓ Idiomatic reading: 'Someone seems to have overeaten.'
✓ Literal reading: 'The hyena seems to have felt the fire.'

Under the assumption that an idiom as a whole must form one constituent at some point of the derivation, its idiomatic interpretation can only survive the apparent displacement of its subparts when such displacement results from movement. The fact that only a literal interpretation is available under Tiriki LD in

(20b) suggests that the configuration in question is indeed HTLD, and that the dislocated object is base-generated in the left periphery.[7]

(20)   Object left dislocation: Idiom chunk

    a.   Lydia a-h-el-e   mw-ana li-chembe
        Lydia 1SM-give-PFV-FV.PST 1-child   5-shovel

        ✓ Idiomatic reading: 'Lydia has given the idle child some work to do.'
        ✓ Literal reading: 'Lydia has given the child a shovel.'

    b.   [li-chembe] , Lydia a-h-el-e   mw-ana
        5-shovel   Ebby 1SM-give-PFV-FV.PST 1-child

        ✗ Idiomatic reading: 'Lydia has given the idle child some work to do.'
        ✓ Literal reading: 'Lydia has given the child a shovel.'

Tiriki HTLD also exhibits no Condition C connectivity. As shown in (21b), the left-dislocation of 'Lydia's child' obviates the Condition C violation, permitting *Lydia* in the dislocated DP and the subject pronoun *ye* to co-refer. This contrast, again, points to a base-generation analysis as an Ā-movement dependency would have resulted in a lower copy of the bound R-expression.

(21)   Object left dislocation: Condition C

    a.   [ye]   a-h-el-e   mw-ana w-a   [Lydia] ma-chungwa
        1-PRO 1SM-give-PFV-FV.PST 1-child   1-ASSOC Lydia   6-orange

        ✗ Bound reading: 'She$_i$ gave Lydia$_i$'s child oranges.'
        ✓ Free reading: 'She$_i$ gave Lydia$_j$'s child oranges.'

    b.   mw-ana w-a   [Lydia] , [ye]   a-*(mu)-h-el-e
        1-child   1-ASSOC Lydia   1-PRO 1SM-1OM-give-PFV-FV.PST
        ma-chungwa
        6-orange

        ✓ Bound reading: 'Lydia$_i$'s child, she$_i$ gave him oranges.'
        ✓ Free reading: 'Lydia$_j$'s child, she$_i$ gave him oranges.'

Beyond connectivity effects, HTLD in Tiriki also shows no Weak Crossover effect (WCO) in (22b) while such an effect is robustly triggered by overt *wh*-movements. I take this to suggest that no movement dependency is involved in Tiriki HTLD, following Cinque (1990).

---

[7]An anonymous reviewer pointed out that alternatively, the loss of idiomatic readings under HTLD might be because idiom chunks cannot be topics (or are in general incompatible with the interpretive effects of Tiriki HTLD). For example, #*As for the bucket, Mary kicked it* cannot be interpreted idiomatically.

(22)　Object left dislocation: Weak Crossover

    a. *wina$_i$ w-a   mw-ana wewe$_i$ y-a-mu-yanz-a
       1-who 1-REL 1-child 1POSS 1SM-PRES-1OM-love-FV
       Intended: 'Who$_i$ does his$_i$ child love?'

    b. Lydia$_i$, mw-ana wewe$_i$ y-a-*(mu)-yanz-a
       Lydia   1-child 1POSS 1SM-PRES-1OM-love-FV
       'Lydia$_i$, her$_i$ child loves her$_i$.'

Other supporting evidence hinges on island constraints (Ross 1967; see (23) on complex NP and adjunct islands) and scope reconstruction (see (24)). While the absence of these effects does not necessarily correlate with an absence of movement, I conclude that the island insensitivity and lack of scope reconstruction of Tiriki HTLD ought to be explained with a base-generation approach rather than independent syntactic properties of Tiriki, considering other diagnostics discussed above.

(23)　Object left dislocation: Island insensitivity

    a. shi-tapu , Hardley a-lol-i   [$_{DP}$ mu-somi yi-v-i]
       7-book   Hardley 1SM-see-FV.PST   1-student 1SM-steal-FV.PST
       'The book, Hardley saw the student who stole (it).'

    b. va-cheni , Alulu a-rhul-i   [$_{TP}$ baada y-a   Hardley
       2-guest   Alulu 1SM-leave-FV.PST   after ASSOC Hardley
       khu-*(va)-shelits-a]
       15-2OM-greet-FV
       'The guests, Alulu left after Hardley greeted (them).'

(24)　Object left dislocation: Scope reconstruction

    a. $\boxed{\text{vuli}}$ mw-ikitsi a-manyiny-e   va-somi $\boxed{\text{veve}}$ shi-tapu
       every 1-teacher 1SM-show-FV.PST 2-student 2-POSS 7-book
       '*Every* teacher showed *his* students a book.'
       ✓ Bound reading: For every teacher x, x showed x's students a book.
       ✓ Free reading: For every teacher x, x showed y's students a book.

    b. va-somi $\boxed{\text{veve}}$, $\boxed{\text{vuli}}$ mw-ikitsi a-*(va)-manyiny-e   shi-tapu
       2-student 2-POSS every 1-teacher 1SM-2OM-show-FV.PST 7-book
       '*His* students, *every* teacher showed them a book.'
       ✗ Bound reading: For every teacher x, x showed x's students a book.
       ✓ Free reading: For every teacher x, x showed y's students a book.

In summary, the data at hand strongly hint at a base-generation analysis of Tiriki HTLD. Therefore, the asymmetry reported at the beginning of this subsection constitutes a real challenge to the analysis advanced in Holmberg et al. (2019) as their account targets Ā-movement—but not non-movement—dependencies.

3.2.2 *Wh*-in-situ in passives

DOMA effect in Tiriki can also result from the interaction between *wh*-in-situ and passivization. Like most Bantu languages, Tiriki allows for both *ex-situ* and *in-situ* strategies for forming *wh*-questions:

(25) Strategies for Tiriki *wh*-questions

a. [sh-ina] sh-a Vusu a-l-il-e
7-what 7-REL Vusu 1SM-eat-PFV-FV.PST

b. Vusu a-l-il-e [sh-ina]
Vusu 1SM-eat-PFV-FV.PST 7-what
'What did Vusu eat?'

Just like how IOs cannot undergo *wh*-movement in DO passives, IOs also cannot stay *in-situ* when DOs are passivized:

(26) Object *wh*-in-situ with passivization

a. *vi-tapu vi-manyiny-w-e w-ina
8-book   8SM-show-PASS-FV.PST 1-who
Intended: 'The books were shown who?'

b. va-somi   va-manyiny-w-e sh-ina
2-student 2SM-show-PASS-FV.PST 7-what
'The students were shown what?'

The fact that *wh*-in-situ is subject to similar DOMA constraints as its *ex-situ* counterpart is suggestive of the involvement of covert movement dependencies.[8] In the paragraphs to come, I describe properties of Tiriki *wh*-in-situ in terms of its island sensitivity, immunity to the focus intervention effect (Beck 2006), and scope reconstruction. I would also like to emphasize that the current investigation into Tiriki *wh*-in-situ is still inconclusive. Much work on whether Tiriki

---

[8]An anonymous reviewer suggested that *wh*-in-situ in Tiriki might not involve movement at all and can be interpreted *in-situ* under an unselective-binding analysis. On this analysis, DOMA can be posited to constrain any syntactic relation that involves a *wh*-feature (with movement or not). I leave this potential analysis of Tiriki *wh*-in-situ for future research.

islands are transparent for Ā-movements and whether covert movements are indeed involved in *wh*-in-situ (based on their interaction with WCO, the licensing of parasitic gaps, etc.) remains to be done.

First, I show that *wh*-in-situ in Tiriki is sensitive to some islands. It is impossible to interpret the *in-situ wh*-question as a matrix question when the *wh*-phrase remains inside a complex NP (in (27a)) or a sentential subject island (in (27b)).[9]

(27)   Object *wh*-in-situ: Island sensitivity

    a.   *Hardley a-landul-e   [$_{DP}$ mu-somi yi-v-i   shina]
        Hardley 1SM-beat-FV.PST   1-student 1SM-steal-FV.PST 7-what
        Intended: 'Hardley disciplined the student who stole what?'
        (*non-echo question*)

    b.   *[$_{CP}$ khuli Franko a-tekh-i   shina] ka-chenyiny-e
            that Frank 1SM-cook-FV.PST 7-what 6SM-surprise-FV.PST
        Maiko
        Michael
        Intended: 'That Frank cooked what surprised Michael?' (*non-echo question*)

However, it is worth noting that *wh*-in-situ seems insensitive to adjunct islands, as illustrated with a temporal adjunct clause.

(28)   Object *wh*-in-situ: Adjunct island

    a.   *wina w-a Lydia a-rhul-i   [$_{CP}$ lw-a Hardley
        1-who 1-REL Lydia 1SM-leave-FV.PST   11-REL Hardley
        a-shelits-e   _____]
        1SM-greet-FV.PST
        'Who did Lydia leave when Hardley greeted?'

    b.   Lydia a-rhul-i   [$_{CP}$ lw-a Hardley a-shelits-e   wina]
        Lydia 1SM-leave-FV.PST   11-REL Hardley 1SM-greet-FV.PST 1-who
        'Lydia left when Hardley greeted who?' (*non-echo question*)

The mixed results of island diagnostics aside, Tiriki *wh*-in-situ also does not exhibit focus intervention effect. Following Beck (2006), I assume that elements such as negation, negative quantifiers, and focus operators intervene between a *wh*-phrase and its licensing C. Specifically, the intervenor blocks the *wh*-phrase's

---

[9]An anonymous reviewer pointed out that the effect of 'island sensitivity' might be because *wh*-in-situ is clause-bounded in Tiriki, like other LF movements (e.g. QR).

projection of focus alternatives and thereby bleeds the semantic association between the *wh*-phrase and the C. Said configuration is schematized below:

(29) Structure of focus intervention
 *[C ...[intervenor [... *wh*-phrase ...]]]

In Tiriki, both *wh*-clefts and *wh*-in-situ can obviate the focus intervention effect. The former is expected as overt *wh*-movements remove the *wh*-phrases from the c-command domain of the intervenors. The fact that no intervention effect is observed in (30b) suggests that *wh*-in-situ in Tiriki may still involve movements, as covert movements have been also shown to achieve similar rescuing effects (Pesetsky 2000, Kotek 2014).

(30) Object *wh*-in-situ: Focus intervention effect

 a. shina sh-a Ebby [w-onyene] a-tekh-i
  7-what 7-REL Ebby 1-only 1SM-cook-FV.PST
  'What did only Ebby cook'

 b. Ebby [w-onyene] a-tekh-i shina
  Ebby 1-only 1SM-cook-FV.PST 7-what
  'Only Ebby cooked what?' (*non-echo question*)

A covert movement analysis of *wh*-in-situ is also supported by scope ambiguities. Apart from the surface scope reading, the *wh*-phrase *shina* can still take wide scope over the quantifier in the subject position without any indication that an overt *wh*-movement has taken place.

(31) Object *wh*-in-situ: Quantifier scope ambiguity
 vuli mu-somi a-som-i shina
 every 1-student 1SM-read-FV.PST 7-what

 'Every student read what?' (*non-echo question*)
 ✓ ∀ ≫ *wh*: For every student x, what did x read?
 ✓ *wh* ≫ ∀: What is x such that every student read x?

Despite the positive evidence for covert movement presented above, I stop short of concluding that *wh*-in-situ in Tiriki does involve covert movement without more systematic investigation into its island sensitivity and properties of *wh*-licensing. There are two ways in which a definitive solution to Tiriki *wh*-in-situ can inform us of the nature of the observed asymmetry: If Tiriki *wh*-in-situ questions are derived via covert movement, the aforementioned asymmetry can then be easily accommodated by the current DOMA analysis as the same constraint

should hold for overt and covert movements alike. If the *wh*-phrase is proved to be *in-situ* at both PF and LF, the DOMA analysis will need to be revised to fully capture the Tiriki patterns.[10]

4 Conclusion

In this paper, I documented and examined different object properties and their (a)symmetry exhibited in Tiriki DOCs. Descriptively, I showed that Tiriki is a largely symmetrical language, allowing both objects to be ϕ-probed and undergo a range of A and Ā-movements. By combining the aforementioned syntactic operations with passivization, I also substantiated the predictions of DOMA (Holmberg et al. 2019): Like Zulu and Lubukusu, it is ungrammatical in Tiriki to object-mark or Ā-extract the IO out of a DO passive. Similar DOMA effects were also replicated in other unattested contexts, namely when the DO is passivized and the IO is left-dislocated or replaced with an *in-situ wh*-phrase. In these contexts, the IO is not involved in apparent movement dependencies, which deviates from the DOMA configurations reported elsewhere.

Analytically, I applied Holmberg et al.'s (2019) account of cross-linguistic DOMA effects and outlined the challenges that Tiriki presents. By revisiting the extant DOMA analysis, I demonstrated how the c-command asymmetry between two objects, the flexible Case-licensing of the high Appl head, and its phasal properties in passives result in unavoidable movement asymmetry. Moreover, I showed that left dislocation and *wh*-in-situ in Tiriki don't necessarily involve movement dependencies and yet are still able to trigger DOMA effects in passives. For future work, a more in-depth understanding of Ā-dependencies and their constraints in Tiriki is needed to derive a more complete account of DOMA facts reported in this paper.

---

[10]An alternative conclusion that can be drawn from the novel DOMA data in Tiriki is that the analytical machinery of the Minimalist Program is not so well-equipped to deal with the observed asymmetry. Crucially, a empirical generalization is lost: Whenever there are more than one grammatical-function-changing operations targeting the objects, passivization, which seeks to promote the grammatical function of the object, *always* targets the thematically more prominent object (IO, or recipient/beneficiary). This robust link between grammatical function hierarchy and thematic prominence is perhaps better accounted for by other frameworks of syntax (cf. discussions in Alsina 1996 and §4 of Zeller 2015). I thank an anonymous reviewer for this insightful comment.

Abbreviations

1	noun class		PST	past tense
1SG	first person singular		PFV	perfective
APPL	applicative		POSS	possessive
FV	final vowel		REL	relativizer
OM	object marker		RP	resumptive pronoun
PASS	passive		SM	subject marker
PRES	present tense			

Acknowledgements

References

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding.* Storrs, CT: University of Connecticut. (Doctoral dissertation).

Adams, Nikki. 2010. *The Zulu ditransitive verb phrase.* University of Chicago. (Doctoral dissertation).

Aldridge, Edith. 2004. *Ergativity and word order in Austronesian languages.* Cornell University. (Doctoral dissertation).

Alexiadou, Artemis. 2017. Left dislocation. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley Blackwell Companion to Syntax*, 2nd edn., 1–34. Hoboken: John Wiley & Sons.

Alsina, Alex. 1996. Passive types and the theory of object asymmetries. *Natural Language & Linguistic Theory* 14(4). 673–723.

Baker, Mark. 1988. Theta theory and the syntax of applicatives in Chichewa. *Natural Language & Linguistic Theory* 16(3). 353–390.

Barss, Andrew & Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17. 347–354.

Bax, Anna & Michael Diercks. 2012. Information structure constraints on object marking in Manyika. *Southern African Linguistics and Applied Language Studies* 30. 185–202.

Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14(1). 1–56.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38(4). 589–644.

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45. 27–89.

Bošković, Željko. 2015. From the Complex NP Constraint to everything: On deep extractions across categories. *The Linguistic Review* 32. 603–669.

Bošković, Željko. 2016. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47. 1–33.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 22(2). 147–186.

Buell, Leston. 2009. Evaluating the immediate postverbal position as a focus position in Zulu. In Masangu Matondo, Fiona McLaughlin & Eric Potsdam (eds.), *Selected Proceedings of the 38th Annual Conference on African Linguistics*, 166–172. Somerville, MA: Cascadilla Proceedings Project.

Carstens, Vicki. 2005. Agree and EPP in Bantu. *Natural Language & Linguistic Theory* 23(2). 219–279.

Cheng, Lisa Lai-Shen & Laura J. Downing. 2012. Against FocusP: Arguments from Zulu. In Ivona Kučerová & Ad Neeleman (eds.), *Contrasts and positions in information structure*, 246–266. Cambridge: Cambridge University Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in linguistics*, 1–52. Cambridge: MIT Press.

Cinque, Guglielmo. 1977. The movement nature of left dislocation. *Linguistic Inquiry* 8(2). 397–412.

Cinque, Guglielmo. 1990. *Types of Ā-dependencies*. Cambridge, MA: MIT Press.

De Guzman, Videa. 1987. Indirect objects in SiSwati. *Studies in African Linguistics* 18. 309–325.

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33(3). 409–442.

Grohmann, Kleanthes K. 2003. *Prolific domains. On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.

Haddican, Bill & Anders Holmberg. 2019. Object symmetry effects in Germanic. *Natural Language & Linguistic Theory* 37(1). 91–122.

Henderson, Brent. 2018. Bantu applicatives and Chimiini instrumentals. In Augustine Agwuele & Adams Bodomo (eds.), *The Routledge handbook of African linguistics*, 1st edn., 262–280. New York: Routledge.

Holmberg, Anders, Michelle Sheehan & Jenneke van der Wal. 2019. Movement from the double object construction is not fully symmetrical. *Linguistic Inquiry* 50(4). 677–721.

Hyman, Larry M. 2003. Suffix ordering in Bantu: A morphocentric approach. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2002*, 245–281. Dordrecht: Kluwer Academic Publishers.

Jerro, Kyle. 2016. *The syntax and semantics of applicative morphology in Bantu.* University of Texas at Austin. (Doctoral dissertation).

Julien, Marit. 2002. *Syntactic heads and word formation.* New York: Oxford University Press.

Kayne, Richard S. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Kimenyi, Alexandre. 1980. *A relational grammar of Kinyarwanda.* Berkeley & Los Angeles: University of California Press.

Kotek, Hadas. 2014. *Composing questions.* Massachusetts Institute of Technology. (Doctoral dissertation).

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer Academic Publishers.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the world.* 2nd edn. Dallas: SIL International.

Lippard, Hannah, Justine Sikuku, Crisófia Lánga da Câmara, Madelyn Colantes, Kang (Franco) Liu & Michael Diercks. Forthcoming. Emphatic interpretations of object marking in Bantu languages. *Studies in African Linguistics.*

Liu, Kang. 2022. *Object marking in Tiriki.* University of Cambridge. (MA thesis).

Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar 1*, 113–151. Stanford: CSLI Publications.

Marlo, Michael R. 2013. Structure and tone in Tiriki. Ms, University of Missouri.

Marlo, Michael R. 2015. Exceptional properties of the reflexive in Bantu languages. *Nordic Journal of African Studies* 24. 1–22.

McGinnis, Martha. 2001. Phases and the syntax of applicatives. In Minjoo Kim & Uri Strauss (eds.), *Proceedings of the North East Linguistic Society 31*, vol. 2, 339–349. Amherst, MA: GLSA.

Nakamura, Masanori. 1997. Object extraction in Bantu applicatives: Some implications for minimalism. *Linguistic Inquiry* 28(2). 252–280.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement* (Linguistic Inquiry Monographs 43). Cambridge, MA: MIT Press.

Pesetsky, David. 2000. *Phrasal movement and its kin.* Cambridge: MIT Press. 132.

Pylkkänen, Liina. 2008. *Introducing arguments* (Linguistic Inquiry Monograph 49). Cambridge, MA: MIT Press.

Ranero, Rodrigo. 2019. Deriving an object dislocation asymmetry in Luganda. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and Description in African Linguistics: Selected Proceedings of the 47th Annual Conference on African Linguistics*, 595–622. Berlin: Language Science Press. DOI: 10.5281/zenodo.3367189.

Rizzi, Luigi. 1990. *Relativized minimality.* Cambridge, MA: MIT Press.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals.* Cambridge, MA: MIT Press.

Ross, John. 1967. *Constraints on variables in syntax.* Massachusetts Institute of Technology. (Doctoral dissertation).

Selvanathan, Naga. 2019. A Case based account of Bantu IAV-focus. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and description in African linguistics: Selected papers from the 47th Annual Conference on African Linguistics*, 623–645. Berlin: Language Science Press. DOI: 10.5281/zenodo.3367191.

Sikuku, Justine & Michael Diercks. 2021. *Object marking in Bukusu: At the interface of pragmatics and syntax.* Ms, Moi University and Pomona College. https://research.pomona.edu/mjkd/papers/.

Sikuku, Justine, Michael Diercks & Michael R. Marlo. 2018. Pragmatic effects of clitic doubling: Two kinds of object markers in Lubukusu. *Linguistic Variation* 18(2). 359–429.

van der Wal, Jenneke. 2006. The disjoint verb form and an empty immediate after verb position in Makhuwa. *ZAS Papers in Linguistics* 36. 233–256.

van der Wal, Jenneke. 2009. *Word order and information structure in Makhuwa-Enahara.* Universiteit Leiden. (Doctoral dissertation).

van der Wal, Jenneke. 2015. Bantu object clitics as defective goals. *Revue Roumaine de Linguistique* 60(2–3). 277–296.

van der Wal, Jenneke. 2017. Flexibility in symmetry: An implicational relation in Bantu double object constructions. In Michelle Sheehan & Laura R. Bailey (eds.), *Order and structure in syntax II: Subjecthood and argument structure* (Open Generative Syntax 2), 115–152. Berlin: Language Science Press.

van der Wal, Jenneke. 2020. The AWSOM correlation in comparative Bantu object marking. In Katharina Hartmann, Johannes Mursell & Peter W. Smith (eds.), *Agree to agree: Agreement in the Minimalist Programme* (Open Gener-

ative Syntax 6), 199–234. Berlin: Language Science Press. DOI: 10.5281/zenodo.3541755.

van der Wal, Jenneke. 2022. *A featural typology of Bantu agreement.* Oxford: Oxford University Press.

Visser, Marianna W. 1986. Cliticization and Case theory in Xhosa. *South African Journal of African Languages* 6(3). 129–137.

Watters, John R. 1979. Focus in Aghem: A study of its formal correlates and typology. In Larry M. Hyman (ed.), *Aghem grammatical structure* (Southern California Occasional Papers in Linguistics 7), 137–198. Los Angeles: University of Southern California.

Zeller, Jochen. 2009. On clitic left-dislocation in Zulu. In Sonja Ermisch Ruediger (ed.), *Focus and topic in African languages* (Frankfurt African Studies Bulletin 18), 131–156. Köln: Rüdiger Köppe Verlag.

Zeller, Jochen. 2012. Object marking in isiZulu. *Southern African Linguistics and Applied Language Studies* 30. 219–325. DOI: 10.2989/16073614.2012.737600.

Zeller, Jochen. 2014. Three types of object marking in Bantu. *Linguistische Berichte* 239. 347–367.

Zeller, Jochen. 2015. Argument prominence and agreement: Explaining an unexpected object asymmetry in Zulu. *Lingua* 156. 17–39. DOI: 10.1016/j.lingua.2014.11.009.

Chapter 16

Negation in languages of the Lotuxo sub-group of Eastern Nilotic

Jonathan Moodie

University of Melbourne

> Lokoya, Lopit and Otuho (Lotuxo) are Eastern Nilotic languages spoken in the area around Torit in Eastern Equatoria province, South Sudan. The three languages are verb initial (VSO) and have a marked-nominative case system.
>
> Negation in these languages is normally expressed through a verbal or verb-like construction. The negative morpheme can be regarded as a verb corresponding to 'not be'. Sometimes the negative verb is marked for person, aspect and modality and sometimes it has the same form for first, second and third person. In all three languages, the standard negative construction shows the same constituent order.
>
> Standard negation involves a negative verb and the lexical verb. The negative verb is clause initial. The lexical verb is marked with the subordinating prefix (except in Otuho) and maintains its person marking. The lexical verb is placed after the subject to give an NEG SVO word order. The subject retains its nominative marking when it is placed before the lexical verb and after the negative verb. It is likely that the negative construction originated as a bi-clausal construction.

1 The languages

Lokoya, Lopit and Otuho (Lotuxo) are Eastern Nilotic languages spoken in the area around Torit in Eastern Equatoria province, South Sudan. They are members of the Lotuxo sub-group of the Eastern-Nilotic languages, which are listed, together with the estimated number of speakers, in Table 1. There are two other languages in this sub-group, Lango and Dongotono. I have no data from these languages, and they are not included in this study. The members of the Eastern Nilotic language family are listed in Figure 1.
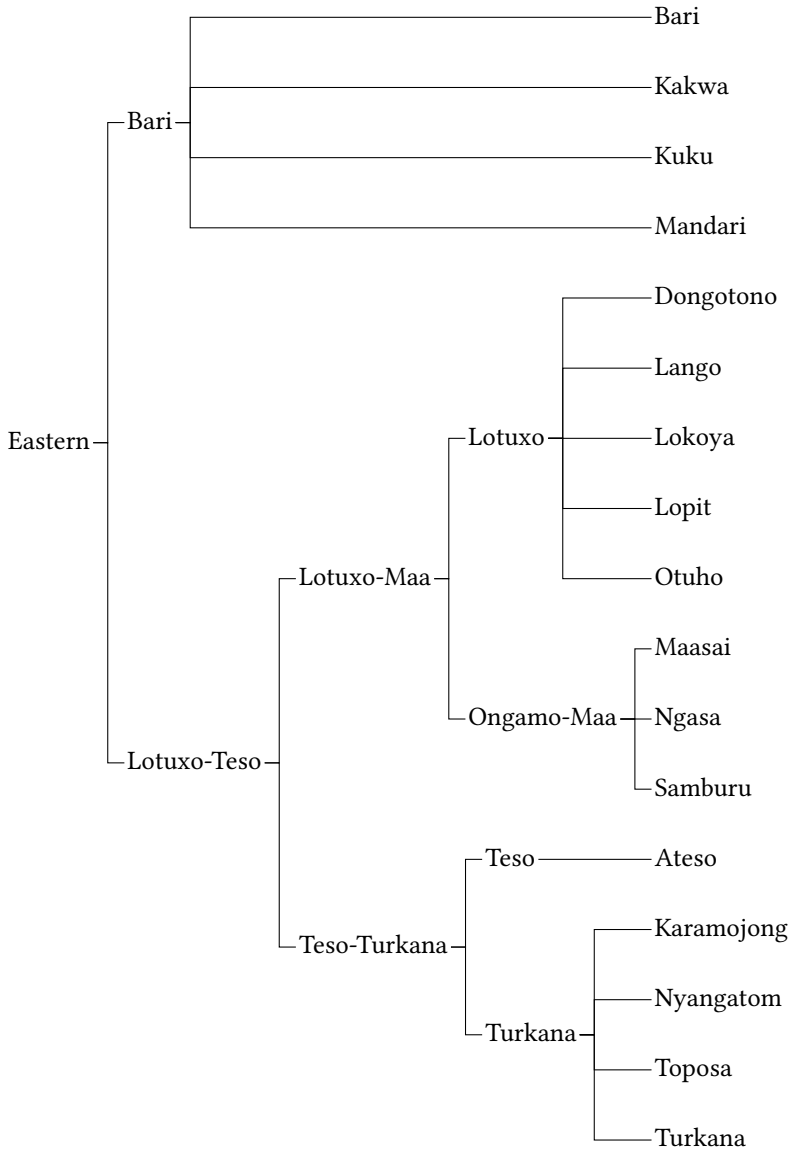
Figure 1: The Eastern Nilotic languages (based on Lewis et al. 2016)

Table 1: The languages of the Lotuxo sub-group.

Language	Speakers
Dongotongo	1,000
Lango	38,000
Lokoya	12,000
Lopit	50,000
Otuho (Lotuko)	140,000

A map showing the location of the speakers of the Nilotic languages is given in Figure 2. The location of Lango is not marked, but is in the area to the south of the Dongotongo area (included in the area marked Lotuxo).

The Lokoya, Lopit and Otuho languages have much in common with each other and with the other Eastern Nilotic languages. The following is a general description of the three languages in this study based on my fieldwork and on the work of Muratori (1938) and Westerman (1944).

Lokoya, Lopit and Otuho have a nine-vowel system and display the Advanced Togue Root (ATR) feature, and a process of vowel harmony based on this feature. There are tones (at least High and Low), used for both lexical and grammatical distinctions. There are two grammatical genders, feminine and masculine. Number marking follows a tripartite system of singulative, plurative and replacement number marking (Dimmendaal 2000). The languages have a marked nominative case system with nominative and absolutive case. The nominative case is differentiated from the absolutive by a change in the tonal pattern across the noun.[1] The case system is similar in Otuho and Lokoya.

The three languages are verb-initial and have an unmarked word order of verb-subject-object (VSO) or verb-agent-patient (VAP).[2] There is bound pronominal marking on the verbs which indexes the subject and, sometimes, the first and second person objects. The pronominal subject marking is shown in Table 2. The first, second and third person singular markers (/a-/, /ɪ-/, /ɛ-/) are widespread in

---

[1]See Moodie & Billington (2020: 256) for a description of the case system in Lopit.

[2]Lokoya differs in word order from the other two languages in that it can have an unmarked SVO order as well as VSO. Often, nominal agents/subjects tend to follow a SVO order whereas pronominal agents/subjects tend to follow a VSO order. This may be explained by the tendency for pronouns to be used when the agent/subject is known or clear from the context. On the other hand, nominal agents/subjects tend to constitute new information and are clause initial.
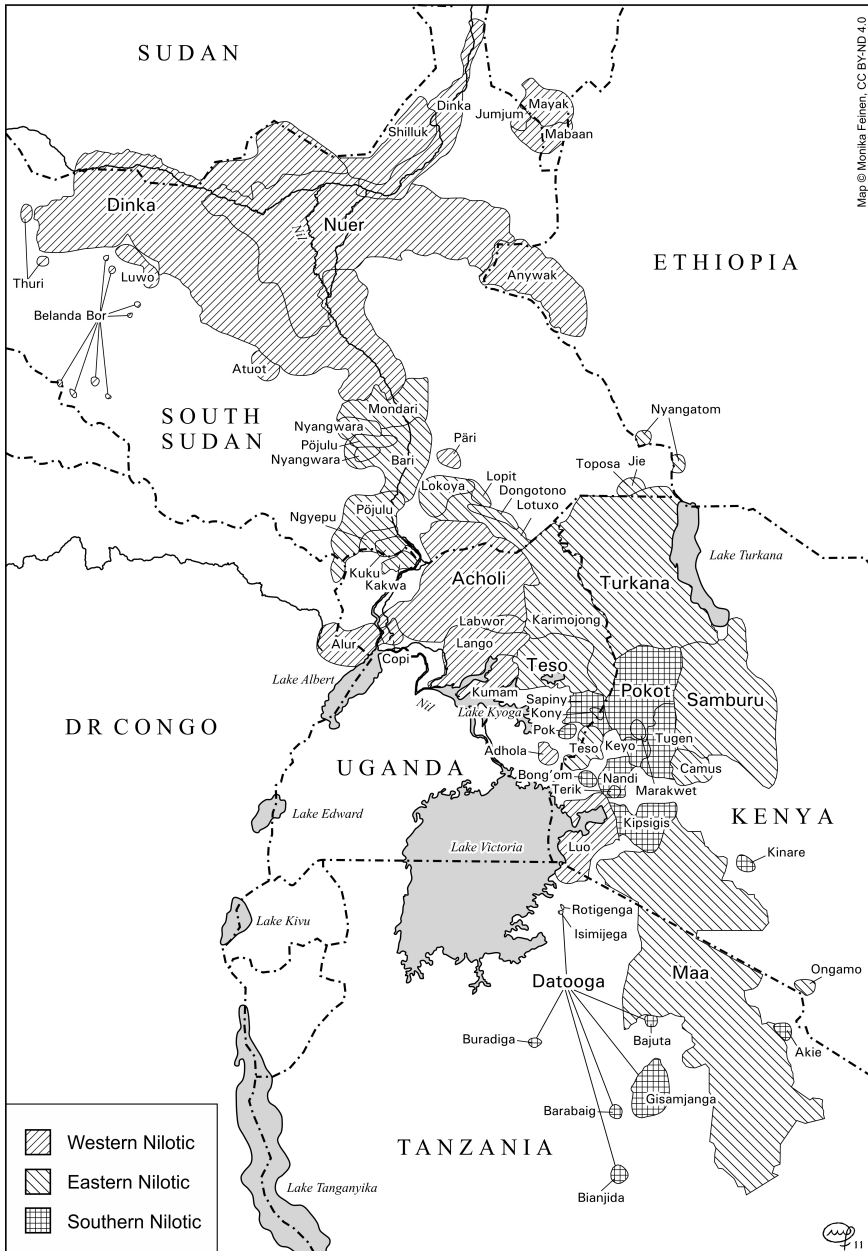
Figure 2: The Nilotic Languages (copyright Monika Feinen, reproduced with permission).

Nilo-Saharan languages (Greenberg 1966: p.86). Note that these are [−ATR] in their underlying form but can be realised differently according to ATR vowel harmony and mid-vowel assimilation processes.[3]

Table 2: Number/person marking for subjects

		singular	plural
	1st	a-	ɛɪ-, ei-
person	2nd	ɪ-, i-	ɪ-, i-
	3rd	ɛ-, e-, ɔ-, o-	ɛ-, e-, ɔ-, o-

Verbs are marked for aspect and mood in Lopit and for aspect in Otuho and Lokoya. Mood in Otuho and Lokoya is expressed with particles or adverbs. There is no grammatical tense and temporal reference is provided by adverbs, prepositional phrases and discourse context. Adjectival concepts are mostly expressed with stative property verbs, often in relative clause constructions.

This study is based on field work with Lopit, Otuho and Lokoya speakers in Melbourne, with Lopit speakers in the Kakuma Refugee camp in northern Kenya and on recordings made with Lokoya speakers in Torit. The Lopit work is based on the grammar of Moodie and Billington (2020) and the Otuho work has been augmented by the work of Muratori (1938, 1948) and Coates (1985). I have also utilized the sketch grammar of Westermann (1944) for Lokoya (or Oxoriok).

In this paper, I use broad phonetic transcription in glossed examples and tables. Abbreviations used in glosses are given at the end of the paper. The transcription of vowels in any example reflects the results of any harmony and assimilation processes present. The in-text examples are also given in broad phonetic transcription. Square brackets are used to indicate what was transcribed. Where words or stems are discussed in general terms rather than in an extract from a transcription, they are transcribed without tones (e.g. /beŋ/, 'not be', Otuho), as the tones in any particular utterance depend on a number of factors such as case, aspect, number.

In this paper, I will first examine standard negation in the three languages (section 2). I will then look at negation in non-declarative clauses (section 3), in stative predications (section 4) and in non-main clauses (section 5). I will then briefly compare negation across the Eastern Nilotic group (section 6). A summary and some conclusions follow in section 7.

---

[3]With some verbs in Otuho and Lokoya, the first-person singular prefix is /e-/ or /ɛ-/.

2 Overview of standard negation

2.1 Introduction

Standard negation is described as "the basic way(s) a language has for negating declarative verbal main clauses" (Miestamo 2005: p.1). Standard negation in the three languages in this study is expressed through verbal constructions. I propose that the negative morpheme is best regarded as a negative verb, even though it does not always have the full range of features found in the verb in the Lotuxo sub-group. The basic declarative clausal structure in these languages is VERB SUBJECT OBJECT. I propose that, in a negated construction, the order is NEG.VERB SUBJECT VERB OBJECT.

I will now present reasons why the negative morpheme should be regarded as a verb, using some examples from the languages studied. As shown in (1), the negative verb [ábéŋ], indexes the person of the subject.

(1) Otuho
     á-béŋ   ní   è-góɲú àmòlòŋ
     1SG-not.be 1SG.NOM 1SG-see baboon.ABS
     'I don't see the baboon.' MB3-001 00:45:48

As shown in (2), the negative verb, /ɲa/, in Lopit can be marked for the irrealis mood with the prefix [ŋaî-] and for the person of the subject.

(2) Lopit
     [á-ŋaî-ɲà   náŋ   l-á-mweí   ][á-ŋaî-ibóɲ náŋ
     1SG-IRR-not.be 1SG.NOM SBO-1SG-be.sick 1SG-IRR-meet 1SG.NOM
     xò=xìjò   xʊ̀ná l-ɔ̀-lʊ́ɲà   ]
     with=people.ABS REL.PL SBO-3-be.many
     'If I weren't sick, I would have met many people.' MB1-014 01:25:40

In these constructions, the subject has nominative case marking. Since these languages are verb initial, I propose that the subjects are, in fact, subjects of the negative verbs. These languages are marked-nominative languages and it is generally accepted that there is no nominative case before the verb (König 2008).[4] If the negative word (e.g. [ábéŋ] or [áŋaîɲà]) were not a verb, one would expect the subject to have absolute case marking since it would proceed the

---

[4]The rule is expressed by König as "In preverbal position the core participants S, A and O occur always in one case form only, namely the morphologically unmarked one" (2008: p.240). The absolutive is the unmarked case (Moodie & Billington 2020: p.262).

verb ([èɡóɲú] or [lámweí]). This supports the hypothesis of the verbal nature of the negative constituent.[5]

However, not all the negative words in these languages share the full range of verbal features. It appears that there could be a pattern of syncretism of the pronominal markers. As is discussed in §2.3 and §2.4, the negation word for first, second and third person is usually expressed as /ɪŋa/ in Lopit and /awaŋ/ in Lokoya respectively. Nevertheless, in these cases, the subjects retain their nominative marking when placed after the negator and before the lexical verb. In addition, in both Lopit and Lokoya, the subordinate marker is used on the lexical verb, which often occurs where there are two verbs in a construction.

For these reasons, I proposed that the negation word in these languages is best regarded as a verb or a verb-like word. I will now present standard negation in each of the three languages before discussing the construction in some detail in §2.5.

2.2 Standard negation in Otuho

In Otuho, the negative constituent can be regarded as a verb, in this case /beŋ/, 'not be'. To show this in detail, I will first describe case marking in this marked nominative language. Examples (3) and (4) show the change in case marking when the word for baboon changes from the absolutive-marked object [àmòlòŋ] in (3) to the nominative-marked subject [ámólóŋ] in (4).[6] Note that the word order changes from VSO to VOS as a result of the prominence hierarchy (Moodie & Billington 2020: p.248).

(3) Otuho
 è-góɲú ní   àmòlòŋ
 1SG-see 1SG.NOM baboon.ABS
 'I see the baboon.' MB3-001 00:40:41

---

[5]This kind of construction also occurs with other words that are derived from auxiliary-like verbs. The word /lɔɟɔ/ or/laɟɔ/ in Lopit is often used to introduce an adverbial clause as in i, where it can be translated as 'when'. It is probably derived from the verb /ɟɔ/, 'say' and could be glossed as /l-a-ɟɔ/, (SBO-1SG-say). It is verb-like in that it indexes the subject of the subordinate clause, which is in the nominative case.

 (i) [làɟɔ́   náŋ   l-á-wú   à=tòrìt][á-bót   náŋ   à=bòlìs]
   when.1SG 1SG.NOM SBO-1SG-go to-Torit 1SG-go.direct 1SG.NOM to.police
   'When I go to Torit, I will go straight to the police.' MB1-160 00:14:20

Similar constructions are found in Otuho (/ette/, 3.go, 'and then' (Muratori 1938: p.156)).

[6]The first-person singular pronoun changes from /ni/ in the nominative to /nani/ in the absolutive. This is unusual and has only been observed for this pronoun in this language.

(4) Otuho

ì-góɲú nàní ámólóŋ
3sg-see 1sg.abs baboon.nom
'The baboon sees me.' mb3-001 00:38:25

Another variation from the unmarked word order can occur when there is clefting, whereby the subject or object is expressed before the verb. In the following example, the speaker is emphasising that it was the baboon (and not a different animal) that saw him. The subject /amoloŋ/ no longer has nominative case marking. This is an example of no nominative case before the verb, as discussed above in §2.1.

(5) Otuho

àmòlòŋ ì-góɲú nàní
baboon.abs 3sg-see 1sg.abs
'(It is) The baboon (who) sees me.' mb3-003 00:19:16

When the utterances in (3) and (4) are negated, as shown in (6) and (7) respectively, the negative verb /beŋ/ is clause-initial and the subject is placed before the lexical verb /goɲu/. The subjects, /ni/ and /amoloŋ/, retain their nominative marking, supporting the hypothesis of the verbal characteristics of the negative constituent.

(6) Otuho

á-béŋ ní è-góɲú àmòlòŋ
1sg-not.be 1sg.nom 1sg-see baboon.abs
'I don't see the baboon.' mb3-001 00:45:48

(7) Otuho

ò-béŋ ámólóŋ ì-góɲú nàní
3sg-not.be baboon.nom 3sg-see 1sg.abs
'The baboon doesn't see me.' mb3-001 00:46:25

2.3 Standard negation in Lopit

This is similar to Otuho and involves the verb or verb-like constituent /ɲa/, 'not be'. The morpheme /ɲa/ can be regarded as the base form or root of the negative verb. In simple negation, the negative constituent /ɪɲa/ does not have the normal pronominal marking and can be regarded as an allomorph of the verb /ɲa/. It is used for 1st, 2nd and 3rd person and I gloss it as 'not.be'. It might be a syncretic

form, possibly coming from the second person form /ɪ-ɲa/. As shown in (9), the negative constituent is clause initial, the subject appears before the lexical verb and maintains its nominative case marking.

(8) Lopit
    eí-wóló mɔ́lɔ́ŋ   nàŋ
    3>1-see baboon.NOM 1SG.ABS
    'The baboon sees me.' MB1-133 00:10:51

(9) Lopit
    íɲà   mɔ́lɔ́ŋ   l-eí-wóló   nàŋ
    not.be baboon.NOM SBO-3>1-see 1SG.ABS
    'The baboon does not see me.' MB1-133 00:12:22

In more complex negation constructions, normal pronominal prefixing is used, as illustrated in (2). In addition, prefixing can be used on the negative verb /ɲa/ to mark modal, subordinate, and sequential constructions. This is discussed in §5. It should be noted that, unlike Otuho, the lexical verb /wolo/ in (9) is prefixed with what I call the (Lopit) subordinate marker, SBO /l-/.

A further example of the verbal nature of the negative constituent is provided in example (10). Clauses in Lopit can be coordinated using the sequential prefix /x-/ on the verb in the second clause (Moodie & Billington 2020: p.351). When the second clause is a negative construction, the sequential prefix combines with the negative verb to form the word /xɔɲa/ in the Dorik dialect and /xɪɲa/ in the Ngutira dialect (Moodie & Billington 2020: p.302). The prefix /x-/ has only been observed as a prefix with verbs.

(10) Lopit
    á-lìxà   náŋ   xòlòŋì wùnìk dè=tìm   x-ɔ̀ɲà
    1SG-hunt.PFV 1SG.NOM days.ABS three   in=forest.ABS SEQ-not.be
    l-á-ròmà   nàbó tiàŋ
    SBO-1SG-find.PFV one.F animal.ABS
    'I hunted for three days in the bush and didn't find one animal.' MB1-314
    00:16:51

2.4 Standard negation in Lokoya

This uses two morphemes, /awaŋ/ and /kaŋ/ in standard negation. From the data so far available, /awaŋ/ is used for the negation of clauses with intransitive and stative verbs. The morpheme /kaŋ/ is used for negation with transitive verbs.

Both /awaŋ/ and /kaŋ/ are often expressed as /aŋ/. In addition, there is another morpheme, /ɪkɛ/, which is used as the negator in some non-verbal predicates (see §4.1). I will first examine /awaŋ/.

The verb /awaŋ/ shows similar syntactic behaviour to its counterparts in Otuho and Lopit. As shown in (12), the negative constituent is clause initial, the subject appears before the lexical verb and maintains its nominative case marking. From the limited work carried out to date, the initial vowel /a/ does not change with the person of the subject and is possibly a syncretic form in the negative verb paradigm. Note that the lexical verb /wak/ in the negative construction is marked with the prefix /x-/. This prefix does not show gender or number agreement and is glossed as the negative subordinate marker (x-, NEG.SBO), in contrast to the normal (non-negative) subordinate marker.[7]

(11) Lokoya
    à-wàk   náŋ   àndén dí=tòrít mòíté
    1SG-want 1SG.NOM INF.go to=Torit tomorrow
    'I want to go to Torit tomorrow.' MB2-005 01:21:29

(12) Lokoya
    áwàŋ náŋ   x-á-wàk   àndén dí=tòrít
    not.be 1SG.NOM NEG.SBO-1SG-want INF.go to=Torit
    'I don't want to go to Torit.' MB2-005 01:22:50

Some examples with /kaŋ/ are given in (13) and (14). The verbs /isara/, 'give' and /ŋoxi/, 'finish', 'consume', are both transitive. The construction with /kaŋ/ is similar to that for /awaŋ/ in that the subject maintains its nominative case and the lexical verb is prefixed with the negative subordinator /x-/, 'NEG.SBO'. There is no initial vowel in this negator. This contrast with all the other negators or negative verbs and is a topic for further research.

(13) Lokoya
    káŋ òtúlò   lé   x-e-ìsárà-k   àxárí   do=oìtó
    not.be man.NOM this.NOM NEG.SBO-3-give-DAT water.ABS to=boy.ABS
    'This man didn't give water to the boy.' MB2-025 00:38:42

(14) Lokoya
    káŋ òtúlò   x-ì-ŋóxí   àɲárí
    not.be man.NOM NEG.SBO-3-finish food.ABS
    'The man didn't finish the food.' MB2-025 00:32:18

---

[7]In non-negative constructions, the Lokoya subordinate markers show agreement for singular gender (l-, SBO.M; and n-, SBO.F) and for plural (x-, SBO.PL). See (48) for an example of /n-/, SBO.F.

The negative morpheme is often expressed as /aŋ/, which appears to be a shortened form of both /awaŋ/ and /kaŋ/ as it is used for both intransitive and transitive constructions, as illustrated in (15) and (16) respectively. It has some verbal characteristics in that it is clause-initial and the subject which follows it is marked nominative.

(15) Lokoya
    áŋ   náŋ   x-à-wón   x=à-xítéŋ   a-ìbòŋ
    not.be 1SG.NOM NEG.SBO-1SG-exist with=F-cow 3-be.white
    'I don't have a white cow.' MB2-025 00:17:40

(16) Lokoya
    áŋ   náŋ   x-á-tèxó   òwárù
    not.be 1SG.NOM NEG.SBO-1SG-see.PFV leopard
    'I didn't see the leopard.' MB2-019 01:25:26

2.5 Discussion of the standard negation construction

The use of a negative and a lexical verb in the negative construction occurs in some of the world's languages, mainly Oceanic, Salish, Yuman and some Paleo-Siberian languages (Payne 1985: pp. 207-222). In his discussion on negative verbs, Payne groups them into auxiliary negative verbs and higher negative verbs. In the first group, the negative verb "acts as a finite auxiliary to the lexical verb which in turn typically occurs in some non-finite form" (Payne 1985: pp. 207). The Lopit negative verb cannot be regarded as an auxiliary verb since the lexical verb maintains its pronominal marking and can also be marked for aspect, as is shown with the verb /iwus/ in the contrast between (17) and (18). That is, the lexical verb is finite.

(17) Lopit
    íŋà   náŋ   l-a-íwús-ò   ʃaî
    not.be 1SG.NOM SBO-1SG-drink-IPFV tea.ABS
    'I'm not drinking tea' MB1-269 00:32:18

(18) Lopit
    íŋà   náŋ   l-a-îwús   ʃaî
    not.be 1SG.NOM SBO-1SG-drink.PFV tea.ABS
    'I didn't drink tea.' MB1-269 00:32:50

In Payne's other group, the negative verbs take a full sentential complement. The Lopit negative verb might belong to this group. For example, in (19), 'I'm not

going to Torit', could be interpreted as something like 'I am not that I'm going to Torit' (or 'I am not the one that I am going to Torit'), where the expression [láwú àtòrít], '(that) I'm going to Torit', could be a relative clause.

(19) Lopit

íɲà	náŋ	l-á-wú	à=tòrít
not.be	1SG.NOM	SBO-1SG-go.IPFV	to=torit.

'I'm not going to Torit.' MB1-163 00:05:54

Under some circumstances, /ɪɲa/ can be translated as 'there is not' or 'there is no' (see §4.4). If this were the case here, one might interpret the construction in (19) as 'It is not me that I am going to Torit' or as 'It's not that I'm going to Torit'. However, since the subject maintains its nominative case marking when the negative is formed, these interpretations are not possible. In addition, the use of the first-person pronominal marking in the verb /wu/ means that (19) could not be interpreted as 'I am not the one (who is) going to Torit'.

An alternative interpretation of the negative construction can be made using the concept of the CONTROL relation (Kroeger 2004: pp. 103-134). We can consider (19) as a bi-clausal construction, [íɲà náŋ], 'I am not', and [láwú àtòrít], '(that I am) going to Torit'. The clause [láwú àtòrít] can be regarded as a SENTENTIAL COMPLEMENT, i.e. a finite subordinate clause which functions as an argument of the verb [íɲà] (Kroeger 2004: p. 109). This clause is finite in that the verb has pronominal and aspect marking. A possible lexical entry for the verb /ɲa/, 'not be', is shown in Table 3.

Table 3: A possible lexical entry for the verb /ɲa/

/ɲa/, 'not.be'	<agent,	proposition >
	|	|
	subject	S-COMP

The subject of the main clause is [náŋ], 'I', and it can be described as the CONTROLLER which CONTROLS the subject of the subordinate clause (the CONTROLLEE). As a result, in (19), the subject of the subordinate clause [láwú àtòrít] is the first-person singular pronoun [náŋ] and is indexed on the verb [láwú] with the first-person pronominal marker /a/.

Similar analyses can be applied to standard negation in both Otuho and Lokoya. It should be noted that all these analyses are tentative and further research is required.

3 Negation in non-declarative clauses

3.1 Negative imperatives

The three languages have a range of constructions for the negative imperative. Both Otuho and Lopit use a negative imperative prefix on the lexical verb. Lopit also uses a separate (auxiliary) verb. Lokoya uses a separate morpheme, which might be a verb or a particle.

In Otuho, the negative imperative is expressed with the verbal prefix [xɛ̀-] as shown for the second person imperative in (20). The prefix is the same in the plural, as shown in (21).

(20) Otuho
      xɛ̀-ló   tòɾìt
      NEG.IMP-go.SG Torit
      'Don't go to Torit!' MB3-004 00:21:49

(21) Otuho
      xɛ̀-fwátà   tòɾìt
      NEG.IMP-go.PL Torit
      'Don't go to Torit!' MB3-004 00:22:06

In Lopit, there is a special construction for negative imperatives. It involves the negative imperative prefix /xaɪ-/ together with the word /idek/, which can be translated as 'leave', 'give up', 'don't do'. Some examples are given in (22) and (23).

(22) Lopit
      ídèk   xaɪ-íɾɔ̀
      not.do.IMP NEG.IMP-speak
      'Don't talk!' MB1-014 00:36:40

(23) Lopit
      ídèk   xai-ísò   nàŋ   xíɾíŋɔ̀
      not.do.IMP NEG.IMP-give 1SG.ABS meat.ABS
      'Don't give me meat!' MB1-324 00:09:38

In Lokoya, the morpheme /ko/ is used, together with the standard imperative, to express the negative imperative. The same form is used for singular and plural, as is illustrated in (24) and (25). I have insufficient data to determine whether /ko/ is a particle or a verb (or has been derived from a verb).

(24) Lokoya

kó   ì-ló
NEG.IMP IMP-go.SG
'Don't go!' MB2-023 00:35:42

(25) Lokoya

kó   ì-fún-ù
NEG.IMP IMP-come.PL-VEN
'Don't come!' MB2-023 00:37:27

3.2 Negative Interrogatives

The Otuho and Lopit languages use prefixes on the negative verb for negative polar questions. I have not yet been able to determine how negative polar questions are expressed in Lokoya.

Polar interrogatives in Otuho use the same construction as the corresponding positive clause except that there is a difference in prosody, with increasing pitch at the end of the question. Negative polar interrogatives can be formed using the negative verb /beŋ/, prefixed with the marker [xɛ́-] which I gloss as a negative interrogative marker. It has the same form as the negative imperative marker, except that is has a High instead of a Low tone. A pair of affirmative and negative polar questions is given in (26) and (27). The word order with the negative question verb is the same as the word order in negative declarative clauses, as shown in (6) and (7), for example.

(26) Otuho

í-túk   íjè   àŋíjò ŋìrjà
2-finish 2SG.NOM VN.eat food.ABS
'Have you finished eating the food?' MB3-004 00:31:55

(27) Otuho

xɛ́-bɛ́ŋ   íjè   í-túk   àŋíjò ŋìrjà
NEG.Q-not.be 2SG.NOM 2-finish VN.eat food.ABS
'Haven't you finished eating the food (yet)?' MB3-004 00:32:20

In Lopit, polar interrogatives use the prefix /x-/ on the lexical verb. Negative polar interrogatives can be formed using the negative verb /ɲa/ together with the question marker /x-/. A pair of affirmative and negative polar questions is given in (28) and (29). As with Otuho, the word order in the negative polar question is the same as the word order in negative declarative clauses, such as in (9). In

(29), the subject [íjé], 'you', is placed in front of the lexical verb, which maintains its pronominal prefix and is also prefixed with the subordinate marker /l-/. The word [xɔ́ɲà] is glossed as 'Q-not.be'. This is somewhat similar to /ɪɲa/, discussed above, except that /ɔ-/ is used, instead of /ɪ-/, for the first, second or third person.

(28) Lopit
x-í-t:ɔ́xɔ̀   íjé
Q-2-finish 2SG.NOM
'Have you finished?' MB1-293 00:20:02

(29) Lopit
x-ɔ́ɲà   xàtí íjé   l-í-t:ɔ́xɔ̀   ´
Q-not.be but 2SG.NOM SBO-2-finish
'Haven't you finished (yet)?' MB1-330 00:39:01

4 Negation in stative predications

There are several different kinds of stative predications and different ways of describing them. I use the terminology of Payne (1997) and of Dryer (2007). The languages have stative predications with and without copula verbs. The copula verbs are similar across the three languages: /aɾa/, 'be' and /w:on/ ~ /wɔn/, 'be, exist'. Sometimes standard negation is used with a copula verb, even if there is no copula verb in the positive construction.

4.1 Identity or nominal predication

There are two main types of identity or nominal predications. The first is illustrated in the Lopit example (30) and can be described as an equative predicate nominal (Payne 1997: p. 114) or as a referential nominal predicate (Dryer 2007: p. 233). The copula verb /ɾa/ has a fused form in most situations. The form /aɾa/ is generally used for first, second and third person in these situations.[8]

---

[8]Like /ɪɲa/, /aɾa/ does not usually show pronominal indexing. However, it can have pronominal and modality marking and thus be considered to be a form of the verb /ɾa/, 'be', as shown with in the following.

(i) Lopit
í-maì-ɾá íjè   xábʊ́ í-maì-lʊáxà-k íjè   íjòxoì
2-COND-be 2SG.NOM chief.ABS 2-COND-help-DAT 2SG.NOM 1PL.ABS
'If you were the chief, you would help us.' MB1-006 00:37:12

(30) Lopit

  árá xàbʊ̀ lɔ̀rɛ̀wâ lìtí

  be chief.NOM husband.ABS my.M.ABS

  'The chief is my husband.' MB1-162 00:15:50

The negation of (30) is given in (31). The construction is similar to standard negation. The negative verb is clause-initial, and the subject is placed after the negative verb and before the copula.

(31) Lopit

  íɲà xàbʊ̀ l-árá lɔ̀rɛ̀wâ lìtí

  not.be chief.NOM SBO-be husband.ABS my.M.ABS

  'The chief is not my husband.' MB1-341 00:27:52

Sometimes, this kind of stative predication is expressed without copula verbs. Demonstratives can be used, as shown in (32). When this verbless copula construction is negated, the copula verb, /ra/, 'be', is introduced with the negative verb, as shown in (33). Once again, this is similar to the standard negation construction.

(32) Lopit

  xábʊ́ ìl:ɛ́ŋ

  chief.ABS this.M.NOM

  'This is the chief.' MB1-341 00:26:38

(33) Lopit

  íɲà ìl:ɛ́ŋ l-árá xábʊ́

  not.be this.M.NOM SBO-be chief.ABS

  'This is not the chief.' MB1-341 00:26:59

In Otuho, stative predications usually have no copula verb, as illustrated in (34). As in Lopit, in example (32), the demonstratives have nominative case and are placed after the noun phrase, which is in the absolutive case. When the negative construction is formed, a copula verb (/ara/ or /w:ɔn/) is usually added, as shown in (35).

(34) Otuho

  xóbú xóxoì òlò

  chief.ABS our.ABS this.M.NOM

  'This is our chief.' MB3-004 00:40:26

(35) Otuho

    ò-béŋ   árá xóbú   xóxoì   òlò

    3SG-not.be be   chief.ABS our.ABS this.M.NOM

    'This is not our chief.' MB3-004 00:40:55

Identity or nominal predications in Lokoya are usually constructed without a copula verb, as shown in (36).

(36) Lokoya

    òxóbú   òtúló   xíndàŋ

    chief.NOM husband.ABS my.M.ABS

    'The chief is my husband.' MB2-025 00:00:56

The negation of (36) is given in (37). A different negator /ɪkɛ/ is used and this is glossed as NEG and can be described as the nominal predicate negator. This negator does not appear to have any verbal characteristics and could be described as a particle. It is only observed in non-verbal nominal predication.

(37) Lokoya

    òxóbú   ìkɛ́ òtúló   xíndàŋ

    chief.NOM NEG husband.ABS my.M.ABS

    'The chief is not my husband.' MB2-025 00:02:01

The second type of identity or nominal predication has been called proper inclusion (Payne 1997: p. 114) or non-referential nominal predication (Dryer 2007: p. 233). In both Otuho and Lopit, a copula verb is used, as shown in (38) and (40) respectively. In the Otuho examples, (39) is the negation of (38). This construction is also similar to standard negation.

(38) Otuho

    árá íɲí   ètíj:énàní

    be 3SG.NOM teacher.ABS

    'He/she is a teacher.' MB3-004 00:44:41

(39) Otuho

    ò-béŋ   íɲí   l-árá ètíj:énàní

    3SG-not.be 3SG.NOM SBO-be teacher.ABS

    'He/she is not a teacher.' MB3-004 00:46:18

Corresponding examples are given for Lopit in (40) and (41). These have the same constructions as Otuho.

(40) Lopit
árá íɲɛ́   xaìtíj:énàní
be 3SG.NOM teacher.ABS
'He/she is a teacher.' MB1-329 00:17:06

(41) Lopit
íɲà íɲɛ́   l-árá xaìtíj:énàní
not.be 3SG.NOM SBO-be teacher.ABS
'He/she is not a teacher.' MB1-329 00:17:12

In Lokoya, a copula verb is not used for proper inclusion or non-referential nominal predication, as shown in (42). When negated, the negator /ɪkɛ/ is used, as shown in (43). This is the same as with equative predicate nominals or referential nominal predicates, as discussed in relation to (37).

(42) Lokoya
lálá mɔ̀ɲɛ̀   ò-írìjá tànání
my.ABS father.ABS M-liria person.ABS
'My father is a Lyria man.' MB2-006 00:23:31

(43) Lokoya
lálá mɔ̀ɲɛ̀   ìkɛ́ ò-írìjá tànání
my.ABS father.ABS NEG M-liria person.ABS
'My father is not a Lyria man.' MB2-025 00:03:36

4.2 Property predications

Property assignment or attributive construction predicates in all three languages are usually expressed with a stative verb construction (or intransitive verbal predicate), as shown in (44) for Lopit. The same constructions are used for both permanent and temporary property assignment.

(44) Lopit
ɛ̀-xálàn íɲɛ́   bínɔ̀
3-be.lazy 3SG.NOM very
'He is very lazy.' MB1-128 00:37:11

This kind of expression is negated using standard negation, as shown in (45).

(45) Lopit
íɲà   náŋ   l-á-xálàn
not.be 1SG.NOM SBO-3-be.lazy
'I am not lazy.' MB1-128 00:36:05

The situation is similar in Otuho, and examples of property predication and its negation are given in (46) and (47). Once again, standard negation is used.

(46) Otuho
ò-múnò   íɲí
3SG-be.happy 3SG.NOM
'He is happy.' MB3-004 00:53:59

(47) Otuho
ò-béŋ   íɲí   ò-múnò
3SG-not.be 3SG.NOM 3SG-be.happy
'He is not happy.' MB3-004 00:54:07

Examples for Lokoya are given in (48) and (49). Note that, as mentioned in footnote 2, nominal subjects tend to follow an SVO order and that the subordinate marker changes to the negative subordinate marker. When SVO order is used in these languages, the subject loses its nominative case marking and is marked with the absolutive case. This is an example of 'no case before the verb' in marked nominative languages as discussed in §2.1. In the negative construction in (49), the subject, [aító], 'girl', is placed after the negative verb (or the verb-like negative word) and hence is marked with the nominative case.

(48) Lokoya
aìtó   n-ò-múnò
girl.ABS SBO.F-3-be.happy
'The girl is happy.' MB2-002 00:26:16

(49) Lokoya
áwàŋ aító   x-ò-múnò
not.be girl.NOM NEG.SBO-3-be.happy
'The girl is not happy.' MB2-002 00:26:45

4.3 Locative predications

Negation in locative predications in Otuho is similar to standard negation. That is, the negative verb /beŋ/ is used in the clause initial position. This is shown in example (51).

(50) Otuho

  ò-wón   àmɔ́lɔ́ŋ   ɔ̀tɔ̀=xídɛ̀ jànì
  3sg-exist baboon.nom on=top   tree.abs

  'The baboon is on (top of) the tree.' mb3-004 00:53:05

(51) Otuho

  ò-béŋ   àmɔ́lɔ́ŋ   ò-wón   ɔ̀tɔ̀=xídɛ̀ jànì
  3sg-not.be baboon.nom 3sg-exist on=top   tree.abs

  'The baboon is not on (top of) the tree.' mb3-004 00:52:14

In Lopit, standard negation can also be used in locative predicate constructions with the copula verb, /w:ɔn/, 'exist', as shown in (52) and (53).

(52) Lopit

  ɔ̀-w:ón mɔ́lɔ́ŋ   dɛ̀=jànì
  3-exist baboon.nom   in=tree.abs

  'The baboon is in the tree.' mb1-162 00:38:14

(53) Lopit

  íɲà   mɔ́lɔ́ŋ   l-ɔ̀-w:ón   dɛ̀=jànì
  not.be baboon.nom sbo-3-exist in=tree.abs

  'The baboon is not in the tree' mb1-171 00:01:45

Lokoya also uses the copula verb, /won/, 'exist', in locative predications, as shown in (54). In contrast to Lopit and Otuho, Lokoya does not use standard negation for the negative of these constructions. Only the negative verb is used, as shown in (55).

(54) Lokoya

  ò-wón àxárí   dì xáŋ
  3-exist water.abs in house.abs

  'There is water in the house.' mb2-025 00:16:05.

(55) Lokoya

  áwáŋ áɲàrí   dì=ðàðá   àxàŋ
  not.be food.abs in=interior house.abs

  'There is no food inside the house.' mb2-019 01:26:35

4.4 Existential predications

Existential predications can overlap with locative predications in some languages (Dryer 2007: p. 240). However, in Lopit, Otuho and Lokoya there are some exis-

tential predicates with no overt attribution of a location. An example of an existential predication is given in (56), where the copula verb /wːɔn/, 'exist', 'be available', is used with a dummy subject. The negative existential predication is given in (57) using the negative verb /ɪŋa/ with no lexical or copula verb. Here, [íɲà] can be translated as 'there is not' or 'there is no'. Negation of this kind of existential (perhaps best described as 'non-locative existential') is different from standard negation. The noun /xɪsʊŋ/ in (56) is in the absolutive case as it is the existential predicate (i.e. not a subject).

(56) Lopit
ɔ̀-wːɔ́n xìsʊ́ŋ   xʊ́ná l-ɔ́-bwɔ̀r
3-exist cows.ABS of.F.PL SBO-3-be.white
'There are white cows (lit. it exists cows which are white)' MB1-341
00:17:51

(57) Lopit
íɲà   xìsʊ́ŋ   xʊ́ná l-ɔ́-ɲórì
not.be cows.ABS of.F.PL SBO-3-be.green
'There are no green cows.' MB1-341 00:18:08

Another example of a non-locative existential is given in (58). Once again, there is no copula verb in the negative construction. However, it is worth noting that, since the negative clause ([lɛ́ɲà dáxá], 'if there is no food') is subordinate, the negative verb has the third person pronominal prefix (referencing a dummy subject).

(58) Lopit
l-ɛ́-ɲà   dáxá   é-jeí xìjò
SBO-3-not.be eat.VN.ABS 3-die people.ABS
'If there is no food, people die.' MB1-341 00:23:35

A pair of utterances demonstrating a positive and a negative existential with the same forms as (56) and (57) is shown in (59) and (60).

(59) Lopit
ɔ̀-wːɔ́n xìfjɔ̂ŋ
3-exist water.ABS
'There is water.' MB1-341 00:21:57

(60) Lopit
ínà   xìfjôŋ
not.be water.ABS
'There is no water.' MB1-256 00:07:49

A consultant was asked if one could say the utterance shown in (61). He said that it "will sound incomplete" (MB1-256 00:22:49) and that an utterance like that in (62) would be more acceptable. Example (62) could be regarded as a locative existential construction. This confirms that standard negation is used in locative predicate constructions but not in non-locative existential constructions.

(61) Lopit
? ínà   xìfjôŋ   l-ɔ̀-wːɔ́n
  not.be water.ABS SBO-3-exist
'There is no water.' MB1-341 00:22:29

(62) Lopit
ínà   xìfjôŋ   l-ɔ̀-wːɔ́n   ìní
not.be water.ABS SBO-3-exist here
'There is no water here.' MB1-341 00:23:03

The constructions are similar in Otuho. The non-locative existential has no copula verb. This is shown in (63) which is the Otuho equivalent of (60). When a locative predication is made from (63), as shown in (64) the copula verb /wɔn/, 'exist', is used.

(63) Otuho
ò-béŋ   nàːɾì
3SG-not.be water.ABS
'There is no water.' MB3-004 00:33:30

(64) Otuho
ò-béŋ   ò-wónì   nàːɾì   ì=xàŋ   xɔ̀xɔí
3SG-not.be 3SG-exist water.ABS in=house.ABS our.ABS
'There is no water in our house.' MB3-004 00:35:28

Non-locative existentials in Lokoya are similar to the other two languages, as illustrated in (65) and (66). The copula verb is used in the positive construction but not in the negative one.

(65) Lokoya
ò-wón àxìsɔ́m   x-à-bóŋ-íxà
3-exist cows.ABS SBO.PL-3-be.white-PL
'There are white cows.' MB2-025 00:14:08

(66) Lokoya
àwáŋ àxìsɔ́m   x-à-lì
not.be cows.ABS SBO.PL-3-be.green
'There are no green cows.' MB2-025 00:14:37

4.5 Possessive predications

Possessive predicate constructions can also be negated with the negative verb. In Lopit, the comitative construction /wːɔn xɔ/, 'exist with', is often used to express possession, as shown in (67). The negative form is shown in (68).

(67) Lopit
ɛí-wːɔ́n íjòxoì   xò=dòŋì?
1PL-exist 1PL.NOM with=drums.ABS
'We have drums (lit. we are with drums).' MB1-329 00:06:09

(68) Lopit
íɲà   íjòxoì   l-ɛí-wːɔ́n   xò=dòŋì?
not.be 1PL.NOM SBO-1PL-exist with=drums.ABS
'We have no drums (lit. we are not with drums).' MB1-329 00:06:40

Lokoya uses a similar construction for possessive predicates and the negative form also uses standard negation, as illustrated in (69) and (70).

(69) Lokoya
í-wːɔ́n   xɔ́xɔì   xɔ=àxàɾì
1PL-exist 1PL.NOM with=water.ABS
'We have water.' MB2-001 00:12:58

(70) Lokoya
awaŋ xɔ́xɔì   x-í-wːɔ́n   xɔ=àxàɾì
not.be 1PL.NOM NEG.SBO-1PL-exist with=water.ABS
'We have no water.' MB2-00:24:05

5 Negation in non-main clauses

Negation in non-main clauses is the same as standard negation in Otuho and Lokoya. It is similar in Lopit except that the normal range of pronominal marking is used on the negative verb /ɲa/.

In Otuho, negative constructions in subordinate clauses are similar to those in declarative main clauses. An example of a conditional clause is shown in (71). Otuho uses adverbs (e.g. /ɲa/, /dwo/) to express the modalities such as the conditional, irrealis and hypothetical (Muratori 1938: p.439). The negation of the clause in (71) is shown in (72). The negative auxiliary is clause-initial, and the lexical verb is placed after the subject. The conditional particle /dwo/ and the subject maintain their positions after the clause-initial verb.

(71) Otuho
á-ŋwaí   dwò   ní
1SG-be.ill COND 1SG.NOM
'If I were sick...' MB3-004 01:04:27

(72) Otuho
á-béŋ   dwò   ní   á-ŋwaí á-ló   dwò   ní   tòrìt
1SG-not.be COND 1SG.NOM 1SG-be.ill 1SG-go COND 1SG.NOM Torit.ABS
'If I wasn't sick, I could have gone to Torit.' MB3-004 01:02:45

In Lopit, modality is normally expressed with modal prefixes on the verbs such as /ŋaɪ-/, irrealis and /tV-/, obligative (Moodie & Billington 2020: p.237). The negative verb can be marked for modality, and, in these constructions, the normal pronominal marking is used on the negative auxiliary. An example is given in the conditional clause using the irrealis prefix /ŋaɪ-/ in (73) where the subordinate clause is [áŋaîɲà náŋ lámweí], 'if I were not sick'. The negative verb [áŋaîɲà] has pronominal marking /a-/, '1SG', and the lexical verb, [lámweí], retains its pronominal marking.

(73) Lopit
[á-ŋaî-ɲà   náŋ   l-á-mweí]   [á-ŋaî-ibóŋ náŋ
1SG-IRR-not.be 1SG.NOM SBO-1SG-be.sick 1SG-IRR-meet 1SG.NOM
xò=xìjò   xʊ̀ná l-ɔ̀-lʊ́ŋà]
with=people.ABS REL.PL SBO-3-be.many
'If I weren't sick, I would have met many people.' MB1-014 01:25:40

In Lokoya, modality is expressed with particles, such as /ara/, which indicate a conditional or irrealis modality. In contrast to Otuho, the conditional particle

precedes the verb. A non-negative conditional construction is shown in (74). In a negative conditional construction, as in (75), the conditional particle maintains its position at the front of the clause and the remainder of the clause has the normal standard negation structure.

(74) Lokoya
      ara   á-díxá   náŋ
      COND 1SG-be.sick 1SG.NOM
      'If I were sick...' MB2-025 00:15:24

(75) Lokoya
      ara aŋ náŋ   x-á-díxá   a-ísúr   náŋ
      COND NEG 1SG.NOM NEG.SBO-1SG-be.sick 1SG-dance 1SG.NOM
      'If I were not sick, I would dance.' MB2-025 00:17:35

6 Negation in Eastern Nilotic languages

Eastern Nilotic languages have different ways of expressing standard negation – verbs, prefixes and particles. Some languages use all three methods. These are shown in Table 4.

Table 4: Methods of negation in Eastern Nilotic languages

	Lotuxo-Maa				Teso-Turkana			Bari
	Lopit	Otuho	Lokoya	Maa	Turkana	Toposa	Ateso	Bari
verb	ɲa	beŋ	waŋ, kaŋ	ɛɪtʊ	mam	me	mam	
prefix				m-	ma-; pe-; ɲi-	ɲ-; pa-		
particle				ḿmὲ	mɛɛrɛ	meere	mamʊ	ti; tine

In Otuho, Lopit and Lokoya, negation is mostly expressed with the negative verbs. From an examination of the work on Eastern Nilotic languages, it appears that prefixes and particles are mostly used in Maa, Bari and the Teso/Turkana languages (Barasa 2017, Dimmendaal 1983, Schröder & Schröder 1984, Spagnolo 1933, Tucker & Mpaayei 1955). However, sometimes negative verbs, translated

as 'lack', are used, particularly in the Teso/Turkana languages. In addition, in Maa, an auxiliary negative verb is used. The morpheme /ɛɪtʊ/ is described as a "negative perfect(ive) aspect auxiliary verb, indicating that something is not yet done or complete" (Payne & Ole-Kotikash 2008). It does not inflect for person.

Where negative verbs are used in the Teso/Turkana languages, they show limited similarities with Otuho, Lopit and Lokoya. Firstly, there is pronominal marking on the negative verb, as shown with the Turkana verb /mamʊ/, 'lack', in (76), and the Ateso verb /mamar/, 'not.have', in (77).

(76) Turkana
      ɛ-màmʊ̀ kà   nège`
      3-lack   from here
      'He is not here.' (Dimmendaal 1983: 456)

(77) Ateso
      nés á-màmàr   ápòrèí   k=éòŋ
      CONJ 1SG-not.have scar.ABS PREP=me
      '… so I have no scars on me.' (Barasa 2017: 248)

It may be possible to distinguish between a negative verb and a negative particle. In Ateso, the negator /mamʊ/ can be described as a particle since the subject is marked with the absolutive, and not the nominative case, when it moves in front of the lexical verb and after the negator (Barasa 2017: p.208). That is, /mamʊ/ cannot be regarded as a verb (or as verb-like), since if it were a verb, the subject would have the nominative case. This is illustrated with [éòŋ], '1SG.ABS', in example (78).

(78) Ateso
      màmʊ̀ éòŋ   á-ŋàdàkì   àpéséí   kón
      NEG   1SG.ABS 1SG-keep.PAST money.ABS your
      'I have not kept your money.' (Barasa 2017: p.220)

However, sometimes Barasa also gives examples where the subject is in the nominative case. An example is given with [èòŋ], '1SG.NOM', in example (79). The use of the nominative case for the subject [èòŋ], would suggest that /mamʊ/ is still regarded by some speakers as a verb or verb-like.

(79) Ateso
      màmʊ̀ èòŋ   á-gòlòk-ìt   ékèkì
      NEG   1SG.NOM 1SG-close-PFV door.ABS
      'I did not close the door.' (Barasa 2017: p.249)

7 Discussion and conclusions

In Otuho, Lopit and Lokoya, standard negation is almost always expressed with a negative verb. The negative construction is similar across the three languages, using the order NEG.VERB SUBJECT LEXICAL.VERB OBJECT. The subject retains its nominative marking when placed before the lexical verb, since it still follows a (negative) verb. Both the negative verb and the lexical verb are finite, in that they have pronominal marking and can also have aspect or modal marking.

Different negative verbs are used in the three languages: /beŋ/ for Otuho, /ɲa/ for Lopit and /waŋ/ or /kaŋ/ for Lokoya. It is especially interesting as there are many common verbs. For example, the copula verbs are almost identical: /aɾa/, 'be' and /w:on/ ~ /wɔn/, 'be, exist'.

The main differences between the languages in standard negation relate to the extent of pronominal marking, the subordinating prefix marking of the lexical verb and the extent of modal marking on the negative verb. The commonalities and differences across the languages are listed in Table 5.

Table 5: Features of standard negation constructions

Feature	Otuho	Lopit	Lokoya
pronominal marking on negative verb	yes	sometimes	no
pronominal marking on lexical verb	yes	yes	yes
subject has nominative marking	yes	yes	yes
subordinating prefix on lexical verb	no	yes	yes
TAM marking on the negative verb	no	yes	no

A feature of these languages is the bi-clausal nature of standard negation. That is, both the negative verb and the lexical verb are finite. The negative verb can have pronominal marking and be inflected for aspect and modality. The lexical verb also has pronominal marking and, at least in Lopit and Lokoya, has aspect marking. Both verbs index the person of the subject. I have suggested that this construction could be interpreted as a relative clause construction or as a control relation. More research is required to understand these constructions.

Generally, negated non-verbal constructions use a copula verb as well as the negative verb, even if the positive non-verbal construction has no copula. These constructions are of the same form as standard negation. The only exceptions appear to be for non-locative existentials (and Lokoya locative existentials) where the negative verb is used without a copula.

Different constructions to standard negation are used for negative imperatives and negative interrogatives. These use verbal prefixes and/or auxiliary verbs which are unrelated to those used in standard negation.

Negation in non-main clauses is the same as standard negation for Otuho and Lokoya. It is similar in Lopit except that the normal range of pronominal marking is used on the negative verb /ɲa/.

The use of negative verbs appears to be much more prominent in Otuho, Lopit and Lokoya than in the other Eastern Nilotic languages, which mostly use negative verbal prefixes or particles.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

3>1	3rd person s with 1st person o	SBO	subordinator
CONJ	conjunction	S-COMP	sentential complement
I	Class I	SEQ	sequential marker
II	Class II	VEN	ventive
R	verb root	VN	verbal noun

References

Barasa, David. 2017. *Ateso Grammar: A descriptive account of an Eastern Nilotic language.* Munich: LINCOM Europa.

Coates, Heather. 1985. Otuho phonology and orthography. *Occasional Papers in the Study of Sudanese Languages* 4. 86–118.

Dimmendaal, Gerrit J. 1983. *The Turkana language* (Publications in African Languages and Linguistics 2). Dordrecht: Foris Publications.

Dimmendaal, Gerrit J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 44(2). 214–261.

Dryer, Matthew S. 2007. Clause types. In Timothy Shopen (ed.), *Language typology and syntactic description: Clause structure*, 2nd edn., vol. 1, 224–275. Cambridge: Cambridge University Press.

Greenberg, Joseph H. 1966. *The languages of Africa.* Bloomington: Indiana University.

König, Christa. 2008. *Case in Africa.* Oxford: Oxford University Press.

Kroeger, Paul. 2004. *Analyzing Syntax: A Lexical-Functional Approach.* Cambridge: Cambridge University Press.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the World, Nineteenth Edition*. Dallas, TX: SIL International.

Miestamo, Matti. 2005. *Standard negation: The negation of declarative verbal main clauses in a typological perspective*. Berlin: De Gruyter Mouton.

Moodie, Jonathan & Rosey Billington. 2020. *A Grammar of Lopit*. Leiden, The Netherlands: Brill.

Muratori, Carlo. 1938. *Grammatica Lotuxo*. Verona: Missioni Africane.

Muratori, Carlo. 1948. *English-Bari-Lotuxo-Acoli vocabulary*. Okaru: Catholic Mission Printing Press.

Payne, Doris L. & Leonard Ole-Kotikash. 2008. *Maa online Dictionary*. http://darkwing.uoregon.edu/~dlpayne/Maa%20Lexicon/lexicon/main.htm (20 January, 2023).

Payne, John R. 1985. Negation. In Timothy Shopen (ed.), *Language typology and syntactic description: Clause structure*, 1st edn., vol. 1, 197–242. Cambridge: Cambridge University Press.

Payne, Thomas. 1997. *Describing Morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

Schröder, Helga & Martin Schröder. 1984. The Toposa verb. *Occasional Papers in the Study of Sudanese Languages* 5. 1–47.

Spagnolo, Lorenzo M. 1933. *Bari grammar*. Verona: Missioni Africane.

Tucker, Archibald Norman & John Tompo Ole Mpaayei. 1955. *A Maasai grammar: With vocabulary*. London: Longmans, Green.

Westermann, Diedrich. 1944. Das Oxoriok: Eine sprachliche Skizze. *Studien zur Auslandskunde/Afrika* 3. 19–46.

Chapter 17

A preliminary phonology of Vale

Kenneth S. Olson

SIL International

This paper is a preliminary phonology of Vale, a Sara-Bongo-Bagirmi language in Central African Republic. Vale's consonant system includes implosives, prenasalized plosives, and prenasalized fricatives, but it lacks labial-velar plosives. Robust voiceless fricatives are limited to /s/ and /h/. Vale has five oral and five nasal vowel phonemes. I found no evidence for the presence of ATR vowel harmony. It has two level tones, and I posit five underlying tone melodies: L, H, LH, HL, and HLH. All of these melodies occur on monomorphemic words. The L in the HLH pattern is realized as a floating L which triggers non-automatic downstep, so the HLH pattern surfaces as HꜜH.

1 Introduction

In this paper, I present a preliminary overview of the phonology of Vale (pronounced [và.lé]), a language spoken by approximately 5,400 people in the north-central part of Central African Republic: in Nana-Grébizi prefecture (Kaga Bandoro subprefecture) and Ouham prefecture (Batangafo and Kabo subprefectures). The two language consultants I worked with consider Vale to have six dialects: Dogu Saki, Tane Ngravo, Doro, Hula, Zabba, and Bbuna. Both of them speak Dogu Saki. The language is classified as part of the Sara-Bongo-Bagirmi (SBB) group within the Nilo-Saharan family. Dimmendaal et al. (2019: 343–344) consider SBB to be essentially the same as "West Central Sudanic." The ISO 639-3 code for Vale is [vae]. More general information can be found in *Ethnologue*[1] (Eberhard et al. 2023) and *Glottolog*[2] (Hammarström et al. 2022).

---

[1]https://www.ethnologue.com/language/vae
[2]https://glottolog.org/resource/languoid/id/vale1250

Little previous research has been done on Vale. The only language data I've found are lists of numbers in Bruel (1905: 100)[3] and Gaudefroy-Demombynes (1906: 60), some additional vocabulary items in the latter source, and a few lexical items in Samarin (1971: 231). There are brief mentions of the language in Tucker (1940: 11) and Tucker & Bryan (1956: 17). SIL conducted a language survey on Vale in the 1990s (Moéhama 2021).

Yanguende (2011) discusses the phonology and nominal morphology of what he calls "Vale de Ndélé," but there is some doubt as to whether that is a variety of Vale (cf. §6).

I worked with two mother-tongue speakers of Vale. Speaker 1 is a 42-year-old male. He grew up in the Vale region, but he has lived off and on in Bangui for many years. He has a *baccalauréat*, as well as training in a couple of trades. He speaks Vale, Sango, and French. Both of his parents are mother-tongue Vale speakers.

Speaker 2 is a 30-year-old male. He grew up in the Vale region, but he moved to Bangui as a teenager, where he finished his *baccalauréat*. He also studied for one year at the University of Bangui. Besides Vale, he speaks Sango, French, Luto, Ngama, and "Sara".[4] His father speaks Vale and his mother speaks Luto.

I worked with the speakers in a three-week phonological analysis workshop at the ACATBA[5] center in Bangui in October 2021. I was assisted by Adelphe Dogue-Syssa, a member of ACATBA. We employed the participatory phonology research method described in Kutsch Lojenga (1996). We processed about 800 words that had been collected in 2020 by Dogue-Syssa. (John Berthelette and Paul Murrell also did some preparatory work for the workshop.) In addition, we collected noun phrase and verb phrase paradigms, as well as a short text (*The North Wind and the Sun*), in order to examine tone behavior.

I discuss consonants in §2, vowels in §3, syllable structure in §4, and tone in §5. I discuss some areas for further research in §6, and I draw conclusions in §7. The data are transcribed using the extant International Phonetic Alphabet (IPA 2006).

2 Consonants

Vale has 25 consonant phonemes, as shown in Table 1.

---

[3]Vale is referred to as "Télé de Guélo" in Bruel (1905).

[4]I was unable to verify which variety of Sara this was.

[5]L'Association centrafricaine pour la traduction de la Bible et l'alphabétisation

Table 1: Consonant phonemes

	Bilabial	Labiodental	Alveolar	Palatal	Velar	Glottal
Plosive	ɓ		ɗ			
	p		t		k	
	b		d		g	
	ᵐb		ⁿd		ᵑg	
Fricative		(f)	s			h
	v		z			
	ᵐv		ⁿz			
Nasal	m		n			
Flap/trill	(ⱱ)		r			
Lateral			l			
Approximant	w			j		

The prenasalized consonants are written with two symbols in Table 1, but they pattern as unitary phonemes in Vale. There is no symbol in IPA for a prenasalized consonant, so they will be written as two consonants, e.g., [mb] in phonetic transcription. Parentheses indicate phonemes that are marginal to the phonological system.

Contrast between consonants within the same broad category of place of articulation is shown in Table 2. Insofar as possible, I provide contrast between nouns in word-initial position with a following /a/. Where this was not possible, I've provided additional data.

The period (full stop) symbol ⟨ . ⟩ indicates a syllable break. Glosses were obtained in French and translated into English for this paper.

I consider the voiceless labiodental fricative /f/ to be a marginal phoneme because it is rare. An exhaustive list of occurrences in our corpus is shown in Table 3. The word for 'hold' is pronounced [kù.hà] by the Dogo Saki group and [kù.fà] by the Doro and Hula people.

The prenasalized labiodental fricative /ᵐv/ is pronounced [ɱv] in Vale. I was consistently able to observe visually the labiodental nasal [ɱ] when the speakers produced words containing the sound.

The Vale speakers produced one ideophone containing a voiced bilabial flap: [ⱱa] 'the sound of a slap in the face'. In the production of the flap, the upper lip is brought into the mouth so that the lower lip flaps against the upper lip instead of the upper teeth (Olson & Hajek 2003: 158). Speaker 1 identified the word as being

Table 2: Consonant contrasts

	Phoneme	Phone	Example	Gloss
Labial	/ɓ/	[ɓ]	[ɓà.ɾà]	rainy season
	/p/	[p]	[pà.ɾí]	bile
	/b/	[b]	[bà.ɗà]	lizard
	/ᵐb/	[mb]	[mbáá]	guest, visitor
	/f/	[f]	[fú.ˈlá]	sacrifice
			[tú.ˈfá]	squeak (wheel) (v.)
	/v/	[v]	[và.ɗú]	warthog, pig
	/ᵐv/	[ɱv]	[ɱvàí]	beard, mane
	/m/	[m]	[mà.ɾà]	crocodile; butcher
	/w/	[w]	[wá.jà]	sand
Alveolar	/ɗ/	[ɗ]	[ɗá.mà]	medicine
	/t/	[t]	[tà.lò]	dew
	/d/	[d]	[dà.ɾì]	locust
	/ⁿd/	[nd]	[ndà.là]	duck
	/s/	[s]	[sàá]	smoke
	/z/	[z]	[zá.màl]	camel
	/ⁿz/	[nz]	[nzà.mè]	squirrel
	/n/	[n]	[nà.ŋgà]	dirt, soil
	/r/	[r]	[ró]	body
			[rí]	name
			[kàm.rá]	sky
		[ɾ]	[ɓà.ɾà]	rainy season
	/l/	[l]	[lá]	dance (n.)
Palatal	/j/	[j]	[jà.ŋgà]	hawk
		[ɟ]	[ɟ̃à.má]	cloth
Velar	/k/	[k]	[kà.zà]	sun, day
	/g/	[g]	[gà.zò]	horn
	/ⁿg/	[ŋg]	[ŋgà.ɓà]	husband
Glottal	/h/	[h]	[hà.ɗù]	light

Table 3: Words containing /f/

Transcription	Gloss
[fú.ˈlá]	sacrifice
[fò.ndò]	plantain
[tú.ˈfá]	squeak (wheel) (v.)
[kù.hà] ~ [kù.fà]	hold

more typical of Sango speech than Vale speech, so it may be best to consider the sound a part of the Sango phonological system instead of Vale.

The phoneme /r/ is trilled [r] in word-initial position and following a consonant. In intervocalic position, it is usually produced with a single contact [ɾ], but I sometimes heard it trilled in that position, particularly when the word was emphasized. This is similar to what Ladefoged (1968: 30) found in Hausa, where a trilled [r] is often produced with a single contact in normal speech. Closer to the Vale region, the same process also occurs in the Nduga variety of Luto (Olson 2013).

The phoneme /r/ does not sound retroflexed to me. In addition, the Vale speakers noted that the sound is not like the Gbanu rhotic, which is usually produced as a retroflex [ɽ] (Murrell 2021).[6] The distribution of [r] and [ɾ] in Vale is formalized in (1).

(1)   Intervocalic tapping of /r/ (optional)
       /r/ → [ɾ] / V_V

The alveolar lateral /l/ sounds like a regular clear [l], with no velarization.

The palatal approximant /j/ is nasalized [j̃] when the following vowel is a nasal phoneme, as shown in (2). This process is similar to what Valenzuela et al. (2001: 283) found for Shipibo, although I did not observe a corresponding nasalization of /w/ as they did.

(2)   Nasalization of /j/
       /j/ → [j̃] / _Ṽ

The resulting allophone [j̃] sounds like a palatal nasal [ɲ], so much so that we initially considered transcribing it as the latter, particularly since cognates in Boyeldieu et al. (2006) are transcribed with ⟨ɲ⟩. However, the Vale speakers noted

---

[6]We studied Gbanu in the same workshop as Vale, so we were able to compare the two rhotics.

that there is no contact between the tongue and the palate in the articulation of the sound.

An alternative analysis could be considered, in which a phonemic nasal palatal approximant /ɟ̃/ contrasts with /j/. In this analysis, the nasalization of /ɟ̃/ would cause a following vowel /V/ to become nasalized [Ṽ]. However, I did not opt for this alternative analysis for several reasons.

First, I did not find a systematic spread of nasalization from nasal consonants to a following vowel.

Second, phonemic nasal vowels do exist in Vale, and they regularly occur following non-nasal consonants, as discussed in §3 below.

Third, native speaker intuitions strongly favor the chosen analysis. The Vale speakers easily recognized the distinction between oral and nasal vowel phonemes, while they were consistent in considering [j] and [ɟ̃] to comprise a single phoneme. Concomitant with this, they utilized a single symbol, ⟨y⟩, to represent both [j] and [ɟ̃] in the working orthography.

Labial-velar plosives are absent in Vale, apparently lost from Proto-SBB (Boyeldieu 2006: 8). These sounds merged with labial plosives in the closely-related language Luto (Olson 2021). It would be a reasonable hypothesis to suggest that this merger likely occurred in Vale as well.

3 Vowels

Vale has five oral vowel phonemes and five nasal vowel phonemes, as shown in Table 4. The inventory of oral vowels is reduced compared to what is typically found in Sara languages. More commonly, one finds seven oral vowels /i e ɛ a ɔ o u/, e.g. Bagiro (Boyeldieu 2000), or six oral vowels /i e a ɔ o u/, e.g. Mbay (Keegan 1997). A reduced central vowel [ə] is also common in Sara languages, e.g. Mbay (Keegan 1997) and Gula (Nougayrol 1999), but this is absent from Vale.

Contrast between the vowels is shown in Table 5. I illustrate contrast between oral vowels immediately following a word-initial /k/ and between nasal vowels following alveolar consonants.

Although rare, near minimal pairs do exist between oral and nasal vowels, as shown in Table 6.

According to my auditory impressions, the mid vowels /e/ and /o/ vary between open-mid and close-mid, without a clear environmental cause. That is, sometimes I heard the phonemes as [ɛ] and [ɔ] instead of [e] and [o], respectively. The International Phonetic Association (IPA 1949: 7) recommends employing the

---

[7]An additional example of [ĩ] is [ɟ̃ĩ.lí] 'melt, be wet (*fondre, être mouillé*)'.

Table 4: Vowel phonemes

	Front	Central	Back
Oral			
	i		u
	e		o
		a	
Nasal			
	ĩ		(ũ)
	ẽ		õ
		ã	

Table 5: Vowel contrasts

	Phoneme	Phone	Transcription	Gloss
Oral				
	/i/	[i]	[kì.lá]	tail
	/e/	[e]	[ké.ɾí]	firewood
	/a/	[a]	[kà.zà]	sun, day
	/o/	[o]	[kò.ɾò]	pain
	/u/	[u]	[kù.là]	work
Nasal				
	/ĩ/	[ĩ]	[tĩ̀ĩ̀][7]	scorpion
	/ẽ/	[ẽ]	[ndẽ́]	few
	/ã/	[ã]	[tã̀ã́]	caterpillar
	/õ/	[õ]	[zì.tṍ]	hill
	/ũ/	[ũ]	[sṹ.sṹ]	story, tale

Table 6: Contrast between oral and nasal vowels

Phoneme	Phone	Transcription	Gloss
/a/	[a]	[kà.jà]	heal, cure (v.)
/ã/	[ã]	[kà.j̃à]	respect (v.)

roman letters in such cases instead of the non-roman ones, particularly if the transcription is intended to be somewhat broad. I point out this apparent free variation for the sake of future research on Vale.

The close back nasal vowel /ũ/ is rare, occurring in only two words in our corpus: [sǘ.sǘ] 'story, tale' and [kã́ù̃] 'fly (insect)'. As a result, I consider it to be a marginal phoneme, and include it in parentheses in Table 4 above.

In addition to the nasal vowel phonemes, there is an optional non-phonemic "bleeding over" of nasalization from a nasal consonant or vowel onto a preceding oral vowel, as formalized in (3).

(3)   Leftward nasal spreading (optional)
      /V/ → [Ṽ] / _X̃

Examples are given in Table 7.

Table 7: Examples of leftward nasal spreading

		Transcription	Gloss
/kè.ᵐvè/	→	[kẽ̀.ɱvè]	find
/kã́ù̃/	→	[kã́ù̃]	fly (insect)

4 Syllable patterns

Four syllable patterns are attested in Vale, which are shown in Table 8.

Table 8: Syllable patterns

Syllable	Transcription	Gloss
V	[à.lí]	bird
CV	[ró]	body
CVV	[ɗóé]	fight
	[ɱvàí]	beard, mane
CVC	[kàm]	stone, pit

The V syllable pattern is only found in word-initial position. In the CVC syllable, the coda consonant is a sonorant,[8] and it does not bear tone.

---

[8]The implosive [ɗ] occurs in the coda in [gáɗ.gó] 'nape of neck', but comparative evidence suggests that that word is a compound (Boyeldieu et al. 2006).

The CVV syllable pattern requires some comment. The exhaustive list of attested vocalic sequences in these syllables is: [ii, ee, aa, oo, uu, ĩĩ, ẽẽ, ãã, õõ, ie, ia, io, ei, ai, au, oi, oe, ua, ĩã, ẽĩ, ãũ]. The question that arises is how to interpret these sequences (Pike 1947; Burquest 1998).

The sequences with two identical vocalic segments ([ii], [ee], etc.) could be interpreted as such, or as long vowels ([iː], [eː], etc.). In sequences containing one close (i.e. high) vocalic segment (e.g. [ia]), the close vocalic segment could be interpreted as a vowel or as an approximant (i.e. [ia] or [ja]).

There is one sequence in Vale, [oe], in which neither vocalic segment is close. Such sounds are generally considered to be "nonsuspicious" or "unambiguous", with an interpretation as an approximant highly unlikely (Pike 1947: 129; Burquest 1998: 156). In addition, the sequence [oe] is robust, occurring in several lexical items: [ɗóé] 'fight', [zòè], a grammatical particle, [kóꜜé ɓó.lé] 'hiss', [ɓòó dóꜜé] ~ [ɓòó déꜜé] 'under, below'. As a result, I do not consider it to be marginal to the phonological system.

The standard procedure is that suspicious sequences are "interpreted phonemically by analogy with the nonsuspicious ones" (Pike 1947: 128). This leads us to construe the close vocalic segments as vowels rather than approximants and the long vocalic segments as a sequence of two identical vowels.

The tonal behavior on these vocalic sequences harmonizes well with this interpretation. The long vowels can bear two tones, e.g. [ŋvàá] 'breast', and the close vowels in non-identical sequences bear tone as well, e.g. [ŋvàí] 'beard, mane'. From the perspective of moraic theory, this means that CVV syllables have two moras and the tone bearing unit (TBU) in Vale is the mora.

5 Tone

Vale has two unambiguous tone levels, which I label as High (H) and Low (L). Words demonstrating the contrast between these two levels are shown in Table 9.

Besides the three bisyllabic tone patterns shown in Table 9 (LL, HL, HH), words with a LH tone pattern are also attested in Vale, e.g. [pà.ɾí] [‿ ¯] 'bile'. Hence, all four of the tone patterns typically expected in a two-tone system occur in Vale.

However, there is a complication in the data that can be interpreted in two very different ways: an additional fifth tone pattern occurs on some bisyllabic words. This additional pattern is comprised of a "Superhigh" (S) tone on the first syllable followed by a High tone on the second syllable, e.g. [kő.né] [¯ ¯] 'year'. For now, I will call this fifth tone pattern "Superhigh-High" (SH). Sample contrasts between SH and HH are shown in Table 10.

Table 9: Tonal minimal pairs between H and L

Transcription	Pitch	Gloss
[vì.jà]	[_ _]	father
[ví.jà]	[‾ _]	pap, mushy food
[kó.lé]	[‾ ‾]	quarrel, argue
[kó.lè]	[‾ _]	luck

Table 10: Contrast between words with [‾ ‾] and [‾ ‾] surface tone

Transcription	Pitch	Gloss
[tó.lé]	[‾ ‾]	kill, extinguish
[tő.lé]	[‾ ‾]	carve
[kó.nő]	[‾ ‾]	thorn
[kő.né]	[‾ ‾]	year

One possible interpretation of these data is to consider that Vale has three tone levels, with a Superhigh tone occurring on the additional SH pattern. The key question that this interpretation raises is why there is only one bisyllabic pattern that involves the Superhigh tone (SH), while the other possible combinations are not attested (SS, SL, HS, LS).

Given this first interpretation of the data, we can posit an initial hypothesis for the tone melodies of Vale, shown in Table 11.

Table 11: Tone melodies (first draft)

Melody	Transcription	Pitch	Gloss
L	[vì.jà]	[_ _]	father
H	[kó.lé]	[‾ ‾]	quarrel, argue
LH	[pà.ɾí]	[_ ‾]	bile
HL	[ví.jà]	[‾ _]	pap, mushy food
SH	[kő.né]	[‾ ‾]	year

Another possible interpretation of these data—and the one we will ultimately adopt—is to consider an analysis that involves non-automatic downstep (Yip

2002: 147–150). This would not be the first case of downstep in SBB, as Boyel-dieu (2000) posited downstep in the related language Bagiro.

In this view, the "Superhigh" tone is actually a H tone raised via the phonetic process H-Raising (Connell & Ladd 1990), which raises the pitch of a H tone immediately preceding a L tone. The non-automatic downstep results from the presence of a floating L tone.[9] I will employ autosegmental phonology (Gold-smith 1976) later in this section to illustrate this.

There is good evidence that downstep occurs in Vale, but elucidating it will require an extended example. Spoken in isolation, the words [j̃ɛ́.lè] 'wind' and [kú.lí] 'cool' have HL [‾ _] and HH [‾ ‾] tone patterns, respectively. However, when they are put together in the noun phrase 'cool wind', the tone pattern becomes [‾ ‾ ‾ ‾]. The accompanying phonetic transcription is: [j̃ɛ́.ˈlé kú.lí].

The L on the second syllable of 'wind' has become floating. This floating L triggers a lowering of register which propagates through to the rest of the phrase. This register lowering is typical of languages with downstep.

The pronunciation of the words [j̃ɛ́.lè] 'wind' and [kú.lí] 'cool' in isolation (i.e. surface representations, SR) suggests the underlying representations (UR) for the words shown in Table 12.

Table 12: Words for 'wind' and 'cool' in Vale

UR	SR	Gloss
/jẽ.le, HL/	[j̃ɛ́.lè]	wind
/ku.li, H/	[kú.lí]	cool

The autosegmental notation for the surface forms of the two words in isolation is shown in (4).

(4) Surface representations (SR) of the individual words 'wind' and 'cool' in autosegmental notation

```
 H  L             H
 |  |             |\
j ẽ. l e         ku. l i
```

As mentioned above, when these two words are put together in a noun phrase, one gets [j̃ɛ́.ˈlé kú.lí] 'cool wind', where a downstep occurs after the first syllable.

---

[9]Odden (1982) provides evidence that in some cases downstep is the result of two underlying H tones.

In autosegmental terms, the L tone on the second syllable of 'wind' is delinked, which allows the H in [kú.lí] to spread leftward onto that syllable. The L is not deleted, but rather remains as a floating tone which triggers the downstep.

An autosegmental rendering of the underlying representation for this phrase is shown in (5).

(5)   Underlying representation (UR) of 'cool wind'
      H L   H

      j ẽ. l e   ku. l i

One possible derivation would then proceed as shown in (6).

(6)   Autosegmental derivation of 'cool wind'

UAC	Spreading	HLH Plateauing
H L H	H L H	H L H
┊ ┊ ┊	│ │ │﹨	│ ‡ ⟋⟋ │﹨
j ẽ. l e ku. l i	j ẽ. l e ku. l i	j ẽ. l e ku. l i

This derivation proceeds as follows. First, the *Universal Association Convention* (UAC) shown in (7) applies, linking the three tones to the first three tone bearing units (TBUs) of the phrase.

(7)   Universal Association Convention (UAC)
      "Associate tones with tone bearing units, one-to-one, left to right."
      (Kenstowicz 1994: 317)

Second, a *Spreading* rule applies, linking the final TBU to the final tone. This is as formalized in (8).

(8)   Spreading
      T
      │﹨
      V V

Finally, the rule *HLH Plateauing* shown in (9) applies, delinking the L tone and spreading the following H tone leftward to the now unassociated TBU.

(9)   HLH Plateauing
      "[A] singly linked L between two H tones delinks, and the second H
      spreads left and has a downstepped register."
      (Cahill 2019: 115; cf. Roberts et al. 2016)

An autosegmental rendering of the surface representation is shown in (10). Note that the L tone is not linked to a TBU, but it nevertheless remains in the representation. The theoretical claim is that this configuration triggers downstep.

(10) Surface representation (SR) of 'cool wind'

```
      H   L     H
      |    ╱   |╲
   j ẽ. l e   ku. l i
```
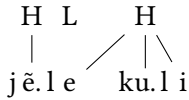
If we extend this analysis to monomorphemic forms, the downstep attested on monomorphemic lexical items could also be attributed to a floating L tone. In this view, a word like [kó.ˈné] 'year' would have an underlying /HLH/ tonal pattern, i.e. /ko.ne, HLH/. An autosegmental derivation would proceed as shown in (11).

(11) Autosegmental derivation of 'year'

UAC   Spreading   HLH Plateauing

```
   H  L  H                             H  L  H
   ┊  ┊                                |  ǂ ╱
   ko.n e              –               ko.n e
```

In this derivation, the UAC once again applies first, linking the first two tones to the two TBUs, and leaving the extra tone unassociated. Since there are no extra TBUs, *Spreading* does not apply. Finally, *HLH Plateauing* applies, delinking the L tone and linking the final H with the delinked TBU.

If we accept this view, Vale has five underlying tone melodies: L, H, LH, HL, and HLH. These are exemplified in Table 13.

Table 13: Tone melodies (revised)

Melody	UR	SR	Gloss
L	/vi.ja, L/	[vì.jà]	father
H	/ko.le, H/	[kó.lé]	quarrel, argue
LH	/pa.ri, LH/	[pà.rí]	bile
HL	/vi.ja, HL/	[ví.jà]	pap, mushy food
HLH	/ko.ne, HLH/	[kó.ˈné]	year

Tone melody inventories like this are not unheard of. Siane (Papua New Guinea) has an inventory of the same five tone melodies as Vale: L, H, LH, HL, and HLH

(James 1994, Kenstowicz 1994: 385). In Africa, Mende (Sierra Leone) also has five tone melodies, with a LHL three-tone melody instead of HLH: L, H, LH, HL, LHL (Leben 1973: 64).

A couple of additional examples of the downstep process in Vale are as follows. When we combine /zo, L/ 'head', /ke.te, HL/ 'before', and /ba.ŋga, HL/ 'sweet potato', one gets [zò ké.ꜜté bá.ŋgà] 'first sweet potato'. When /ko.jo, HL/ 'mother' and /se.ge, HLH/ 2PL.POSS are combined, one gets *two* downsteps [kó.ꜜjó sé.ꜜgé] 'your (pl.) mother'.

While the analysis of Vale presented here appears to be motivated, it also has the surprising result that the language has an underlying form /HLH/ that is dispreferred on the surface (Cahill 2007) and hence is not realized as such.

6 Discussion

Vale is geographically close to the Sara group of SBB. In addition, typological features provide some support for considering Vale to be a part of the Sara group: The language exhibits no evidence of ATR vowel harmony, it has a robust set of phonemic nasal vowels, and voiceless fricatives are limited to /s/ and /h/ (Dimmendaal et al. 2019: 347–348).

However, there is one important exception: Vale has two phonemic level tones (H and L), whereas most Sara languages have 3-tone systems (Dimmendaal et al. 2019: 349), including the nearby varieties of Luto (Olson 2021). More research needs to be done to situate Vale within the SBB group, but placing it within Sara would seem to be a reasonable first hypothesis.

Boyeldieu (2006: 10–11) reconstructed a two-tone system for Proto-SBB and posited six bisyllabic nominal proto tone patterns: *LL, *HH, *LH, *HL(a), *HL(b), and *HL(c).[10] Four of the tone melodies in Vale correspond directly to the first four patterns posited by Boyeldieu. Table 14 shows the tone correspondences, and Table 15 provides sample data supporting the correspondences. I've included data from the Nduga variety of Luto, which is geographically adjacent to Vale. The Proto-SBB and Nduga data—as well as the reference numbers—are taken from Boyeldieu et al. (2006). (The Nduga data were collected by Pierre Nougayrol.)

What remains to be determined is whether the underlying HLH melody in Vale (realized as HꜜH) corresponds to either of the additional two Proto-SBB tone patterns, *HL(b) or *HL(c). The data at the bottom of Table 15 show that cognates do exist, but the proto forms for the tones are unclear. Resolving this is left for further research.

---

[10]Boyeldieu employed numbers in his labeling: *11, *22, *12, *21a, *21b, and *21c.

Table 14: Bisyllabic nominal tone correspondences between Proto-SBB, Vale, and Nduga

Proto-SBB	Vale	Nduga
*LL	LL	ML
*HH	HH	MM
*LH	LH	LL
*HL(a)	HL	HM
—	HꜜH	HM

The main topic for further research in this study of Vale is word patterns. I was able to identify monomorphemic words in Vale with one or two syllables without much difficulty, but when I looked at forms that appeared to be longer than two syllables, further investigation indicated that they were likely multimorphemic.[11] As a result, I was unable to determine how the five tone melodies were realized on monomorphemic words of three or more syllables.

When I discussed the HLH melody with other linguists, the first question that arose was, "Are you sure that these bisyllabic words aren't compounds?" Snider (2018: 33–37) mentions this caveat explicitly and offers two diagnostics for identifying compounds: (1) Is the tonal pattern rare compared to the other ones? and (2) Do native speakers recognize the words as compounds?

Concerning the first diagnostic, words with a HLH pattern appear to be relatively common in Vale. Concerning the second diagnostic, sometimes the speakers could identify compounds, but not always.

One additional check I did was to compare the Vale words containing a HLH pattern with Proto-SBB forms (cf. the bottom of Table 15). In all cases where there was a resemblance between these two forms, the Proto-SBB form was monomorphemic. This increased my confidence that we were actually dealing with monomorphemic words.

Finally, I compared the phonological description of "Vale de Ndélé" in Yanguende (2011) with my own research. A large number of the lexical items in Yanguende's *mémoire* resemble those from my own data, so it's clear that the language is a Sara variety. However, Yanguende posits seven oral vowels /i e ɛ a ɔ o u/, five nasal vowels /ĩ ɛ̃ ã ɔ̃ ũ/, and three level tones /H M L/. All three of

---

[11]An anonymous reviewer pointed out that languages in the region often have a "disyllabic maximum" on words.

Table 15: Sample bisyllabic nominal correspondences between Proto-SBB, Vale, and Nduga

SBB segments	Tones	Vale	Nduga	Gloss	No.
*talɔ	*LL	tà.lò	tālù	dew	N/040
*mama	*LL	mà.mà	māmà	python	N/013
*kARKɛ; *kiRKɔ; *kVRKV	*LL	kò.kè	kākè	iron, hoe	N/029
*kaga	*HH	ká.gá	kāgā	wood, tree	N/199
*kumu	*HH	kú.mú	kūmū	navel	N/200
*ngeri; *Keri	*HH	ké.rí	kīrī	firewood	N/206
*baɗu	*LH	và.ɗú	vàɗù	warthog	N/094
*Suɓu	*LH	kù.ɓú	kùɓù	oil	N/107
*Sali, *Salu, *SOlI	*LH	à.lí	yìlì	bird	N/093
*kORngO	*HL(a)	kó.ŋgò	kóngō	cliff, hill	N/166
*bisi; *ɓisi; *Cisi	*HL(a)	bí.sì	bísī, bísì	dog	N/160
*ngulu; *ngulɛ; *nguɭu; ... *nguɭɛ	*HL(a)?	ŋgú.lù	ngúlū	yam	N/671
*putu; *piti	—	hú.ꜜtí	fítī	flower	N/767
*Cuɓa; *Nuɓa; *ɭuɓa	—	nú.ꜜɓá	núɓā	God	N/264
*ɓɔlɔ	—	ɓó.ꜜló	ɓólĪ	be bitter	V/370
*aɗa	—	tá.ꜜzó	tāz	count (v.)	V/194
*igbo; *OgbO; *nɗigbo; ... *nɗOgbO	—	tí.ꜜbó	īgbò	wash (v.)	V/142
*ɗiyo; *ɗVyV; *ɗV; *Ciyo; ... *CVyV; *CV	—	zíꜜó	zyó	two	X/005
*ndOmO; *ndumi; *ndimi	—	ndó.ꜜmó	—	forget (v.)	V/296

these features are characteristics of Luto (Olson 2021) rather than Vale. Two varieties of Luto—Ndoka and Wad—are spoken in and around Ndélé, while the Vale region is 300 kilometers west of Ndélé. This suggests to me that the speech variety Yanguende studied is likely Ndoka or Wad. It is common for Central Africans to migrate for work, education, marriage, or other reasons, so it's possible that there is a community near Ndélé that self-identifies as Vale but has adopted the local langauge.

7 Conclusion

This paper makes a significant contribution to our understanding of a language that has received very little attention in the literature. I've provided information about Vale consonants, vowels, syllable structure, and tone.

At the same time, it is a modest first step in the large task of documenting a language. There are topics of further research that could prove fruitful. These include: (1) confirming Vale's status as a Sara language, (2) clarifying the morphemic structure of multisyllabic forms, which could inform, in particular, the tone analysis, (3) examining tonal behavior in the tense-aspect-mood system, and (4) comparing the extant Vale forms with Proto-SBB to understand the historical development of the language, including identifying the process that led to the loss of labial-velar consonants, as well as determining the source of the HLH tone pattern.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

H	High tone	SBB	Sara-Bongo-Bagirmi
L	Low tone	SR	Surface representation
M	Mid tone	UAC	Universal Association
n.	Noun		Convention
No.	Number	UR	Underlying representation
S	Superhigh tone	v.	Verb

Acknowledgements

Kenneth S. Olson

References

Boyeldieu, Pascal. 2000. *La langue bagiro (République Centrafricaine): Systématique, textes et lexique* (Research in African Studies 4). Frankfurt: Peter Lang.

Boyeldieu, Pascal. 2006. *Présentation des langues Sara-Bongo-Baguirmiennes.* Paris: CNRS-LLACAN. https://llacan.cnrs.fr/SBB/.

Boyeldieu, Pascal, Pierre Nougayrol & Pierre Palayer. 2006. *Lexique comparatif historique des langues Sara-Bongo-Baguirmiennes.* Paris: LLACAN-CNRS. https://llacan.cnrs.fr/SBB/.

Bruel, Georges. 1905. *Le cercle du Moyen-Logone.* Paris: Comité de l'Afrique Française. DOI: 10.5479/sil.626490.39088010252781.

Burquest, Donald A. 1998. *Phonological analysis: A functional approach.* 2nd edn. Dallas: Summer Institute of Linguistics.

Cahill, Michael. 2007. More universals of tone. *SIL Electronic Working Papers* 7. https://www.sil.org/resources/archives/7816.

Cahill, Michael. 2019. Tone, orthographies, and phonological depth in African languages. In Samson Lotven, Silvina Bongiovanni, Phillip Weirich, Robert Botne & Samuel Gyasi Obeng (eds.), *African linguistics across the disciplines: Selected papers from the 48th Annual Conference on African Linguistics*, 103–123. Berlin: Language Science Press. DOI: 10.5281/zenodo.3520575.

Connell, Bruce & D. Robert Ladd. 1990. Aspects of pitch realisation in Yoruba. *Phonology* 7. 1–29. DOI: 10.1017/S095267570000110X.

Dimmendaal, Gerrit J., Colleen Ahland, Angelika Jakobi & Constance Kutsch Lojenga. 2019. Linguistic features and typologies in languages commonly referred to as "Nilo-Saharan". In H. Ekkehard Wolff (ed.), *The Cambridge handbook of African Linguistics. Part II: Comparative and descriptive African Linguistics*, 326–381. Cambridge: Cambridge University Press. DOI: 10.1017/9781108283991.011.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2023. *Ethnologue: Languages of the world.* 26th edn. Dallas: SIL International. www.ethnologue.com.

Gaudefroy-Demombynes, Maurice. 1906. *Documents sur les langues de l'Oubangui-Chari: (Extrait du tome II des Actes du XIV Congrès International des Orientalistes).* Paris: Ernest Leroux. http://hdl.handle.net/2027/wu.89085118743.

Goldsmith, John A. 1976. *Autosegmental phonology.* Cambridge, MA: MIT. (Doctoral dissertation). http://dspace.mit.edu/handle/1721.1/16388.

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2022. *Glottolog 4.7*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://glottolog.org.

IPA. 1949. *The principles of the International Phonetic Association*. London: International Phonetic Association. DOI: 10.1017/S0025100311000089.

IPA. 2006. IPA news. *Journal of the International Phonetic Association* 36(1). 133–135. DOI: 10.1017/S0025100306002465.

James, Dorothy J. 1994. Word tone in a Papuan language: An autosegmental solution. *Language and Linguistics in Melanesia* 25(2). 125–148. https://www.langlxmelanesia.com/llm-archive.

Keegan, John M. 1997. *A reference grammar of Mbay* (LINCOM Studies in African Linguistics 14). München: Lincom.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Oxford: Blackwell.

Kutsch Lojenga, Constance. 1996. Participatory research in linguistics. *Notes on Linguistics* 73. 13–27. https://www.sil.org/resources/archives/4087.

Ladefoged, Peter. 1968. *A phonetic study of West African languages*. 2nd edn. (West African language monograph series). London: Cambridge University Press.

Leben, William R. 1973. *Suprasegmental phonology*. Cambridge, MA: MIT. (Doctoral dissertation). https://dspace.mit.edu/handle/1721.1/16364.

Moéhama, Elisée. 2021. Rapport d'enquête sociolinguistique: Première évaluation parmi les Vale. *Journal of Language Survey Report* 2021-025. https://www.sil.org/resources/publications/entry/88747.

Murrell, Paul. 2021. *An overview of Gbanu phonology*. Tech. rep. Bangui: SIL International.

Nougayrol, Pierre. 1999. *Les parlers gula: Centrafrique, Soudan, Tchad: Grammaire et lexique*. Paris: CNRS.

Odden, David. 1982. Tonal phenomena in KiShambaa. *Studies in African Linguistics* 13(2). 177–208. DOI: 10.32473/sal.v13i2.

Olson, Kenneth. 2013. *A sketch of Lutos phonology*. Unpublished ms. http://www.sil.org/resources/archives/60749.

Olson, Kenneth S. 2021. Labial-velar to labial sound changes in Luto: Paper presented at the Annual Conference on African Linguistics 51–52, 8–10 April 2021. University of Florida, Gainesville, FL.

Olson, Kenneth S. & John Hajek. 2003. Crosslinguistic insights on the labial flap. *Linguistic Typology* 7(2). 1–30. DOI: 10.1515/lity.2003.014.

Pike, Kenneth L. 1947. *Phonemics: A technique for reducing languages to writing*. Ann Arbor, MI: The University of Michigan Press.

Roberts, David, Stephen L. Walter & Keith Snider. 2016. Neither deep nor shallow: A classroom experiment testing the orthographic depth of tone marking in Kabiye (Togo). *Language and Speech* 59(1). 113–138. DOI: 10 . 1177 / 0023830915580387.

Samarin, William J. 1971. Adamawa-Eastern. In Thomas A. Sebeok (ed.), *Linguistics in Sub-Saharan Africa* (Current Trends in Linguistics 7), 213–244. The Hague: Mouton De Gruyter. DOI: 10.1515/9783111562520-009.

Snider, Keith. 2018. *Tone analysis for field linguists*. Dallas: SIL International.

Tucker, Archibald Norman. 1940. *The Eastern Sudanic languages*, vol. 1. London: Oxford University Press (for the International Institute of African Languages & Cultures).

Tucker, Archibald Norman & Margaret Arminal Bryan. 1956. *The Non-Bantu languages of North-Eastern Africa (with a supplement on the Non-Bantu languages of Southern Africa, by E. O. J. Westphal). Handbook of African Languages, Part III*. London, New York, Cape Town: Oxford University Press (for the International African Institute).

Valenzuela, Pilar M., Luis Márquez Pinedo (Yoi Sani) & Ian Maddieson. 2001. Shipibo. *Journal of the International Phonetic Association* 31(2). 281–285. DOI: 10.1017/S0025100301002109.

Yanguende, Simplice. 2011. *Phonologie et morphologie des nominaux du vale de Ndélé en Centrafrique*. Bangui: Université de Bangui. (Mémoire de Master 2 en linguistique, Filière Professionnelle). https://www.sil.org/resources/archives/93213.

Yip, Moira. 2002. *Tone* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139164559.

Chapter 18

Verb-marked reciprocals in Wolof

Sofiya Ros
Utrecht University

Giada Palmieri
Utrecht University

This paper provides a description of reciprocal morphemes in Wolof. We present three verbal affixes associated with reciprocal interpretations, and we propose that they reflect different strategies: *-ante* is a valence-reducing morpheme that turns transitive verbs into reciprocal verbs; *-e* and *-oo* mark predicates with an inherent reciprocal meaning and do not operate on the verbs' argument structure.

1 Introduction

Wolof is a Niger-Congo Atlantic language spoken in Senegal, Gambia and Mauritania. It is an agglutinative language with a rich verbal and nominal morphology (Ka 1982, Buell & Sy 2005). Verbal derivations use distinct suffixes which may attach to a verb root and permit alterations to the category, valence and semantics of a verbal base. Valence-changing suffixes may derive structures containing new arguments with different thematic roles. In (1b) the applicative suffix *-al* changes the argument structure of the verb *togg* 'to cook' in (1a), adding an argument with the semantic role of a beneficiary. In (1c) the applicative suffix *-e* adds an argument with the semantic role of an instrument.

(1)   a. Ñu ngi   togg jën.
          PRST.3PL cook fish
          'They cook fish.'

b. Ñu ngi   togg-al   Khady jën.
    PRST.3PL cook-BEN Khady fish
    'They cook fish for Khady.'

c. Ñu ngi   togg-e   jën wi ak   ndox.
    PRST.3PL cook-INS fish the with water
    'They cook fish with water.'

Reciprocity in Wolof is also derived with the use of verbal suffixes that attach to a verbal base. In the literature, three different verbal suffixes are described as expressing reciprocity: *-ante*, *-e* and *-oo* (Church 1981, Voisin 2002, Creissels & Nouguier-Voisin 2008, Diouf 2009, Ka 1981, 1982). In (2), the suffix *-ante* is added to the verbal base *bëgg* 'to love' to make each of the individuals in the denotation of the subject (Khady and Fatou) occupy both the role of agent and patient. Similarly, in (3) and (4) the affixes *-e* and *-oo* are added to the verbal bases *gis* 'to see' and *xul* 'to argue' respectively, leading to reciprocal configurations.

(2) Khady ak   Fatou dañu   bëgg-ante.
     Khady and Fatou FOC.V.3PL love-RECP
     'Khady and Fatou love each other.'

(3) Khady ak   Fatou dañu   gis-e.
     Khady and Fatou FOC.V.3PL see-RECP
     'Khady and Fatou met.'

(4) Khady ak   Fatou dañu   xul-oo.
     Khady and Fatou FOC.V.3PL argue-RECP
     'Khady and Fatou argue with each other.'

Verbal morphology in Wolof has been central to some extensive studies (Voisin 2002, Church 1981), but little attention has been dedicated to reciprocal affixes specifically (Creissels & Nouguier-Voisin 2008). A crucial remaining question is whether reciprocal morphemes in Wolof operate on the argument structure.

The goal of this paper is to fill this gap, providing an overview of Wolof reciprocal morphology. We rely on novel data and we propose a different treatment for the morpheme *-ante* as opposed to *-e/-oo*: we propose that *-ante* is a productive reciprocal morpheme operating on the argument structure of the verbs it combines with, while *-e* and *-oo* are lexicalized markers of verbs that have an inherent reciprocal meaning.

The paper is structured as follows. In §2, we first lay down some terminology, introducing two strategies that express reciprocity cross-linguistically (§2.1) and

we then review previous works on reciprocity in Wolof (§2.2). In §3 we present our proposal, and we empirically support it by illustrating the distributional and morphological properties of Wolof reciprocal morphemes. Finally, in §4 we discuss the significance of our findings and we draw some general conclusions.

2 Previous studies

2.1 Lexical vs grammatical reciprocity

Cross-linguistically, two different strategies are associated with reciprocal interpretations: grammatical reciprocity and lexical reciprocity (Haspelmath 2007).

  Grammatical reciprocity is a productive strategy, where reciprocity is expressed by an element – like a pronoun or a derivational morpheme. For example, in English and Russian (Indo-European), grammatical reciprocity is expressed by the pronouns *each other* (5) and *drug druga* (6), respectively. Such pronouns might occupy the object position and can be used to express reciprocity with virtually all transitive verbs which may have no reciprocal meanings on their own. In Chichewa (Bantu) grammatical reciprocity is expressed by the derivational morpheme *-an-*. Within the VP, this affix does not occupy the object slot, but that of derivational affixes, after the verb stem (7).

(5)   Kim and Alex praised each other.

(6)   Devuški   blagodaryat drug druga.
      girls.PL.NOM thank   RECP
      'The girls thank each other.'

(7)   Galu ndi mwana a-na-lum-an-a.
      dog   and child   3PL-PST-bite-RECP-FV
      'The dog and the child bit each other.'
      (Mchombo & Ngalande 1980:570-571)

  Lexical reciprocity refers to the strategy by which reciprocity is expressed by the inherent meaning of a closed class of verbs. This strategy is not productive, but restricted to verbs denoting "naturally reciprocal" events, defined by Kemmer (1993) as "events that are either necessarily (e.g. 'meet') or else very frequently (e.g. 'fight', 'kiss') semantically reciprocal" (p.102).

  In some languages, lexical reciprocity is expressed without any grammatical marking. For instance, in English some predicates express reciprocity in their

intransitive entry (8); lexical reciprocity may be found in predicates with a transitive alternate (8a), as well as in predicates without a direct object (8b). In other languages, lexical reciprocals require non-productive markers, such as *-sya* in Russian (9). Lexical reciprocal verbs may undergo 'semantic drift': they may have a reciprocal meaning that is different from the meaning of the verbal base. For example, in (9b) the lexical reciprocal formed with a verbal base *drat* 'to tear' gets the new meaning 'fight'.

(8) a. Kim and Alex hugged.

    b. Kim and Alex argued.

(9) a. Alex i   Anna celuut-sya.
       Alex and Anna kiss-RECP
       'Alex and Anna kiss.'

    b. Alex i   Max derut-sya.
       Alex and Max tear-RECP
       'Alex and Max fight.'

2.2 Wolof reciprocal markers

In the literature, it is argued that Wolof exhibits three verbal suffixes whose uses include reciprocity: *-ante*, *-oo*, *-e*.

In Voisin 2002 and Creissels & Nouguier-Voisin 2008 the authors describe five distinct suffixes that are combined in a group of so-called 'co-participative' suffixes: *-ante*, *-oo*, *-e*, *-andoo* and *-aale*. In this paper we address only three of them, because *-andoo* and *-aale* are not strictly associated with reciprocal meanings. The affix *-aale* expresses a relation of simultaneity between the event represented by the verb and another event ('at the same time') or, in its lexicalized use, it carries a meaning of co-participation (10). The suffix *-andoo* implies a plurality of participants involved in the same event with the same role (11).

(10) *nekk* 'be somewhere' > *nekk-aale* 'live together'
     (Creissels & Nouguier-Voisin 2008:302)

(11) Ñoom ñaar ñepp toog-andoo ci   lal bi.
     3PL   TWO all   sit-COPART LOC bed DEF
     'They both sat on the bed together.'
     (Creissels & Nouguier-Voisin 2008:303)

Moving on to reciprocity, the authors describe *-e* as a reciprocal marker of a limited class of verbs, namely verbs denoting 'naturally reciprocal events' (12). Such events are defined as "two participant events in which the exchange of roles is not absolutely obligatory, but nevertheless constitutes the normal situation" (Creissels & Nouguier-Voisin 2008 :298). The morpheme *-ante* is presented as the most productive reciprocal marker, which, unlike *-e*, is not restricted to specific classes of predicates (13). Finally, *-oo* is described as a marker of sociative and reciprocal events (14).[1]

Diouf (2009) provides a list of Wolof verbal morphemes, among which *-ante* is described as being a reciprocal affix. The author also lists seven different *-e* suffixes with different functions; however, bearing in mind the polysemy of this marker, in the current paper we solely focus on its reciprocal use. Ka (1982) provides an extensive overview of Wolof verbal affixes, also reporting *-ante* and *-oo* to have a reciprocal meaning.

Church (1981) points out some semantic differences between the morphemes *-ante* and *-oo*, providing examples where the same verb root leads to different reciprocal interpretations with different suffixes, as illustrated in (15) and (16).[2] However, it is unclear whether these contrasts are systematic and whether they also extend to the affix *-e*.

(12) Ñu ngi doon xeex-e   ci koñu   kër ñoom Paate.
     PRST.3PL PST fight-RECP LOC street.CONN house 3PL Pathé
     'They were fighting on the Pathe street.'
     (Voisin 2002:345)

(13) Rey-ante nañu.
     kill-RECP PRF.SBJ.3PL
     'They killed one another.'
     (Creissels & Nouguier-Voisin 2008:304)

(14) Seen   wax yi wor-oo   nañu.
     POSS.2PL word DEF betray-RECP PRF.SBJ.3PL
     'Your declarations are contradictory.' (lit. 'betray one another')
     (Creissels & Nouguier-Voisin 2008:303)

---

[1] Some authors consider *-oo* to be a bimorphemic affix. A plausible origin of reciprocal *-oo* in Wolof is the combination of the middle marker *-u* with the ancient marker of co-participation *-e* (Creissels & Nouguier-Voisin 2008:303, Diouf 2009:57).

[2] Our own translation from French.

(15) *dog* 'to cut'

    a. *dog-ante* 'to cut each other'

    b. *dog-oo* 'to separate from each other'

    (Church 1981:183)

(16) *dég* 'to hear'

    a. *dég-ante* 'to hear from each other'

    b. *dég-oo* 'to get along'

    (Church 1981:175)

The works presented here provide a clear picture of the existence of three different reciprocal affixes in Wolof, but the difference between them is still unclear: the properties and distribution of these morphemes are not explicitly described in the literature. The claim that *-e* marks 'naturally reciprocal' events is based on the meanings traditionally associated with predicates of this class, but it is not independently supported by properties specific to this morpheme. Very little attention is paid to the constraints on the use of each affix. Moreover, it is not clear what the role of the suffix *-oo* is and how it differs from *-ante* and *-e*.

Based on the existing literature, it is not possible to determine the differences in distribution of these three reciprocal morphemes, nor what morphological processes they reflect; the aim of the current paper is to fill this gap.

3 Wolof reciprocal markers: new evidence

In order to shed light on the distributional and morphological properties of reciprocal morphemes in Wolof, we conducted a series of interviews with two native speakers. One of the speakers lives in Senegal and the interviews took place online. The second speaker lives in the Netherlands, but uses Wolof daily and often travels to Senegal. The interviews with the second speaker took place in person in the Netherlands. Wolof is the first language for both informants and both speak urban Dakar Wolof (McLaughlin 2001). Our interviews included translation tasks and grammaticality judgment tasks. We asked each informant to translate sentences from French or Dutch. We also presented Wolof sentences to participants, inviting them to rate them as grammatically acceptable or unacceptable. The judgments and the translations were consistent for both speakers.

On the basis of the data collected, we propose that the different Wolof verbal affixes reflect different morphological processes. We argue that *-ante* is a

productive reciprocal morpheme, expressing *grammatical reciprocity*. We propose that this morpheme operates on the argument structure of verbs, reducing the valency: it makes the predicate intransitive, removing the object position and leading to a reciprocal interpretation. Treatments of reciprocal affixes as intransivizing morphemes have been proposed for a number of languages where reciprocity is expressed by affixes that attach to transitive verb bases, including Chichewa (Bantu; Dalrymple et al. 1994), Malagasy (Austronesian; Keenan & Razafimamonjy 2004), and Passamaquoddy (Algonquian; Bruening 2006). This is illustrated in (17) with a Malagasy example: a transitive verb (17a) can be affixed with the reciprocal morpheme *-if-* (17b), leading to what is proposed to be an intransitive reciprocal verb (Keenan & Razafimamonjy 2004). We propose that the same pattern can be identified in Wolof, where the morpheme *-ante* may productively turn any transitive verb into a reciprocal verb.

(17) a. M-an-enjika   an-dRabe Rakoto.
          PRS-ACT-chase ACC-Rabe Rakoto
          'Rakoto is chasing Rabe.'

      b. M-if-an-enjinka   Rabe sy   Rakoto.
          PRS-RECP-ACT-chase Rabe and Rakoto
          'Rabe and Rakoto are chasing each other.'
          (Keenan & Razafimamonjy 2004:177)

By contrast, we propose that the Wolof suffixes *-e* and *-oo* are associated with *lexical reciprocity*. We argue that they are markers on predicates that are lexicalized as reciprocals. The distinction that we are proposing is based on different distributional and morphological properties of these three morphemes that we will present in the rest of this section.

3.1 Productivity

A first contrast in the distribution of Wolof reciprocal morphemes regards their productivity. The suffix *-ante* is productive and it leads to reciprocal interpretations with virtually any transitive verb. The morphemes *-e/-oo*, on the other hand, have a restricted use. Let us illustrate this contrast with the verb *bañ* 'to hate'. This predicate takes a direct object, as shown in example (18). The suffix *-ante* can attach to this predicate, leading to a reciprocal interpretation: in (19a) the object position is removed and the individuals in the subject are interpreted as both agent and patient of the action denoted by the verb. In other words, Khady hates Idrissa and Idrissa hates Khady. By contrast, the suffixes *-e* and *-oo* lead to

ungrammaticality if applied to the transitive verb base (19b). The same holds for other transitive verbs, for example: *rey* 'to kill' (20); *dimbali* 'to help'; *nuyu* 'to greet'; *lekk* 'to eat'; *bëgg* 'to love'; *ragal* 'to be scared of', and *fóon* 'to kiss'.

(18) Khady bañ na   Idrissa.
      Khady hate PRF.3SG Idrissa
      'Khady hates Idrissa.'

(19) a. Khady ak   Idrissa dañu   bañ-ante.
          Khady and Idrissa FOC.V.3PL hate-RECP
          'Khady and Idrissa hate each other.'

      b. *Khady ak   Idrissa dañu   bañ-e/ bañ-oo.
          Khady   and Idrissa FOC.V.3PL hate-RECP

(20) a. Khady rey na   muus mi.
          Khady kill PRF.3SG cat   the
          'Khady killed the cat.'

      b. Idrissa ak   Oumar rey-ante nañu.
          Idrissa and Oumar kill-RECP PRF.SBJ.3PL
          'Idrissa and Oumar killed one each other.'

      c. *Idrissa ak   Oumar rey-oo/ rey-e nañu.
          Idrissa   and Oumar kill-RECP   PRF.SBJ.3PL

This suggests that *-e* and *-oo* do not freely operate on just any verb, but they are instead markers of a closed class of transitive predicates. As pointed out in the literature (Voisin 2002, Creissels & Nouguier-Voisin 2008), predicates marked by *-e* generally denote 'naturally reciprocal' events, in the terminology by Kemmer (1993). We note that this observation also extends to *-oo*: some representative examples include *xuloo* 'to argue', *taggoo* 'to say goodbye', and *booloo* 'to unite'.

The restricted use of *-e* and *-oo*, limited to verbs that are typically associated with reciprocal configurations, lends support to a treatment of these affixes as markers of lexical reciprocal entries: we propose that they do not lead to reciprocal interpretations themselves, but mark predicates with an inherent reciprocal meaning. We propose that *-ante*, on the other hand, is directly responsible for reciprocal interpretations, turning transitive verbs into reciprocal verbs.

3.2 Intransitive verbs

Another piece of evidence comes from the fact that *-e* and *-oo*, unlike *-ante*, can mark predicates without a transitive entry. Let us provide an example from the

verb *dëkk* 'to live'. As illustrated in (21), this predicate is intransitive: it cannot take a direct object, but only a PP. This intransitive verb can be marked by *-e* or *-oo*, denoting a reciprocal relation between the individuals in the subject (22a). However, the same verb leads to ungrammaticality with *-ante*, as shown in (22b). The same pattern can be identified with other intransitive verbs, for example *fécca* 'to dance', *mer* 'to be angry' or *xul* 'to argue'. They cannot take a direct object and in order to denote reciprocity they are marked with *-e* or *-oo*, while they lead to ungrammaticality with *-ante*.

(21) a. Khady mu ngi   dëkk moom   rekk.
          Khady PRST.3SG live   PRO.3SG alone
          'Khady lives alone.'

      b. Khady mu ngi   dëkk ak/si   Fatou.
          Khady PRST.3SG live   with/on Fatou
          'Khady lives with Fatou.'

(22) a. Khady ak   Fatou ñoo   dëkk-oo/dëkk-e.
          Khady and Fatou FOC.V.3PL live-RECP
          'Khady and Fatou are neighbours.'

      b. *Khady ak   Fatou ñoo   dëkk-ante.
          Khady   and Fatou FOC.V.3PL live-RECP

In (22a) there is no direct object to be removed, therefore we can propose that *-e* and *-oo* do not operate on the argument structure the same way that *-ante* does. Moreover, these examples show that *-ante* is not productive with intransitive verbs.

3.3 Semantic drift

Verbs that combine with the affixes *-e* or *-oo* may get interpretations that do not preserve the meaning of the verb stem. This process, referred to as *semantic drift*, is illustrated in examples (23)-(25) below. In combination with the verbal affix *-oo*, the verb *dog* 'to cut' denotes a break-up (23b). Similarly, when combining with the affix *-e*, the verb *degg* 'to hear' refers to keeping in contact (24b), whereas the verb *gis* 'to see' denotes a meeting (25b). The same holds for the verb *daj* 'to find', denoting a meeting when combining with the affix *-e* and for the verb *dëkk* 'to live' that denotes being neighbours when combined with *-e* or *-oo* (22a).

By contrast, verbs reciprocalized by *-ante* always keep the meaning of the transitive verb stem. For instance, unlike in (24b), the predicate *degg* 'to hear' can only

denote a reciprocal hearing event when bearing the morpheme *-ante* (26). Similarly, the predicate *gis* 'to see' univocally leads to a mutual 'seeing' configuration with *-ante* (27), unlike its counterpart with *-e* in (25b).

(23) a. Khady mu ngi   dog mburu.
          Khady PRST.3SG cut bread
          'Khady cuts the bread.'

     b. Khady ak   Fatou dañu   dog-oo.
          Khady and Fatou FOC.V.3PL cut-RECP
          'Khady and Fatou broke up.'

(24) a. Khady degg na   Fatou.
          Khady hear PRF.3SG Fatou
          'Khady heard Fatou.'

     b. Khady ak   Fatou ñu ngi   degg-e.
          Khady and Fatou PRST.3PL hear-RECP
          'Khady and Fatou keep in contact.'

(25) a. Khady gis na   Fatou.
          Khady see PRF.3SG Fatou
          'Khady saw Fatou.'

     b. Khady ak   Fatou gis-e   nañu.
          Khady and Fatou see-RECP PRF.SBJ.3PL
          'Khady and Fatou met.'

(26) Khady ak   Fatou nañu   degg-ante.
       Khady and Fatou PRF.SBJ.3PL hear-RECP
       'Khady and Fatou heard each other.'

(27) Khady ak   Fatou dañu   gis-ante.
       Khady and Fatou FOC.V.3PL see-RECP
       'Khady and Fatou saw each other.'

In the literature, there is shared consensus that only lexicalized reciprocals can undergo a semantic drift (Kemmer 1993, Haspelmath 2007, Nedjalkov 2007, Siloni 2012). The pattern illustrated above further supports our proposal that reciprocal verbs marked by *-e* and *-oo* must be lexicalized entries.

3.4 Lexical reflexivity

Reflexivity in Wolof is expressed by the NP *bopp* 'head' and a possessive deter-
miner, which may occupy the object position of any transitive verb (28). How-
ever, verbs denoting grooming or body-related actions that fall into the cate-
gorization of 'naturally reflexive' events (Kemmer 1993) may express reflexivity
with the verbal suffix *-u* (29). Note that unlike the NP strategy, the affix *-u* is
not productive; we refer to the verbs that can bear this affix as *lexical reflexives*.
Verbs from this class may have a transitive entry (30): we rely on the assumption
that such predicates may have a transitive alternate and an intransitive reflexive
alternate, marked by *-u*.

(28) Ñun da   ñoo bañ sunu bopp.
     1PL AUX.FOC 1PL hate 1PL   head
     'We hate ourselves.'
     (Tamba 2008:4)

(29) Khady sang-u   na.
     Khady wash-REFL PRF.3SG
     'Khady washed.'

(30) Khady mu ngi   sang muus mi.
     Khady PRST.3SG wash cat   the
     'Khady washes the cat.'

Crucially, Wolof verbs with a lexical reflexive entry can only express reci-
procity with the affix *-ante* (31a), while they lead to ungrammaticality with *-e*
and *-oo* (31b).

(31)   a. Khady ak Fatou ñu ngi   sang-ante.
         Khady and Fatou PRST.3PL wash-RECP
         'Khady and Fatou washed each other.'
     b. *Khady ak   Fatou ñu ngi   sang-e/ sang-oo.
        Khady   and Fatou PRST.3PL wash-RECP

Essentially, verb stems that are used to form lexical reflexives cannot appear
with the affixes *-e* and *-oo*, which we propose to be markers of lexical reciprocity.
We take this pattern as support for the idea that lexical reflexives and lexical re-
ciprocals are two distinct classes with no overlapping entries (Reinhart & Siloni
2005). In English, for instance, lexical reflexives and lexical reciprocals have the

same surface form, as they are both realized with zero morphology (32). Yet, there is no overlap in the meaning: (32a) unambiguously denotes a reflexive configuration and (32) a reciprocal configuration.

(32) a. John shaved.

b. John and Joe hugged.

In Wolof, the distinction between these two classes of verbs is detectable in their morphological realization. Verb stems that combine with the lexical reflexive marker *-u* systematically lead to ungrammaticality with *-e* and *-oo* (33 - 35). Conversely, *-u* cannot attach to verbs that may express reciprocity with *-e* and *-oo*, as illustrated in (36).

(33) *wat* 'to shave'

a. *wat-u* 'to shave (oneself)'

b. **wat-oo, *wat-e*

(34) *sol* 'to dress'

a. *sol-u* 'to dress (oneself)'

b. **sol-oo, *sol-e*

(35) *sang* 'to wash'

a. *sang-u* 'to wash (oneself)'

b. **sang-oo, *wash-e*

(36) *gis* 'to see'

a. **gis-u*

b. *gis-oo, gis-e* 'to meet (each other)'

4 Conclusion

In order to study the properties of reciprocal morphemes in Wolof, we collected novel data through a series of interviews with native speakers. The data substantiate a distinct treatment for the morpheme *-ante*, as opposed to the morphemes *-oo/-e* : we argue that *-ante* is a productive morpheme that turns transitive verbs into intransitive reciprocal verbs, while *-oo* and *-e* mark only a restricted class of verbs, without operating on the argumental structure. This twofold analysis boils down to the the distinction between *grammatical* and *lexical* reciprocity introduced in §2.1: while *-ante* is responsible for grammatical reciprocity, *-oo* and

-e are markers of predicates with a lexical reciprocal entry. This treatment is in line with the observation that *-e* is a polysemous marker (Diouf 2009:57): cross-linguistically, it is not uncommon for non-productive middle markers to be associated with an array of interpretations, including lexical reciprocity (Kemmer 1993).

Our proposal relies on a number of contrasts between *-ante* and *-e/-oo*. We have illustrated that *-ante* is productive with transitive verbs (§3.1), but it cannot reciprocalize predicates without a direct object (§3.2). By contrast, *-e* and *-oo* can only combine with a restricted class of predicates that denote 'naturally reciprocal' events (Kemmer 1993), regardless of whether they have a transitive entry or not (§§3.1-3.2). We have also seen that only verbs marked by *-e* and *-oo* may undergo a 'semantic drift' (§3.3), while only *-ante* can reciprocalize predicates with a lexical reflexive entry (§3.4).

We would also like to draw attention to the typological peculiarity of reciprocity in Wolof. Cross-linguistically, it is common for languages with an overt morphological distinction between grammatical and lexical reciprocity to express the former with pronominal elements, like *each other* in English, and the latter with a verb-marked form, like zero-morphology in English (8), the suffix *-sya* in Russian (9), the verbal affix *-óz-* in Hungarian (Uralic; Rákosi 2008) or the *hitpael* template in Hebrew (Semitic; Doron 2003), *inter alia*. In Wolof, grammatical and lexical reciprocity are overtly distinguished, but both expressed by means of verbal affixes. This language nonetheless lends support to Kemmer (1993)'s observation that productive markers are morpho-phonologically heavier than non-productive middle-related markers.

In this paper, we have considered *-e* and *-oo* as two variations of a morpheme with the same function. We have not encountered cases where these two affixes lead to different interpretations with the same verb stem, nor have we found empirical grounds to motivate a different treatment of *-e* and *-oo*. However, further research might reveal some morphological, semantic or distributional differences between these two morphemes, which may explain the claim that *-oo* and not *-e* marks both sociative and reciprocal events (Voisin 2002: 281). Future works may also focus on the semantics of Wolof reciprocals, to find out whether there are differences or restrictions in the way *-ante*, *-e* and *-oo* may express the different kinds of reciprocal configurations.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

ACT	active voice	POSS	possessive
BEN	benefactive case	PRO	pronominal base
CONN	connective	PRST	presentative
COPART	co-participation	SG	singular
FV	final vowel		

Acknowledgements

References

Bruening, Benjamin. 2006. The Morphosyntax and semantics of verbal reciprocals. Unpublished ms. University of Delaware.

Buell, Leston & Mariame Sy. 2005. A fixed hierarchy for Wolof verbal affixes. In Yuni Kim & Rebecca Cover (eds.), *Proceedings of the thirty-first annual meeting of the Berkeley Linguistics Society* (Special Session on Languages of West Africa), 25–36. DOI: 10.3765/bls.v31i2.816.

Church, Eric. 1981. *Le système verbal du Wolof.* Dakar: Université de Dakar.

Creissels, Denis & Sylvie Nouguier-Voisin. 2008. Valency-changing operations in Wolof and the notion of co-participation. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 289–306. Berlin & New York: De Gruyter Mouton.

Dalrymple, Mary, Sam A Mchombo & Stanley Peters. 1994. Semantic similarities and syntactic contrasts between Chicheŵa and English reciprocals. *Linguistic Inquiry* 25(1). 145–163.

Diouf, Jean-Léopold. 2009. *Grammaire du Wolof contemporain.* Revue et complétée. Paris: L'Harmattan.

Doron, Edit. 2003. Agency and voice: The semantics of the Semitic templates. *Natural Language Semantics* 11(1). 1–67.

Haspelmath, Martin. 2007. Further remarks on reciprocal constructions. In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions*, vol. 4, 2087–2115. Amsterdam: John Benjamins.

Ka, Omar. 1981. *La dérivation et la composition en Wolof* (Les langues nationales au Sénégal). Dakar: Centre de Linguistique Appliquée de Dakar (CLAD).

Ka, Omar. 1982. *La syntaxe du Wolof: Essai d'analyse distributionelle.* Thèse de 3 cycle. Dakar: Université de Dakar. (Doctoral dissertation).

Keenan, Edward & Jean-Paulin Razafimamonjy. 2004. Reciprocals in Malagasy. *Oceanic Linguistics* 43(1). 177–207.

Kemmer, Suzanne. 1993. *The middle voice* (Typological Studies in Language 23). Amsterdam: John Benjamins.

Mchombo, Sam A. & Rosemary M. Ngalande. 1980. Reciprocal verbs in Chicheŵa: A case for lexical derivation. *Bulletin of the School of Oriental and African Studies* 43(3). 570–575.

McLaughlin, Fiona. 2001. Dakar Wolof and the configuration of an urban identity. *Journal of African Cultural Studies* 14(2). 153–172.

Nedjalkov, Vladimir P. 2007. Overview of the research: Definitions of terms, framework, and related issues. In Vladimir P. Nedjalkov (ed.), *Reciprocal constructions* (Typological Studies in Language 71), 3–114. Amsterdam: John Benjamins.

Rákosi, György. 2008. The inherently reflexive and the inherently reciprocal predicate in Hungarian: Each to their own argument structure. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and Typological Explorations*, vol. 192, 411–450. Berlin & New York: De Gruyter Mouton. DOI: 10.1515/9783110199147.411.

Reinhart, Tanya & Tal Siloni. 2005. The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry* 36(3). 389–436.

Siloni, Tal. 2012. Reciprocal verbs and symmetry. *Natural Language & Linguistic Theory* 30(1). 261–320. DOI: 10.1007/s11049-011-9144-2.

Tamba, Khady. 2008. Anaphora questionnaire response - Wolof. Version 1.3. In Ken Safir (ed.), *Afranaph website*. New Brunswick: The State University of New Jersey. https://afranaphproject.afranaphdatabase.com/wolof-casemenu-227.

Voisin, Sylvie Nouguier. 2002. *Relations entre fonctions syntaxiques et fonctions sémantiques en Wolof.* Université Lumière Lyon2. (Doctoral dissertation).

Chapter 19

A Rizzian analysis of the left periphery in Sewama Mende

Jason D. Smith

Michigan State University

Rizzi (1997, 2001) proposes a universal structure for the left periphery. In this paper I show that data from Mende, a Mande language spoken in Sierra Leone, supports Rizzi's analysis. I show that the Mende left periphery is bracketed by Force and Finite Phrases, with Focus, and Topic constructions between. I further propose that all the functional heads in the left periphery in Mende are head-initial with surface variations resulting from movement of focused and topicalized constituents into the specifier positions of their respective heads.

1 Introduction

Rizzi (1997, 2001) argues that Larson's (1988) theory of an articulated VP and Pollock's (1989) work leading to an expanded inflectional level should likewise be extended to the complementizer layer. He proposes that the complementizer system looks upward, expressing force, that is whether the sentence is a question, exclamation, statement, etc., while also looking downward, expressing finiteness, which interacts with the expression of tense, mood, agreement, etc. in the TP. In this paper I show that the CP structure of Mende, a Mande language spoken in Sierra Leone, supports Rizzi's analysis. While there is a small body of research on the syntactic structure of Mande languages (c.f. Mahou: Koopman 1984, Bambara: Koopman (1992), Kpelle: Travis (1989), Dafing: Sande et al. (2019)), this work is the first generative syntax analysis of the left periphery of a language in the Mande family. As such it provides an important contribution to the understanding of the left periphery and provides a baseline for future research on Mande languages.

Adamou et al. (2018) note that Information Structure categories, including focus and topic, have been assumed to be universal, being expressed through a variety of language-specific means. In light of this, a more detailed description and analysis of African languages, and even more specifically Mande languages, is warranted.

It has been claimed that the Mande languages have a strict SOVX word order (Creissels 2005, Nikitina 2009).[1] Mende is canonically SOVX with the subject and object preceding the verb, and with adjuncts surfacing post-verbally. In (1) the subject *Mary* and direct object *nikeisia* 'the cows' precede the verb *gbɛ* 'chase,' while the locative phrase *kpaa hun* 'on the farm' and temporal phrase *gboi* 'yesterday' follow the verb.

(1) **S**   **O**   **V**   **X**
    Mary nike-i-sia   gbɛ-nga   kpaa hun gboi
    Mary cow-DEF-PL chase-PFV farm on   yesterday
    'Mary chased the cows on the farm yesterday.'

Focus in Mende can be marked either in-situ or via movement to the left periphery. In-situ focus (glossed as FOC.I) is indicated by the morphological marker *lɔ*, which follows the entity that it focuses. Left peripheral focus (glossed as FOC.L) is characterized by movement of the focused constituent to the left periphery, where it is focused by *mia*, which I argue is a syntactic head. I discuss this in more detail in §4.2. In cases of left peripheral focus, when a lexically appropriate pronoun is available, a resumptive pronoun is found in the focused constituent's canonical position. Both arguments (2) and adjuncts (3–4) can be focused.

(2) Argument Focus

    a. Mary *nike-i-sia*   **lɔ**   gbɛ-nga   kpaa hun gboi
        Mary cow-DEF-PL FOC.I chase-PFV farm on   yesterday
        'Mary chased THE COWS on the farm yesterday.'

    b. *nike-i-sia*   **mia**   Mary ti   gbɛ-nga   kpaa hun gboi
        cow-DEF-PL FOC.L Mary 3PL chase-PFV farm on   yesterday
        'Mary chased THE COWS on the farm yesterday.'

---

[1] The argument that Mande languages are **strictly** SOVX (Creissels 2005, Nikitina 2009) does not hold for Mende. In addition to the presence of adverbs in pre-verbal positions (see (9) below), in Smith (2022) I show a variety of contexts in which direct objects appear in post-verbal positions in Mende. I propose, instead, that Mende is **canonically** SOVX.

In (2a) the direct object *nikeisia* 'the cows' is focused in-situ, while in (2b) it is focused in the left periphery. Note the third person resumptive pronoun *ti* in the object's pre-movement position.

(3)   Adjunct Focus with Resumption

    a.   Mary nike-i-sia   gbɛ-nga *kpaa hun* **lɔ**
       Mary cow-DEF-PL chase-PFV farm on FOC.I
       'Mary chased the cows ON THE FARM.'

    b.   *kpaa hun* {**mia**} / {***lɔ**} Mary nike-i-sia   lɔ-nga   na
       farm on   FOC.L/FOC.I Mary cow-DEF-PL see-PFV LOC
       'It is on the farm that Mary chased the cows.'

(4)   Adjunct Focus without Resumption
    gboi   **mia** Mary nike-i-sia gbɛ-nga kpaa hun
    yesterday FOC.L Mary cow-DEF-PL chase-PFV farm on
    'It is yesterday that Mary chased the cows on the farm.'

In (3a) the locative phrase *kpaa hun* 'on the farm' is focused in-situ by *lɔ*, while in (3b), it appears in the left periphery where it is followed by *mia*. The locative resumptive pronoun *na* 'there' surfaces in the phrase's canonical position. In (4) the temporal phrase *gboi* 'yesterday' appears in the left periphery, and there is no resumptive pronoun in its pre-movement position.

Focus fronting, like that found in Mende, is likewise found in many Kwa, Bantu, and Chadic languages (Aboh et al. 2007). The focus markers *lɔ* and *mia* in Mende resemble the particle *ne* used in Kikuyu to mark focus in ex-situ constructions, as detailed by Schwarz (2007). He further argues for a focus-phrase analysis in which *ne* either is the head of the focus-phrase or the spell-out of a focus feature, an analysis that likewise seems plausible for Mende.

Contrasting with intonation languages which use stress to indicate focus, it has been proposed that some tone languages instead use syntactic transformations. The suggestion is that the presence of tone mitigates the effectiveness of intonation and stress in marking focus (Aboh et al. 2007). In my investigation of Sewama Mende, I am currently still researching the role of tone in the language. To the extent that there appears to be lexical tone, it does not mirror the results reported in the literature.[2] Regardless of the prominence of tone in

---

[2]Innes (1967) and Spears (1967) both documented tone in the language, and their work was used by Leben (1973) and Goldsmith (1976) in establishing autosegmental phonology. Innes's (1969) dictionary uses Kɔɔ Mende, while it is unclear which dialect Spears uses. Given my ongoing research, I do not mark tone in this paper. It should also be noted that Mende orthography does not mark tone.

Sewama Mende, it does seem that it patterns like tonal languages in utilizing syntactic transformations above phonological processes in marking focus.

Using data from two native speakers in Bo, Sierra Leone, in this paper I lay out the structure of the left periphery in Sewama Mende, arguing that it aligns with Rizzi's proposed analysis for a universal hierarchy. The structure of the paper is as follows. §2 is a brief introduction to the language, while §3 introduces Rizzi's analysis. In §4 I lay out the structure of the left periphery in Mende, making a correlated argument that focused and topicalized constituents move (that is they are not base-generated) into the left periphery. §5 is a brief conclusion.

2 Background

Mende (ISO 639-2 *men*) is spoken by about two million speakers in the southern and eastern parts of Sierra Leone and Liberia (Eberhard et al. 2020). Williamson & Blench (2000) argue that the Mande languages are an early offshoot of the Niger-Congo family. There are 4 major dialects: *Kɔɔ* (eastern Sierra Leone), *Kpa* (southwestern Sierra Leon), *Sewama* (south-central Sierra Leone), and *Waanjama* (southeastern Sierra Leone and Liberia). While most previous research has focused on *Kɔɔ* (c.f. Innes 1967), this research examines *Sewama* Mende, spoken in and around Bo, the largest city in the Mende area of Sierra Leone. I am not aware of any research that focuses exclusively on the Sewama dialect. Innes (1961) proposes that the dialects differ little lexically, and while I have noticed distinctions, they are not relevant to the present discussion of the left periphery.

To this point there has been very little syntactic research on the language, with most previous work focused on tone and Consonant Mutation (Dwyer 1969, Conteh et al. 1986, Tateishi 1990). Descriptive grammars include Aginsky (1935), Crosby (1944), Spears (1967), and the substantial work of Innes (1961, 1967, 1969). Sengova (1981) is a dissertation by a native speaker considering tense and aspect in the language. Nearly all previous analyses of tone and consonant mutation are based on data from Spears and Innes.

3 Rizzi's Analysis

Much has been written in response to the analysis set out in Rizzi (1997, 2001) (c.f. É. Kiss 1998, Benincà & Poletto 2004, Cinque & Rizzi 2008), though data from the Mande language family has not yet been considered under this framework. In this section I briefly introduce his analysis of focus and topic within the articulated complementizer system before turning to the Mende data.

Rizzi observes that topicalized and focused constituents both appear to the left of a canonical clause in Italian. He also observes that there is a consistency in how these constructions surface. He argues that a topic construction consists of the topic itself, which is preposed and expresses old information. Typically, it is set off by a comma intonation. The part of the sentence that is not the topic is the *comment*, which expresses new information.

Rooth (1985, 1992) describes the focus semantic value of a sentence as contrasting with its ordinary semantic value. He further suggests that the focus value of a sentence is the set of alternatives from which the actual semantic value is selected. While Rizzi argues that a focused constituent is preposed and receives focal stress, the part of the clause that is not focused is called the *presupposition*, and it is information presumably shared by the interlocutors (Jackendoff 1972).

Subsequent research has shown that focus constructions need not be preposed, nor receive focal stress. This is particularly germane to African languages. While information structure categories (including focus) have been assumed to be universal, much of what is known comes from well-described languages, with lesser known languages at times challenging previously held assumptions (Adamou et al. 2018). This is, in fact, the case with Mende, where focus is indicated via morphology and syntax without any focal stress.

Rizzi argues that topic and focus interact with a number of other functional heads in the left periphery, and, crucially, that the left periphery is structured the same cross linguistically. His proposal for the universal structure of the left periphery is shown in (5). The force head connects the clause to supraordinate material and precedes the focus head. The focus head precedes the finiteness phrase. The finiteness phrase connects the CP with the lower TP. Topic phrases can surface between the various other functional heads in the articulated CP. In the remainder of this paper, I show how the Mende data supports this analysis.

(5)   Force {Top*} Foc {Top*} Fin IP/TP

4 The left periphery of Mende

4.1 Finiteness Phrase (FinP)

In this section I describe Mende's left periphery, working from the bottom upwards, looking first at the finiteness phrase.

In Mende plural subjects are obligatorily followed by a *subject marker*. Before discussing the position of the subject, I first consider subject markers, as they
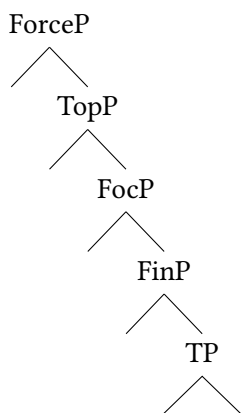
Figure 1: Rizzi's Proposed Left Periphery

offer crucial insight into the position of the clausal subject. In (6a) the 3rd person plural subject marker *ti* follows the subject *ndupuisia kpɛlɛ* 'all the children.' By using a quantified subject, we can rule out that the subject is really a topic (Rizzi 1997). In (6b) the subject marker remains for the plural subject, even without the quantifier. In (6c) the sentence is ungrammatical with the expected 3rd person singular subject marker *ngi*.[3] The data in (6d) and (6e) show that subject markers also surface for non-human and non-living plural subjects.

(6)   a.   ndupu-i-sia kpɛlɛ *(**ti**) mangu-i-sia   mɛ-i   lɔ
           child-DEF-PL all   3PL mango-DEF-PL eat-PST NM
           'All the children ate the mangoes.'

      b.   ndupu-i-sia *(**ti**) mangu-i-sia   mɛ-i   lɔ
           child-DEF-PL 3PL mango-DEF-PL eat-PST NM
           'The children ate the mangoes.'

      c.   ndupu-i   (***ngi**) mangu-i-sia mɛ-i   lɔ
           child-DEF.SG 3SG   mango-DEF-PL eat-PST NM
           'The child ate the mangoes.'

---

[3]In the following data set and future sets, note the presence of the particle *lɔ* in a post-verbal position. I am uncertain of the precise role that *lɔ* plays in these construction, and, as such, I gloss it as a neutral marker (NM), as its presence indicates that there is no other focused constituent in the clause. This *lɔ* particle can also surface as a lengthening of certain vowels, including *a*, as in (7).

    d. nike-i-sia   *(**ti**) mangu-i-sia   mɛ-i   lɔ
        cow-DEF-PL 3PL   mango-DEF-PL eat-PST NM
        'The cows ate the mangoes.'

    e. windo-i-sia   *(**ti**) wɔ-ngɔ
        window-DEF-PL 3PL break-STAT
        'The windows are broken.'

The same subject markers are used for simple future tense constructions.

(7)   a. ndupu-i-sia *(**ti**) mangu-i-sia   mɛ-ma a
        child-DEF-PL 3PL mango-DEF-PL eat-FUT NM
        'The children will eat the mangoes.'

    b. ndupu-i   (***ngi**) mangu-i-sia   mɛ-ma a
        child-DEF.SG 3SG   mango-DEF-PL eat-FUT NM
        'The child will eat the mangoes.'

Interestingly, a different subject marker surfaces for habitual / present constructions. In this construction, the third person plural subject marker is *ta* while the singular subject marker is *a*.

(8)   a. nyapu-i-sia *(**ta**) mahe-i   male lɔ tatovo   gbi ma
        girl-DEF-PL 3PL chief-DEF meet NM Monday all   on
        'The girls meet the chief every Monday.'

    b. nyapu-i   *(**a**) mahe-i   male lɔ tatovo   gbi ma
        girl-DEF.SG 3SG chief-DEF meet NM Monday all   on
        'The girl meets the chief every Monday.'

In light of the data in (6) through (8), I propose that these subject markers are polymorphemic, consisting of a series of syntactic heads. In each of the previous data sets, it is [t] that marks 3rd person plural, and I propose that there is a null morpheme [∅] that marks 3rd person singular. This morpheme surfaces as the head of a Subject Phrase (SubjP), agreeing with the subject in person and number. In (8) habitual aspect is marked by [a], and, therefore, *t-a* marks 3rd person plural – habitual aspect. Given this analysis, the subject marker *a* in (8b) is really ∅-a, signifying 3rd person singular – habitual tense. The [a] surfaces as the head of a Hab(itual)P. Since in both the simple past and future the subject marker surfaces

as *ti*, it cannot be the case that [i] marks tense.[4] Instead, I suggest that since it is not possible for [t] to surface as a string by itself, [i] surfaces as a default vowel.

Even though they have traditionally been written as a single orthographic unit, I propose that the subject markers have an articulated structure similar to that shown in Figure 2 for the 3rd person plural habitual subject marker *ta*.
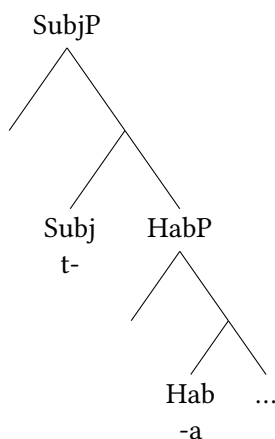
Figure 2: Mende Subject Marker

Turning now to the position of the subject, a crucial component of the previous analysis of subject markers is that they agree in person and number with the subject. This would point to some type of agreement relationship, however, the subject and subject marker do not form a constituent. This can be seen when an adverb intervenes between the two, as in (9) in which the modal adverb *gbolenhunwe* 'clearly' intervenes between the subject *ndupuisia* 'the children' and the 3rd person plural subject marker *ti*.

(9) ndupu-i-sia {*ti} gbolenhunwe {ti} nike-i   huma-nga
     child-DEF-PL 3PL clearly   3PL cow-DEF steal-PFV
     'The children have clearly stolen the cow.'

---

[4]The subject marker *ti* is also used in a variety of other tense/aspect constructions, such as the perfective.

(i) ndupu-i-sia ti mangu-i-sia mɛ-nga
     child-DEF-PL 3PL mango-DEF-PL eat-PFV
     'the children ate the mangoes'

In light of this, I suggest that the subject moves through SpecSubjP, triggering agreement before moving to a higher position. The most obvious candidate for a landing spot is the specifier of FinP, which I would argue has a null head. In this position it connects the left periphery with the lower portion of the clause (Rizzi 1997, 2001), in this case via an agreement relationship between the subject and subject marker. It also aligns with Cardinaletti's (1997) proposal that SpecFinP is filled with the 'subject of predication.'

Figure 3 shows what the structure including the FinP subject marker of the sentence in (10) looks like under this analysis.[5]

(10) ndupu-i-sia *(ta) mangu-i-sia   mɛ lɔ folo gbi
      child-DEF-PL 3PL mango-DEF-PL eat NM day all
      'The children eat the mangoes every day'

4.2 Focus Phrase (FocP)

As noted in (2), focus in Mende can be marked either in-situ or via movement to the left periphery. In the left periphery examples, the subject (11b), object (11c), and locative phrase (11d) can all be focused. Note the subject-object asymmetry – when the subject moves into the left periphery there is no resumptive pronoun, while movement of the object utilizes a resumptive.

(11) Left Peripheral Focus
      a. ndupu-i-sia ti   mangu-i-sia   mɛ-nga kpaa hun
         child-DEF-PL 3PL mango-DEF-PL eat-PFV farm on
         'The children ate the mangoes on the farm.'

      b. *ndupu-i-sia* **mia** ti   mangu-i-sia   mɛ-nga kpaa hun
         child-DEF-PL FOC.L 3PL mango-DEF-PL eat-PFV farm on
         'It is the children that ate the mangoes on the farm.'

      c. *mangui-i-sia* **mia** ndupu-i-sia ti   ti   mɛ-nga kpaa hun
         mango-DEF-PL FOC.L child-DEF-PL 3PL 3PL eat-PFV farm on
         'It is the mangoes that the children ate on the farm.'

      d. *kpaa hun* **mia** ndupu-i-sia ti   mangu-i-sia   mɛ-nga na
         farm on   FOC.L child-DEF-PL 3PL mango-DEF-PL eat-PFV LOC
         'It is on the farm that the children ate the mangoes.'

---

[5]Given the focus of this paper on the left periphery and space constraints, I do not make any particular claims about the clausal structure below the subject marker. Further research is necessary to flesh this out in sufficient detail.
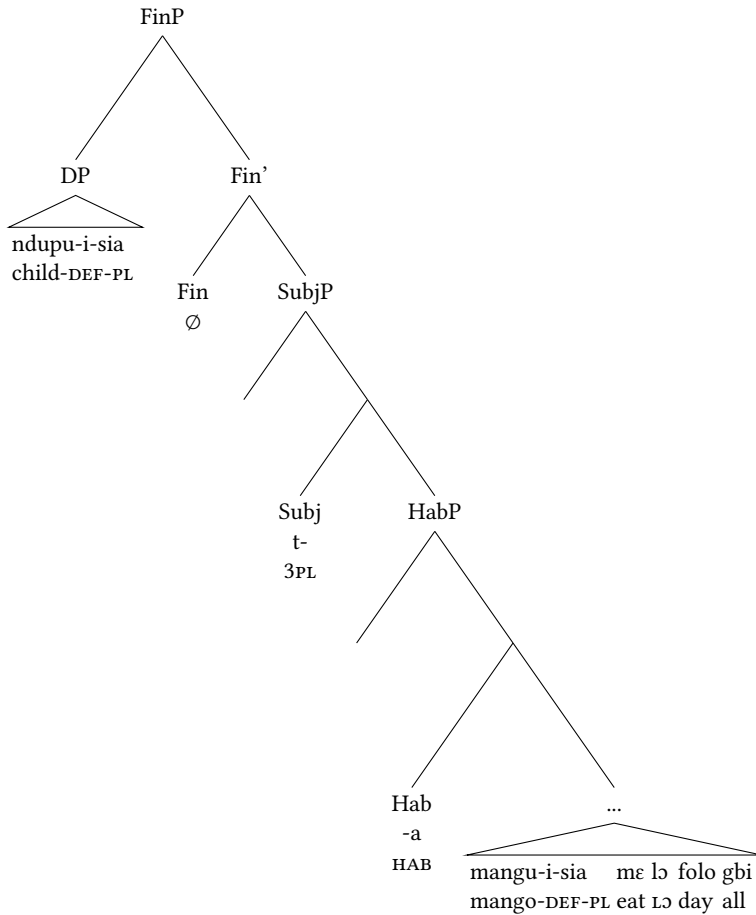
Figure 3: Mende Finite Phrase

The same subject (12a), object (12b), and locative phrase (12c) can also be focused in-situ.

(12) In-Situ Focus

a. *ndupu-i-sia* **lɔ** ti mangu-i-sia mɛ-nga kpaa hun
child-DEF-PL FOC.I 3PL mango-DEF-PL eat-PFV farm on
'THE CHILDREN ate the mangoes on the farm.

b. ndupu-i-sia ti *mangu-i-sia* **lɔ** mɛ-nga kpaa hun
child-DEF-PL 3PL mango-DEF-PL FOC.I eat-PFV farm on
'The children ate THE MANGOES on the farm.'

  c. ndupu-i-sia ti mangu-i-sia mɛ-nga *kpaa hun* **lɔ**
    child-DEF-PL 3PL mango-DEF-PL eat-PFV farm on FOC.I
    'The children ate the mangoes ON THE FARM.'

As seen in the following example, the focus marker *mia* cannot occur outside of the left periphery (13a), and the focus marker *lɔ* cannot occur within the left periphery (13b).

(13) a. *ndupu-i-sia ti *mangu-i-sia* **mia** mɛ-nga kpaa hun
     child-DEF-PL 3PL mango-DEF-PL FOC.L eat-PFV farm on
     'The children ate THE MANGOES on the farm.'

   b. *mangui-i-sia* **lɔ**   ndupu-i-sia ti ti mɛ-nga kpaa hun
     mango-DEF-PL FOC.I child-DEF-PL 3PL 3PL eat-PFV farm on
     'It is the mangoes that the children ate on the farm.'

It is ungrammatical to have two focus constructions within the same clause.

(14) *ndupu-i-sia* **mia** ti mangu-i-sia **lɔ**   mɛ-nga kpaa hun
   child-DEF-PL FOC.L 3PL mango-DEF-PL FOC.I eat-PFV farm on
   'It is the children that ate THE MANGOES on the farm.'

Turning to left peripheral focus, the focused constituent must appear before the clausal subject, which I have argued is in SpecFinP; it cannot follow the clausal subject, as in (15a). It is also ungrammatical to have two focused constituents in the left periphery, as in (15b)

(15) a. FinP   FocP   TP
     *ndupu-i-sia **kpaa hun** *mia* ti mangu-i-sia mɛ-nga na
     child-DEF-PL farm on FOC.L 3PL mango-DEF-PL eat-PFV LOC
     'It is on the farm that the children ate the mangoes.'

   b. FocP$_1$ FocP$_2$ FinP   TP
     *mangu-i-sia kpaa hun mia ndupu-i-sia ti ti mɛ-nga na
     mango-DEF-PL farm on FOC.L child-DEF-PL 3PL 3PL eat-PFV LOC
     Intended: 'It is the mangoes on the farm that the children ate.'

To this point I have asserted that focused constituents are not base generated in the left periphery, rather, that they move there. Using the following data, I argue for a movement analysis to explain how focused constituents surface in the left periphery. The evidence includes quantifier float and reconstruction effects.

Quantifier float occurs when a DP raises into a higher position in the clause, stranding its quantifier. The surface position of the quantifier indicates a position through which the DP has transited (Sportiche 1988, Fitzpatrick 2006).

The examples in (16) illustrate quantifier float in Mende. (16a) is a canonical SOV sentence with *nikeisia* 'the cows' as the direct object. In (16b) the quantifier *kpɛlɛ* 'all' quantifies the DP direct object. In (16c) the direct object *nikeisia* is fronted with the quantifier, while in (16d) the quantifier is stranded. If the DP object can be fronted and the quantifier can remain in the canonical direct object position, we need a story for how the two can be separated. The most natural story is that the DP object moved via Ā-movement into the left periphery, stranding the quantifier.[6] Interestingly, it is also possible for the quantifier to be stranded in a post-verbal position.[7]

(16) a. Peter **nike-i-sia** ngeya-nga
     Peter cow-DEF-PL buy-PFV
     'Peter has bought the cows.'

    b. Peter **nike-i-sia** *kpɛlɛ* ngeya-nga
     Peter cow-DEF-PL all buy-PFV
     'Peter has bought all the cows.'

    c. **nike-i-sia** *kpɛlɛ* mia Peter { ti yeya-nga / ngeya-nga }
     cow-DEF-PL all   FOC.L Peter { 3PL buy-PFV / buy-PFV   }
     'It is all the cows that Peter has bought.'

    d. **nike-i-sia** mia Peter *kpɛlɛ* (*ti) ngeya-nga
     cow-DEF-PL FOC.L Peter all   3PL buy-PFV
     'It is all the cows that Peter has bought.'

    e. **nike-i-sia** mia Peter *(ti) yeya-nga *kpɛlɛ*
     cow-DEF-PL FOC.L Peter 3PL buy-PFV all
     'It is all the cows that Peter has bought.'

---

[6]Note that the resumptive pronoun is optional when the quantifier is pied-piped into the left periphery or surfaces in a pre-verbal position. It is ungrammatical for the resumptive pronoun to surface when the object has moved to the left periphery and the quantifier is stranded in a post-verbal position. Further research is necessary in order to better understand these discrepancies.

[7]This points to the direct object merging and/or moving via A-movement through a post-verbal position at some point in the derivation. In Smith (unpublished manuscript) I suggest (following Kayne 1994) that all Mende verb phrases are head-initial, with OV word order derived via leftward movement.

A second argument for movement utilizes reconstruction effects, that is a context in which a constituent is in one position on the surface but behaves as if it were in a lower position. For the sake of space, I introduce only one instance of reconstruction effects, namely ideophones. Ideophones, which have been described as vivid sensory words, are fairly common in African languages (Dingemanse 2018, Downing 2019). They are very similar to adverbs, and there is a very strong selectional relationship between the ideophone and the verb, and as such ideophones typically appear with just one, or at most a few verbs (Tamba et al. 2012, Torrence 2013).

Consider the example of the Mende ideophone *kpe*, meaning 'clean through.' While (17a) indicates that Kpana *ngulii lewenga* 'has cut the branch,' the presence of the ideophone *kpe* in (17b) specifies that he cut it 'clean through', as opposed to e.g. 'cutting a notch into the branch'. In (17c) the verb *bɔ* 'shoot' is used instead. While in English the statement 'I shot the arrow clean through the target' is perfectly acceptable, (17d) indicates that in Mende *kpe* 'clean through' cannot be used with *bɔ*. Given this tight selectional relationship, the ideophone typically directly follows the verb.

(17)   a.   Kpana nguli-i   lewe-nga
            Kpana tree-DEF cut-PFV
            'Kpana has cut the branch.'

       b.   Kpana nguli-i   lewe-nga **kpe**
            Kpana tree-DEF cut-PFV   clean.through
            'Kpana has cut the branch clean through.'

       c.   Kpana kɔli-i   bɔ-nga
            Kpana leopard-DEF shoot-PFV
            'Kpana has shot the leopard.'

       d.   *Kpana kɔli-i   bɔ-nga   **kpe**
            Kpana leopard-DEF shoot-PFV clean.through
            'Kpana has shot the leopard clean through.'

(18b) introduces another ideophone that can go with *lewe* 'cut,' that is *fiikifiiki*, which describes cutting 'with a sawing (or back and forth) motion'. Canonically, the ideophone *fiikifiiki* immediately follows the verb *lewe* 'cut.' (18c) is a left peripheral focus construction. Note that the ideophone can surface in the left periphery, separate from its verb, with the meaning, 'It is with a sawing motion that Peter has cut the branch.' In order to account for the separation between the ideophone and the verb, the most plausible explanation is that the ideophone merged adjacent to the verb, before moving to the focus position in the left periphery.

(18)   a. Peter nguli-i   lewe-nga
             Peter tree-DEF cut.PFV
             'Peter has cut the branch.'

        b. Peter nguli-i   lewe-nga **fiikifiiki**
             Peter tree-DEF cut.PFV   sawing.motion
             'Peter has cut the branch with a sawing motion.'

        c. **fiikifiiki**   mia Peter nguli-i   lewe-nga
             sawing.motion FOC.L Peter tree-DEF cut.PFV
             'It is with a sawing motion that Peter has cut the branch.'

    Having argued that quantifier float and reconstruction effects point to Ā-move-ment to the left periphery, I now clarify into what position the focused con-stituents move. Recall from (15) that focused constituents must surface in a posi-tion to the left of the finite phrase and that *mia* is always used to indicate focus in the left periphery. I propose that *mia* is the head of FocP and it attracts a +Fo-cus constituent to its specifier (Chomsky 1993). The structure and position of the focus construction is shown in (19)/Figure 4. The focused constituent *kpaa hun* 'on the farm' moves into the specifier position of the focus head *mia*, while the resumptive pronoun *na* remains in the locative's pre-movement position.

(19)   kpaa hun mia ndupu-i-sia ti   mangu-i-sia   mɛ-nga na
       farm on   FOC.L child-DEF-PL 3PL mango-DEF-PL eat-PFV LOC
       'It is on the farm that the children have eaten the mangoes.'
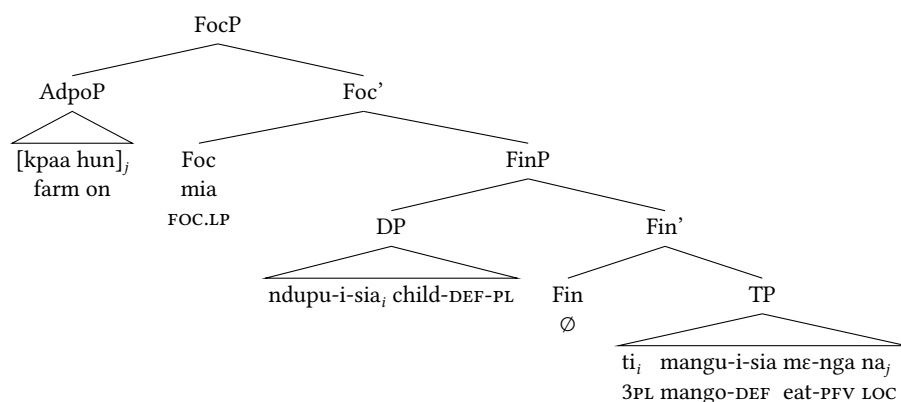
Figure 4: Mende Focus Phrase

Similar to focused constituents, wh-words in Mende can also surface in-situ or in the left periphery. In-situ they are marked with *lɔ*, while in the left periphery, they immediately precede *mia*. As such, I argue that they move into the same position in the left periphery, that is the specifier of the focus head *mia*. In (20b) the locative phrase is moved into the focus position, while in (20c) the wh-word *mindo* 'where' moves into the same position. Note that the same resumptive pronoun *na* is also used in both examples.

(20) a. ndupu-i-sia ti   mangu-i-sia   mɛ-nga **kpaa hun**
          child-DEF-PL 3PL mango-DEF-PL eat-PFV farm on
          'The children have eaten the mangoes on the farm.'

      b. **kpaa hun** *mia* ndupu-i-sia ti   mangu-i-sia   mɛ-nga na
          farm on   FOC.L child-DEF-PL 3PL mango-DEF-PL eat-PFV LOC
          'It is on the farm that the children have eaten the mangoes.'

      c. **mindo** *mia* ndupu-i-sia ti   mangu-i-sia   mɛ-nga na
          where FOC.L child-DEF-PL 3PL mango-DEF-PL eat-PFV LOC
          'Where have the children eaten the mangoes?'

It is ungrammatical to have both a focused constituent and a wh-word in the left periphery, whether they each have their own focus head (21b) or both precede the focus head (21c). This indicates that the wh-word and focused constituent are in the same position.

(21) a. ndupu-i-sia ti   mangu-i-sia   mɛ-nga **kpaa hun**
          child-DEF-PL 3PL mango-DEF-PL eat-PFV farm on
          'The children have eaten the mangoes on the farm.'

      b. *****kpaa hun** *mia* **gbɛ**   *mia* ndupu-i-sia ti   ti   mɛ-nga na
          farm on   FOC.L what FOC child-DEF-PL 3PL 3PL eat-PFV LOC
          'It is on the farm that the children have eaten what?'

      c. *****kpaa hun gbɛ**   *mia* ndupu-i-sia ti   ti   mɛ-nga na
          farm on   what FOC.L child-DEF-P 3PL 3PL eat-PFV LOC
          'It is on the farm that the children have eaten what?'

The same syntactic structure is generated for questions (22)/Figure 5 as is generated for focus constructions (19). The wh-word *mindo* 'where' surfaces in the specifier of the focus head *mia*, with the resumptive pronoun *na* in the pre-movement position.

(22) mindo mia   ndupu-i-sia ti mangu-i-sia   mɛ-nga na
      where FOC.LP child-DEF-PL 3PL mango-DEF-PL eat-PFV LOC
      'Where have the children eaten the mangoes?'
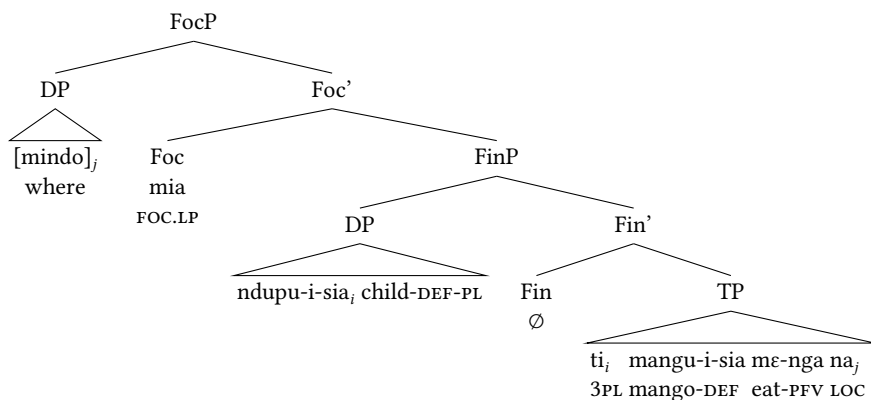
Figure 5: Mende Left-Peripheral wh-Phrase

4.3 Topic Phrase (TopP)

I now turn to topic phrases. There are four characteristics of topic phrases to highlight. First, topic phrases in Mende can appear in a variety of internal positions within the articulated left periphery. In (23a) the TopP *manguisia (va)* '(as for) the mangoes' precedes the FinP *ndupuisia* 'the children.' In (23b) it follows the wh-phrase *mindo mia* 'where is it' in the FocP position, while in (23c), it precedes the same FocP. This lines up with Rizzi's proposal that topic phrases can surface in various positions in the left periphery. Second, note that *va* 'as for' is optional in each of the examples. In discussions with my language consultants, the use of *va* 'for' seems quite arbitrary. Third, there is a comma intonation that sets apart topic phrases. When clause initial, as in (23a) and (23c), the comma follows the phrase, while in (23b) we see the comma intonation both precedes and

follows the topic phrase. Finally, we can see that the 3rd person plural resumptive pronoun *ti* remains in the topic's pre-movement position.[8]

(23) a. **mangu-i-sia$_i$ (va)**, ndupu-i-sia$_j$ ti$_j$ ti$_i$ mɛ-nga
mango-DEF-PL for, child-DEF-PL 3PL 3PL eat-PFV
'As for the mangoes, the children ate them.'

b. mindo mia, **mangu-i-sia$_i$ (va)**, ndupu-i-sia$_j$ ti$_j$ ti$_i$ mɛ-nga na
where FOC.L, mango-DEF-PL for, child-DEF-PL 3PL 3PL eat-PFV LOC
'Where is it, as for the mangoes, the children ate them?'

c. **mangu-i-sia$_i$ (va)**, mindo mia   ndupu-i-sia$_j$ ti$_j$ ti$_i$ mɛ-nga na
mango-DEF-PL for,   where FOC.L, child-DEF-PL 3PL 3PL eat-PFV LOC
'As for the mangoes, where is it that the children ate them?'

Unlike focus constructions (c.f. 15b), in Mende it is possible to have more than one topic phrase in a clause. In (24) there are two topics *manguisia* 'the mangoes' and *gboi* 'yesterday.' Both of them can be introduced by the topic head *va* 'for,' as in (24a), while it is also possible for neither to be introduced by *va* (24b). This lines up with Rizzi's (1997) argument that recursion of the topic-comment structure is possible. We can conclude that the topic head can optionally be expressed as *va* or be phonologically null.

---

[8]In Mende, singular entities can be topicalized. When a subject is topicalized, a resumptive pronoun obligatorily surfaces, as in (i).

(i) Peter va, mindo mia   ngi nike-i   lɔ-nga   na
Peter for, where FOC.L 3SG cow-DEF see-PFV LOC
'As for Peter, where is it that he saw the cow?'

As is the case for object pronouns more broadly, topicalized singular objects utilize a re-sumptive pronoun when the subject is human (ii), while a phonologically null resumptive is used for non-human objects (iii). Note that the absence of a phonologically realized resumptive pronoun does not trigger consonant mutation on the verb.

(ii) nyapu-i va, mindo mia Peter ngi   lɔ-nga   na
girl-DEF for, where FOC.L Peter 3SG.HUM see-PFV LOC
'As for the girl, where is that Peter saw her?'

(iii) nike-i   va, mindo mia   Peter ∅   tɔ-nga   na
cow-DEF for, where FOC.L Peter 3SG.NONHUM see-PFV LOC
'As for the cow, where is that Peter saw it?'

(24) a. mangu-i-sia   va, mindo mia, gboi   va ndupu-i-sia ti *ti*
         mango-DEF-PL for where FOC.L yesterday for child-DEF-PL 3PL 3PL
         mɛ-nga na
         eat-PFV LOC
         'As for the mangoes, where is it, yesterday, the children ate them?'

      b. mangu-i-sia,   mindo mia, gboi,   ndupu-i-sia ti *ti* mɛ-nga
         mango-DEF-PL where FOC.L yesterday child-DEF-PL 3PL 3PL eat-PFV
         na
         LOC
         'The mangoes, where is it, yesterday, the children ate them?'

Similar to focus phrases, when a pronoun is available, it surfaces in the canonical position of the topicalized constituent. This explains the presence of the resumptive pronoun *ti* (indicated in italics) in (24). Since there is no pronoun to represent a temporal phrase, there is no resumptive in the canonical position of *gboi* 'yesterday.'

Concluding this section, the structure of the clause in (25) is shown in Figure 6. The topicalized phrase *manguisia* 'the mangoes' moves into the specifier position of the topic head *va* to check its [+topic] feature. In this example the topic phrase precedes the focus phrase, though as shown above, it can also follow it.

(25) mangu-i-sia   (va), mindo mia   ndupu-i-sia ti ti mɛ-nga na
      mango-DEF-PL for,   where FOC.L, child-DEF-PL 3PL 3PL eat-PFV LOC
      'As for the mangoes, where is it that the children ate them?'

4.4 Force Phrase (ForceP)

The final phrase in the left periphery that I consider is the Force Phrase. The ForceP in Mende is headed by the declarative complementizer *kɛ* 'that' or the interrogative complementizer *ina* 'if.' The force head *kɛ* introduces embedded clausal complements (26a), while *ina* introduces embedded questions (26b).[9]

(26) a. Peter hungɛ-nga [kɛ ndupu-i-sia ti mangu-i-sia   mɛ-nga]
         Peter explain-PFV that child-DEF-PL 3PL mango-DEF-PL eat-PFV
         'Peter explained that the children ate the mangoes'

---

[9]Note that even though Mende has canonical OV word order, CP complements never appear in a pre-verbal position. I argue in Smith (in press) that in canonical constructions the object raise for case. Since CPs do not need case (Stowell 1981), CP objects in Mende remain in a post-verbal position
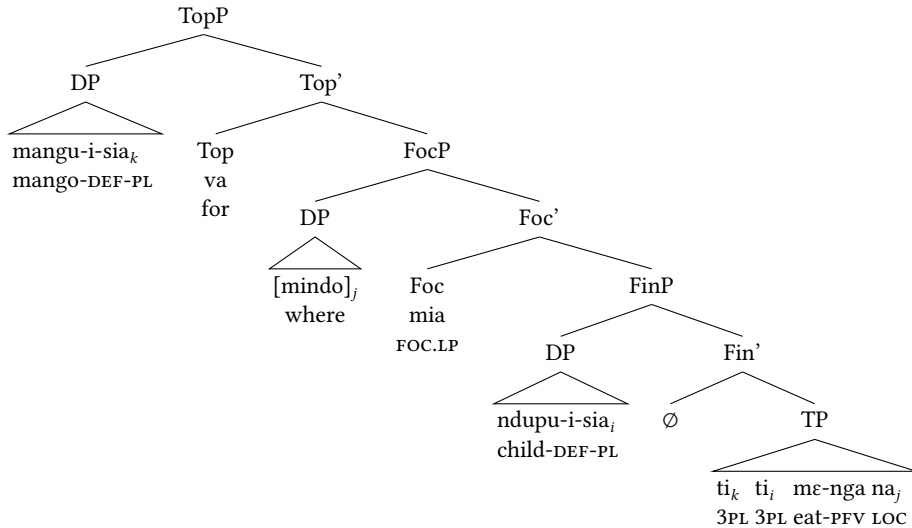
Figure 6: Mende Topic Phrase

b. Peter mɔli-nga [ina ndupu-i-sia ti   mangu-i-sia   mɛ-nga]
   Peter ask-PFV if   child-DEF-PL 3PL mango-DEF-PL eat-PFV
   'Peter asked if the children ate the mangoes'

Rizzi asserts that the Force head looks upward connecting the phrase with supraordinate structure. In these examples the matrix verb *hungɛ* 'explain' in (26a) can take a clausal complement, which is obligatorily headed by the declarative complementizer *kɛ*. In (26b) the matrix verb *mɔli* 'ask' takes an embedded question, headed by the complementizer *ina*. In both cases the Force head takes as its complement the remainder of the clause.

It has been argued previously that Mende is a head-final language (c.f. Rice & Cowper 1984 and Tateishi 1990 who indicate a series of complement-head constructions). As laid out in this paper, I would argue that Mende is a head-initial language. In previous work I have argued that Mende's canonical OV word order is derived from an underlying head-initial verb phrase (Smith 2022). Here I extend this argument to the various functional heads in the left periphery, arguing that each of its functional heads is head-initial. The Finiteness Phrase has a null head and the sentential subject moves into its specifier. The Focus Phrase is headed by *mia*, with the focused constituent moving into its specifier. Topic Phrases are optionally headed by *va* with the topicalized constituent in the specifier. The Force Phrase headed by *kɛ* or *ina* takes the remainder of the clause as its complement.

   This structure is shown in example (27). Each of the functional heads in Figure 7 is marked in bold, with the constituents that have moved into their specifier marked in italics. This analysis accounts for the proposed head-initial structure Mende in the left periphery.

(27)   Peter kabande-nga [ina mangu-i-sia   va, mindo mia ndupu-i-sia   ti
       Peter wonder-PFV if   mango-DEF-PL for where FOC child-DEF-PL 3PL
       ti   mɛ-nga na]
       3PL eat-PFV LOC
       'Peter wondered, if, as for the mangoes, where it is the children had eaten them.'
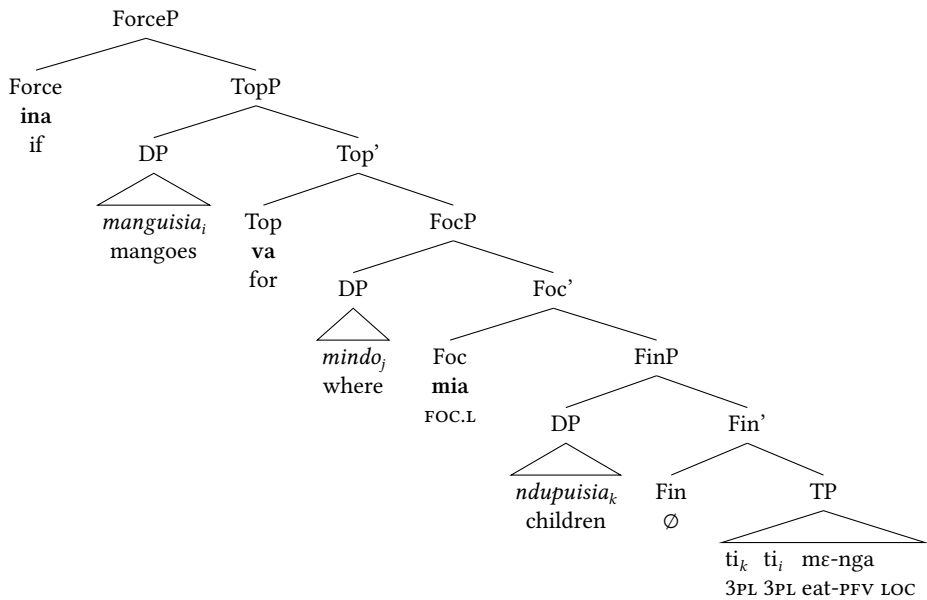
Figure 7: Mende Force Phrase

5 Conclusion

To this point there has been no analysis of the left periphery of a Mande language. In this paper I have shown that the structure of the left periphery in Mende provides cross-linguistic support for the argument in Rizzi (1997, 2001) for a universal hierarchy in the left periphery. This can be seen in comparing the tree in

Figure 7, showing the Mende left periphery, with the tree in Figure 8, showing Rizzi's proposed hierarchy, repeated from Figure 1.

(28)   Force {Top*} Foc {Top*} Fin IP/TP
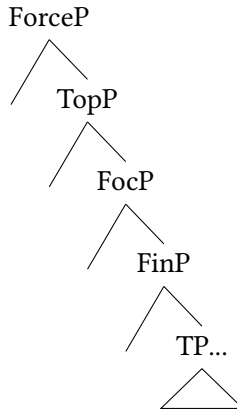
Figure 8: Rizzi's Proposed Left Periphery

From the bottom up, I show how the structure of the left periphery in Mende can be analyzed as being bracketed by a Force Head and Finite Head with Topic, and Focus heads within the articulated CP structure. I also argued for a movement analysis for focused and topicalized constituents, as opposed to a base-generation analysis of focused and topicalized constituents in the left periphery.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

FIN	finite	NM	neutral marker
FOC.I	in-situ focus	STAT	stative
FOC.L	left peripheral focus	TP	tense phrase
IP	inflectional phrase		

Jason D. Smith

Acknowledgements

References

Aboh, Enoch, Katharina Hartmann & Malte Zimmermann. 2007. Focus and grammar: The contribution of African languages. In Enoch Oladé Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 1–14. Berlin, New York: De Gruyter Mouton.

Adamou, Evangelina, Katharina Haude & Martine Vanhove. 2018. Investigating information structure in lesser-known and endangered languages. In Evangelina Adamou, Katharina Haude & Martine Vanhove (eds.), *Information structure in lesser-described languages*, 1–14. Amsterdam: John Benjamins Publishing Company.

Aginsky, Ethel. 1935. A grammar of the Mende language. *Language* 3(11). 7–111.

Benincà, Paola. & Cecilia Poletto. 2004. Topic, focus, and V2. In Luigi Rizzi (ed.), *The structure of CP and IP: The cartography of syntactic structures*, vol. 2 (Oxford Studies in Comparative Syntax), 52–75. Oxford: Oxford University Press.

Cardinaletti, Anna. 1997. Agreement and control in expletive constructions. *Linguistic Inquiry* 2. 521–533.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Cinque, Guglielmo & Luigi Rizzi. 2008. The cartography of syntactic structures. *Studies in Linguistics* 2. 42–58.

Conteh, Patrick, Elizabeth Cowper & Keren Rice. 1986. The environment for consonant mutation in Mende. In Gerrit J. Dimmendaal (ed.), *Current approaches to African linguistics*, vol. 3, 107–116. Dordrecht: Foris.

Creissels, Denis. 2005. S-O-V-X constituent order and constituent order alternations in West African languages. In Yuni Kim & Rebecca Cover (eds.), *Proceedings of the Thirty-first Annual Meeting of the Berkeley Linguistics Society, Special session on languages of West Africa*, 37–51. Berkeley: Berkeley Linguistics Society.

Crosby, Kenneth. 1944. *An introduction to the study of Mende.* Cambridge: W. Heffer & Sons.

Dingemanse, Mark. 2018. Redrawing the margins of language: Lessons from research on ideophones. *Glossa* 3(1). DOI: 10.5334/gjgl.444.

Downing, Laura. 2019. Tumbuka prosody: Between tone and stress. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and description in African linguistics: Selected papers from the 47th Annual Conference on African Linguistics*, 75–94. Berlin: Language Science Press. DOI: 10.5281/zenodo.3367128.

Dwyer, David. 1969. *Consonant mutation in Mende.* Michigan State University. Department of Linguistics, Oriental and African Languages. (MA thesis).

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74(2). 245–273.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2020. *Ethnologue: Languages of the world.* 23rd edn. Dallas: SIL International.

Fitzpatrick, Justin. 2006. *The syntactic and semantic roots of floating quantification.* Cambridge, MA: MIT. (Doctoral dissertation). https://dspace.mit.edu/handle/1721.1/37420.

Goldsmith, John A. 1976. *Autosegmental phonology.* Cambridge, MA: MIT. (Doctoral dissertation). http://dspace.mit.edu/handle/1721.1/16388.

Innes, Gordon. 1961. *The structure of sentences in Mende.* London: University of London, School of Oriental & African Studies.

Innes, Gordon. 1967. *A practical introduction to Mende.* London: University of London, School of Oriental & African Studies.

Innes, Gordon. 1969. *A Mende-English dictionary* (West African Language Monograph Series). Cambridge: Cambridge University Press.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar.* Cambridge, MA: The MIT Press.

Kayne, Richard S. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Koopman, Hilda J. 1984. *The syntax of verbs: From verb movement Rules in the Kru languages to Universal Grammar* (Studies in Generative Grammar 15). Dordrecht: Foris.

Koopman, Hilda J. 1992. On the absence of case chains in Bambara. *Natural Language & Linguistic Theory* 10(4). 555–594.

Larson, Richard. 1988. On the double object construction. *Linguistic inquiry* 19(3). 335–391.

Leben, William R. 1973. *Suprasegmental phonology.* Cambridge, MA: MIT. (Doctoral dissertation). https://dspace.mit.edu/handle/1721.1/16364.

Nikitina, Tatiana. 2009. The syntax of postpositional phrases in Wan, an 'SOVX' language. *Studies in Language* 33(4). 910–33.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic inquiry* 20(3). 365–424.

Rice, Keren & Elizabeth Cowper. 1984. Consonant mutation and autosegmental morphology. In Joseph Drogo, Veena Mishra & David Testen (eds.), *Papers from the twentieth regional meeting of the Chicago Linguistic Society*, vol. 20, 309–320. Chicago, IL: Chicago Linguistic Society.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Springer, Kluwer Academic Publishers.

Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi* (North Holland Linguistic Series: Linguistic Variations 59), 287–296. Amsterdam: Elsevier.

Rooth, Mads Edward. 1985. *Association with focus (Montague grammar, semantics, only, even).* Amherst, MA: University of Massachusetts, Amherst. (Doctoral dissertation).

Rooth, Mads Edward. 1992. A theory of focus interpretation. *Natural language Semantics* 1(1). 75–116.

Sande, Hannah, Nico Baier & Peter Jenks. 2019. The syntactic diversity of SAuxOV in West Africa. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and description in African linguistics: Selected papers from the 47th Annual Conference on African Linguistics* (Contemporary African Linguistics 4), 667–701. Berlin: Language Science Press. DOI: 10.5281/zenodo.3367195.

Schwarz, Florian. 2007. Ex-situ focus in Kikuyu. In Enoch Oladé Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 139–160. Berlin, New York: De Gruyter Mouton.

Sengova, Matthew. 1981. *A classification of tense, aspect, and time specification in the verb system of Mende.* Madison, Wisconsin: University of Wisconsin dissertation. (Doctoral dissertation).

Smith, Jason D. 2022. Case and word order in Sewama Mende. In *Proceedings of the Fifty-seventh Annual Meeting of the Chicago Linguistic Society*, 373–386. Chicago, IL: Chicago Linguistic Society.

Spears, Richard. 1967. *Basic course in Mende.* Evanston, IL: Northwestern University.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19(3). 425–449.

Stowell, Timothy A. 1981. *Origins of phrase structure.* Cambridge, MA: MIT. (Doctoral dissertation).

Tamba, Khady, Harrold Torrence & Malte Zimmermann. 2012. Wolof quantifiers. In Edward Keenan & Denis Paperno (eds.), *Handbook of quantifiers in natural language*, 891–939. New York: Springer.

Tateishi, Koichi. 1990. A morphological analysis of Mende consonant mutation. In John Hutchison & Victor Manfredi (eds.), *Current approaches to African linguistics*, vol. 7, 99–110. Dordrecht: Foris.

Torrence, Harold. 2013. A promotion analysis of Wolof clefts. *Syntax* 16(2). 176–215.

Travis, Lisa. 1989. Parameters of phrase structure. In Mark Baltin & Anthony Kroch (eds.), *Alternative conceptions of phrase structure*, 263–279. Chicago, IL: University of Chicago Press.

Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African Languages: An introduction*, 11–42. Cambridge: Cambridge University Press.

Chapter 20

Determining word boundaries in *afaan Oromoo* (Oromic)

Tamam Youssouf

University of Toronto

Oromic, the language of the Oromo people, a Cushitic language spoken mainly in Ethiopia, has officially been using a Latin script-based orthography since 1991. The orthography is widely accepted by Oromic speakers. With the expansion of the application of this orthography to more than literacy and elementary education, the necessity for some updates on the rules has been felt at times. One such required update is the rules of determining word boundaries encompassing affixes, clitics, and words. This issue is very important in Oromic to minimize ambiguity and manage the length of the words. This paper proposes which morphemes should be affixed and which ones should be written separately, based mainly on the criteria proposed by Kutsch Lojenga (2014), and explains how this helps disambiguation.

1 Introduction

The Oromo are a Cushitic people who mainly live in the Oromia region of Ethiopia, constituting 35.8% of the total 2021 estimated population of 110.9 million (CIA, World Fact Book, Africa: Ethiopia). The Oromo also live in parts of Kenya, Somalia and Tanzania (see Janko 2007; Clamons et al. 1999). Their language is considered the third most widely spoken Afro-Asiatic language in the world after Arabic and Hausa, the second most widely used indigenous mother tongue in Africa, and a lingua franca for many groups in Ethiopia (Gragg 1982: viii). Despite its widespread use in speech, the language did not have an official writing system until 1991, when the new Ethiopian Peoples Revolutionary Democratic Front (EPRDF) government of Ethiopia at the time approved the use of the current Latin script-based orthography, referred to as *Qubee*.

The orthography is based on the work of a committee formed in the neighbouring country of Somalia during the end of the 1970s. This committee agreed on basic conventions but left out the rest of the rules, apparently assuming that the English language conventions were to be followed. As time passed with no systematic follow up on the work of this committee, the need to introduce some additional conventions became apparent. One such convention is the determination of word boundaries.

This paper presents the challenges that the lack of rules on the word boundaries posed, and suggests how to overcome these challenges using criteria developed by Kutsch Lojenga (2014). As the language has not had a standard version yet,[1] this work is largely based on Eastern Oromia (Oromoland), the author's native dialect.

2 What is a word?

Before discussing word boundaries, we need to define the term "word" itself, which is a morphological concept that is difficult to precisely define.[2] However, for the purpose of this paper, Lojenga's (2014) definitions will be used. According to her, a *word* is a unit which 1) has meaning all by itself; 2) is able to move around and be put in different places in a sentence; and 3) can be split from other words by another independent word, such as the two words in *the book* being separated by another word *big* as in *the big book* (Kutsch Lojenga 2014: 137, see also Eaton & Schroeder 2012 and Sinha & Talukdar 2013). This definition distinguishes words from other morphemes which are dependent and have to remain bound.

In writing, word boundary placement relates to the attempt to locate the appropriate place to separate units of a language by a space. It is uncommon to see written languages not have units separated by spaces. However, the determination of these units is not given enough attention. Detailed rules are not set in the development of orthographies. The importance of word boundary placement may not have been foreseen.

Generally, the words that we see separated by spaces contain more than one morpheme. However, whether all morphemes that we see conjoined should be written that way or not has to be determined. The problem is that there are no universal criteria applicable to all languages detailing when to write certain morphemes or group of morphemes conjunctively or disjunctively and why. Therefore, one of the primary issues to be addressed in this paper is establishing these

---

[1]A committee was formed in 1992 to standardize the language but did not last long.

[2]For detailed argumentation regarding the indeterminacy of word segmentation, please see Haspelmath (2011).

criteria regarding the Oromo orthography, with the intention of minimising ambiguity.

The other motive to address criteria for identifying word boundaries is to check word length. Most Oromic[3] words, especially verbs and nominals, are complex word forms (i.e., they consist of more than one morpheme). The Oromic base is what Haspelmath & Sims (2010: 21) call a "bound stem" that cannot also function as a word form. Hence there is an inbuilt tendency towards longer word forms. Oromic is an agglutinating language and most of the words contain more than one morpheme.[4]

As observed a long time ago by Tutschek (1845), all primitive verbs

> have the property of producing, by affixing syllables, new verbs which are different modifications of the primitive signification of the radical verb. The number of members belonging to each of the verbal families so produced is, however, extremely various, and depends on the nature of the radical verb; whence it arises that in some verbs, singular forms are altogether wanting in the series: others are confined to only a few branches: and others again are capable of being extended to the sixth and even to the eighth link of the chain (Tutschek 1845: 10-11)

This number of links in the chain is further multiplied because as Tucho et al. (1996: 155) observed, a transitive verb root can have as many as four forms derived from it, namely active, passive, autobenefactive, and causative.

3 The criteria for identifying a word

Let us start by summarising Lojenga's (2014: 91-99) set of criteria, which is the primary reference in this paper for word boundary placement. She identifies three types of morphemes: independent word, a clitic and a bound morpheme. She also provides three sets of criteria for identifying words, which are 1) syntactic, 2) phonological, and 3) semantic.

---

[3]I prefer to use *Oromo* for the people and *Oromic* for the language following Arab — Arabic and Amhara — Amharic respectively, both of which are Afro-Asiatic.

[4]Words like the following (22 letters, 7 syllables, and 8 morphemes) are common:

(i) *gu~gurgur-sii-f-ách-uu-dhaa-f*
    RED~sell-CAUS-DAT-ABEN-INF-0-GEN
    'in order to have sb sell sth piece by piece for one's own benefit'

The set of *syntactic* criteria are a) mobility: the ability for a morpheme to be in different places in a sentence, b) separability: the morpheme can separate from the neighbouring lexical morpheme by the insertion of another morpheme, and c) substitutability: a grammatical morpheme can substitute a lexeme in the same syntactic slot (paradigmatic substitution).

The set of *phonological* criteria include a) pronounceability in isolation, b) phonological unity: influence from the lexical morpheme to grammatical morphemes and/or vice versa by phonological processes involving tone, vowel harmony, and the like; and c) phonological bridging: how to divide words in a phrase or sentence when in speech everything is blended.

The final set of criteria Lojenga identifies is the *semantic* one, which include a) referential independence: whether the morpheme can communicate meaning in isolation, b) conceptual unity: if words acquire a new meaning when placed together, and c) minimal ambiguities: if writing disjunctively or conjunctively helps disambiguate.

Applying these criteria to the issue of Oromic word boundaries that Eaton & Schroeder (2012: 229) call *word break placement decisions*, we find that, in the current orthography, there are morphemes which sometimes are written as an independent word, and sometimes written together with the following or preceding word. Such examples are not regulated or consistent even within the same piece of literature by the same author. There is no explanation for writing the morphemes adjoined or disjoined. However, this act of joining or disjoining at will is not free of implications for the language.

As Lojenga explains, "(w)hen it comes to word recognition in the reading process, it is easier to learn to recognize a word that has a constant visual image, *especially at the beginning of the word*" (Kutsch Lojenga 2014: 84; emphasis added). Thus, the use of prefixes can make the visual image of a word unstable, slowing down the reading process in a language like Oromic, where most of the prefixes could be written as separate words (I return to this issue below). Furthermore, scholars like Gasser (2012) have observed that Oromic is predominantly a suffixing language instead of prefixing and/or infixing one. Avoiding prefixing thus yields some advantages in the writing system as it may facilitate word recognition in reading. Accordingly, arbitrary prefixing of a morpheme can be a burden to the process of word recognition.

Besides, listing prefixed words in the dictionary and locating them can pose problems. Will the word be inserted only in its prefixed form or in both? If we insert the word with the prefix, the prefix has to go in with all the possible words to which it is prefixed. If we do not include the prefix in the dictionary entry, there is no good reason to prefix it in writing either.

The next issue is identifying those morphemes that can make sense either standing alone or joined. This identification applies mainly to clitics, encompassing most of the prefixed and some suffixed Oromic morphemes. This group is different from the inflectional or derivational morphemes that have to be suffixed.

Regarding this difference between affixes and clitics, Zwicky & Pullum (1983) write:

> The primary difference between them [affixes and clitics] is that word-clitic combinations are governed mainly by syntactic considerations. In contrast, the conditions governing the combining of stems with affixes are morphological and/or lexical, concerned with the substructure of a finite set of words. (Zwicky & Pullum 1983: 503)

However heterogeneous these clitics are, our Oromic examples overwhelmingly suggest that all or most of the morphemes that pose a challenge, as to whether to write them as bound or independent morphemes, are clitics. Their heterogeneity may contribute to the lack of spontaneous or intuitive consensus in writing them either adjoined or disjoined. Below I list the most salient morphemes around which there is disagreement on how to associate them with the adjacent words.

4 Morphemes sometimes conjoined as prefixes

Table 1 presents a list of common morphemes sometimes written as prefixes. When dialectal differences occurs, I begin with the Eastern, separating the others by the forward-slash (/). The symbol is also used to separate gender if represented by different forms, beginning with the unmarked masculine followed by the feminine.[5]

The primary justification for prefixing these morphemes is that they are phonologically a part of the following word, or when we pronounce them, we pronounce the two words without a pause. Additionally, it is argued that since the motto of the Oromic orthography is "write as you pronounce", we have to conjoin

---

[5]It is necessary to use IPA because of discrepancies in the choice of the letters and digraphs, such as <c, q, x, ph> and the unregulated case of digraph gemination. I will use *italics* for *Qubee* and square brackets for IPA. Geminates will be written as CC instead of C:. To save space, I skip using the *Qubee* version in the tables, and only use the IPA conventions there.

[6]Exclusively used in the East before 1991. In the other parts of Oromia *hín* or *ín* was used, distinguished from the negative only by the high tone.

Table 1: Some common morphemes sometimes used as prefixes.

Morpheme	Gloss	Category	Added on	Sample
[hin]	not	negation	_VP	hin ɗuf-ín.
				not come-NEG
				'don't come.'
[if]/[of]/[uf]	self	reflexivity	_VP	if múr-é.
				self cut-3SG.PST
				'(he/I) cut (my/him)self.'
[ní][6]	FOC	focus	_VP	ní de:m-a.
				FOC go-3SG.IPFV
				'he goes.'
[wal]/[wol]	RECP	reciprocity	_VP	wal fit'-án.
				RECP finish-3PL.PST
				'(they) killed each other.'

them. However, the motto seems to stress the contrast with the English spelling, where there are more graphemes in a word than the represented sounds. The syntactic and semantic criteria of creating ambiguity and other aspects were not anticipated. Besides, it is sometimes difficult to find a pause between a series of what are separate words in a phrase. This argument of pause does not even satisfy the phonological criterion to adjoin.

The other argument is the tonal distinction that a native speaker applies when pronouncing such complex words, which has its disadvantages. To use the tonal difference in reading, one must see the context, which involves looking at the following word before reading the word with an appropriate tone, which slows down the speed. Again, the orthography does not mark tone, which is a crucial component of the language. As it stands now, it is easier to use space than to introduce tone marking, which involves adding a new element to the orthography.

The tendency to adjoin the above morphemes as prefixes is more predominant (although inconsistent) in literature published inside Oromia, Ethiopia. It is not uncommon to find the same morphemes sometimes written disjunctively and sometimes conjunctively, even within the same document. In the literature produced abroad, these clitics are generally written disjunctively. The Oromo Liberation Front (OLF) and affiliated bodies were the main producers of literature in Qubee. All literature produced by this body uses the disjunctive format. One possible cause of the predominance of affixing the clitics in Oromia is the

influence of the Ethiopic abugida writing system, used by Amharic, the official language.

In Ethiopic script-written languages, many monosyllabic words, especially adpositions, are written conjunctively to the grammatical words they precede or follow. As Griefenow-Mewis (2001: 57) observed, the Ethiopian writing system does not allow writing a one-syllable word separately. From the abugida (Rogers 2005: 205) nature of the system, one syllable consists of one grapheme unless it is a closed syllable, which is rare. Since most of the users of the current orthography for Oromic were originally trained in Amharic, it is hard to deny the influence.

If we assume the motto "writing as we speak" and the Ethiopic abugida as possible factors behind conjoining the morphemes in Table 1, we may also need to present the case for writing them disjunctively. Let us see if the Kutsch Lojenga (2014) criteria support the idea of writing the morphemes in Table 1 as independent words, taking them as examples.

The first sample morpheme is the negative marker *hin* '(do) not'. The form is *hin* + V$_{\text{root}}$ + AGR. We have the conjugation in Table 2 taking an imperfect form for *deem-* 'go'. The suffix after the hyphen marks the agreement.

Table 2: Non-past conjugation of *deem-* 'go' with *hin* 'not'

Person	Negative	Root	AGR
1sg	hin	dé:m-	-u
1pl	hin	dé:m-	-nu
2sg	hin	dé:m-	-tu
2pl	hin	dé:m-	-tan
3sg.f	hin	dé:m-	-tu
3sg.m	hin	dé:m-	-u
3pl	hin	dé:m-	-an

Let's take the syntactic criterion of substitutability. The negative *hin* can be substituted in the sentences in Table 2, which are in the imperfect form, by the jussive *haa* 'let', for example, in the same slot and the sentence will still be correct. Moreover, the mobility criteria can also apply without changing the meaning. Thus the negative imperative can be conveyed with relocated *hin* as in (1) below:[7]

---

[7]There is a very subtle difference in the mood of the instructor when the recipient hesitates and the sender changes his mind with hidden disappointment and withdraws the initial instruction. But, it conveys the same message, 'don't go!'

(1) Mobility of *hin* as the syntactic criterion to be written separately

  a. *Hin deemin*
     hin de:m-í-n
     not go-IMP.2SG-NEG
     'Do not go!'

  b. *Deemuu hin baddin*
     dé:m-ú: hin badd-í-n
     go-INFIN not try-IMP.2SG-NEG
     'Do not try to go!'

As for the semantic criteria of minimal ambiguities, the words in (2) have *hin* as their initial integral syllable. Thus it is difficult to guess whether the words with prefixed *hin* are prefixed complex or simple words unless one knows the meaning beforehand.

(2) Possible ambiguity created by conjoining *hin* to the next word

  a. *hinaaffaa*   [hiná:ffá:ʔ][8]   'jealousy'
  b. *hinniyyuu*   [hinníjjú:]   'negative competition'
  c. *Hindii*   [híndí:]   'India'
  d. *hirriiba*   [hirrí:ba][9]   'sleep (n)'
  e. *hilleensa*   [hille:nsa]   'rabbit'
  f. *hirroo*   [hirró:]   (said to incite bull to mate)
  g. *hirre*   [hírre]   'we distributed'
  h. *hinnaa*   [hinná:]   'red dye from plant leaves'
  i. *hinxilif*   [hint'ilíf]   'snatch (ideophone)'
  j. *hiyy-*   [hiyy-]   'poverty; become poor'
  k. *hincinnii*   [hinc'ínní:]   'tiny particles from harvest that irritate human skin'

Additionally, the semantic criterion of conceptual unity also applies. Unlike *irra* and *itti,* to be discussed later, there is no change in the meaning if we write *hin* adjoined or disjoined.

Similarly, two of the phonological criteria, namely pronounceability in isolation and phonological bridging, also apply. Thus, *hin* can be pronounced separately and, as there is a pause in speech between *hin* and the verb it negates, it can be separated by a hyphen, for example.

The negative particle *hin* typically precedes the verb it negates as a proclitic. Subsequently, there is an assimilatory change to the final /n/ which seems to

---

[8]Phonetically all long-vowel final words have a faint ʔ at the end, which is not written.

[9]*Hirriiba* 'sleep' (n) can be confused with a form of the verb *riib-*, namely *hirriiba* 'he does riib' (where assimilation has affected /hin + riiba/).

encourage writing *hin* conjunctively with the following verbs. This entices the proponents of writing *hin* conjunctively to invoke the presence of phonological unity. In connected speech, statements such as *hin láalu* 'he/I do not look' is pronounced [hillá:lu], and according to Oromic orthography, "you write what we hear". Additionally, if we go with this criterion, *hin* is closer to a preceding word, where in fast speech *inní hin láalu* 'he does not look' is pronounced as [inníllá:lu]. On the other hand, words like *gad* 'down', *ʔol* 'up' also initiate such assimilation with the following verbs as in *gad + táaʔi* 'sit down' → [gattá:ʔi]; *ʔol + réebi* 'chase up' → [ʔorré:bi]. But no one writes these and other words of this category such as *birá* 'near', *ɟalá* 'under', *duubá* 'behind' conjunctively with the verbs that follow. Thus, attaching *hin* in any form, based on this relation of phonological unity, will require us to conjoin these other words as well, and no writer currently conjoins these words.

The next common morpheme that is sometimes prefixed is the monosyllabic reflexive marker *if* 'self'. It precedes the VP: [if + VP]. There is also a reciprocal in the form of *wal/wol*[10] 'each other' in Table 1, that requires a plural verb. Both these terms are inflected for certain cases, just like nouns, as indicated in Table 3.

---

[10]The spelling *wol* is mainly in the SE dialect.

Table 3: The inflection of *if* and *wal* for different cases followed by suffixes or postpositions.

	Followed	Gloss	Sentence	Sentence (IMP)
ʔif	k'ab-	hold	ʔif k'abi	hold yourself!
	-i:f	for (DAT)	ʔifi:f k'abi	hold for yourself!
	-i:n	by (INST)	ʔifi:n k'abi	use to hold yourself!
	-(i)tti	to	ʔifitti k'abi	hold to yourself, hug!
	-(i)rra:	from	ʔifirra: k'abi	hold from yourself!
	bira	near	ʔif bira k'abi	hold by your side!
wal	k'ab-	hold	wal k'aba:	hold each other, wrestle!
	-i:f	for (DAT)	wali:f k'aba:	hold for each other!
	-i:n	by (INST)	wali:n k'aba:	hold together!
	-(i)tti	to	walitti k'aba:	hold to each other, gather (TR)!
	-(i)rra:	from	walirra: k'aba:	hold from each other!
	ɟala	under	wal ɟala k'aba:	hold under each other!

The syntactic criterion of substitutability applies in this case. Nominals in the accusative case can be substituted in that slot in all conjugated forms of the verb where *if* and *wal* apply, as in *muc'aa k'aba*: 'hold the child', *isii k'abaa* 'hold her' where *muc'aa* 'child' and *isii* 'her' substitute for *wal* and so on.

The separability and the minimal ambiguity criteria require us to write *if* disjunctively. Regarding separability, postpositions can come between *if* and the following verb. The same holds true for *wal* as well. Thus in (3b), (3c), (3d), (3f) and (3g), postpositions such as *dura* 'in front', *gad* 'down', the particle *hin*, and the single phoneme dative marker *-f* intervene between *if* and *wal* and the following verb.

(3)   The separation of *if* and *wal* from the verb phrase to which they are cliticised

    a. *If qabi.*
        ʔif   k'áb-i
        self hold-IMP.2SG
        'volunteer!' (lit. hold yourself)

    b. *If dura qabi*
        ʔif   dura k'ab-i
        self front hold-IMP.2SG
        'hold in front of you!'

    c. *If hin qabinaa.*
        ʔif   hin k'ab-í-n-a:
        self not hold-IMP-NEG-2PL
        'Do not hold yourselves!'

    d. *If gad hin qabinaa.*
        ʔif   gad   hin k'ab-í-n-a:
        self down not hold-IMP-NEG-2PL
        'Don't hold yourselves down!'

    e. *Wal qabaa.*
        wal   k'áb-a:
        each.other hold-IMP.2PL
        'wrestle (lit. hold each other)!'

    f. *Wal jala qabaa*
        wal   ɟala k'áb-a:
        each.other under hold-IMP.2PL
        'don't hold under each other!'

g. *Waliif qabaa.*
   wal-iː-f   k'áb-a:
   each.other-LV-DAT hold-IMP.2PL
   'Hold for each other!'

Additionally, the semantic criterion of minimal ambiguity also supports writing *if* and *wal* disjunctively. Some words have *if* or *wal* as their integral initial syllables, including the following in (4) and (5).[11]

(4) Possible ambiguity created by writing *if* conjunctively
     a. *Ifa*   [ʔifá]   'light'
     b. *Iftaan*   [ʔiftáːn]   'after tomorrow'
     c. *Iftiina*   [ʔiftíːna]   'light'
     d. *Ifaajii*   [ʔifáːɟíː]   'energy, wealth and time spent on sth.'

(5) Possible ambiguity created by writing *wal* conjunctively
     a. *Wal'aan*   [walʔáːn-]   'medically treat (TR)'
     b. *Walaawwal*   [walaːwwál-]   'being indecisive (INTR)'
     c. *Walaloo*   [walalóː]   'poem'
     d. *Wallee*   [walléː]/[wálleː]   '(love) song/ possibly?'
     e. *Walakkaa*   [walákkaː]   'centre'
     f. *Wallaal*   [walláːl-]   'lose knowledge (TR)'
     g. *Waleensuu*   [waleːnsúː]   'type of tree'
     h. *Walaba*   [walábá]   'independent'
     i. *Waldaya*   [waldájá]   'association'
     j. *Walalaa*   [walálá:]   'liquified honey'
     k. *Warrana!*   [warrána]   'hey family! (VOC)'

The next morpheme in the list of those sometimes prefixed is the emphatic positive focus marker *ni*, the opposite of *hin*. See also the note on *hin* and *ni* in Table 1. Syntactically, it is *ni* + VIPFV. It can be substituted by negative *hin*, locatives like *bira* 'near', or the temporal adverb *amma* 'now' in the same slot, as shown in (6). We will take an example of the first criterion of substitutability first.

(6) Substitutability of *ni*
     a. *Ni taa'an.*
        ní   taːʔ-an
        FOC sit-3PL.IPFV
        'They sit.'

---

[11]The hyphens used in (5) indicate that [walʔáːn-] and [walaːwwál-] are bound roots.

b. *Hin taa'an.*
hin tá:ʔ-an
NEG sit-3PL.IPFV
'They do not sit.'

c. *Bira taa'a.*
birá ta:ʔ-a
near sit-3SG.IPFV
'He sits near.'

d. *Amma teessi.*
amma te:ss-i
now sit-3SG.F.IPFV
'She sits now.'

Coming to the semantic criteria, we find ambiguity by conjoining *ni* to the verb next to it because there are words that have *ni* as their integral initial syllable, including those in (7).

(7) Words with *ni* as their integral initial syllables
   a. *Nigirtii*   [nigírtí:]   'quarrel, dispute'
   b. *Niitii*   [ni:tí:]   'wife'
   c. *Niis*   [ni:s]   'as well as'
   d. *Niin*[12]/*nan*   [ni:n]/[nan]   1SG indicator

In general, most of the criteria that apply to the negative marker *hin*, such as substitutability, mobility, minimizing ambiguity, conceptual unity, and pronounceability in isolation apply to *ni* as well.

5 Morphemes sometimes unnecessarily suffixed

Unlike the morphemes that are sometimes prefixed, which are clitics, suffixed morphemes are numerous. Moreover, most of them are generally accepted as proper inflectional or derivational suffixes. Oromic is an agglutinating language, and the affixes follow the root or the stem, except in the case of reduplication. Thus, most of the activity takes place on the right side of the root.

These suffixed morphemes have different shapes: single segments, mono- or multisyllabic suffixes. There is no question as to the single segment morpheme's

[12]Exclusively Eastern dialect. This is *ni + n* (emphatic *ni* and 1SG marker suffix *-n*) but written as one word.

position, and they are unanimously written conjunctively with the base they relate to.

Even though there is a consensus in writing the single segment morphemes as suffixes, there are cases where the accompanying morphophonology that requires long vowels before these suffixes create complications.[13] One such case is when the stem already has a long final vowel (actually /ʔ/-final, which is not spelt).[14] To add one of these suffixes on an already long-vowel-final base, we cannot directly add the morpheme but need an additional vowel. Adding this long vowel creates vowel hiatus, and it is necessary to resolve hiatus with an epenthesis where normally [ɗ], spelled *dh* is epenthesised.[15]

The introduction of this epenthesis also comes with a challenge. There is controversy whether *dha* is a copula or not. I have argued elsewhere (Youssouf 2019) that *dha* is not a copula, but it is a *dha* epenthesised to avoid vowel hiatus with the predicative case marker /a/ that follows it. This explanation brings us to the position of *dha* itself. Example (8)[16] below is one case. This confusion emanates from the lack of a rule to geminate digraph-represented sounds, that is, both [ɗ] and [ɗɗ] are spelled *dh.* I argue that it is easier to write *dha* disjunctively, e.g. *Boruun jabaa dha* for (8a), than making a new rule for the gemination of the digraphs (e.g. *dhdh*). Otherwise, the two statements in (8a) and (8b) cannot be distinguished. Note that the epenthetic consonant is glossed as Ø to indicate that it has no meaning.

(8)   One case for the separation of *dha*

    a.   *Boruun jabaadha*
        borúː-n   ɟábaː-ɗ-a
        boru-NOM strong-Ø-PCM
        'Boru is strong.'

    b.   *Boruun jabaadha*
        borúː-n   ɟabaː-ɗɗ-a
        tomorrow-1SG strong-ABEN-1SG.IPFV
        'I become strong tomorrow.'

---

[13]In the Western part of Eastern Oromia and Central Oromia, the vowel is short.

[14]The existence of this /ʔ/ after long vowel-final nouns was noted, among others, by Andrzejewski (1957).

[15]It may be the case that it has to do with the final 'ʔ' rather than epenthesis. In the Western dialect of Eastern Oromia and part of Central Oromia, the final vowel length is absent. Thus *déːmu: ɗáːn* → [déːmu: ɗán]. Note also there is no sequence of two or more different vowels in Oromic. Thus Oromia → *Oromiyaa*; out → *awti*; oil → *oyli.*

[16]I coined PCM following Banti (1988: 28) who uses "case" for the nominals used as predicate, and I added "marker".

The [ɗ] in this example is an epenthetic consonant or "empty morpheme" (Lloret-Romanyach 1988), inserted between the otherwise adjoined predicative case marker /a/, as in *Boruu-n dardár-**a*** 'Boru is an adult'. Separating this single phoneme clitic from its host can be challenged. However, other factors also encourage writing *dha* disjunctively. The first is the phonological criteria related to the pause that always occurs before /ɗa/ when it is in the final position.

The final problem with [ɗa] involving affixation is the lack of consistency. For example, some writers disjoin [ɗa] but conjoin [ɗaː], sometimes itself with suffixes like the instrumental *-n*, as in [ɗaːn], and disjoining when [ɗ] is followed by the predicative or accusative case marker *-a* suffixed as in [ɗa]. It is common to see [ɗa] written separately as <dha>, as in (9a), while [ɗaːn], the longer one, is joined as <-dhaan> as in (9b). This tendency to write short [ɗa] alone disjunctively could come from the lexical copula concept especially the Amharic *nɜw* 'is' which is written disjunctively.

(9) Inconsistency in writing [ɗ]-epenthesised words disjunctively

    a. *Deemuu dha fedha.*
        déːm-uː ɗ-á   feɗ-a
        go-INF 0-PCM want-3SG.M.IPFV
        '[It] is to go that he wants.'

    b. *Deemuudhaan fedha.*
        déːm-uː-ɗáː-n   feɗ-a
        go-INF-PCM-1SG want-1SG.IPFV
        'What I want is to go.'

It is not always the case that all multi-phoneme suffixes should better be written disjoined from the preceding word. Some are better written conjunctively. Those in this group are mainly consonant-initial. However, unlike the monophonemic ones, they need not be preceded by a long vowel. They just attach to an existing vowel of the host word form or are separated by an epenthetic short *i* in case the host word form is consonant-final. These morphemes are shown in the first six rows in Table 4. Some writers write them separately by adding or prothesising [i] at the beginning. These morphemes include conjunctive -[mmoː], locative case markers -[tti] 'to' and -[rra] 'on', and ablative case marker -[rraː] 'from'; -[jjuː] 'and', -[lleː] 'even', and the accusative focus marker -[jji], mainly in the SW Oromia dialect.

---

[17]Exclusively used in the SW Oromia dialect.

Table 4: List of some multi-phoneme morphemes sometimes suffixed.

Morpheme (Dial. Area)	Gloss	Category	Added on	**Example**
[(a/i)mmo:]	as for	Conj	NP_	is-á-mmó: 'as for him'
[(i)lle:]	as well	Conj/FOC	NP_	is-á-llé: 'him as well'
[(i)rra]	on	Postp	NP_	is-á-rrá 'on him'
[(i)tti]	to	Loc/Postp	NP_	is-á-ttí kénni 'give to him'
[(i)jji] (SW)[17]	it is	Focus	NP_	is-á-jjí gaafate 'it's him that he asked'
[(i)jju:]	even	Focus	NP_	is-á-jjú: ɗawe 'he hit even him'
-[uma]	just/only	Focus	NP_	/is-a-uma wa:mi/ [is-uma wa:mi] (E) [isa-ma wa:mi] (SW) 'call just him'

If we write these morphemes disjunctively with [i] prosthesis, we create ambiguity as shown in (10) by the following word-form pairs, where one is written adjoined, and the other disjoined.

(10)   Morphemes for which writing conjoined or disjoined creates minimal pairs

   a. *Isatti kenni.*
      ʔis-á-ttí   kénn-i
      he-ACC-LOC give-2SG.IMP
      'Give (something) to him!'

b. *Isa itti kenni.*
ʔis-á   ʔitti kénn-i
he-ACC LOC give-2SG.IMP
'Give him to (sth)!'

c. *Isarra kaayi.*
ʔis-á-rrá   ka:j-i
he-ACC-on put-2SG.IMP
'Put (sth) on him!'

d. *Isa irra kaayi.*
ʔis-á   ʔirra ká:j-i
he-ACC on   put-2SG.IMP
'Put him on (sth)!'

e. *Isarraa fuudhi.*
ʔis-á-rra-a: fú:ɗ-i
he-ACC-on-ABL take-2SG.IMP
'Take (sth) from him!'

f. *Isa irraa fuudhi.*
ʔis-á   ʔirra-: fú:ɗ-i
he-ACC on-ABL take-2SG.IMP
'Take him from (sth)!'

g. *Ganamollee arke.*
ganámó-llé: árk-e
N–as.well   see-1SG/3SG.PST
'I/he saw Ganamo as well.'

h. *Ganamo illee arke.*
ganámó ʔíllé: ʔárk-é
N   jaw.bone see-3SG.M.PST
'Ganamo saw the jaw bone.'

i. *Ganamoyyuu dhagaye.*
ganámó-jjú: ɗagáj-é
N-even   hear-3SG.M.PST
'Even Ganamo heard.'

j. *Ganamo iyyuu dhagaye.*
ganámó ʔíjj-ú: dhagáj-é
N   shout-INF hear-1SG/3SG.M.PST
'I/he heard Ganamo shouting.'

Note that the suffixed morphemes have initial geminate consonants, and the short [i] put before these suffixed morphemes are not part of them but epenthesised, as mentioned above. This [i]-epenthesis can be seen in inflectional or derivational processes where consonant-initial suffixes have to be affixed to consonant-final stems. In that case, a cluster of more than two consonants is created, as in the case of *arkite* 'you/she saw' where [i] is epenthesised to avoid an [rkt] cluster so that /ark + te/ becomes [arkíte] not *[arkte]. It should be added that native speakers use tone also to disambiguate. However, since tone has not yet been marked in the orthography, as stated above, it is safer to disambiguate by using a space. Also note that the [i]-initial versions like *illee, irra, irraa, itti* and *ijjuu* are categorically different words.[18]

This group is not entirely uniform. For example, the case of -[mmo:], seems different from the group because the preceding vowel alternates between [a] in the East and [i] in the Western dialect. It is possible that underlyingly it is /a/-initial and segmental haplology applies, whereby one of two similar consecutive segments, a vowel, in this case, is deleted. Still, it seems to constitute no problem if written conjunctively like the rest in Table 4.

Again, there are other clitics consisting of multiple segments, with different phonological properties. These include -[icca], -[itti:] and -[uma]. They require the noun stems that precede them to lose their final vowel(s), if any, before attaching. Owens (1985: 96) calls -[icca]/-[itti:] particulative. They are close to the English definite article. Some people nowadays tend to use -[icca] in the English definite *the* sense, attaching it to every noun. These morphemes are normally cliticized on non-proper nouns, like collective or generic nouns.

When they attach to such nouns, they distinguish gender. For example, on [sare:] 'dog', [k'a:llu:] 'traditional religious authority', [faranɟi:] 'a white person', gender is distinguished by attaching *-icca* for the masculine and *-itti:* for the feminine. *Tokko* 'one (M)' or *takka* 'one (F)' is added to make them indefinite. Thus [faranɟicca tokko] means 'a white man, any white man'. For stems with short final vowels, [t] is epenthesised. Thus [araba] 'Arab' becomes [arabticca] not *[arabicca]; [ʃe:ka] 'sheikh' becomes [ʃe:k+ticca] which changes to [ʃe:jticca][19] not *[ʃe:kicca].

---

[18]For example, detached *illee* 'jaw bone' is a noun, *ijjuu* 'to make noise' is a verb. Similarly, *itti* disjoined in *muc'aa itti kanni* 'give the boy to' modifies the verb while conjoined in *muc'atti kanni* 'give to the boy' is a postposition.

[19]In the Eastern dialect, velars become palatal approximants before coronals. Example: /da:k+ta/ → [da:jta]. Palatal approximants undergo total assimilation before a coronal. Example: /da:j + ta/ → [da:tta].

The two *-icca* and *-itti:* particulative morphemes also apply to the cardinal number *tokko/takka* 'one'. Similarly, *-an* is added on the other single digits and the last digits in the tens; 10, 20, 30 and so on. If the number is vowel-final, the vowel is omitted except for *lama* 'two'. Table 5 summarises the particulative morphemes on the numbers.

Table 5: Numbers and their particulative suffixes

		-icca M -itti: F	-an	-en
1	tokk-o M takk-a F	tokkicca takkitti:		
2	lama		lama:n	(lame:n)
3	sad-ih		sadan	(sade:n)
4	afur		afran	
5	ʃan		ʃanan	
6	ɟah-a		ɟahan	
7	torb-a		torban	
8	sadde:t		sadde:ttan	
9	sagal		saɟlan	
10	kuɗan		kurnan	

The next morpheme in the list being suggested to be conjoined in Table 4 above is -[uma], 'only, just'. It is added word-finally, and the final vowel of the word it is added to is dropped (see Owens 1985: 92). When the vowel(s) that precede this morpheme is/are deleted, ambiguity is created. For example, when it is suffixed on the positive focus marker *ni* and on *nu* '1PL.ACC', the [i] and [u] on the respective words are dropped, and the new term for both becomes [numa]. Thus the statement [numa ɲa:ta] can mean either 'he only/just eats' or 'he only/just eats us'. Even the tone distinction is not audible. Ambiguity generally arises when one morpheme or word is written or pronounced one way and has more than one meaning (homography or homophony). As can be seen in (11), when the vowels of focus marker *ni* and that of 1PL accusative case *nu* are deleted and *-uma* is suffixed, this results in homophones and homographs.

(11) Ambiguity created by suffixes that delete vowels before them

  a. *ni-uma = numa* FOC

  b. *nu-uma = numa* 1PL.ACC

However, not all dialects follow this rule. For example, in the Southeastern Oromia dialect, the suffixes lose their initial vowels instead of the stems losing their final vowels. Thus while in other dialects *tokko* 'one (M)' becomes [tokkicca], in the Southeast dialect it becomes [tokkocca]; *takka* 'one (F)' becomes [takkáttí:] not [takkíttí:] as elsewhere. Similarly, while *namá* 'human/person' becomes [námuma] elsewhere, in the Southeastern dialect, it becomes [námama]. This again creates another ambiguity as *-ma* is the passive voice suffix in Oromic. For example, [ɗabá] is 'loss/lack'. Then 'just/only loss' will be [ɗabama] which also means 'disappearance' unless we mark the tone, to distinguish [ɗábama] 'just/only loss', and [ɗabáma] 'disappearance', which cannot be marked in the current orthography. Writing *ma* alone disjunctively without its initial vowel may solve this while keeping with the Southeastern dialect.

An additional dialectal issue regarding *-uma* is that in some parts of the West, they will take a different path to disambiguate. If we suffix *-uma* to [k'ará] 'sharp' and [k'árá:] 'stalk', for example, both will become [k'áruma]. On the other hand, the examples in (12)-(13) illustrate the two words in equative sentences.

(12) *Billawni kun qara.*
      billaw-ní   kun   k'ár-a
      dagger-NOM this.M sharp-PCM
      'This dagger is sharp.'

(13) *Kun qaraa dha.*
      kun   k'ára: ɗ-a
      this.M stalk   0-PCM
      'This is a stalk.'

Thus when *-uma* is suffixed to *k'áraa*, which has a final long vowel, the epenthetic [ɗ] without the following *a* is added, leading to [k'ára:ɗuma]. However, this eliminates one ambiguity and creates another one. When you pronounce [k'ára:ɗuma] there is a gap between [k'ára:] and [ɗuma] which also means 'being finished'. Besides, this does not disambiguate the case of short vowel-finals like *ni+ -uma* and *nu+ -uma* both of which end up as [numa], as indicated in (11) above. This ambiguity further supports adopting the Southeastern dialect that leaves the stem intact but deletes *-u* from *-uma* and suffix only *-ma*. Thus, /k'ára + -uma/ → [k'árama], /k'ára: + -uma/ → [k'ára:ma].[20]

---

[20]This version also has its own ambiguity because *-ma* is a passive voice marker on verbs unless written separately.

6 Conclusion

Oromic is an agglutinating language with affixing on the right side of words with virtually no prefixes or infixes. Yet, one of the pertinent undecided issues in the Oromic orthography is word boundary determination, meaning which morphemes to write conjunctively as affixes and which ones to write separately, with space between the affixes and the host words. This decision mainly concerns clitics, which are phonologically attached to a given host word even when they stand individually from a morphosyntactic viewpoint.

Establishing the difference between clitics that could better be written disjoined rather than adjoined to the host word has been addressed using the criteria put forward by Kutsch Lojenga (2014) as the main tool. A sample list of the clitics randomly written conjunctively with the host words has been provided, namely: *hin* 'not', *íf* 'self', *ní* FOC, and *wal* 'each other'.

The motive of this article is to suggest strategies to minimize ambiguity, which is the primary goal of standardizing orthography of a language. Possible causes of lack of convention on whether to write these clitics jointly or disjunctively are put forward. The original motto "we write as we speak", intended to emphasize the phonetic nature of the orthography, and the Ethiopic script used to write Amharic and Tigrinya as well as Oromic in the pre-1991 era, is put forward as possible causes.

Suggestions have been made, with sufficient arguments, to write the above mentioned four clitics as distinguished from the inflectional and derivational suffixes as well as single phoneme morphemes that are invariably written conjunctively. As lack of tone marking also has a share in perpetuating this ambiguity, the need to address that is also alluded to.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

E	Eastern dialect	ABEN	Autobenefactive
W	Western dialect	CONJ	Conjunction
S	Southern dialect	LV	Long vowel
SE	Southeastern dialect	PCM	Predicative case marker
SW	Southwestern dialect	RED	reduplicant

Acknowledgements

References

Andrzejewski, Bogumił W. 1957. Some preliminary observations on the Borana dialect of Galla. *Bulletin of the School of Oriental and African Studies* 19(2). 354–374. Note: This reference contains a word for the Oromo language that is now considered derogatory by the Oromo people and is no longer used.

Banti, Giorgio. 1988. Two Cushitic systems: Somali and Oromo nouns. In Harry van der Hulst & Norval Smith (eds.), *Autosegmental studies on pitch accent*, 11–50. Dordrecht: Foris.

Clamons, Robbin, Ann E. Mulkern, Gerald Sanders & Nancy Stenson. 1999. Functionalism and formalism in linguistics: Volume II: Case studies. In Michael Darnell, Edith Moravcsik, Frederick Newmeyer, Michael Noonan & Kathleen Wheatley (eds.), *Functionalism and formalism in linguistics: Volume II: Case studies*, 59–76. Amsterdam: John Benjamins.

Eaton, Helen & Leila Schroeder. 2012. Word break conflicts in Bantu languages: Skirmishes on many fronts. *Writing Systems Research* 4(2). 229–241. DOI: 10.1080/17586801.2012.744686.

Gasser, Mike. 2012. *HornMorpho 2.5 user's guide*. Research Group: Human Language Technology and the Democratization of Information. Indiana University.

Gragg, Gene B. 1982. *Oromo dictionary*. East Lansing, MI: African Studies Center, Michigan State University.

Griefenow-Mewis, Catherine J. 2001. *A grammatical sketch of written Oromo*. Cologne: Rüdiger Köppe.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80.

Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding morphology*. 2nd edn. London: Hodder Education.

Janko, Kebede H. 2007. Causative verb and palatalization in Oromo: Evidence from Oromo dialects and related Cushitic languages. *Journal of Oromo studies* 14(2). 89–104.

Kutsch Lojenga, Constance. 2014. Basic principles for establishing word boundaries. In Michael Cahill & Keren Rice (eds.), *Developing orthographies for unwritten languages*, 73–106. Dallas, TX: SIL, International.

Lloret-Romanyach, Maria-Rosa. 1988. *Gemination and vowel length in Oromo morphophonology*. Indiana University. (Doctoral dissertation).

Owens, Jonathan. 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Helmut Buske.

Rogers, Henry. 2005. *Writing systems: A linguistic approach*. Malden, MA: Blackwell Publishing.

Sinha, Suchismita & Partha H. Talukdar. 2013. Word and syllable boundary of Sylheti phonemes/ syllables. *International Journal of Computing, Communications and Networking* 2(1). 1–5.

Tucho, Yigazu, R. David Zorc & Eleanor C. Barna. 1996. *Oromo newspaper reader, grammar sketch and lexicon*. Kensington, MD: Dunwoody Press.

Tutschek, Charles. 1845. *A grammar of the Galla language*. Munich: F. Wild. Note: This reference contains a word for the Oromo language that is now considered derogatory by the Oromo people and is no longer used.

Youssouf, Tamam. 2019. The copula in the Oromo language (Oromic). *International Journal of Innovative Research and Development* 8. 46–55.

Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English N'T. *Language* 59(3). 502–513.

Chapter 21

The reflexive in Kalenjin: Its syntactic status and semantic functions

Joost Zwarts
Utrecht University

Andrew Chelimo
Moi University

Prisca Jerono
University of Nairobi

Kalenjin *kee(y)* is typically treated as an arity-reducing verbal suffix with reflexive (and reciprocal) meaning. However, we show that it is actually a separate word with a typologically unusual combination of two functions: it can be the object of a verb or preposition (contributing an arity-reducing 'middle' meaning that is broader than just reflexive and reciprocal) or it can modify pronouns (with an intensifying meaning).

1 Introduction

Kalenjin is a cluster of languages, classified as Southern Nilotic and spoken mainly in Kenya (Rottland 1982). It has a morpheme, pronounced as [kɛː], [kɛːj], or [kɛːx] across the different languages, with reflexive and reciprocal meaning. We refer to this item with *kee(y)* in this article, with an optional final glide, although this does not, strictly speaking, cover the Pökoot form [kɛːx] with its final velar fricative [x].

There is no specific study of *kee(y)*, but the typical view seems to be that it is an arity-reducing suffix deriving reflexive and reciprocal verbs (§2). However,

we argue in this paper that it is not a suffix (§3), that it is not restricted to verbs (§4), that it is not only reflexive and reciprocal (§5), and that it is often but not always arity-reducing (§6). In the resulting picture (§7), *kee(y)* shows a typologically unusual combination of two functions: it can contribute a 'middle' meaning, as the object of a verb or preposition, or it can be an intensifying modifier of pronouns. Our argumentation is based on corpus data from Endo-Markweta, complemented with constructed examples from Keiyo and Tugen, but we suspect our conclusions to be valid for Kalenjin in general.

2 Kalenjin *kee(y)*

Kalenjin has a verb-initial sentence structure, with the subject (marked tonally with nominative case) and object (unmarked, absolutive case) following in a relatively free order (Creider & Creider 1983). Various valency-increasing suffixes (like applicative and instrumental) allow oblique arguments to function as objects of the verb, but without requiring adjacency to the verb.

Kee(y) is typically adjacent to the verb and this might be the reason that, it is usually discussed in the literature as part of the verbal morphology, as a suffix. That is what we find in Rottland's overview of Kalenjin, but also in the grammars of Kipsigis (Toweett 1979) and Cherang'any (Mietzner 2016) and the specific morphological studies of Tugen (Jerono 2018) and Keiyo (Sitienei Jepkoech 2018). However, in their description of Nandi, Creider & Creider (1989) treat *kee(y)* as a 'particle' and write it separately from the verb, while Rottland also considers a clitic status for *kee(y)*, given that its vowel does not harmonise with the verb in advanced tongue root (ATR) (Rottland 1982: 229). The orthographies of Kalenjin languages show similar divergences. To illustrate, the Kalenjin Union Bible has no space before *kee(y)*, unlike the Bible translations into Marakwet, Pökoot, and Sabaot. Clearly, there is no agreement about the morphosyntactic status of *kee(y)* among writers and scholars of Kalenjin.

The literature also specifies that *kee(y)* can have both reflexive and reciprocal uses. In his overview of reflexive/reciprocal polysemies in African languages, Heine (2000) also mentions Kalenjin *kee(y)*, not only with these two meanings, but also with a *middle* function (Kemmer 1993). He refers to Toweett (1979: 336), but no middle examples can be found there, unfortunately. On the other hand, Heine does not count Kalenjin among the languages where the reflexive has an additional *emphatic* use (as in English, in *he did it* HIMSELF), but Mietzner (2016: 76) points to 'reflexive personal pronouns' like *íɲéé-kêy* 'he himself'. Clearly, some empirical clarification is needed with respect to the range of functions of

kee(y), because only then can we be sure how Kalenjin fits into the typology of African (and other) languages in this respect.

For these reasons, we started to explore the variety of uses of *kee(y)* in a fairly large corpus of one particular Kalenjin variety, namely the New Testament translated into Endo-Marakwet (the northern variety of Marakwet), and determined some properties in more detail through constructed examples in two other varieties (Keiyo and Tugen, spoken by the second and third author, respectively). The reflexive is *keey* in Endo-Marakwet and Tugen and *kee* in Keiyo.

All examples are given in the orthography of the Endo-Marakwet Bible. Tone is not represented. The consonants are spelled as in Swahili, where <ch> is used for [c], <ng'> for [ŋ], <ny> for [ɲ], and <y> for [j], but <j> (for ɟ), <b>, <g>, and <d> are never used, because it is always the underlying voiceless phoneme (<ch>, <p>, <k>, <t>, respectively) that is represented. The short vowels <a>, <e>, <i>, <o>, and <u> have long variants (<aa>, <ee>, <ii>, <oo>, <uu>). ATR is represented by a macron on the non-high vowels (<ā>, <ē(ē)>, <ō(ō)>); the long +ATR version of <a> is pronounced and written as <oo>. As a result, readers will sometimes encounter '+ATR words' in this paper without any macrons, like *kimwoochi* and *tiipik* in (24). Also, in line with our analysis, the morpheme *kee(y)* is written as a separate word, and not with a hyphen, unless it is clearly part of a complex word.

The New Testament in Endo-Marakwet contains 1815 relevant occurrences of the string *keey*, mostly written as a separate word (1653 tokens).[1] In the remaining 162 non-separate occurrences, *keey* is part of the verb *karkeey* 'be like' and the nominal stems *cheepaykeey* 'prostitute(s)' and *peerkeey* 'virgin(s)', that we will consider later on. We did not specifically search for the variant *kēēy* with ATR vowel *ēē* (represented through the macron), because it is known that *kee(y)* is opaque for ATR vowel harmony (e.g., Lodge 1995) (although there are a few relevant occurrences of *kēēy*, e.g., in a word like *kārkēēyin* 'likeness', that we will return to in the next section).

3 *Kee(y)* is not a suffix

If *kee(y)* were a suffix, then it is always the last one, following all other suffixes. Of course, one suffix has to be the final one, but there are two important reasons

---

[1]The word *kumwoochikeey* 'they said to one another' is accidentally spelled without a space, and should have been *kumwoochi keey*. The forms *keeyaat* (*kee-yaat*) and *kikeeyeeng'* (*ki-kee-yeeng'*) are the only two hits in which the string *keey* does not correspond to the reflexive morpheme, but these are irrelevant for the current analysis because they involve the morpheme *kee-* before *y*.

not to treat *kee(y)* as the final suffix of the verb. The first reason is that words can come between the verb and *kee(y)* (1). The adverb *nyuun* 'then' in (1a) provides the clearest example, but the full subject pronoun *aneen* 'I' in (1b) also illustrates the point because *aneen* is neither a suffix itself, nor is this a case where *aneen* is emphatically modified by *keey* (because *aneen* and *keey* are two distinct arguments of the verb here).

(1) Endo-Marakwet

    a. kaa-kee-syaak nyuun keey
        RP-1PL-judge then REFL
        'we judged ourselves then'

    b. a-kuskuus   aneen keey
        1SG-make.weak 1SG   REFL
        'I make myself weak'

The second indication that *kee(y)* is not a suffix is its opacity for vowel harmony (e.g., Lodge 1995). Notice first how the 1PL object suffix *-eech* behaves in *neet-eech* 'teach us' and *toorēt-ēēch* 'help us'. In the latter word, +ATR spreads from the verb root *toorēt* 'help' to *-eech* and makes it *-ēēch*. With the –ATR verb *neet* 'teach', the suffix remains *-eech*. In contrast, with the same two verbs, *keey* keeps its –ATR vowel not only in *neet keey* 'teach oneself', but also in *toorēt keey*. This non-spreading behaviour was the reason that Rottland (1982: 229) considered a clitic status for *kee(y)*.[2] *Kee(y)* is similar in this respect to pronominal forms accompanying verbs as subjects, in (2a), or as possessive (2b) and demonstrative (2c) pronominals with nouns:

(2) Endo-Marakwet

    a. *āmāraa* 'I want', *kichāmēēchaan* 'we love', *āchāmēēkwaan* 'you (pl.) love', *keemwaanyēēn* 'you (sg.) say'

    b. *wēēchiikwaak* 'their brothers', *mālāktiing'waang'* 'their reward', *koonēētiisyeenyiin*[3] 'his teaching'

    c. *kookeelyaanoonēē* 'that star', *taapukeechoochēē* 'those flowers', *paannyēēnyi* 'this moment'.

---

[2]"Der Vokal war schon im PK [Proto-Kalenjin] harmonieneutral, d.h., das Suffix (Enklitikon?) gehörte nicht eindeutig zum Wortverband." [The vowel was already in Proto-Kalenjin harmony neutral, i.e., the suffix (enclitic?) did not unambiguously belong to the word domain.] If *kee(y)* is in fact a clitic, an equals sign would be the appropriate boundary symbol in interlinear glosses.

[3]Notice that the morpheme *-nyiin* is not just opaque, but its -ATR feature even affects the preceding vowel.

These examples are not meant to suggest that *kee(y)* is pronominal, but to illustrate that morphemes can be very close to the verb or noun without being in its ATR domain, whatever the precise phonological characterisation of that domain may be.[4] *Kee(y)* contrasts in this respect with the contemporative/sociative suffixes *-yō* and *-sōōt*, which do share their +ATR value with the preceding verb root: e.g., *ng'ēētyō* and *ng'ēētsōōt* are derived from the -ATR verb *ng'eet* 'stand up'. A verb with one of these suffixes is plural in the sense that it is only compatible with a plural subject (Zwarts et al. 2023).

Lexicalisation might create words that contain the reflexive *kee(y)* as a part of them. Given its meaning, it is not inconceivable that the word *karkeey* 'be like' originally had *keey* contributing a reciprocal meaning. The combination is functioning as one word now, with a non-transparent meaning, allowing derivations like *karkayiit* 'become like' and *kārkēēyin* 'likeness'. *Keey* is no longer a separate word here and it fully participates in lexical phonological processes affecting its vowel (shortening and ATR harmony, respectively). Also, the word itself can take *kee(y)* as an object, as shown in the Keiyo example (3).[5]

(3) Keiyo
    i-kerkeéy   keè   laak-o-chu
    CL2-be.like REFL child-PL.DEF-these
    'These kids are copying each other (to look the same).'

Other potential examples of words with 'incorporated' *kee(y)* are *cheepaykeeyaan* 'prostitute' (maybe with *pay* 'feed'), *peerkeeyaan* 'virgin' (maybe with *peer* 'keep whole'), and *kimekeyaan* or *kipekeyaan* 'selfish person' (with unidentifiable verbs).

Even though *kee(y)* is not an affix, it is not an independent word either, i.e. it cannot be used to start a sentence, or as a full utterance all by itself. Unlike the English reflexive and reciprocal pronouns, for instance, it cannot be used as the answer to a question, but it needs to be part of a sentential answer (4). This is in line with *kee(y)* being enclitic: even though postverbal constituents have a relatively free word order, the position of *kee(y)* is rather rigidly right-adjacent to the verb (apart from a few discourse markers).

---

[4]Whether *kee(y)* is pronominal or not, and whether and how we could decide that, is a separate discussion that we do not address in this paper.

[5]The class 2 prefix *i-* (Rottland 1982: 123) marks the verb as causative and transitive here. Note also the segmental and tone differences between the two *kee(y)*'s here.

(4) Keiyo

    a. Ki-ng'war ng'oo cheepyoos-ee? Ki-ng'war kee.
       DP-scratch who woman-DEF DP-scratch REFL
       'Who did the lady scratch? (She scratched) herself.'

    b. Ki-iim ng'oo piichooto? Ki-iim kee.
       DP-annoy who people-those DP-annoy REFL
       'Who did those people annoy? (They annoyed) each other.'

4 *Kee(y)* is not only for verbs

Not only is *kee(y)* not a suffix, it is not exclusively verbal either. It can also occur with nominalised verbs (5). The morpheme *-aa(p)* in these examples functions like a preposition ('of').

(5)   a. Endo-Marakwet
       las-at-aa keey
       praise-NMLZ-of REFL
       'praise of oneself, i.e. pride'

    b. Endo-Marakwet
       riip-ot-oo keey
       guard-NMLZ-of REFL
       'guarding of oneself, self-control'

    c. Tugen
       toorēt-ēēt-aap keey
       help-NMLZ-of REFL
       'the helping of each other, collaboration'

    d. Tugen
       wiirun-ēēt-aap keey
       throw-NMLZ-of REFL
       'the falling, lit. throwing of oneself'

Kee(y) is external to the nominalisation in (5). The example in (6), where *kee(y)* is *internal* to the nominalisation *-nat*, seems to point to a suffixal status of *kee(y)* at first sight.

(6) Tugen
las-keey-nat-ēēt
praise-REFL-NMLZ-TH.DEF
'praise of oneself, i.e. pride'

However, as we already pointed out, with strongly lexicalised forms, this is what we might expect. The combination *las keey* 'praise oneself' is such a form with a non-transparent meaning ('boast').

Kee(y) is even possible with nouns that are not deverbal, to indicate a possessive (7).[6]

(7) Tugen
ki-sooman Kiptuum ak   Kipēēt kitaapuu-syek-aap keey.
DP-read   Kiptum and Kibet book-PL.DEF-of   REFL
'Kiptum and Kibet read their own/each other's books.'

We also find *kee(y)* with free-standing prepositions, at least with *po* 'of', as illustrated in (8):

(8) Tugen
ma-po keey chii.
NEG-of REFL person
'No man is an island (lit. of himself).'

However, we do not find it with other prepositions or conjunctions, like *ākōō* 'and, with':[7]

(9) Endo-Marakwet
  a. a-riir-ee   keey ākōō laak-ōō-k-wook
      1SG-weep-APPL REFL and   child-PL-DEF-2PL
      'Weep about yourself and your children'
  b. *a-riir-ee   laak-ōō-k-wook ākōō keey
      1SG-weep-APPL child-PL-DEF-2PL and   REFL
      'Weep about your children and yourself'

---

[6]The other way to form reflexive and reciprocal possessives is illustrated in (20).

[7]An anonymous reviewer suggested that the correct generalisation for the examples in this section is that *kee(y)* occurs in 'possessive' constructions. This might be a possibility, provided it is clear enough what we mean by 'possessive'.

As Mietzner (2016: 76) already showed, we also find *kee(y)* (in a different type of role, emphatic) with pronouns (10).[8] We will say more about that use in §6.

(10) Endo-Marakwet

    a. inyēēn keey.
       3SG   REFL
       '(Who came?) Only him.'

    b. a-ku-ng'ālool-chi akwaaneek keey
       and-3-talk-DAT   3PL   REFL
       'and he spoke to them privately'

5 *Kee(y)* is not only reflexive and reciprocal

We already know that *kee(y)* has both reflexive and reciprocal uses (11).

(11) Endo-Marakwet

    a. ā-ng'ālool-ēē keey.
       1SG-talk-APPL REFL
       'I am talking about myself.'

    b. a-taakwees keey
       2PL-greet   REFL
       'greet each other!'

We will not enter into the discussion of whether this is a matter of ambiguity or vagueness (see, for instance, Palmieri 2020). What is important here is that there is a broader spectrum of meaning(s) of which the reflexive and reciprocal uses are a part. Kemmer (1993) showed that languages can express these meanings through their middle voice, taken broadly as a way of marking where a verb is "intermediate in transitivity between one-participant and two-participant events" (Kemmer 1993: 3). This marking can cover not only reflexive and reciprocal meanings, but many more. For instance, the German object pronoun *sich* is treated as a middle voice marker for a range of situation types beyond the ordinary reflexive and reciprocal (12).[9] Syntactically, the verbs in (12) are transitive (taking *sich* as object), but semantically they are intransitive.

---

[8]In Kipsigis the form is *keen* with pronouns and *kee* with verbs. This final *n* might be an additional suffix. Thanks to an anonymous reviewer for pointing this out to us.

[9]Examples from Kemmer (1993), sometimes slightly adapted.

(12) a. natural reflexive, grooming: *sich anziehen* 'dress (oneself)'

   b. natural reciprocal: *sich küssen* 'kiss (each other)'

   c. emotion: *sich fürchten* 'become afraid'

   d. nontranslational motion: *sich verbeugen* 'bow'

   e. change in body posture: *sich hinlegen* 'lie down'

   f. collective: *sich sammeln* 'gather'

   g. spontaneous: *sich auflösen* 'dissolve'

   h. impersonal: *Hier tanzt es sich gut* 'One can dance well here'

Reflexive and reciprocal situations are called *natural* when they typically happen for oneself or with one another, respectively. English tends to drop the reflexive and reciprocal pronouns in these situations, as shown by the parentheses around *oneself* and *each other* in (12a) and (12b), respectively. Some middle meanings, like (12f) and (12g), are also known as 'anticausative' (but this is not a term that Kemmer used in her overview of middle meanings).

Note that we use the term *middle* here in Kemmer's sense (common in the typological literature) and not in the more specific sense that it has in the generative literature, for impersonal constructions like *This book reads well*, which do not seem relevant for *kee(y)*. It is also important to stress that languages differ in the range of uses that they express with a middle marker and the productivity with which verbs are involved in these uses. The German verbs with *sich* illustrate one particular instantiation of Kemmer's middle voice, but they are not intended to characterise a 'prototypical' middle.

With Kalenjin *kee(y)* we also find examples that are middle-like. For instance, the combination *neet keey* literally means 'teach oneself', but it usually has the non-deliberate meaning of 'learn'. This illustrates the difference between a fully transitive construction, with agent and patient role distinguished, and a construction that has essentially one role, although it is based on a transitive verb. (13a-f) presents more examples like this, with Kemmer's categories and labels.

(13) Endo-Marakwet

   a. spontaneous: *takus keey* 'drown', *wiiru keey* 'fall'

   b. collective: *rum keey*, 'gather', *ruruuk keey*, 'gather' *tuuyo keey*, 'gather', *pēēsyō keey* 'separate'

   c. non-translational motion: *walak keey* 'turn', *ng'uruuk keey* 'bend'

   d. translational motion: *wēēchi keey* 'go', *ng'unta keey* 'walk quickly'

   e. emotion keey: *iim keey* 'worry', *las keey* 'boast'

   f. grooming: *uun keey* 'wash', *laak keey* 'dress'

(14) Keiyo
  a. change of posture: *tēēp kee* 'sit down', *teleel kee* 'stand up'
  b. body activity: *sus kee* '(body) itch', *liil kee* '(teeth) ache'

These uses might be very similar to the reflexive, but they are different. An animal drowning can be described by *takus keey* (lit. 'drown oneself'), but that does not mean that the animal is agent and patient at the same time, performing an action on itself. Rather *kee(y)* marks that a transitive verb is used to describe an event that has only one participant. The broader set of uses of *kee(y)* also involves a broader type of antecedent. While antecedents of *keey* are typically animate in reflexive and reciprocal uses, we also find inanimate antencedents in its anticausative uses, for which a reflexive analysis, with agent and patient roles linked to one inanimate participant, does not make sense.

(15) Keiyo
  a. roong' kee kārātiik
    pour REFL blood-DEF
    'the blood gushed'
  b. choor kee asiis
    steal REFL sun
    'the sun rose'
  c. tum kee pēy
    pour REFL water
    'water poured'
  d. mil kee koyin
    turn REFL stone-PL
    'stones fell'

We can conclude that Kalenjin *kee(y)* has middle functions, already claimed by Heine (2000) and Toweett (1979), who, however, did not provide supporting examples. The range of examples covered by the middle varies from language to language. In Kalenjin, natural reciprocity does not seem to be expressed using *kee(y)* but with what are called the "contemporative" suffixes (Rottland 1982: 127) *-yō* and *-sōōt* (16).[10]

---

[10]An anonymous reviewer pointed out that *-yō* might be cognate with a suffix with antipassive/middle functions in Bari and other Nilotic languages. Even if this is the case, it does not seem to have those functions in Kalenjin now, see Zwarts et al. (2023).

(16) Endo-Marakwet

choomnyō 'reconcile', *chuunchuunnyō* 'disagree', *tuupchō* 'be brothers', *pēēsyō* 'disperse', *pooryō* 'fight', *tēēniityō* 'be equal', *tuuyō* 'meet'

These verbs also have transitive, causative alternants which can then be used with *kee(y)* (17).

(17) Endo-Marakwet

a. si-mē-ē-tēēniit-yō   keey nkōō chiitō   aka
    PURP-NEG-1SG-make.equal-CONT REFL with person other
    'so that you do not compare yourself with somebody else'

b. kii nēē   kaa-kuu-tuu-yō keey pi-choochēē
    DP when RP-3-meet-CONT REFL people-those
    'when those people had assembled'

In line with the broader middle semantics of *kee(y)*, (17b) can be seen as an anticausative (like German *sich sammeln* 'come together', but literally 'collect oneself').

By claiming that *kee(y)* has middle functions we do not wish to imply that it productively derives those functions with each transitive verb (like it does for the reflexive and reciprocal functions). For instance, *kee(y)* does not derive the "anticausative" alternant of every transitive verb in a productive way; only some verbs have an alternant marked with *kee(y)* that can be characterized as such.

6 *Kee(y)* is arity-reducing, but not always

Most of the uses of *kee(y)* that we have seen are arity-reducing in the following general sense. The combination $P_2+kee(y)$ of a two-place predicate P_2 with *kee(y)* results in a one-place predicate and the combination $P_3+kee(y)$ of a three-place predicate P_3 is effectively a two-place predicate. This arity-reduction is illustrated in (18).

(18) Endo-Marakwet

a. si-ku-ngʼālool-chi keey
    PURP-3-talk-DAT   REFL
    'so that he talks to himself'

b. ku-ungʼ-a   keey akwaaneek
    3-hide-APPL REFL 3PL
    'he hid himself for them'

c. chēēr-ēēn-ook   keey
    strengthen-APPL-2PL REFL
    'strengthen oneself with you'

The verb *ng'ālool-chi* '*x* talk to *y*' in (18a) is a two-place predicate and *kee(y)* saturates the internal argument *y* of this predicate, reducing it to a one-place predicate ('talk to oneself'). The verb *ung'-a* '*x* hide *y* for *z*' in (18b) is a three-place predicate. *Kee(y)* saturates argument *y* and *akwaaneek* 'them' saturates *z*. (18c) differs from (18b) in the order in which *kee(y)* and the other object saturate arguments of the verb. The verb *chēēr-ēē(n)* '*x* strengthen *y* with *z*' first gets *z* saturated with the object pronoun and then *y* with *kee(y)*. The examples also illustrate the role of verbal suffixes like dative *-chi* and applicative *-a/-ee(n)* in creating argument positions that *kee(y)* can saturate, thereby effectively reducing the arity of the verb. However, since *kee(y)* is not a verbal suffix, but a separate object, it does not operate on the verb's lexical-semantic argument structure, but it reduces arity by saturating a syntactic argument position.

    Kee(y) does not combine with one-place predicates; it is always one of the internal arguments (objects) that is saturated. Even the translational motion cases like (13d) above always seem to have a suffix that adds a non-subject slot for *kee(y)* to saturate. The basic motion verbs *wō* 'go (sg.)' and *pa* 'go (pl.)' (that are one-place) only allow *kee(y)* when the suffix *-chi* is there to license it as an object.

(19)   Endo-Marakwet

a. a-ku-wēē-chi   keey
    and-3-go.SG-DAT REFL
    'and he went away'

b. paani kaa-ku-pēē-chi keey saang'
    when RP-3-go.PL-DAT REFL outside
    'when they had gone outside'

c. ku-pa   saang'
    3-go.PL outside
    'to go outside'

(19c) shows that the goal *saang'* 'outside' is already licensed by the verb root and that *-chi* in (19) must therefore be instrumental in licensing *kee(y)*. What exactly happens in these motion descriptions with *kee(y)* is a matter for further study, but it is clear that we have no exception here to the generalisation that *kee(y)* reduces the arity of a predicate by saturating an argument position of that predicate.

The same is true for the reciprocal verbs illustrated in (17) above. Although verbs with contemporative *-yō* are usually intransitive, here they are transitive and there is an object argument that *kee(y)* saturates in (17).

Kee(y) can also saturate possessor arguments (20).

(20)   Keiyo

    a. ki-ng'wār-chin-i   kee John patay.
        DP-scratch-DAT-IPFV REFL John back

        'John was scratching his own back.'

    b. ng'wār-chin-i   kee pātooy.
        scratch-DAT-IPFV REFL backs

        'They were scratching each other's backs.'

Unlike the English pronouns, however, and the example in (7), *kee(y)* in (20) does not directly saturate the possessor argument of the noun for 'back', but it targets the additional argument created by *-chi* as part of a three-place verb *ng'wār-chin-i* '*x* scratch *y*'s *z*'.

Before turning to the non-arity-reducing use of *kee(y)*, we want to give one more piece of evidence here for our argumentation that *kee(y)* does not reduce the arity of a verb by directly operating on the verb (i.e., as a suffix), but by saturating an internal syntactic argument of the verb (i.e., by being its separate object). The reciprocal meaning of *kee(y)* can be made explicit by adding a non-ambiguous reciprocal expression, essentially a plural pronoun conjoined with itself (21).

(21)   Tugen

    a. ki-chām-ē   keey acheek eng' acheek.
        1PL-love-IPFV REFL we   and we

        'We love each other.'

    b. o-chām-ē   keey okweek eng' okweek.
        2PL-love-IPFV REFL you   and you

        'You love each other.'

    c. chām-ē   keey icheek eng' icheek.
        love-IPFV REFL they   and they

        'They love each other.'

The transitive verb *cham* 'love' requires *kee(y)* in (21). On the other hand, these reduplicated reciprocals can also occur without *kee(y)* (22), when the predicate is intransitive.

(22) Tugen

ak o-tēpii   eng' kaalyeet okweeke eng' okweeke
and 2PL-stay in peace   you   and you

'and stay in peace with one another'

The phrase *okwege eng' okwege* 'you and you' accompanies the intransitive pred-icate *otēpii eng' kaalyeet* 'stay in peace'. There is no object position here that *ok-wege eng' okwege* could saturate; it must function as an adjunct. It makes sense then that the reduplicated reciprocals in (21) are adjuncts too, while *kee(y)* is the argument of the transitive verb *cham*.

While the core uses of *kee(y)* reduce a predicate's arity by saturating an in-ternal argument, this is obviously not true for emphatic *kee(y)* in combination with pronouns (10). Unlike verbs and prepositions, pronouns do not have an ar-gument position that can be satisfied. In that use, *kee(y)* is an adjunct that adds an emphatic meaning to the pronoun.[11]

Intensifying *kee(y)* can be paraphrased as 'alone', 'by oneself', 'without help', 'personally' (König & Siemund 2000). The different paraphrases suggest a lexical field with subtle distinctions (23) that we will not explore further.

(23) Keiyo

Ā-ām-iisyēy   ani-kee.
1SG-eat-INTR-IPFV 1SG-REFL

'I am eating by myself/alone/without help.'

The modification of a pronoun with *kee(y)* also allows for more regular reflex-ive and reciprocal uses, as shown in (24), similar to the pronoun+*self* forms in English, for instance.

(24) Tugen

a. Ki-mwoo-chi Mary tiip-ik   akopo ichee-keey.
   DP-tell-DAT Mary girl-DEF about 3SG-REFL

   'Mary told the girls about themselves.'

---

[11]It is important to distinguish the emphatic *kee(y)* from a demonstrative marker that is often deceptively similar in some Kalenjin languages. In Cherang'any it is *kɛy* and treated as a 'post-final referential marker', marking a referent that is 'just mentioned' (Mietzner 2016: 165). It is part of a larger series of demonstratives with a temporal dimension. In Endo-Marakwet its form is *kay*. Given that there has been a conflation of short *ɛ* with *a* in Endo-Marakwet, we can conclude that there are two distinct items in Kalenjin: reflexive/emphatic *kee(y)* and demon-strative *key*). What also distinguishes the two is that the emphatic *kee(y)* seems to be restricted to pronouns, while the demonstrative *key* combines with noun phrases more generally.

b. ki-sooman kitapuu-syek-aap ichee-keey.
   DP-read   book-PL.DEF-of   3SG-REFL
   'They read each other's books.'

7 Conclusion

We have shown that the syntactic status of Kalenjin *kee(y)* is not that of a suffix, but a separate word, although probably an enclitic. It has two semantic functions. It is usually an argument of a verb, contributing a reflexive, reciprocal, or (other) middle meaning, and sometimes an adjunct of a pronoun, with an emphatic/intensifying meaning. The resulting empirical picture clarifies the representation of *kee(y)* in the literature about Kalenjin and it has consequences for the broader typology of reflexivity, reciprocity, and intensification.

It seems that Kalenjin does something unusual by using one and the same item for three meanings: reflexive, middle, and emphatic. We already saw that German *sich* is used reflexively, but also as a middle marker, while it has a different emphatic item (*selbst*). On the other hand, Juang (Austroasiatic) uses its reflexive morpheme also both emphatically and reciprocally (Patnaik & Subbarao 2000: 843), but middle meanings are not mentioned. In her typological overview of Oceanic, Moyse-Faurie (2017) makes clear that these languages adhere to the generalisation made in König & Siemund (2000: 59): "If a language uses the same expression both as intensifier and reflexive anaphor, this expression is not used as a marker of derived intransitivity." *Derived intransitivity* is what we called *middle.* Also Kemmer (1993: 196) writes: "It appears that either emphatic reflexive/direct reflexive polysemy or direct reflexive/middle polysemy is permitted, but not a three way polysemy involving all of these situation types." Clearly, the polysemy of Kalenjin *kee(y)* requires a reconsideration of these claims, as well as a closer look into the situation of Kalenjin, both synchronically and diachronically.

The grammaticalisation perspective on the broader middle domain (Kemmer 1993, Heine 2000) would lead us to expect that *kee(y)* originates from a noun (e.g., for 'body') that has developed into a middle marker through an intermediate emphatic and then reflexive use and that increasing grammaticalisation also forces the emphatic to be renewed again. Unfortunately, we have no clues about a nominal origin of *kee(y)*. Outside of Kalenjin (but still within Southern Nilotic), both Datooga and Omotik have a singular/plural contrast (Rottland 1982: 151,191). However, this does not necessarily suggest a nominal source for *kee(y)*, because

categories other than nouns show number contrasts too.[12] There are indications that emphatic *kee(y)* is being renewed, if we can interpret various competing items with similar emphatic meanings (*kipaat*, *ākityōōn*) as such, at least.

Our study of Kalenjin *kee(y)* is also relevant for what Safir & Selvanathan (2016) propose for "transitive reciprocal constructions" (as they call it), in Niger-Congo, which involve an ambiguous object (reflexive/reciprocal) that gets its reciprocal interpretation from a marker (overt or covert) in the verbal morphology ("little *v*"). Since we have argued *kee(y)* to be an object with that type of ambiguity, the question arises how this Nilo-Saharan element would fit into this proposal. We have found no relevant cases in Kalenjin where *kee(y)* expresses reciprocity together with reciprocal verbal morphology. The examples in (17) are irrelevant, because *kee(y)* is not reciprocal there, but middle-like. For a better understanding of transitive reciprocals/reflexives, it would be useful to include both Niger-Congo and Nilo-Saharan languages to explore a hypothesis like that of Safir and Selvanathan.

Another important question for future research is how our observation that *kee(y)* has middle voice functions fits into a broader and deeper analysis of verb alternations in Kalenjin, including the recent work of Kouneli (2021) on such alternations in Kipsigis. The empirical breadth of such an analysis should involve the division of labor between *kee(y)* and the suffix *-ak* ('stative' and 'potential passive' in Rottland 1982, but treated as a middle by Kouneli) and the classification of verbs as class 1 or 2. With multiple morphosyntactic devices involved in transitive/intransitive alternation, the question becomes acute how their operations might differ. The theoretical depth of such an analysis needs to address how these devices are syntactically implemented, not only in syntactic heads like little *v* or Voice, but also in other ways, doing justice to the richness of verbal alternations in Kalenjin.

Finally, with its narrow focus on one single word in Kalenjin this paper already has general typological ramifications. Nevertheless, it would be useful to zoom out and include other Nilotic languages, which have different systems of marking intensifying, reflexive, reciprocal, and middle meanings to get a richer picture of the connections and encodings of these meanings, in African languages, and beyond.

---

[12]Thanks to Gertrud Schneider-Blum and Alice Mitchell for pointing out to us the complexities of the issue here.

Abbreviations

Abbreviations in this chapter follow the Leipzig Glossing Rules, with the following additions.

CL2	class 2	RP	recent past
CONT	contemporative	TH	theme vowel
DP	distant past		

Acknowledgements

References

Creider, Chet A. & Jane Tapsubei Creider. 1983. Topic-comment relations in a verb-initial language. *Journal of African Languages and Linguistics* 5. 1–15.

Creider, Chet A. & Jane Tapsubei Creider. 1989. *A grammar of Nandi* (Nilo-Saharan: Linguistic Analyses and Documentation 4). Hamburg: Helmut Buske.

Heine, Bernd. 2000. Polysemy involving reflexive and reciprocal markers in African languages. In Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reciprocals: Forms and functions*, 1–29. Amsterdam: John Benjamins.

Jerono, Prisca. 2018. Passive and antipassive in Tugen. *International Journal of Language and Linguistics* 5(3). 189–197. DOI: 10.30845/ijll.v5n3p19.

Kemmer, Suzanne. 1993. *The middle voice* (Typological Studies in Language 23). Amsterdam: John Benjamins.

König, Ekkehard & Peter Siemund. 2000. Intensifiers and Reflexives: A typological perspective. In Zygmunt Frajzyngier & Traci S. Curl (eds.), *Reflexives: Forms and functions* (Typological Studies in Language 40), 41–74. Amsterdam: John Benjamins.

Kouneli, Maria. 2021. Low (in)transitivity: Evidence from Kipsigis. In Alessa Farinella & Angelica Hill (eds.), *Proceedings of the Fifty-First Annual Meeting of the North East Linguistic Society (NELS 51)*, 53–66. Amherst: GLSA.

Lodge, Ken. 1995. Kalenjin phonology and morphology: A further exemplification of underspecification and non-destructive phonology. *Lingua* 96(1). 29–43.

Mietzner, Angelika. 2016. *Cherang'any: A Kalenjin language of Kenya*. Cologne: Rüdiger Köppe.

Moyse-Faurie, Claire. 2017. Reflexives markers in Oceanic languages. *Studia Linguistica* 71(1-2). 107–135. DOI: 10.1111/stul.12065.

Palmieri, Giada. 2020. On the ambiguity between reflexivity and reciprocity in Italian. In Mariam Asatryan, Yixiao Song & Ayana Whitmal (eds.), *Proceedings of the Fiftieth Annual Meeting of the North Eastern Linguistic Society (NELS 50)*, vol. 2, 279–289. Amherst: GLSA.

Patnaik, Manideepa & K. V. Subbarao. 2000. An initial note on lexical anaphors and pronouns in Juang. In Barbara Lust, Kashi Wali, James W. Gair & K. V. Subbarao (eds.), *Lexical anaphors and pronouns in selected South Asian languages: A principled typology*, 841–860. Berlin, New York: Mouton de Gruyter.

Rottland, Franz. 1982. *Die Südnilotischen Sprachen: Beschreibung, Vergleichung und Rekonstruktion* (Kölner Beiträge zur Afrikanistik 7). Berlin: Reimer.

Safir, Ken & Naga Selvanathan. 2016. Niger-Congo transitive reciprocal constructions and polysemy with reflexives. In Doris L. Payne, Sara Pacchiarotti & Mokaya Bosire (eds.), *Diversity in African languages: Selected papers from the 46th Annual Conference on African Linguistics* (Contemporary African Linguistics 1), 495–512. Berlin: Language Science Press.

Sitienei Jepkoech, Nancy. 2018. *The morphosyntax of the Keiyo verb*. Eldoret: Moi University. (MA thesis).

Toweett, Taaitta. 1979. *A study of Kalenjin linguistics*. Nairobi: Kenya Literature Bureau.

Zwarts, Joost, Andrew Chelimo & Prisca Jerono. 2023. Two verbal plurals in Endo-Marakwet. *Arusha Working Papers in African Linguistics* 5(1). 51–67.

Name index

ACAL in SoCAL

This volume contains a selection of papers that were presented at the *53rd Annual Conference on African Linguistics*, which was held virtually at the University of California San Diego. There are 21 papers covering phonology, morphology, syntax, lexical semantics, sociolinguistics, typology and historical linguistics. The volume features a keynote paper that proposes a novel community-based approach to language documentation. African languages investigated in detail include Wolof, Mende, Dangme, Kusaal, Nzema, Anii, Nigerian Pidgin, Tunen, Nyokon, Vale, Lokoya, Lopit, Otuho, Kalenjin, Tiriki, Oromo, Tigrinya, Asá, Qwadza, and Ikalanga.