Morphosyntactic variation in East African Bantu languages

Descriptive and comparative approaches

Edited by

Hannah Gibson

Rozenn Guérois

Gastor Mapunda

Lutz Marten

Contemporary African Linguistics 8

language science press

Contemporary African Linguistics

Editors: Michael R. Marlo, Laura J. Downing

In this series:

1. Payne, Doris L., Sara Pacchiarotti & Mokaya Bosire (eds.). Diversity in African languages: Selected papers from the 46th Annual Conference on African Linguistics.
2. Persohn, Bastian. The verb in Nyakyusa: A focus on tense, aspect and modality.
3. Kandybowicz, Jason, Travis Major & Harold Torrence (eds.). African linguistics on the prairie: Selected papers from the 45th Annual Conference on African Linguistics.
4. Clem, Emily, Peter Jenks & Hannah Sande (eds.). Theory and description in African Linguistics: Selected papers from the 47th Annual Conference on African Linguistics.
5. Lotven, Samson, 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.
6. Sibanda, Galen, Deo Ngonyani, Jonathan Choti & Ann Biersteker (eds.). Descriptive and theoretical approaches to African linguistics: Selected papers from the 49th Annual Conference on African Linguistics.
7. Akinlabi, Akinbiyi, 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.
8. Gibson, Hannah, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.). Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches.

ISSN: 2511-7726

Morphosyntactic variation in East African Bantu languages

Descriptive and comparative approaches

Edited by

Hannah Gibson

Rozenn Guérois

Gastor Mapunda

Lutz Marten

language science press

Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.). 2024. *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches* (Contemporary African Linguistics 8). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/383
© 2024, the authors
Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):
http://creativecommons.org/licenses/by/4.0/
ISBN: 978-3-96110-433-8 (Digital)
978-3-98554-091-4 (Hardcover)

ISSN: 2511-7726
DOI: 10.5281/zenodo.10453704
Source code available from www.github.com/langsci/383
Errata: paperhive.org/documents/remote?type=langsci&id=383

Cover and concept of design: Ulrike Harbort
Typesetting: Yanru Lu, Sebastian Nordhoff
Proofreading: Alena Witzlack Amir Ghorbanpour, Elliott Pearl, Jeroen van de Weijer, Keira Mullan, Lotta Aunio, Ludger Paschen, Mary Anna Walter, Nina van der Vlugt, Patricia Cabredo, Prisca Jerono, Raquel Benítez Burraco, Tom Bossuyt, Troy E. Spier
Fonts: Libertinus, Arimo, DejaVu Sans Mono, Source Han Serif JA
Typesetting software: XeLaTeX

Language Science Press
xHain
Grünberger Str. 16
10243 Berlin, Germany
http://langsci-press.org

Storage and cataloguing done by FU Berlin

Contents

Author biographies **iii**

1 Introduction
Lutz Marten, Hannah Gibson, Rozenn Guérois & Gastor Mapunda **1**

I The nominal domain

2 Demonstratives in Chiyao: An analysis of their form, distribution and functions
Julius Taji **17**

3 The structure, distribution and function of demonstratives in Runyankore-Rukiga
Allen Asiimwe **43**

4 Concord and agreement in Eastern Bantu: The augment and noun classes in Nyakyusa
Amani Lusekelo **85**

5 The morphosyntax of locative expressions in Kiwoso
Aurelia Mallya **109**

6 A comparative study of the locative system in South-Tanzanian Bantu languages
Gastor Mapunda & Fabiola Hassan **129**

II The verbal domain

7 Multiple-object constructions in Ganda
Nobuko Yoneda & Judith Nakayiza **155**

8 Object marking in four Mozambican Bantu languages
Armindo Ngunga & Crisófia Langa da Câmara **191**

9 Verb extensions and morphosyntactic variation in Bantu: The case of Sumbwa
Kulikoyela Kahigi **229**

10 Tense and aspect marking in Bantu languages of the Morogoro region, Tanzania
Malin Petzell & Peter Edelsten **277**

III Comparative and historical approaches

11 A comparative sketch of TA markers in Kilimanjaro Bantu: In search of the directionality of semantic shift and micro-parametric correlation
Daisuke Shinagawa **325**

12 The historical development of the reflexive-reciprocal polysemy in Hehe
Lengson Ngwasi & Abel Mreta **351**

13 Morphosyntactic variation in Old Swahili
Lutz Marten, Hannah Gibson, Rozenn Guérois & Kyle Jerro **373**

Index **421**

Author biographies

Allen Asiimwe ***(ORCID 0000-0003-1906-519X)***

Allen Asiimwe is a Senior Lecturer in the Department of African Languages, School of Languages, Literature and Communication at Makerere University in Uganda. Her research interests lie in morphology and syntax, and the interface between information structure and syntax in the Bantu language cluster of Runyankore-Rukiga. She was one of the collaborators on the project on 'Bantu Syntax and Information Structure' (BaSIS) (https://bantusyntaxinformationstructure.com/) investigating the influence of information structure on the syntax of Rukiga (Bantu, JE 14). She also has a passion for documentation and preservation of cultural heritage of the Bakiga. She recently documented Rukiga personal names and conducted a linguistic analysis of these names. The ultimate goal of this project was to come up with an encyclopaedic collection of Rukiga personal names preserved for future reference, useful to scholars in linguistics, gender studies, and anthropology, but also to contribute to the debate on African onomastics.

Peter Edelsten

Peter Edelsten is a Research Associate in Linguistics at SOAS University of London. He has a background in mathematics, agriculture, finance and linguistics. His principal linguistic publication is as co-author, with Chiku Lijongwa, of a *Grammatical Sketch of Chindamba*, an endangered language of southern Tanzania (Cologne: Köppe, 2010). Since 2016 he has provided mathematical and computing assistance to the project 'Morphosyntactic variation in Bantu: Typology, contact and change' housed in the Linguistics Department at SOAS under the direction of Lutz Marten, and contributed to publications generated by the project.

Hannah Gibson ***(ORCID 0000-0003-2324-3147)***

Hannah Gibson is a Professor of Linguistics at the University of Essex. She holds a PhD in Linguistics from SOAS University of London. Before joining Essex, she

was a Japan Society for the Promotion of Science postdoctoral researcher at Osaka University in Japan and a British Academy Postdoctoral Research Fellow at SOAS University of London. Her research examines linguistic variation with a focus on the morphosyntax of the Bantu languages of Eastern and Southern Africa. She also works on language contact, language change and multilingualism. Hannah is involved in a number of externally-funded research projects, including as Principal Investigator on the project 'Grammatical variation in Swahili: contact, change and identity' (2021-2025) and 'Microvariation in youth language practices in Africa' (2021-2024). She has carried out research in Tanzania, Kenya and South Africa.

Rozenn Guérois ***(ORCID 0000-0003-3602-8622)***

Rozenn Guérois is a Researcher at the laboratory LLACAN (Langage, Langues et Cultures d'Afrique) of the French National Centre for Scientific Research (CNRS), and an affiliated researcher at the Department of Linguistics of the University of KwaZulu-Natal (UKZN). She received her PhD from the University of Lyon 2 in 2015. Before joining the CNRS in 2020, she was a postdoctoral researcher at SOAS University of London (2015-2018) and Ghent University (2018-2020). Her research interests include the description of Bantu languages spoken in Mozambique, morphosyntactic variation, typology, and comparative and historical linguistics. She has recently obtained an ANR Grant (French National Research Agency) for the 'OriKunda' project (2023-2027). Since 2020, she has been a member of the committee of the journal *Linguistique et Langues Africaines* (LLA).

Fabiola Hassan ***(ORCID 0000-0003-3407-7025)***

Fabiola F. Hassan is a Lecturer of Kiswahili and Bantu linguistics in the Department of Kiswahili at the University of Dodoma, Tanzania. She holds a PhD from the University of Dar es Salaam. Her areas of specialization include phonology, morphology, syntax, semantics, and lexicography. Her doctoral thesis is titled *Upinduzi wa Kimahali katika Lugha ya Kiswahili* ('Locative Inversion in Swahili'). She has published a number of articles and books on historical and comparative linguistics, psycholinguistics, applied linguistics and descriptive linguistics. Currently her duties include teaching, research, and consultancy in Tanzania. She is interested in conducting research on Kiswahili and other Bantu languages.

Kyle Jerro ***(ORCID 0000-0002-1467-2524)***

Kyle Jerro is a Senior Lecturer at the University of Essex. They are interested in semantics, syntax, and the interfaces between syntax, semantics, and pragmatics. In particular, their focus has been on the question of possible verb meanings and how the meaning of a verb derives argument realization, especially on the semantics of valency changing morphology like applicatives and causatives in Bantu languages. They have also conducted research in the areas of aspect and temporal reference, information structure, the semantics of location and directed motion, reference and fictional worlds, and the psycho-semantic reality of noun classes in Bantu languages.

Kulikoyela Kahigi

Kulikoyela Kahigi is Associate Professor of Linguistics in the Department of Languages and Linguistics, Saint Augustine University of Tanzania (SAUT), Mwanza. Before moving to SAUT, he taught Swahili language and linguistics at the University of Dar es Salaam (UDSM) for many years. While at UDSM, he was, at various times, Chief Editor of the journals *UTAFITI*, *KIOO CHA LUGHA* and *KISWAHILI*. He was also involved in the Microsoft Kiswahili localisation project as Chief Moderator and Language Expert (2004-8), and was part of the Management Committee, Research Team and Editorial Panel of the 'Languages of Tanzania' (LOT) project which published 20 dictionaries and the *Atlasi ya Lugha za Tanzania* ('Atlas of the languages of Tanzania', Dar es Salaam: LOT, 2009). His current interests are in Bantu linguistics (synchronic and diachronic), discourse analysis, stylistics, and poetry. His publications include *Sema Kiswahili: Furahia Tanzania* (Dar es Salaam: US Peace Corps Tanzania, 1997) as consultant and co-author, 'Structural and cohesion dimensions of style: A considerations of some texts in Maw 1974' (*Journal of Linguistics and Language Education* 3: 1-79, 1997), *Misingi ya Kusoma* ('Foundations of Reading') (Dar es Salaam: Children's Book Project Tanzania, 2003), *Sumbwa-English-Swahili/English-Sumbwa-Swahili dictionary* (Dar es Salaam: LOT, 2008), *Kahe-English-Swahili/English-Kahe-Swahili lexicon,* Dar es Salaam: LOT, 2008); *Misingi ya Semantiki na Pragmatiki ya Kiswahili* (Dar es Salaam: KAUTTU, 2019). His most recent collections of poetry include *Mageuzi* (Dar es Salaam: KAUTTU, 2019) and *New Horizons* (Dar es Salaam: E & D Vision Publishers, 2018).

Crisófia Langa da Câmara ***(ORCID 0000-0001-5786-991X)***

Crisófia Langa da Câmara is a researcher at the Department of Language and Communication of the Centro de Estudos Africanos at the Universidade Eduardo Mondlane, Mozambique. Her research focuses on the description of Mozambican Bantu languages, minority and endangered language documentation, morphology and syntax of Bantu languages, language policy, multilingualism and gender. Her research work also involves the development of teaching manuals and bilingual education teacher training materials. She has published one book, several journal articles and book chapters, and edited one book.

Amani Lusekelo ***(ORCID 0000-0001-6275-237X)***

Amani Lusekelo is an Associate Professor of Linguistics at the University of Dar es Salaam. He holds a PhD (African Languages and Literature) from the University of Botswana and MA (Linguistics) from University of Dar es Salaam in Tanzania. He teaches graduate courses covering the structure of African languages, contact linguistics, field methods, and advanced morphology. His research areas include the noun phrase, lexical borrowing, ethnobotany, linguistic landscape, and tense and aspect, focusing mainly on Akie, Datooga, Hadzabe, Nyakyusa, Nyamwezi, Sandawe, Sukuma and Swahili. Since 2013, he has supervised to completion ten doctoral candidates and currently has six ongoing PhD students at the University of Dar es Salaam and University of Dodoma. His selected authored and co-authored publications include *Plant Nomenclature and Ethnobotany of the Hadzabe Society of Tanzania* (Cape Town: University of Western Cape, 2023), 'Nominal morphology and syntax in Rwa-Meru' (in Blasius Achiri-Taboh (ed.), *The Bantu Noun Phrase: Issues and Perspectives*, 78-95. London: Routledge, 2023), 'Portrayal of the COVID-19 pandemic in political cartoons in Tanzania' (*Cogent Arts and Humanities* 10(1), 2023), 'The V and CV augment and exhaustivity in Kinyakyusa' (*Studies in African Linguistics* 51(2): 323-345, 2022), 'Locating the Hadzabe in the wilderness' (*Utafiti* 17(1): 130–154, 2022), 'Linguistic and social outcomes of interaction of Hadzabe and Sukuma in north-western Tanzania' (*Utafiti* 15(1): 348–373, 2020), and 'African linguistics in eastern Africa' (in H. Ekkehard Wolff (ed.), *A History of African Linguistics*, 133–152. Cambridge: CUP, 2019).

Aurelia Mallya ***(ORCID 0000-0001-7842-1022)***

Aurelia Mallya is a Senior Lecturer in the Department of Foreign Languages and Linguistics, College of Humanities, University of Dar es Salaam, Tanzania. Her research focus is in syntax, morphology, lexical semantics, and their interfaces,

particularly in Bantu languages. She is also interested in language documentation and descriptive linguistics. Dr Mallya completed her Bachelor's and Master's degrees at the University of Dar es Salaam, and her PhD at Stellenbosch University in South Africa. Some of her recent publications include 'Passive and anticausative constructions: A new perspective to morphosyntax and lexical semantic interfaces' (*Southern African Linguistics and Applied Language Studies* 37(4): 315-324, 2019) and 'Locative-subject alternation constructions in Kiwoso' (*Ghana Journal of Linguistics* 9(2): 1-21, 2020). She is currently the Head of the Department of Foreign Languages and Linguistics, University of Dar es Salaam.

Gastor Mapunda ***(ORCID 0000-0001-5683-0175)***

Gastor Mapunda is Associate Professor of English and Linguistics in the Department of Foreign Languages and Linguistics, University of Dar es Salaam. He completed his Bachelor's and Master's degrees at the University of Dar es Salaam, and his PhD at the Universities of Dar es Salaam and Bristol on a split-site arrangement. His areas of research are applied linguistics, sociolinguistics, and discourse analysis. His doctoral research was on classroom interaction in multilingual communities. He has also researched extensively on language contact. He has served in various positions, including deputy chair of the Governing Council of the Tanzania Institute of Education (TIE) (2017-2021), member of the National Technical Committee for Curricula Review (2022-2023), and convenor of the African Humanities Conference 2023 held at the University of Dar es Salaam. His current research ties are with colleagues from the Universities of Dar es Salaam, Essex, and St. John's University of Tanzania, among others. His most recent publications include '"You must be crazy!" Teacher corrective feedback and student uptake in two Tanzanian secondary schools' (*Journal of Modern Research in English Language Studies* 10(4): 1-20, 2023) and 'A morphological analysis of Kemunasukuma personal names' (*Linguistik Online* 123(5): 95–113, 2023).

Lutz Marten ***(ORCID 0000-0002-2289-5701)***

Lutz Marten is Professor of General and African Linguistics at SOAS University of London, where he has served as Dean of the Faculty of Languages and Cultures and as Head of the SOAS Doctoral School. He is interested in linguistic theory, comparative and historical linguistics, and questions of language and identity. Most of his work focuses on African languages and he has conducted research in Eastern and Southern Africa. His publications include *At the Syntax-Pragmatics Interface* (Oxford: OUP, 2002), *A Grammatical Sketch of Herero* (with

Jekura Kavari and Wilhelm Möhlig, Cologne: Köppe, 2002), *The Dynamics of Language* (with Ronnie Cann and Ruth Kempson, Elsevier, 2005), and *Colloquial Swahili* (with Donovan McGrath, London: Routledge, 2003/2012). He is the Founding Chair of the International Conference on Bantu Languages and he is the current editor of the *Transactions of the Philological Society*, the oldest scholarly journal devoted to the general study of language and languages with an unbroken tradition.

Judith Nakayiza ***(ORCID 0000-0002-7302-1932)***

Judith Nakayiza is a researcher with a specific interest in sociolinguistics, language planning and policy (ideology, beliefs and identity), language practices in diverse or multilingual spaces, the sociolinguistics of Bantu languages and Ugandan languages. She previously worked as a lecturer in the School of Languages, Literature and Communication, Makerere University, Uganda for over 15 years. She has also conducted post-doctoral research at SOAS University of London, funded by the Commonwealth Scholarship Commission. She is currently an Independent Researcher with affiliations to SOAS University of London, UK and Makerere University, Uganda. Some of her publications present research on the sociolinguistic situation of English, implementing language rights in multilingual Uganda, language attitudes and Ideologies in Uganda, the sociolinguistics of multilingualism in Uganda, Luganda in Uganda and literacy acquisition in selected Ugandan rural primary schools.

Armindo Ngunga

Armindo Saúl Atelela Ngunga, PhD in Linguistics (from UC Berkeley), Chairperson of the Northern Integrated Development Agency, is Professor of Linguistics at the Eduardo Mondlane University in Mozambique, where he taught general linguistics, phonetics, phonology and morphology of Bantu languages since 1984. He has supervised several graduate and Master's dissertations and PhD theses. He served as Secretary of State in Cabo Delgado Province (2020-2021), Vice-Minister of Education and Human Development (2015-2020), Director of the Centre of African Studies (2007-2015) and Dean of the Faculty of Arts and Social Sciences at the Eduardo Mondlane University (1999-2007). He has published tens of journal articles and book chapters, and has written and edited books on grammar and lexicography of Mozambican languages. He has edited a book series called *As nossas línguas* ('Our languages'), with twelve titles since 2009.

Lengson Ngwasi ***(ORCID 0000-0001-9114-6491)***

Lengson Ngwasi is a Lecturer in Linguistics in the Department of Foreign Languages and Linguistics at the University of Dar es Salaam, Tanzania. He received his PhD in 2021 from the University of Gothenburg, Sweden. He also has a Master degree in Linguistics from the University of Dar es Salaam (received in 2016), as well as a Bachelor's degree in Arts with Education from the same university that he received in 2013. In 2016, during his MA studies, he secured a Linneaus-Palme Exchange position for six months at the University of Gothenburg. His research interest includes historical and comparative linguistics, grammaticalization, language variation, and language typology. His research works are his MA dissertation on *Reflexive marking in Kihehe* and his PhD dissertation *The multiple functions of the reflexive prefix in Hehe, Sukuma, Nilamba, and Nyaturu.* He is currently working on a paper entitled 'The typology of reciprocal marking in Hehe, Sukuma, Nilamba, and Nyaturu'.

Malin Petzell ***(ORCID 0000-0002-0774-5131)***

Malin Petzell is a Reader (Docent) in African Languages at the Department of Languages and Literatures at the University of Gothenburg, Sweden. Her speciality lies in the documentation, description and analysis of both under-described and unwritten languages, and particularly Tanzanian Bantu languages. Her work focuses primarily on verbal morphosyntax including tense and aspect, and aspectual classification of verbs. Two other areas of interest are comparative studies of closely related languages, and field methods and data management.

Daisuke Shinagawa ***(ORCID 0000-0002-9320-6520)***

Daisuke Shinagawa is Associate Professor at the Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies. He has been carrying out his descriptive field linguistic research on underdescribed Kilimanjaro Bantu languages including Rwa (E621A), Siha (E621C), Uru (E622D), and Rombo-Mkuu (E623C). His research interests are based on structural diversity of Bantu languages and typological principles lying behind. He has also been working on the structural variations of Swahili, particularly focusing on urban varieties generally known as Sheng. His current publications include 'A micro-parametric survey on typological covariation related to focus marking strategies, based on the Bantu Morphosyntactic Variation database' (with Lutz Marten, *Linguistique et langues africaines* 9(1): 1-25, 2023), *Descriptive Materials of Morphosyntactic Microvariation in Bantu Vol.2: A microparametric survey of*

morphosyntactic microvariation in Southern Bantu languages (co-edited with Seunghun J. Lee and Yuko Abe, Tokyo: ILCAA, 2021), *Variation in Swahili, special issue of Swahili Forum 26* (co-edited with Nico Nassenstein, 2019), and *Descriptive Materials of Morphosyntactic Microvariation in Bantu* (co-edited with Yuko Abe, Tokyo: ILCAA, 2018).

Julius Taji ***(ORCID 0000-0003-0516-0583)***

Julius Taji is a Senior Lecturer in the Department of Foreign Languages and Linguistics, University of Dar es Salaam, Tanzania and a Research Associate in the Department of Linguistics and Language Practice at University of the Free State, South Africa. He holds a PhD (Linguistics), an M.A. (Linguistics), and a B.A. (Education) from the University of Dar es Salaam. His areas of specialization include morphosyntax of Bantu languages, lexicography and sign language linguistics. Since the completion of his PhD in 2017, Julius has participated in collaborative research projects with academics from several universities, including SOAS and the University of Essex in the UK, the Research Institute for Languages and Cultures of Asia and Africa (ILCAA) in Japan, and the University of the Free State, South Africa. He has also won several research awards, including the PANGeA Early Career Fellowship Award hosted by the Graduate School of the Faculty of Arts and Social Sciences at Stellenbosch University, South Africa in 2019, a Full-Time Postdoctoral Research Fellowship in the Faculty of Humanities, University of the Free State, South Africa in 2020 and the African Humanities Programme (AHP) Fellowship in 2021. Currently, Julius is involved in a collaborative research project on grammatical variation in Swahili with colleagues from SOAS, the University of Essex, and Kenyatta University.

Nobuko Yoneda

Nobuko Yoneda started her doctoral course at the Graduate School of Tokyo University of Foreign Studies in 1995, majoring in African Linguistics, and received her PhD in 2000. She worked as a Lecturer at Osaka Jogakuin University from 2000 to 2002. She has worked at Osaka University since 2002, first as Associate Professor (2008-2010) and then since 2010 as Professor, teaching Swahili grammar, Bantu linguistics, and African sociolinguistics. Her research interests are descriptive linguistics in Bantu languages and sociolinguistics. She has been conducting fieldwork in Africa, mainly in Tanzania and Namibia, since 1993. Her main publications include 'Noncausal/causal verb alternations in Swahili' (*Linguistique et Langues Africaines* 8(2), 2022), 'Noun-modifying constructions in

Swahili and Japanese' (in Prashant Pardeshi and Taro Kageyama (eds.), *Handbook of Japanese Contrastive Linguistics*, 433-452. Berlin: de Gruyter Mouton, 2018), 'Conjoint/disjoint distinction and focus in Matengo (N13)' (in Larry Hyman and Jenneke van der Wal (eds.), *The Conjoint/Disjoint Alternation in Bantu*, 14-60. Berlin: de Gruyter Mouton, 2017) and 'Word order in Matengo (N13): Topicality and informational roles' (*Lingua* 121(5): 754-771, 2011).

Chapter 1

Introduction

Lutz Marten[a], Hannah Gibson[b], Rozenn Guérois[c,d] & Gastor Mapunda[e]

[a]SOAS University of London [b]University of Essex [c]LLACAN CNRS [d]University of KwaZulu-Natal [e]University of Dar es Salaam

1 Background

The present volume *Morphosyntactic Variation in East African Bantu Languages* has, as indicated in the title, three interacting foci of interest: Variation in morphosyntax, the study of Bantu languages, and a regional focus on East Africa. Each of these foci deserves a little bit of discussion to illuminate the motivation for this book.

Morphosyntactic variation is a comparatively recent field of study in the wider domain of comparative linguistics. Both phonological and lexical comparative studies have been an established part of the field for a long time, especially in historical linguistics. With respect to Bantu languages, a proto language had been reconstructed by the end of the nineteenth century, based on phonological and lexical reconstruction, although even early studies of Bantu languages reflected an interest in morphology and, to a lesser extent, syntax (e.g. Bleek 1862–1869, Meinhof 1906). However, recent years have seen an impressive growth in research examining morphosyntactic variation in Bantu languages, highlighted, for example, in the volume on comparative Bantu grammar edited by Mchombo 1993). Work in this tradition includes studies which look at specific construction types from a cross-Bantu perspective or in a given language, such as the examinations of double object constructions (Bresnan & Moshi 1990, Rugemalira 1993), locative inversion (Demuth & Mmusi 1997, Morimoto 2000, Khumalo 2010) and object marking (Beaudoin-Lietz et al. 2004, Riedel 2009).

Lutz Marten, Hannah Gibson, Rozenn Guérois & Gastor Mapunda. 2024. Introduction. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 1–13. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663761

A particular approach to the study of morphosyntactic variation in Bantu, following Marten et al. (2007), employs a set of surface parameters or variables to conduct systematic comparative studies often involving a larger group of languages (e.g. Bax & Diercks 2012, Petzell & Hammarström 2013, Marten & van der Wal 2014, Zeller & Ngoboka 2015, Mtenje 2016, Shinagawa & Abe 2019, Shinagawa & Marten 2021). A comprehensive list of such parameters is developed in Guérois et al. (2017) and the approach is also adopted by some of the papers in the current volume. Two further recent studies in the field of morphosyntactic variation and comparative grammar are the edited volumes by Bostoen et al. (2022), which presents historical-comparative reconstructions for a number of Bantu grammatical structures, and Bloom Ström et al. (forthcoming), which contains chapters on variation from across the Bantu-speaking area.

The approximately 500 Bantu languages spoken across vast areas of Central, Eastern and Southern Africa provide the wider empirical backdrop to the present volume. Bantu languages are united by the presence of a number of broad typological similarities, including, for example, complex noun class systems, agglutinative verbal morphology with a rich array of verbal affixes, and basic SVO word-order which is subject to pragmatically motivated word-order variation. However, within this similarity, the languages also exhibit a high degree of fine-grained micro-variation across all linguistic domains. This micro-variation results in part from independent diachronic developments such as processes of grammaticalisation and reanalysis, and in part from language contact both between Bantu languages and between Bantu languages and neighbouring non-Bantu languages. The high number of different Bantu languages and lects, the geographic density of the Bantu-speaking area, and the specific and often multilingual ecologies in which Bantu languages are spoken make the language group an important area of research for our understanding of developments and processes of morphosyntactic variation.

The regional focus of the volume is East Africa. Our conception of East Africa is inclusive, and the papers in the volume report on research on languages spoken in Kenya, Tanzania, Uganda, Malawi, Mozambique and the Democratic Republic of the Congo. Linguistically speaking, East Africa is a diverse area, in which languages of four African linguistic families are spoken and which is also home to the Rift Valley linguistic area (Kießling et al. 2007, Harvey et al. 2023).

There is a long tradition of comparative work on Eastern Bantu languages. Eastern Bantu was the subject of the lexicostatistical survey of Nurse & Philippson (1980), while Hinnebusch et al. (1981) examined several phonological and morphological features of languages in this area. Nurse (1985) then addresses the question of phonological areal features in North-Eastern Bantu, and (Nurse

& Muzale 1999) examined changes in tense and aspect in the same linguistic zone. A comprehensive historical-comparative reconstruction of Proto-Sabaki was developed in Nurse & Hinnebusch (1993). More recently, Nicolle has developed three studies concentrated on Eastern Bantu languages, two of which deal with discourse strategies in narrative texts (Nicolle 2014, 2015), and a third which compares the expression of information structure (Nicolle 2016). In terms of Bantu linguistic classification, however, the area has not been identified with a specific sub-group of the family. Rather, in most lexically-based classifications, East African Bantu languages are grouped together with Southern and Central Bantu languages as part of the larger 'Osthochland Gruppe' (Heine et al. 1977) or the 'Eastern' group (Grollemund et al. 2015). On the other hand, in terms of non-lexical data, the languages of the East African region in particular have been noted to share high degrees of structural similarity (e.g. Hinnebusch et al. 1981). More recently, Edelsten et al. (2022) identify several morphosyntactic aspects which may serve to distinguish East Bantu languages from non-East Bantu languages. These include a symmetric pattern in ditransitive constructions, negation marking in dependent clauses by a post-verbal negative marker, widespread subject inversion constructions, and the co-occurrence of formal and semantic locative inversion constructions.

2 Origins of the book

The book, in part, has its origins in two closely related projects. The first was the Leverhulme Trust funded project 'Morphosyntactic Variation in Bantu: Typology, contact and change' (RPG-2014-208), which was led by Professor Lutz Marten and based at SOAS University of London (2014-2018). The project aimed to investigate linguistic similarities in a sample of Bantu languages, with a view to better understanding how the structures of different Bantu languages have been shaped by the interaction of processes of historical innovation, language contact, and universal functions of human language.

One of the key outputs of the project was the development of a list of 142 descriptive level parameters of morphosyntactic variation (Guérois et al. 2017), and the creation of a database which enabled the storage and representation of data related to languages of the sample with respect to these parameters – the Bantu Morphosyntactic Variation (BMV) database (Marten et al. 2018). The project set up partnerships and collaborations with researchers in Africa, Europe and Asia. Additionally, as part of the project, a workshop on approaches to morphosyntactic variation in Bantu was held at the University of Dar es Salaam on 20/21 July 2016, in which some of the work reported in this volume was first discussed.

The second project was a collaboration between researchers at SOAS University of London and the University of Dar es Salaam, and was led by Professor Lutz Marten and Dr Gastor Mapunda. This project was funded by the British Academy and was entitled 'Parametric approaches to morphosyntactic variation in Eastern Bantu languages' (2017-18). The goal of the project was to build on aspects of morphosyntactic variation which had previously been described for Bantu languages and on the emerging parametric approach, and to extend this to grammatical variation in Eastern Bantu. The project sought to build on original empirical evidence from some twenty Bantu languages of Eastern Africa, with a view to contributing to a better understanding of the historical, comparative and typological patterns that have shaped the linguistic landscape of East Africa. In specific terms, the project enabled a second workshop which was held at the University of Dar es Salaam 13-15 September 2017, bringing together researchers working on East African Bantu in a practical, interactive session to explore the parametric approach. The exchange also enabled collaborators at the University of Dar es Salaam to visit SOAS.

Many of the chapters in the current volume have their origins in the workshops that took place as part of this project. However, we were also fortunate enough to have received a number of contributions from those who did not attend the original workshops, which have further strengthened and broadened the empirical coverage and the theoretical and methodological breadth of the volume.

3 Chapters in the volume

Against this background, the present volume includes chapters which offer both comparative and descriptive accounts of Bantu languages spoken in East Africa. Several of the languages discussed (e.g. Shinyiha, Runyankore-Rukiga, Kiwoso, Kihehe and Sumbwa) have not been the subject of previous extensive descriptions. The volume thus presents new empirical data, improving the descriptive status of the languages discussed in the volume. In addition, and including other more well-known Bantu languages (e.g. Swahili, Nyakyusa, Ciyao and Sena), the contributions present new or little-treated aspects of their morphosyntax, as well as providing novel data. The results presented in the volume enable close morphosyntactic comparison between languages spoken in the specific geographic area, some of which are in contact with each other.

The volume consists of twelve chapters, in addition to this introduction (Chapter 1). These chapters are grouped into three sections, devoted to the examination

of the nominal domain, the verbal domain, and analyses adopting comparative and historical approaches. A number of the chapters address the question of morphosyntactic variation through an in-depth examination of a single morphosyntactic phenomenon in a small sample of languages. Others embrace a wider perspective with more parameters of variation throughout the northeastern region.

Chapter 2, by Julius Taji, presents a description of the form, function and distribution of demonstratives in Chiyao. The form of the Chiyao demonstrative is shown to be determined by the location of the referent in relation to the speaker or hearer. In terms of distribution, the demonstrative can appear either in pre- or post-nominal position. Ciyao also employs circumdemonstratives in which a demonstrative appears both before and after the nominal. The demonstrative is shown to exhibit a range of grammatical and communicative functions: in addition to expressing the location of an entity in relation to the interlocutors, it can also express emphasis, definiteness and encode anaphoric reference.

Chapter 3, by Allen Asiimwe, examines demonstratives in Runyankore-Rukiga, a language of Uganda. The chapter explores the more common functions of demonstratives observed across Bantu, such as encoding proximity of the referent to the speaker/hearer. However, the study also reports on less frequently examined features such as the nominal and verbal properties of demonstratives and their use to express manner. Pragmatic properties of Runyankore-Rukiga demonstratives, which are divided into exophoric and endophoric demonstratives (the former used for non-anaphoric functions), are also explored. The study draws on data from written literary sources, spontaneous speech and data gathered via elicitation.

Chapter 4, by Amani Lusekelo, discusses the distribution and function of the augment and object markers in Nyakyusa, spoken in Tanzania. Adopting a parametric approach (Guérois et al. 2017), and including both bare nouns and complex noun phrases, the chapter focuses on the (non-)occurrence of the V-augment and CV-particle, the role of demonstratives, and the word-order within the noun phrase. It shows that the main role of the CV-particle is to indicate contrastive focus of the referent, while for the anaphoric demonstrative *-la* 'that/those', the augment and object marking are related to the indication of definiteness. In addition, it is shown that object marking can be optional or obligatory, depending on the verb.

Chapter 5, by Aurelia Mallya, discusses morphological and syntactic properties of locative expressions in the Tanzanian language Kiwoso. The study provides an account of the locative forms and their properties in relation to nominal and verbal morphology. While the locative class 17 prefix *ku-* is used to show agreement on all nominal and verbal modifiers in Kiwoso, the nominal locative prefixes *ku-*,

pa-, *mu-* are unproductive in the language. Kiwoso also exhibits two post-final locative enclitics (=*ho* and =*u*) which are used to cross-reference locative objects. The chapter contributes to the understanding of locatives in Kiwoso as well as locatives in the Bantu language family in general.

Chapter 6, by Gastor Mapunda and Fabiola Hassan, presents a comparative analysis of locative expressions in Bena, Ngoni, Makhuwa, and Yao, i.e. four languages spoken around the Ruvuma River which separates Tanzania and Mozambique. The chapter shows that while these languages share similar features in terms of locative morphology, Makhuwa slightly departs from the three other languages in several respects. This study constitutes an exploration of micro-variation among languages which are typologically very similar.

Chapter 7, by Nobuko Yoneda and Judith Nakayiza, focusses on the Ugandan language Ganda and examines how object noun phrases and object markers behave in multiple-object constructions, focussing in particular on the contexts in which asymmetry between these objects may emerge. The chapter makes two crucial contributions. First, Ganda allows three object markers, not only two as previously reported by other studies on this topic. Second, the symmetricity of the language is sometimes mitigated by semantic and phonological factors, suggesting that there can be variation between constructions within a language, and that "(a)symmetry" is probably not a parameter determined by language, but is more fine-grained.

Chapter 8, by Armindo Ngunga and Crisófia Langa da Câmara, presents an exploration of object marking in four languages of Mozambique. The study adopts six of the parameters of variation detailed in Marten & Kula (2012) and examines the properties of these four languages with regard to these features. The study shows that the four languages exhibit three different patterns with regard to the obligatoriness of object marking and the co-occurrence of object markers. The authors also propose an additional parameter of variation which examines the obligatoriness of the presence of an object marker with certain transitive verbs. The study is micro-comparative in nature and furthers the descriptive status of these four languages as well as our understanding of variation in object marking in Bantu.

Chapter 9, by Kulikoyela Kahigi, describes the verbal extensions in the Tanzanian language Sumbwa and their valence in the context of Bantu comparative data, adopting the parametric approach of Guérois et al. (2017). The study reveals that the verb derivational strategies in Sumbwa closely follow those mapped out by the Proto-Bantu reconstructions, excluding a few innovations among the minor extensions (e.g. *-agan-*, *-agil-*). It also shows that the causative and instrumental share extensions, that the associative markers include the post-verbal

-an- and the pre-verbal *-i-*, that the applicative conveys benefactive, directive, location, and reason meanings, and that there is no systematic fixed order of extensions, except that in all co-occurrences, the passive comes last.

Chapter 10, by Malin Petzell and Peter Edelsten, presents a review of the tense-aspect systems of five Bantu languages of Morogoro region in Tanzania: Kagulu, Luguru, Kami, Ndamba and Pogoro. The study shows that the languages investigated show significant diversity in TAM marking ranging from only two tenses (past and non-past) and limited aspectual distinctions to a system with multiple pasts and futures. The chapter discusses the distribution and meaning of these morphological distinctions, the abundance versus scarcity of specific tense-aspect markers, and the methods of expressing negation, thereby highlighting an unusual diversity in both the distribution and meaning of tense-aspect marking as well as negation across Bantu languages.

Chapter 11, by Daisuke Shinagawa, presents a comparative overview of the tense and aspect systems in Kilimanjaro Bantu languages, including those from which comprehensive information about the tense-aspect system has not been previously made available in the literature. The chapter presents data from eight varieties of Kilimanjaro Bantu, namely Rwa, Siha, Mashami, Kibosho, Uru, Vunjo, Rombo-Mkuu, and Gweno. The data show a general picture of geographical distribution and formal correspondences of shared tense-aspect markers. The chapter also examines systematic correspondences – grammaticalisation chains – and explores the historic processes of change which has given rise to shared tense-aspect markers. The chapter is micro-parametric and comparative in nature and provides possible typological generalisations which might lie behind the variation found in the Kilimanjaro Bantu tense-aspect system.

Chapter 12, by Lengson Ngwasi and Abel Mreta†, describes the historical development of reflexive-reciprocal polysemy in Kihehe by employing the three stages of the Overlap Model of Grammaticalisation Theory proposed by Heine (1993). The paper discusses an interesting feature of reflexive-reciprocal polysemy encoded by the originally reflexive prefix, a polysemy type reported in the grammar of several Bantu languages, but not in great detail (except for Bostoen forthcoming on South-Western Bantu languages). The present paper is, therefore, an important contribution to the field, in that it provides an account of this polysemy type in a little-described language. It also contributes to our understanding of variation in (East) Bantu languages because this morphosyntactic phenomenon deviates from the common Bantu situation in which reflexivity and reciprocity are encoded by two different verbal morphemes.

Chapter 13, by Lutz Marten, Hannah Gibson, Rozenn Guérois and Kyle Jerro, compares written poetic texts of Old Swahili with present-day Standard Swahili

in terms of morphosyntactic features developed in Guérois et al. (2017). Results of this study show significant differences between the two varieties. In particular, it shows that the relation between Old Swahili and Standard Swahili is characterised by a loss of morphosyntactic forms and a loss of variability. The authors argue that these results are likely to reflect the processes of standardisation and regularisation involved in the development of Swahili as a language of wider communication. The findings of the study shed new light on morphosyntactic variation since they show the effect of standardisation and a particular trajectory of morphosyntactic development.

Overall, the chapters brought together in this volume provide a snapshot of the state of the art in the study of morphosyntactic variation in the region, drawing on a range of languages and providing novel empirical data for many of them. The papers give a good impression of the variation encountered, and how different aspects of this play out in different languages. We hope that this contribution will lead to further work on the morphosyntax of East African Bantu languages, where much work remains to be done.

4 Editorial considerations

We make two further notes and observations here which are relevant for the volume as a whole. Firstly, we did not adopt a prescriptive approach to language naming and have instead left chapter authors to employ the naming conventions they consider fit for their chapters and which are used in the local context. This means that in some instances languages appear with their prefix – e.g. *Kiswahili* – while in other contexts they may appear without the prefix – e.g. *Swahili*. We acknowledge that there are different arguments for one convention or the other, and we are also aware that the question of names and naming of languages often has context-specific historical and political significance.

Secondly, we took a similar approach to the representation of data, glossing and abbreviations. We asked authors to be internally consistent in terms of how they represent data and the terms and abbreviations they use, in all cases encouraging the adaptation of the Leipzig Glossing Rules. However, in some chapters, authors have needed to use additional abbreviations, and have, thus, followed different styles that are widespread in their local contexts.

5 Next steps

It is our hope that the present volume serves as a reference point for those interested in Bantu languages in particular, as well as those interested in variation

in morphosyntax and the East African region more broadly. It combines a number of different methodological approaches and insights, but also furthers the descriptive status of a number of the languages examined and mentioned herein, thereby serving as a reference point for future work.

We hope that other scholars might be inspired by the work contained herein, as well as being exposed to data and findings which makes them reassess or revisit current work and emerging ideas, and so that the volume contributes to the further academic investigation and public awareness of African languages.

6 Acknowledgements

As editors, we would like to extend our thanks to all the authors whose work contributes to the volume, as well as the attendees of the two workshops where this project started. We are also grateful to the British Academy and the Leverhulme Trust who generously funded the two workshops, as well both the project 'Morphosyntactic variation in Bantu: typology, contact and change' and 'Parametric approaches to morphosyntactic variation in Eastern Bantu languages'.

We are grateful to all of the reviewers for their detailed comments and input they made on earlier versions of the chapters. Their comments served to strengthen and improve the clarity of the work. We are also grateful to the editors of the *Contemporary African Linguistics* series for their initial and ongoing support of this project. We are also immensely grateful to the team at Language Science Press and here Sebastian Nordhoff must receive a special mention for being an unwavering source of support, enthusiasm and advice during the process of putting the book together.

7 Dedication

We would like to dedicate this volume to the late Dr Abel Mreta. Dr Mreta was a central member of the Department of Foreign Language and Linguistics at the University of Dar es Salaam, at one time the Head of the Department, and a staunch supporter of linguistic research on East African languages. He generously shared his knowledge and experience and he was involved in the training of a large proportion of the linguists working in Tanzania today.

Dr Mreta was born in Kilimanjaro Region, Tanzania. He completed his BA Education and MA Linguistics at the University of Dar es Salaam, and then his PhD at the University of Bayreuth, Germany. Dr Mreta was employed as a tutorial assistant at the University of Dar es Salaam in May 1987. His areas of interest were morphology, historical and comparative linguistics, sign language, and language

documentation. In his career, he rose to the rank of Senior Lecturer in May 2008. He worked as a visiting lecturer at Leiden University (the Netherlands), Gothenburg University (Sweden) and Hankuk University of Foreign Studies (South Korea). Dr Mreta taught and mentored most of the linguists currently working in Tanzania and beyond, including many of the contributors to this volume.

In September 2017, Dr Mreta participated in the workshop that discussed the parametric approach to the study of morphosyntactic variation in Bantu languages jointly organised by the Department of Foreign Languages and Linguistics (University of Dar es Salaam) and SOAS University of London, and it was at this workshop where Dr Mreta, together with other scholars, started working on their chapter contributions. Dr Mreta was a central contributor and supporter of the overall collaborative project and this book project. This contribution is in part reflected in the co-authored chapter by Dr Mreta and Lengson Ngwasi (Chapter 11). Dr Mreta passed on while the manuscripts were still in the review process. He will be remembered fondly by all who knew him and his legacy will live on through his work.

References

Bax, Anna & Michael Diercks. 2012. Information structure constraints on object marking in Manyika. *Southern African Linguistics and Applied Language Sciences* 30(2). 185–202.

Beaudoin-Lietz, Christa, Derek Nurse & Sarah Rose. 2004. Pronominal object marking in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 175–188. Cologne: Köppe.

Bleek, Wilhelm H. I. 1862–1869. *A comparative grammar of South African languages*. London: Trübner.

Bostoen, Koen. Forthcoming. Reflexive-reciprocal polysemy in South-Western Bantu: Distribution, typology and origins. In Eva-Marie Bloom Ström, Hannah Gibson, Rozenn Guérois & Lutz Marten (eds.), *Current approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.

Bostoen, Koen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.). 2022. *On reconstructing Proto Bantu grammar* (Niger-Congo Comparative Studies 4). Berlin: Language Science Press. DOI: 10.5281/zenodo.7560553.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2). 147–185.

Demuth, Katherine & Sheila Mmusi. 1997. Presentational focus and thematic structure in comparative Bantu. *Journal of African Languages and Linguistics* 18(1). 1–19. DOI: 10.1515/jall.1997.18.1.1.

Edelsten, Peter, Hannah Gibson, Rozenn Guérois, Gastor Mapunda, Lutz Marten & Julius Taji. 2022. Morphosyntactic variation in Bantu: Focus on East Africa. *Journal of the Language Association of East Africa* 1(1). 1–24.

Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences* 112(43). 13296–13301.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list.* Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.

Harvey, Andrew, Hannah Gibson & Richard Griscom. 2023. Preverbal clitic clusters in the Tanzanian Rift Valley revisited. *Journal of African Languages and Linguistics*. 175–239. DOI: 10.1515/jall-2023-2010.

Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. Oxford: Oxford University Press.

Heine, Bernd, Hans Hoff & Rainer Vossen. 1977. Neuere Ergebnisse zur Territorialgeschichte des Bantu. In Wilhelm J. G. Möhlig, Franz Rottland & Bernd Heine (eds.), *Zur Sprachgeschichte und Ethnohistorie in Afrika*, 57–72. Berlin: Reimer.

Hinnebusch, Thomas J., Derek Nurse & Martin Mould. 1981. *Studies in the classification of Eastern Bantu languages*. Hamburg: Helmut Buske.

Khumalo, Langa. 2010. Passive, locative inversion in Ndebele and the unaccusative hypothesis. *South African Journal of African Languages* 30(1). 22–34.

Kießling, Roland, Maarten Mous & Derek Nurse. 2007. The Tanzanian Rift Valley area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 186–227. Cambridge: Cambridge University Press. DOI: 10 . 1017 / CBO9780511486272.007..

Marten, Lutz, Peter Edelsten, Hannah Gibson & Rozenn Guérois. 2018. *Bantu Morphosyntactic Variation (BMV) database*. London: SOAS University of London.

Marten, Lutz & Nancy Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30(2). 237–253.

Marten, Lutz, Nancy Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105. 253–338.

Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318–368.

Mchombo, Sam A. (ed.). 1993. *Theoretical aspects of Bantu grammar*. Stanford: CSLI.

Meinhof, Carl. 1906. *Grundzüge einer vergleichenden Grammatik der bantusprachen*. Berlin: Reimer.

Morimoto, Yukiko. 2000. *Discourse configurationality in Bantu Morphosyntax*. Stanford: Stanford University.

Mtenje, Atikonda. 2016. *A comparative analysis of the phonology and morphosyntax of Cisukwa, Cindali and Cilambya*. University of Cape Town. (Doctoral dissertation).

Nicolle, Steve. 2014. Discourse functions of demonstratives in eastern Bantu narrative texts. *Studies in African Linguistics* 43(2). 125–144.

Nicolle, Steve. 2015. *Comparative study of eastern Bantu narrative texts* (SIL Electronic Working Papers 2015-003).

Nicolle, Steve. 2016. Variation in the expression of information structure in eastern Bantu languages. In Sara Pacchiarotti Doris L. Payne & Mokaya Bosire (eds.), *Diversity in African languages*, 377–395. Berlin: Language Science Press.

Nurse, Derek. 1985. Dentality, areal features, and phonological change in northeastern Bantu. *Studies in African Linguistics* 16(3). 243–279.

Nurse, Derek & Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history*. Berkeley: University of California Press.

Nurse, Derek & Henry R. T. Muzale. 1999. Tense and aspect in great lakes languages. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives*, 517–544. Stanford, California: Center for the Study of Language & Information.

Nurse, Derek & Gérard Philippson. 1980. The Bantu languages of East Africa: A lexicostatistical survey. In Edgar C. Polomé & C. P. Hill (eds.), *Language in Tanzania*, 26–67. London: Oxford University Press for the International African Institute.

Petzell, Malin & Harald Hammarström. 2013. Grammatical and lexical comparison of the greater ruvu Bantu languages. *Nordic Journal of African Studies* 22. 129–157.

Riedel, Kristina. 2009. *The syntax of object marking in Sambaa: A comparative Bantu perspective*. Leiden: LOT Publications.

Rugemalira, Josephat M. 1993. Bantu multiple "object" constructions. *Linguistic Analysis* 23. 226–252.

Shinagawa, Daisuke & Yuko Abe (eds.). 2019. *Descriptive materials of morphosyntactic microvariation in Bantu*. Tokyo: Research Institute for Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.

Shinagawa, Daisuke & Lutz Marten. 2021. Micro-typological covariation of negation and focus marking in Bantu languages. 言語研究 *Gengo Kenkyu* 160. 215–248.

Zeller, Jochen & Jean Paul Ngoboka. 2015. On parametric variation in Bantu, with particular reference to Kinyarwanda. *Transactions of the Philological Society* 113(2). 206–231.

Part I

The nominal domain

Chapter 2

Demonstratives in Chiyao: An analysis of their form, distribution and functions

Julius Taji
University of Dar es Salaam

This chapter aims to provide a description of Chiyao demonstratives with a particular focus on their form, distribution, and functions. It is shown that, morphologically, the Chiyao demonstrative consists of an initial element, which is always the vowel *a-,* followed by an agreement marker, and ends with a final element which changes according to the location of the referent in relation to the speaker or hearer. In cases of demonstrative doubling, the post-nominal demonstrative drops the initial vowel. Thus, it is proposed that the initial vowel in the Chiyao demonstrative is optional. Within the broader classification of demonstratives in Bantu languages, the Chiyao demonstratives appear in four main types. These include pronominal demonstratives, which are used as independent pronouns; adnominal demonstratives, which modify nouns; adverbial demonstratives, which modify verbs and locative nouns; and finally, identificational demonstratives, which occur in copula and non-verbal clauses. With regard to their distribution, it is established that the Chiyao demonstrative can occur either post-nominally, or in both pre-nominal and post-nominal positions simultaneously. Finally, it is shown that demonstratives serve various grammatical and communicative functions, including expressing the location of an entity in relation to interlocutors, showing emphasis, indicating definiteness, and encoding anaphoric reference. These findings contribute to our further understanding of the behavior of demonstratives in Bantu languages and inform future descriptive and typological works in this area.

1 Introduction

Studies across Bantu languages have generally indicated that demonstratives perform spatial-deictic roles referring to entities in three main dimensions, namely

Julius Taji. 2024. Demonstratives in Chiyao: An analysis of their form, distribution and functions. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 17–41. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663763

proximal (near the speaker), non-proximal or medial (near the addressee), and distal (distant from the speaker or addressee) (Nicolle 2007a,b, Asiimwe 2014). Despite this uniformity in the dimensions reflected by demonstratives, there is significant variation among Bantu languages in terms of syntactic ordering of demonstratives within the NP and the functions they serve, in addition to their spatial-deictic uses. For example, with regard to syntactic ordering, Swahili allows both pre-nominal and post-nominal demonstratives as shown in (1).[1]

(1) Swahili

a. **m-geni yule** a-li-fika mapema
1-visitor 1.DEM.DIST SM1-PST-arrive early
'That visitor arrived early.'

b. **yule m-geni** a-li-fika mapema
1.DEM.DIST 1-visitor SM1-PST-arrive early
'That visitor arrived early.'

In contrast to Swahili, Sukuma allows only post-nominal demonstratives (2). Placing the demonstrative before the head noun in Sukuma is ungrammatical, as shown in (3).

(2) Sukuma (Nyanda p.c.)

a. **u-ng'w-ana uyu** a-tog-ilwe βu-gali
PPX-1-child DEM.PROX SM1-like-PFV 14-ugali
'This child likes ugali.'

b. **a-ma-shamba ayo** ga-lim-ilwe
PPX-6-farm DEM.NON_PROX SM6-cultivate-PFV
'Those farms have been cultivated.'

(3) Sukuma (Nyanda p.c.)

a. * **uyu u-ng'w-ana** a-tog-ilwe βu-gali
DEM.PROX PPX-1-child SM1-like-PFV 14-ugali
Int: 'This child likes ugali.'

b. * **ayo a-ma-shamba** ga-lim-ilwe
DEM.NON_PROX PPX-6-farm SM6-cultivate-PFV
Int: 'Those farms have been cultivated.'

[1]Unless otherwise indicated, all Swahili examples are from the author as a native speaker.

A similar pattern is attested in Makhuwa where demonstratives occur after the head nouns they modify in their canonical adnominal function as illustrated in (4) below.

(4) Makhuwa (van der Wal 2010: 184)

a. mwálpwá **olé** o-hoó-wa
1.dog 1.DEM.NON_PROX SM1-PERF.DJ-come
'The/that dog came.'

b. ki-kúm-íh-é-ni nipúró **nna** vá
1SG.OM-exist-CAUS-OPT-PLA 5.place 5.DEM.PROX 16.DEM.PROX
'Get me out of this place.'

Makonde exhibits a more free order as it permits multiple occurrence of demonstratives in both pre-nominal and post-nominal positions simultaneously, as in (5) below. As it will be revealed in the subsequent sections, several other languages of southern Tanzania also display this order.

(5) Makonde (Makanjila 2019: 171)
a-yu mu-ana **a-yu** a-ka-pilikan-a **a-yu**
DEM.PROX 1-child DEM.PROX SM1-NEG-payheed-FV DEM.PROX
'This child is stubborn.'

In view of the variation displayed by demonstratives across Bantu, this chapter provides a description of demonstratives in Chiyao, a Bantu language of southern Tanzania, with a focus on their forms, syntactic distribution, and functions.

The Chiyao data in this paper are based on the Masasi dialect and come from two main sources, namely grammaticality judgments and oral narratives. Through grammaticality judgments, a list of constructions with demonstratives in different orders was provided to ten native speakers of Chiyao who then gave feedback on whether they were acceptable or not. As for oral narratives, constructions with demonstratives were extracted from three traditional stories narrated by three different native speakers of Chiyao. The constructions were then analyzed to determine the form, distribution and function of demonstratives. Unless otherwise indicated, all data in this paper are from Chiyao.

The chapter proceeds as follows: §2 provides a brief account of the Chiyao language; §3 discusses the forms of the Chiyao demonstratives; §4 outlines the types of demonstratives; §5 presents the syntactic distribution of demonstratives; §6 focuses on the functions of demonstratives; and §7 provides a conclusion.

2 The Chiyao language

Chiyao is a cross-border Bantu language which is spoken in Southern Malawi, north-western Mozambique, and southern Tanzania. The language is also referred to as Ciyao (Ngunga 1997), Ciyawo (Dicks & Dollar 2010), and Yao (Whiteley 1966). According to Ngunga (2002), there are about three million Chiyao speakers residing in these three countries. A significant number of Chiyao speakers, mainly emigrants from Malawi, are also found in Zambia and Zimbabwe. The present study is based on the Tanzanian variety of Chiyao, which is mainly spoken in the southern regions of Ruvuma (Tunduru District), and Mtwara (Masasi District). The number of Chiyao speakers in Tanzania is estimated at 400,000 (Languages of Tanzania Project, 2009). Nurse & Philippson (1980) classified Chiyao under the Ruvuma Bantu branch of the Rufiji-Kilombero Bantu zone. In Guthrie's updated list by Maho (2009), the language is coded as P21 and is found in the Yao group, along with Mwera (Shimwela) P22, Makonde (Chimakonde) P23, Ndonde (P24), and Mavia (P25).

3 Forms of the Chiyao demonstrative

It has been argued that demonstratives in many Bantu languages start with the "initial element" or "initial vowel" (Wald 1973, Du Plessis 1978, Du Plessis & Visser 1992). The initial vowel may take various forms depending on the vowel of the agreement prefix, such as *a-, e-,* or *o-* (Visser 2008). However, across languages, not all demonstratives have an initial vowel (Asiimwe 2014). Asiimwe (2014), observed that in Runyankore-Rukiga, the initial vowel *a-* is realised in the proximal and medial demonstratives, but it is absent in the distal demonstratives (but see also Asiimwe (2024 [this volume])).

The situation in Chiyao seems to support the analysis that not all demonstratives take an initial vowel. There are cases where a demonstrative occurs without this initial vowel, notably when a demonstrative occurs as a circumdemonstrative as illustrated in (6) below where the post-nominal demonstrative form does not host the initial vowel *a-* (see also §5.3).

(6) a. **ali** lí-koswé **li**
5.DEM.PROX 5-rat 5.DEM.PROX
'This rat'

b. **achila** chi-téla **chíla**
7.DEM.DIST 7-tree 7.DEM.DIST
'That tree (over there)'

Given this scenario, we can therefore posit that the demonstrative in Chiyao take different shapes depending on the noun class, as shown in Table 1.

As Table 1 below indicates, the proximal demonstrative optionally begins with *a-* and ends with *-u, -a*, or *-i*. The non-proximal demonstrative optionally begins with *a-* and ends with *-o*. The distal demonstrative optionally begins with *a-* and ends with *-la*. Generally, all demonstratives in Chiyao have an agreement marker which, as it is proposed here, can be regarded as a base, and a final element. The agreement marker and the final element change in response to noun class and deixis. Thus, the proximal demonstrative appears in the form *(a)-* + agreement marker + *-u/-a/-i* while the non-proximal demonstrative takes the form *(a)-* + agreement marker + *-o*. As for the distal demonstrative, the form is *(a)-* + agreement marker + *-la*. From a phonological point of view, it can be observed that proximal and non-proximal demonstratives take the form of (V)C(C)V while distal demonstratives take the form of (V)CVCV.

Generally, the above scenario seems to suggest that the form of the demonstratives in Chiyao is dependent upon two key factors, namely noun class and deixis. Thus, a demonstrative changes its shape in response to the noun class of the noun it modifies, as well as the spatial relationship between the interlocutors and the referent. It follows that in relation to the spatial relationship, the demonstrative takes one of the three forms for each noun class, namely close to the speaker (proximal demonstrative), near the addressee (non-proximal demonstrative), and far from both the speaker and the addressee (distal demonstrative). The examples in (7) further illustrate this three-way distinction of Chiyao demonstratives and their forms by using a class 7 noun *chiteéngu* 'chair.' Example (7a) illustrates a proximal demonstrative, (7b) illustrates a non-proximal demonstrative, and (7c) shows a distal demonstrative.

(7) a. chi-teengú **achí**
7-chair 7.DEM.PROX
'This is a chair'

b. chi-teengú **achó**
7-chair 7.DEM.NON_PROX
'That is a chair (near you)'

c. chi-teengú **achilá**
7-chair 7.DEM.DIST
'That is a chair (over there)'

In addition to noun class and deixis, the shape of the demonstrative seems to be further influenced by syntactic and pragmatic factors. Pragmatic factors specifi-

Table 1: : Chiyao demonstratives in relation to noun classes and deixis

Class	Prefix	Example	Gloss	demonstratives			
				PROX	NON_PROX	DIST	Example
1	mu-	mundu	person	(a)ju	(a)jo	(a)jula	mundu aju/ajo/ajula ‘this/that person’
2	va-	vandu	people	(a)va	(a)vo	(a)vala	vandu ava/avo/avala ‘these/those people’
3	m-	mteela	tree	(a)wu	(a)wo	(a)wula	mteela awu/awo/awula ‘this/that tree’
4	mi-	miteela	trees	(a)ji	(a)jo	(a)jila	miteela aji/ajo/ajila ‘these/those trees’
5	li-	lindanda	egg	(a)li	(a)lyo	(a)lila	lindanda ali/alyo/alila ‘this/that egg’
6	ma-	mandanda	eggs	(a)ga	(a)go	(a)gala	mandada aga/ago/agala ‘these/those eggs’
7	chi-	chipuula	knife	(a)chi	(a)cho	(a)chila	chipula achi/acho/achila ‘this/that knife’
8	i-/y-	ipuula	knives	(a)yi	(a)yo	(a)yila	ipula ayi/ayo/ayila ‘these/those knives’
9	n	ndembo	elephant	(a)ji	(a)jo	(a)jila	ndembo aji/ajo/ajila ‘this/that elephant’

Table 1: Chiyao demonstratives in relation to noun classes and deixis (continued)

Class	Prefix	Example	Gloss	demonstratives			
				PROX	NON_PROX	DIST	Example
10	n	ndembo	elephants	(a)si	(a)syo	(a)sila	ndembo asi/asyo/asila 'these/those elephants'
11	lu-	lusulo	river	(a)lu	(a)lo	(a)lula	lusulo alu/alo/alula 'this/that river'
12	ka-	kajuni	small bird	(a)ka	(a)ko	(a)kala	kajuni aka/ako/akala 'this/that small bird'
13	tu-	tujuni	small birds	(a)tu	(a)to	(a)tula	tujuni atu/ato/atula 'these/those small birds'
14	u-	ukana	liquor	(a)u	(a)wo	(a)wula	ukana au/awo/awula 'this/that liquor'
15	ku-	kulya	eating	(a)ku	(a)ko	(a)kula	kulya aku/ako/akula 'this/that eating'
16	pa-	paasi	down	(a)pa	(a)po	(a)pala	paasi apa/apo/apala 'here/there down'
17	ku-	kumchiji	left	(a)ku	(a)ko	(a)kula	kumchiji aku/ako/akula 'here/there on the left'
18	mu-	mulisimbo	inside the pit	(a)mu	(a)mo	(a)mula	mulisimbo amu/amo/amula 'here/there inside the pit

cally determine the presence or absence of the initial vowel in the demonstratives. Thus, the demonstrative in its full form with an initial vowel commonly occurs in copula constructions to indicate the location of an entity within the interlocutor(s)' visibility (see §4.2) – this is usually accompanied by a pointing gesture, as in (8). On the other hand, the reduced form (without the initial vowel) modifies nouns in subject or object position and functions to refer to the entity mentioned earlier in the discourse (8).

(8) m-chanda **ájúlá**
1-boy 1.DEM.DIST
'That is a boy (over there)'

(9) m-chanda **júla** a-iíche
1-boy 1.DEM.DIST SM1-arrive.PERF
'That/the boy has arrived.'

The copula construction illustrated in (8) above cannot take a reduced demonstrative, thus explaining the ungrammaticality of (10). Similarly, placing a full demonstrative after a noun in a non-copula construction is ungrammatical, as in (11) below:

(10) * m-chanda **júla**
1-boy 1.DEM.DIST
Int: 'A boy is over there'

(11) * m-chanda **ájúlá** a-iíche
1-boy 1.DEM.DIST SM1-arrive.PERF
Int: 'That/the boy has arrived.'

The form of the demonstrative in (11) looks similar to (7) in that they both contain initial vowels. However, (11) is ungrammatical because the demonstrative occurs with the NP and thus functions as a modifier while in (7) it occurs within the VP as a subject complement. It can therefore be posited that the initial vowel can be present when Noun + DEM are produced in isolation, but it is absent when this sequence is used in a clause as subject or object.

4 Classification of demonstratives

There are various approaches to the classification of demonstratives. Included amongst these are those proposed by Diessel (1999), Dixon (2003), and Van de

Velde (2005). The first two are general typological works while the latter focuses specifically on Bantu. The analysis in this chapter is based on the approaches of Diessel (1999) and Dixon (2003). Van de Velde's (2005) approach is referred to in the section on syntactic distribution of demonstratives since it focuses on the positioning of demonstratives in relation to other elements within a Bantu clause.

Diessel (1999) analyses demonstratives based on the type of elements with which they occur in a clause. His analysis is based on the view that demonstratives can occur independently as arguments, or as modifiers of nouns, or as arguments of copula verbs. On the basis of this view, Diessel establishes four types of demonstratives: pronominal demonstratives, adnominal demonstratives, adverbial demonstratives, and identificational demonstratives. These are discussed below with reference to Chiyao examples.

4.1 Pronominal demonstratives

Pronominal demonstratives can be used as independent pronouns, as is the case with the word *this* in the English sentence *This is nice*. Since they may take the place of noun phrases, they have also been referred to as demonstrative pronouns (Diessel 1999: 72). In Chiyao, pronominal demonstratives show grammatical distinctions also displayed by nouns. This means that they show agreement in terms of noun class and number (see Table 1 above). These demonstrative pronouns can also be cross-referenced on the verb with a subject or object marker. In (12) below, the demonstratives *acho* 'that', and *aga* 'these' are pronominal as they stand alone and take the place of noun phrases.

(12) a. nné ngú-chi-sáká **achó**
I SM1-OM7-want 7.DEM.NOM_PROX
'I want that (one).'

b. **agá** ga-teméche
6.DEM.PROX SM6-break.PFV
'These have been broken.'

The demonstrative *achó* in (12a) refers to a class 7 singular noun, such as *chitengu* 'chair' while *agá* in (12b) stands for a class 6 plural noun, such as *majela* 'hoes', thus evidencing the fact that demonstratives exhibit noun class agreement. The examples in (12a–12b) further illustrate that pronominal demonstratives can occur in both subject position (12b) and object position (12a). In both cases, the nouns which they refer to are also cross-referenced on the verb with either a subject marker or an object marker.

4.2 Adnominal demonstratives

Unlike pronominal demonstratives, which can function like nouns, adnominal demonstratives modify nouns, and across Bantu can appear either before nouns (e.g. in Kirundi) or after nouns (e.g. in Luganda and Ruhaya) (Van de Velde 2005). In Chiyao, adnominal demonstratives can occur either post-nominally (13) or in both pre-nominal and post-nominal positions (cf. §5.3), as in (14). The prenominal or postnominal occurrence of the adnominal demonstrative can also be conditioned by other adnominals, as is the case in Ikyaushi (Spier 2022).

(13) mbwá á**sílá**
10.dog 10.DEM.DIST
'Those dogs'

(14) **asila** mbwá **síla**
10.DEM.DIST 10.dog 10.DEM.DIST
'Those dogs'

4.3 Adverbial demonstratives

Adverbial demonstratives indicate location. These are exemplified in English with words such as *here* and *there*. Adverbial demonstratives are used to indicate the location of an event or situation expressed by the corresponding verb. In Chiyao, adverbial demonstratives occur both pronominally and adnominally. When they occur pronominally, they stand as independent pronouns and can function as subjects (15a) or as verbal modifiers (15b). When the adverbial demonstratives occur adnominally, they co-occur with the locative nouns they modify, as in example (16).

(15) a. **akulá** ku-talíche
17.dem.dist SM17-be.far
'That place (over there) is far.'

b. a-kú-támá **apa**
1sm-prs-stay 16.DEM.PROX
'He stays here.'

(16) a. ku-chi-jíji **kula** ku-talíche
17-7-village 17.DEM.DIST PRS-befar.PFV
'There at the village it is far.'

b. li-jóká lí-jinjíle mu-lisimbo **mo**
5-snake SM5-enter.PFV 18-hole 18.DEM.NON_PROX
'A snake has entered that hole.'

The basic forms of the demonstratives in (16a) and (16b) are *akula* and *amo* respectively. However, it is worth noting that in these examples the adverbial demonstratives have lost the initial vowel, which is a common tendency for adverbial demonstratives occurring post-nominally in this language (see also §3).

The dropping of the initial vowel in adverbial demonstratives which occur post-nominally, illustrated in the examples above, further supports the argument put forward in §3 that the initial vowel does not constitute the core unit of the demonstrative. Moreover, the examples above seem to suggest that the type and syntactic occurrence of the demonstrative can impose some restrictions on the form of the demonstrative – i.e. when the adverbial demonstrative occurs post-nominally, it drops the initial vowel.

4.4 Identificational demonstratives

Identificational demonstratives, or, as Diessel (1999) calls them, "demonstrative identifiers" occur in copula constructions. They function to identify a referent in a speech situation by introducing a new discourse topic or drawing the interlocutors' attention to some existing discourse entity (Payne & Peña 2007). Identificational demonstratives have alternatively been referred to as predicative demonstratives (Denny 1982, Heath 1984), deictic predicators (Schuh 1977), predicative pronouns (Marconnes 1931), existential demonstratives (Benton 1971), pointing demonstratives (Rehg & Sohl 1981), deictic identifier pronouns (Carlson 1994), and presentational pronouns (Moltmann 2013). Identificational demonstratives normally occur pronominally, in subject position, as in the following example from Ewondo, a Bantu language of southern Cameroon (Diessel 1999: 19).

(17) Ewondo (Diessel 1999: 19, glosses added)
káádá **ɲɔ**
1.crab 1.DEM.PROX
'This is a crab'

In Chiyao, identificational demonstratives can be found in copula constructions which are formed using the copula *ni*, as in the following examples.

(18) a. **ajú** ní ambusánga jwenu
1.DEM.PROX is friend 1SG.POSS
'This is your friend.'

b. **aulá** ní m-gunda wénu
3.DEM.DIST is 3-farm 3SG.POSS
'That is your farm.'

In the context where the interlocutors have prior knowledge of the referent, the identificational demonstrative can occur in both positions, i.e. before and after the copula verb, as shown in (19).

(19) a. **ajú** ní ambusánga jwenu **júla**
1.DEM.PROX is friend 1SG.POSS 1.DEM.DIST
'This is that friend of yours (that we talked about).'

b. **aulá** ní m-gunda wénu **ula**
3.DEM.DIST is 3-farm 3SG.POSS 3.DEM.DIST
'That is the farm of yours (that we talked about).'

The demonstratives *aju* 'this,' and *aula* 'that' in (18a) and (18b) respectively are identificational as they serve to identify the nouns with which they occur. The occurrence of these demonstratives in the preverbal position without nouns seems to suggest that a subject nominal that would appear before the demonstrative is "silent" (or omitted, depending on the theoretical approach adopted) since Chiyao widely allows pro-drop.

5 Syntactic distributions of demonstratives

Van de Velde (2005) identified three major patterns for the positioning of demonstratives in Bantu languages in relation to the nouns with which they occur. These are i) the pre-nominal position ("preposed demonstratives" in Van de Velde's terminology), ii) the post-nominal position ("postposed demonstratives"), and iii) both before and after the head noun ("circumdemonstratives"). I discuss each of these in the context of Chiyao below.

5.1 The pre-nominal position

The pre-nominal occurrence of demonstratives has been observed in several Bantu languages, including Nkore, Rundi, Ha, Bemba, Zulu, Xhosa (Van de Velde 2005), and Runyankore-Rukiga (Asiimwe (2024 [this volume])). In Xhosa, the prenominal demonstrative performs the emphatic role when it precedes a noun with an augment, as in (20) below.

(20) Van de Velde (2005: 14), glosses added

a. **loo** m-ntu (non-emphatic)
1.DEM.DIST 1-person
'that person'

b. **lowo** u-m-ntu (emphatic)
1.DEM.DIST AUG1-person
'that person'

However, unlike the aforementioned languages, which allow demonstratives to occur before their head nouns, Chiyao does not permit pre-nominal occurrence of demonstratives without an accompanying post-nominal demonstrative form (i.e. the demonstrative without the initial vowel *a-*). Thus, the examples in (21) below are ungrammatical while those in (22), which employ the circum-demonstrative demonstrative are grammatical.

(21) a. * **aju** mú-ndú
1.DEM.PROX 1-person
'This person'

b. * **aú** m-gunda
3.DEM.PROX 3-farm
'This farm'

(22) a. **aju** mú-ndú **ju**
1.DEM.PROX 1-person 1.DEM.PROX
'This person'

b. **aú** m-gundá **u**
3.DEM.PROX 3-farm 3.DEM.PROX
'This farm'

5.2 The post-nominal position

Chiyao also allows demonstratives to occur in the post-nominal position as illustrated in the following examples.

(23) a. li-jóká **álílá**
5-snake 5.DEM.DIST
'A snake is over there.'

b. chi-sotí **achó**
7-hat 7.DEM.NON_PROX
'A hat is over there (near you)'

c. ku-chi-jíji **ákúlá**
17-7-village 17.DEM.DIST
'There at the village'

When other dependents are present in the NP, the post-nominal demonstrative occupies the final slot in the Chiyao NP template. But it is worth noting that only the reduced form of the demonstrative, without the initial vowel is allowed to occur after other dependents, as shown in (24). These examples further suggest that, canonically, the post-nominal demonstrative occurs last in the Chiyao NP. In (24) below, the demonstratives *síla* (24a) and *lila* (24b) are separate from the head noun, while other modifiers optionally occur between the nominal form and the demonstrative.

(24) a. mbwá syetu syékúlúngwá **síla** sya-júv-ilé m-ma-úkútu.
10.dog 10.POSS 10.big 10.DEM.DIST SM10-hide-PFV 18-6-bush
'Those big dogs of ours hid in bushes.'

b. li-símú ly-épilíu ly-ákúnyákapala **líla** ly-a-jígeelé
5-ogre 5-black 5-ugly 5.DEM.DIST SM5-PST-carry.PST
li-váágó.
5-axe
'That black ugly ogre carried an axe.'

When a demonstrative occurs before other modifiers (e.g. the possessive or an adjective), the sentence becomes ungrammatical as in (25) below.

(25) a. * mbwá **síla** syetu syékúlúngwa sya-júv-ilé m-ma-úkútu.
10.dog 10.DEM.DIST 10.POSS 10.big 10SM-hide-PFV 18LOC-6-bush
Int. 'Those big dogs of ours hid in bushes.'

b. * li-símú **lila** ly-épilíu ly-ákúnyákapala ly-a-jígeelé
5-ogre 5.DEM.DIST 5-black 5-ugly SM5.PST-carry.PST
li-váágó.
5-axe
Int. 'That black ugly ogre carried an axe.'

In (25a) the distal demonstrative *sila* 'those' precedes the modifiers *syetu syekulungwa* 'our big' while in (25b) the distal demonstrative *vala* 'that' precedes the modifiers *vapiliu vakunyakapala* 'black ugly'. In both cases, the resulting sentences are ungrammatical.

As many as five modifiers can appear between the head noun and the demonstrative, including possessives, numerals, adjectives, intensifiers and relative clauses, as long as the demonstrative occurs after the modifiers. Example (26) presents a Chiyao sentence with a series of modifiers of these types followed by a demonstrative.

(26) chi-pula chángu chi-mo ché-kúlúngwa nnope chi-ná-chi-súm-ile
7-knife 7.POSS 7-one 7-big INT 7REL-SM1-buy-PERF
lisó **chila**
yesterday 7.DEM.DIST
'My one very big knife which I bought yesterday'

Considering the position of the demonstrative in relation to other nominal dependents in the Chiyao NP, I propose the following NP template for Chiyao:

(27) (DEM) + N + NUM + ADJ/ASSOC + REL + DEM

Thus, as the above template shows, the post-nominal demonstrative is the most final element in the Chiyao verb template.

5.3 The pre-nominal and post-nominal position

A number of languages contain demonstratives that occur in sort of pairs; one before the head noun and the other after the head. (see Lyons 1999). Due to their nature of occurrence – before and after the head- they are often referred to as circumdemonstratives (see Van de Velde 2005: 6). In Irish, for example, the simultaneous pre-nominal and post-nominal occurrence of demonstratives is evident is such forms as ***an*** *leabhar* 'the book', ***an*** *leabhar* ***seo*** 'this book', ***an*** *leabhar* ***sin*** 'that book', and ***an*** *leabhar* ***úd*** 'yonder book' where the post-nominal particle is obligatory, and the pre-nominal determiner element is the definite article (Lyons 1999: 117). A similar pattern is attested in Chiyao, where a demonstrative occurs first before the head noun in its full form and is then repeated after the noun in a reduced form without the initial vowel. In (28) the distal demonstrative occurs in both post-nominal and pre-nominal positions.

(28) a. n-tu-saídie yákutí pa-ku-m-okóa **ajula** mu-ndu **júla**
SM1-2OM-help how 16-INF-save 1.DEM.DIST 1-person 1.DEM.DIST
'Help us on how to save that person.'

b. nambó **alila** li-kwáta **lila** vá-á-tité va-ngáli
but 5.DEM.DIST 5-dance 5.DEM.DIST SM2-PST-say.PFV SM2-NEG
ma-véngwa a-ka-ika
6-horn SM2.NEG-come
'But that dance, they said the one who doesn't have horns should not attend.'

The circumdemonstrative is the most frequently used demonstrative in Chiyao and in several neighbouring languages, including Makonde (P23) (see Makanjila 2019), Shimwela (P22) (see Taji & Mreta 2017), and Makhuwa (P31) (see van der Wal 2010). Below are examples from these languages:

(29) Makonde (Makanjila 2019: 171)
ai timu ai inamala kung'ana mpila namene
a-i timu **a-i** i-na-mal-a ku-ng'an-a m-pila
9.DEM.PROX 9.team 9.DEM.PROX SM9-TAM-know-FV INF-play-FV 3-ball
namene
very
'This team plays football extremely well.'

(30) Shimwela (Taji & Mreta 2017)
aji mi-kóngo **ji**
4.DEM.PROX 4-tree 4.DEM.PROX
'These trees'

(31) Makhuwa (van der Wal 2010: 201)
ólé nthíyán' **uule** kh-oóthá aa-páh
1.DEM.DIST 1.woman 1.DEM.DIST NEG-SM1.IMPF-lie SM1.IMPF-light
ólumwenku
14.world
'This woman didn't just lie, she set the world on fire.'

In sum, the preceding discussion of the distribution of demonstratives has revealed three main properties relating to Chiyao demonstratives. Firstly, Chiyao does not permit a pre-nominal demonstrative to occur alone without being complemented by a post-nominal demonstrative both of which together form a circumdemonstrative; secondly, the post-nominal demonstrative occupies the most final slot in the NP, after all other modifiers; thirdly, the simultaneous occurrence of demonstratives (also known as the circumdemonstrative) in both pre-nominal and post-nominal positions is the most favoured pattern in Chiyao. In the section

that follows, I discuss different grammatical and communicative roles served by demonstratives in Chiyao.

6 Functions of demonstratives

Demonstratives serve various linguistic roles. In addition to the commonly attested spatial-deictic function whereby they are used to indicate objects in three dimensions (proximal, distal, and non-proximal), demonstratives have also frequently been analyzed as definite markers, as in Bemba, Zulu and Xhosa (Greenberg 1978); relative pronouns, for example in German (Lehmann 1984); third person pronouns, for example in French (Harris 1978, Lambrecht 1981); sentence connectives, as in Hixkaryana in Brazil (Derbyshire 1985); and possessive markers, as in Supyire in Mali (Carlson 1994). It is important to note that these functions are by no means universal. The following sub-sections discuss the functions of demonstratives in Chiyao.

6.1 Spatial-deictic role

Demonstratives found in most languages are deictically contrastive. Most languages have a proximal demonstrative denoting closeness to the deictic center and a distal demonstrative denoting some relative distance from the deictic center (Payne & Peña 2007). Some languages have more elaborate demonstrative systems than others. Spanish, for example, has *este*, *ese* and *aquel* which correspond to proximal, medial and distal demonstratives, respectively (Payne & Peña 2007). There are also other languages that make more than a two-way distinction of demonstratives. Tlingit (North West American) and Samal (Philippines), for example, have four deictic dimensions of demonstratives, while Malagasy (Austronesian) has six deictic dimensions (Levinson 1983). Chiyao makes a three-way distinction in deictic demonstratives, namely close to the speaker (proximal), far from the speaker but closer to the addressee (non-proximal), and far from both the speaker and the addressee (distal). The Chiyao examples in (32) are illustrative of this system.

(32) a. proximal
li-jela ly-ángu **ali**
5-hoe 5-mine 5.DEM.PROX
'My hoe is this one.'

b. non_proximal
li-jela ly-ángu **alyo**
5-hoe 5-mine 5.DEM.NON_PROX
'My hoe is that one.'

c. distal
li-jela ly-ángu **alila**
5-hoe 5-mine 5.DEM.DIST
'My hoe is that one (over there).'

The examples in (32) above illustrate the spatial-deictic role of demonstratives in Chiyao by using the class 5 noun *lijela* 'hoe'. Thus, the demonstrative *ali* (32a) refers to a proximal entity, *alyo* (32b) refers to the non-proximal referent, and *alila* (32c) refers to a distal entity.

6.2 Emphatic role

Cross-linguistically, demonstratives are known for their property of encoding emphasis. In many languages, the expression of emphasis through demonstratives is done by reinforcing the existing demonstrative morphologically, through for example, the addition of further deictic particles, or through reduplication (Lyons 1999; Asiimwe (2024 [this volume])). In Swahili, for example, reduplicated forms of demonstratives such as *hikihiki* 'this', and *kilekile* 'that' are widely attested (Lyons 1999: 116). Like Swahili, Chiyao widely employs demonstrative reduplication to indicate emphasis of some parts of a message conveyed. The emphatic demonstrative is alternatively referred to as the confirmative demonstrative, as it induces the meaning 'the very (same)' (Floor 1998). The Chiyao examples in (33) below are illustrative of this phenomenon.

(33) a. va-lendo vá-aiché u-síku **úla** **úla**
2-guest SM3-arrive.PAST 14-day 14.DEM.DIST 14.DEM.DIST
'The guests came on that very day (not any other day).'

b. tw-ápité li-tálá **líla** **líla**
2-SM-pass 5-way 5.DEM.DIST 5.DEM.DIST
'We used the same way (not any other way)'

As is evident in (33) above, the demonstratives *ula* (33a) and *lila* (33b) are reduplicated for emphatic purposes. The reduplicated demonstratives drop the initial vowel, a common tendency for post-nominal non-deictic demonstratives in Chiyao. The full forms of these demonstratives would be *aula* (33a) and *alila* (33b).

The emphatic role of demonstratives has also been reported in Makhuwa, as in (34) (van der Wal 2010: 186).

(34) Makhuwa (van der Wal 2010: 186)
yoólé mpákhá wa-ámútsy' aáwe
1.dem.dist until 16-2.family 2.POSS
'She/the same went to family's place.'

6.3 Definiteness role

It is generally assumed that demonstratives are universally definite, and that definiteness exists in all languages (Lyons 1999). Lyons (1999: 2) associates definiteness with such properties as familiarity and identifiability. Thus, something definite is familiar to and identifiable among interlocutors. In many Bantu languages, demonstratives have been observed to perform a function similar to definite articles in languages which have articles (Van de Velde 2005). In this way, the demonstrative is used to refer to a referent which is identifiable to both speaker and hearer. This role of demonstratives has been described in a number of Bantu languages, including Chaga (E62), Nyamwezi (F22), and Dciriku (K62) (Van de Velde 2005). Similarly, Taji (2020: 53) reported that demonstratives are important indicators of definiteness in Chiyao. In the examples below, the post-nominal demonstratives *úla* (35b) and *líla* (36b) show that the nouns they modify are familiar among the interlocutors, and thus definite. These contrast with the nouns in (35a) and (36a), which are not modified by demonstratives, suggesting that they are indefinite.

(35) Taji (2020: 53)

a. m-kologo u-jitíche
3-alcohol SM3-be spilt
'Alcohol has been spilt.'

b. m-kologo **úla** u-jitíche
3-alcohol 3.DEM.DIST SM3-be spilt
'That/the alcohol has been spilt.'

(36) Taji (2020: 53)

a. m-ka-jigále li-jela
1SM-FUT-take 5-hoe
'Go and bring a hoe.'

b. m-ka-jigále li-jela **líla**
1SM-FUT-take 5-hoe 5.DEM.DIST
'Go and bring that/the hoe.'

Demonstratives that are used to express definiteness appear in three deictic distinctions – proximal, i.e. closer to the speaker (37), non-proximal, i.e closer to the hearer (38), and distal, i.e. far from both speaker and hearer (39).

(37) **achi** chí-téélá **chi**
7.DEM.PROX 7-tree 7.DEM.PROX
'This tree (near me, speaker)'

(38) **acho** chí-téélá **cho**
7.DEM.NON_PROX 7-tree 7.DEM.NON_PROX
'That tree (near you, hearer)'

(39) **achila** chí-téélá **chíla**
7.DEM.DIST 7-tree 7.DEM.DIST
'That tree (far from both of us)'

In (37–39), the NPs are definite as they are modified by demonstratives. The definite reading of these sentences is further induced by the spatial deictic nature of the demonstratives used, which show that the referents are within the interlocutors' visibility.

As one would expect, there are structural and semantic differences between the sentences in (35–36) and those in (37–39) above. The examples in (35–36) contain only post-nominal demonstratives, and these refer to entities away from the interlocutors' line of vision, while those in (37–39) each contain both pre-nominal and post-nominal demonstratives, and these refer to entities within the interlocutors' vision. As such, they may be accompanied with a pointing gesture. This interpretation is consistent with Taji (2020) who observed that, in Chiyao, demonstrative doubling is related to deictic definite NPs as in (37–39) above while single occurrence of demonstratives is associated with anaphoric reference as in (35–36). In anaphoric reference, demonstratives are used to refer to an entity with which the hearer is familiar not from the physical situation but from the broader linguistic context. The hearer is familiar with the entity because it was mentioned earlier in the text or discourse. One important aspect of anaphoric reference that deserves a separate section as an independent role of demonstratives is tail-head linking, and this is discussed in §6.4 below.

6.4 Tail-head linking

Tail-head linking involves repetition of some part (usually the last – the tail) of the previous sentence in the immediately following sentence (van der Wal 2010: 201). This function of demonstratives is commonly encountered in narratives where an entity or character that has been mentioned in the last part of the previous sentence is co-referenced in the following sentence through a demonstrative. In (40) below, the first sentence introduces the location *mumbugu* 'inside a cave' which is referred to in the following sentence through a post-nominal demonstrative.

(40) ni a-jáwile ku-li-sísa mú-mbugu. Ambano mú-mbugu
and SM1-go.PST INF-REFL-hide 18LOC-9.cave now 18LOC-9.cave
múla mw-álijí ni méési.
18.DEM.DIST 18-contain.PST with water
'And he/she went to hide himself/herself inside a cave. Now, inside that cave, there was water.'

Tail-head linking through demonstratives has also been reported in Makhuwa (van der Wal 2010: 201). In (41a) below, a woman is introduced into the story and she is referred to in the next sentence through a doubled demonstrative (41b):

(41) a. o-m-phwánya nthíyáná m-motsá
SM1.PERF.DJ-1-meet 1.woman 1-one
'He met a woman.'

b. **ólé** nthíyán' **uule** kh-oóthá aa-páh
1.DEM.PROX 1.woman 1.DEM.PROX NEG-SM1.IMPF-lie SM1.IMPF-light
ólumwenku
14.world
'This woman didn't just lie, she set the world on fire.'

7 Conclusion

This chapter has provided a description of Chiyao demonstratives, focusing on their form, distribution, and functions. It has revealed that in terms of form, the Chiyao demonstrative is comprised of an initial element, which is always the vowel *a-*, followed by an agreement marker, and ends with a final element which changes according to the location of the referent in relation to the speaker or

hearer. The demonstrative may be doubled, thus occurring in both pre- and post-nominal positions. In such cases, the post-nominal demonstrative drops the initial vowel but the pre-nominal demonstrative remains intact. The dropping of the initial vowel in the post-nominal demonstrative suggests that the initial element in the Chiyao demonstrative is optional.

The chapter has further provided a syntactic description of demonstratives in Chiyao. The findings in this aspect have shown that, within the broader classification of demonstratives in Bantu languages, the Chiyao demonstratives appear in four main types. The first type includes pronominal demonstratives, which are used as independent pronouns, and the second type encompasses adnominal demonstratives, which modify nouns. The third type of demonstrative that are encountered in Chiyao is that of adverbial demonstratives. These function to modify verbs and locative nouns. The last type includes identificational demonstratives, which occur in copula and non-verbal clauses.

With regard to syntactic distributions of demonstratives, it is established that the Chiyao demonstrative can occur either post-nominally, or in both pre-nominal and post-nominal positions simultaneously.

Finally, the description of functions of demonstratives in Chiyao has shown that demonstratives serve various linguistic and communication purposes, including expressing the location of an entity in relation to interlocutors, emphasis, definiteness, and tail-head linking in narratives. It is hoped that these findings will contribute to further understanding of the behavior of demonstratives in Bantu languages and will inspire further research in this area.

Acknowledgements

I appreciate the inputs I received from the team of researchers of the project Morphosyntactic Variation in Bantu at the manuscript development workshop held in Dar es Salaam in September 2017. I am also grateful to the anonymous reviewers whose comments helped to improve this work to its present form.

Abbreviations

ADJ	Adjective	DIST	Distal
ASSOC	Associative	FUT	Future
AUG	Augment	IMPF	Imperfective
DEM	Demonstrative	INF	Infinitive
DJ	Disjoint	INTENS	Intensifier

LOC	Locative	PLA	Plural addressee
NEG	Negative	POSS	Possessive
NON_PROX	Non-proximal	PRES	Present
NUM	Numeral	PROX	Proximal
OM	Object marker	PST	Past
OPT	Optative	REL	Relative clause
PFV	Perfective	REFL	Reflexive

References

Asiimwe, Allen. 2014. *Definiteness and specificity in Runyankore-Rukiga*. Stellenbosch: Stellenbosch University of South Africa. (Doctoral dissertation). http://hdl.handle.net/10019.1/95926 (16 March, 2015).

Asiimwe, Allen. 2024. The structure, distribution and function of demonstratives in Runyankore-Rukiga. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 43–83. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663765.

Benton, Richard A. 1971. *Pangasinan reference grammar*. Honolulu: University of Hawaii Press.

Carlson, Robert. 1994. *A grammar of Supyire*. Berlin: Mouton de Gruyter.

Denny, J. Peter. 1982. Semantics of the Inuktitut (Eskimo) spatial deictics. *Journal of American Linguistics* 48. 359–384.

Derbyshire, Desmond C. 1985. *Hixkaryana*. Amsterdam: North-Holland.

Dicks, Ian D. & Shawn Dollar. 2010. *A practical guide to the understanding of Chiyawo*. Zomba: Kachere Series.

Diessel, Holger. 1999. The Morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology* 3. 1–49.

Dixon, Robert Malcolm Ward. 2003. Demonstratives: A cross-linguistic typology. *Studies in language* 27(1). 61–112.

Du Plessis, Jacobus Albertus. 1978. *IsiXhosa 4*. Bellville: Via Afrika.

Du Plessis, Jacobus Albertus & Marianna Visser. 1992. *Xhosa syntax*. Pretoria: Via Africa.

Floor, Sebastian. 1998. Confirmative demonstratives. Ms., SIL Nampula.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language*, 47–82. Stanford, CA: Stanford University Press.

Harris, Martin. 1978. *The evolution of French syntax: A comparative approach*. London: Longman.

Heath, Jeffrey. 1984. *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.

Lambrecht, Knud. 1981. *Topic, antitopic and verb agreement in non-Standard French*. Amsterdam: John Benjamins.

Lehmann, Christian. 1984. *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Narr.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Maho, Jouni Filip. 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 (20 March, 2013).

Makanjila, Dominick. 2019. *The internal syntax of the Chimakonde determiner phrase*. Stellenbosch: Stellenbosch University. (Doctoral dissertation).

Marconnes, Francisque A. 1931. *A grammar of Central Karanga: The language of Old Monomotapa as at present spoken in Central Mashonaland, Southern Rhodesia* (Supplement to Bantu Studies 5). Johannesburg: Witwatersrand University Press.

Moltmann, Friederike. 2013. Identificational sentences. *Natural Language Semantics* 21. 43–77.

Ngunga, Armindo. 1997. *Lexical phonology and morphology of the Ciyao verb system*. Berkeley, CA: University of California at Berkeley. (Doctoral dissertation).

Ngunga, Armindo. 2002. *Elementos de gramatica da lingua Yao*. Maputo: Universidade Eduardo Mondlane.

Nicolle, Steve. 2007a. Metarepresentational demonstratives in Digo. In Randi Alice Nilsen, Nana Aba Appiah Amfo & Kaja Borthen (eds.), *Interpreting utterances: Pragmatics and its interfaces: Essays in honour of Thorstein Fretheim*, 127–146. Oslo: Novus.

Nicolle, Steve. 2007b. Textual functions of Chidigo demonstratives. *SOAS Working Papers in Linguistics* 15. 159–171.

Nurse, Derek & Gérard Philippson. 1980. The Bantu languages of East Africa: A lexicostatistical survey. In Edgar C. Polomé & C. P. Hill (eds.), *Language in Tanzania*, 26–67. London: Oxford University Press for the International African Institute.

Akan Demonstratives. 2007. In Doris L. Payne & Jaime Peña (eds.), *Selected proceedings of the 37th Annual Conference on African Linguistics*, 134–148. Somerville, MA: Cascadilla Proceedings Project.

Rehg, Kenneth L. & Damian G. Sohl. 1981. *Ponapean reference grammar*. Honolulu: University Press of Hawaii.

Schuh, Russell G. 1977. Bade/Ngizim determiner system. *Afroasiatic Linguistics* 4(3). 1–74.

Taji, Julius. 2020. Definiteness in Chiyao. *Ghana Journal of Linguistics* 9(2). 44–64.

Taji, Julius & Abel Mreta. 2017. The structure of a noun phrase in Shimwela. *Journal of Linguistics and Language in Education* 11(1). 19–41.

Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics with special focus on Bantu and Mande*, 425–441. Tervuren: Royal Museum for Central Africa.

van der Wal, Jenneke. 2010. Functions of demonstratives in Makhuwa narratives. *Africana Linguistica* 16. 183–213.

Visser, Marianna. 2008. Definiteness and specificity in the isiXhosa determiner phrase. *South African Journal of African Languages* 28(1). 11–29.

Wald, Benji. 1973. Syntactic change in the lake languages of Northeast Bantu. *Studies in African Linguistics* 4(3). 237–268.

Whiteley, Wilfred Howell. 1966. *A study of Yao sentences*. London: Oxford University Press.

Chapter 3

The structure, distribution and function of demonstratives in Runyankore-Rukiga

Allen Asiimwe
Makerere University, Uganda

The aim of this chapter is to give an overview of the morpho-syntactic characteristics and functions of demonstratives in Runyankore-Rukiga (Bantu JE.13-JE.14, Uganda). The study shows that a typical Runyankore-Rukiga demonstrative is comprised of the core demonstrative morpheme prevalent in many Bantu languages, the noun class concord and a suffix. The suffix indicates the position of the referent from the speaker and/or hearer. There are other forms of the demonstrative discussed, such as the demonstrative *-nu* which is used mainly in narratives to mark a continuation topic, and the locative copulative demonstrative realised by the nasal *n-*, which expresses a more specific location. A much-neglected form of demonstrative *-ti* with both nominal and verbal properties, whose primary role is to express manner, is also discussed. In terms of position, data indicate that basic demonstrative types can either precede or follow the noun but it appears that there is no clear-cut connection between the position of the demonstrative and the role it plays. The chapter also discusses pragmatic roles of demonstratives categorized into exophoric for the non-anaphoric uses and endophoric for discourse functions of demonstratives. The study draws data from written literary sources, spontaneous speech and elicitation. The study is a pointer to specific and detailed studies that need to be conducted on demonstratives in Runyankore-Rukiga and other related Bantu languages.

1 Introduction

Runyankore-Rukiga (JE.13–JE.14) is an interlacustrine Bantu language cluster spoken as one of the main indigenous languages of Uganda. Indigenous lan-

Allen Asiimwe. 2024. The structure, distribution and function of demonstratives in Runyankore-Rukiga. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 43–83. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663765

guages of Uganda are classified according to their corresponding ethnic groups (Simons & Fennig 2017). As such, Rukiga is regarded as a language of the Bakiga and Runyankore as a language of the Banyankore. The population of these two ethnic groups is estimated to be a combined total of about six million, based on the 2014 Uganda National Population and Housing Census report. Due to the high level of mutual intelligibility between Runyankore and Rukiga, which is estimated at approximately 94% lexical similarity (Ladefoged et al. 1972, Simons & Fennig 2017), and due to their grammatical affinity, they are linguistically regarded as dialects of the same language hence the overall label Runyankore-Rukiga.

Runyankore and Rukiga belong to the Nyoro-Ganda group (following the classification provided by Maho 2009). Other Ugandan Bantu languages in the same group include Runyoro (JE.11), Rutooro (JE.12), Luganda (JE.15) and Lusoga (JE.16), among others. In the same way that Runyankore-Rukiga are often referred to under this joint label, Runyoro and Rutooro are also sometimes termed Runyoro-Rutooro due to their mutual intelligibility with approximately 93% shared lexicon (Ladefoged et al. 1972, Lewis et al. 2013). Runyankore-Rukiga and Runyoro-Rutooro together form "Runyankitara" (JE.10) – a "language" which is taught in some universities in Uganda (Bernsten 1998).[1]

Runyankore-Rukiga exhibits a concord system which is characteristic of Bantu languages. Elements that modify the noun agree with it in terms of grammatical gender and number. The augment is part of the nominal morphology of Runyankore-Rukiga, which is conditioned by syntactic but also discourse-pragmatic factors (e.g. see Asiimwe et al 2023). The example in (1) illustrates the concord system and the augment in the nominal domain.[2]

(1) ekyo (é)kikópo (é)kihango (é)kiríkutukura
e-ki-o (e)-ki-kopo (e)-ki-hango (e)-ki-riku-tukur-a
DEM-7-MED AUG-7-cup AUG-7-big AUG-7-IPFV-red-FV
'that big cup which is red'

This chapter offers a general descriptive overview of the demonstrative system in Runyankore-Rukiga. Various forms, the distribution, and functions of the demonstrative are discussed. The demonstrative in Runyankore-Rukiga is generally comprised of three parts: the initial demonstrative morpheme, the noun

[1] Runyakitara is not a language but a label used to teach the four mutually intelligible languages.

[2] Otherwise indicated, examples used are not inclined towards a particular language variety. When an example is drawn from an individual language variety say Rukiga or Runyankore, this is indicated.

class concord and the morpheme that shows distance although, in the proximal demonstrative form, distance is zero-marked. Like most other Bantu languages, Runyankore-Rukiga uses a three-way demonstrative system to mark the distance of a referent in relation to the position of the speaker and the hearer (see also §2).[3] In relation to the position of the referent, the following terms are adopted for the current description: proximal for an entity close to the speaker, medial for a referent that is near to the hearer and distal to refer to an entity that is far from both the speaker and the hearer (2) as indicated in the glosses.

(2) a. egy' énju
e-gi e-n-ju
DEM-9.PROX AUG-9-house
'this house'

b. enjw' égyo
e-n-ju e-gy-o
AUG-9-house DEM-9-MED
'that house'

c. enjw' ériya
e-n-ju ∅-e-riya
AUG-9-house DEM-9-DIST
'that house (far from both speaker and hearer but visible)'

The initial element in the morphology of the demonstrative is analysed here as the core demonstrative morpheme, which is underlyingly *a-*. The demonstrative core morpheme *a-* has also been described in a number of other Bantu languages (cf. Wald 1973, Du Plessis et al. 1992, Visser 2008). The analysis of the existence of this initial core morpheme in Runyankore-Rukiga is further motivated in §2 (see also Asiimwe 2014, 2016).

As an adnominal modifier, the demonstrative occurs either before or after the head noun (3a–3b). Runyankore-Rukiga also allows double demonstratives to modify a single noun (3c) for emphasis or particularisation (§4.3). The syntax of demonstratives is discussed in further detail in §3.

(3) a. Egi njojo ní nto
e-gi n-jojo ni n-to
DEM-9.PROX 9-elephant COP 9-young
'This elephant is young.'

[3]There are however some Bantu languages that present a four-way demonstrative system for example, Nande (JD.42) (Valinande 1984 via Van de Velde 2019), Chidingo (E.73) (Nicolle 2007) and Fwe (K402) (Gunnink 2018).

b. Enjoj' égi ni nto
e-n-jojo -*e-gi* ni n-to
9-elephant DEM-9.PROX COP 9-young
'This elephant is young.'

c. Egy' (é)njojw' égi ní nto
e-gi (é)-n-jojo é-gi ni n-to
DEM-9.PROX AUG-9-elephant DEM-9.PROX.COP is 9-young
'This (particular) elephant is young.'

Cross-linguistically, demonstratives play two key semantic roles: (i) to mark a referent that is present in the physical environment relative to the deictic centre and (ii) to refer to an entity already established in discourse (Diessel 1999: 7). Hence, demonstratives are generally used to mark familiar and accessible entities (Lyons 1999). Additionally, a demonstrative may serve to activate old knowledge assumed to be familiar to both the speaker and the hearer. This latter type of demonstrative has a "recognitional" role according to Himmelmann (1996) and Diessel (1999), as exemplified in (4). The analysis of pragmatic functions of demonstratives follows Diessel's (1999) classification of demonstratives into "exophoric" and "endophoric" uses, discussed in §4.

(4) Eshááha yaaw' ériy' énkúr' erahe?
e-shaaha ya-a-we ∅-*e-riya* e-n-kuru e-rahe?
AUG-9.watch 9-CONN-your DEM-9-DIST AUG-9-old 9-where
'Where is that old watch of yours?'

Data for this study come from folktales, newspaper texts, spontaneous speech, and a number of constructions were elicited. The challenge at present is that there is no substantial accessible corpus for Runyankore-Rukiga. Through reading various materials written in Runyankore-Rukiga, and analysing radio and conversation recordings, I identified sentences or parts of sentences or chunks of discourse which are relevant for the study. As a native speaker of Rukiga, I used introspection but also held consultations with other native speakers of both Runyankore and Rukiga on grammatical judgements of given constructions.[4,5]

[4]I wish to thank the following native speakers of Runyankore-Rukiga: Fridah Katushemererwe, Justus Turamyomwe, Aron Turyasingura, Celestino Oriikiriza, Misah Natumanya and Emmy Rwomushana for discussing parts of the data analysed in this chapter.

[5]Since there was no accessible corpus for Runyankore-Rukiga at the time of writing this chapter, it was not possible to present a quantitative analysis of the demonstrative system and the distribution of the forms within a corpus. This will remain an avenue for future research.

2 The morphology of demonstratives

This section discusses the various forms of the demonstrative in Runyakore-Rukiga. The basic form of the demonstrative is composed of three elements: the core demonstrative *a-* (§2.1), the agreement morpheme and the suffix that marks the position of the referent from the speaker and/or hearer (§2.2). There are other forms of the demonstrative discussed in this section, including the demonstrative *–nu* (§2.3) and the 'identificational' *n-* (§2.4). In addition, the locative (§2.5), as well as manner demonstrative forms (§2.6), are discussed.

2.1 Argument for the demonstrative root *a-*

Previous studies, such as Morris & Kirwan (1972) and Taylor (1985: 137–138), suggest that the initial morpheme of the demonstrative in Runyankore-Rukiga is an initial vowel.[6] For instance, Taylor (1985: 137) argues that the morpheme is an augment which is obligatory in the first degree of distance, that is, the proximal demonstrative in the present analysis. Taylor (1985) further points out that this morpheme is deleted when it appears within the scope of a negative operator just like any other augment of any other nominal under the same syntactic conditions. He further states that it is possible for the augment to be retained if the demonstrative is placed before the lexical noun such that, instead, it is the augment of the head noun that is affected by the negative operator. Taylor (1985) again notes that the demonstrative retains its augment when the head noun is implicit. Since Taylor (1985) observes contexts in which the initial morpheme of the demonstrative is retained, it appears that it is indeed a central element in the morphology of the demonstrative.

Wald (1973), Du Plessis et al. (1978) and Du Plessis et al. (1992) have established that the canonical core demonstrative morpheme which is attested in many Bantu languages is underlyingly the morpheme *a-*. As also seen in a number of other Bantu languages, the initial element of the demonstrative in Runyankore-Rukiga appears to be the core of the demonstrative although in some languages such as Chiyao (P.21), the initial element of the demonstrative is not indispensable since it can be dropped depending on the syntactic and pragmatic factors (see Taji (2024 [this volume])). Visser (2008: 28) notes that this morpheme may appear allomorphically as *a-*, *e-*, or *o-* depending on the vowel of the agreement prefix of head noun, following the rules of vowel harmony, identical to the form of the respective augment. This is observed for Runyankore-Rukiga as well.

[6] Other terms such as augment or preprefix are often used. The term augment is chosen for the current discussion.

The quality of the vowel of the class prefix determines the shape of the initial morpheme of the demonstrative. As shown in (5a-c), the initial morpheme appears as: /a/ if the vowel of the class concord is /a/; if the vowel of the noun class prefix is /u/, the core morpheme manifests phonologically as /o/, while /e/ is as a result of having /i/ or /o/ as the vowel of the noun class prefix. The examples in (5a–5c) show the proximal and medial forms. For the distal form in (5d), the core demonstrative *a-* is morphologically zero-marked.

(5) a. amat'áaga
a-ma-te a-ga
AUG-6-milk DEM-6.PROX
'this milk'

b. omut' óogu/ogwo
o-mu-ti o-gu/o-gu-o
AUG-3-tree DEM-3.PROX/DEM-3-MED
'this tree/that tree'

c. e-mit' éegi/egyo
e-mi-ti e-gi/e-gi-o
AUG-4-tree DEM-4.PROX/DEM-4-MED
'these/those trees'

d. omuti gúri(yà)
o-mu-ti ∅-gu-ri(ya)
AUG-3-tree DEM-3-DIST
'that tree (far from speaker and hearer)'

In the remaining part of this subsection, I give evidence to support the claim that the initial morpheme of the demonstrative is not an augment as has been claimed in a number of previous studies but, instead is the core morpheme of the demonstrative. I demonstrate that the initial morpheme of the postnominal adnominal demonstrative following a negative verb is not affected by the negative operator the same way as the augment of other nominals. The examples below show that nouns (6a) and nominal modifiers such as adjectives (6b), possessives (6c) and some quantifiers (6d) following a negative verb can appear without the augment but the initial element of the demonstrative is not affected by negation as shown in (6e).

(6) a. Tíbaareeba (é)ngagi
ti-ba-aa-reeb-a (e)-n-gagi
NEG-2.SBJ-N.PST-see-FV AUG-10-gorilla
'They have not seen (the) gorillas.'

b. Tíbaareeba (é)mpango
ti-ba-aa-reeb-a (e)-n-hango
NEG-2.SBJ-N.PST-see-FV AUG-10-big
'They did not see (the) big ones.'

c. Tíbaareeba (é)zaawe
ti-ba-aa-reeb-a (e)-za-awe
NEG-2.SBJ-N.PST-see-FV AUG-10-your
'They did not see yours.'

d. Tíbaareeba (é)zindi
ti-ba-aa-reeb-a (e)-zi-ndi
NEG-2.SBJ-N.PST-see-FV AUG-10-other
'They did not see (the) others.'

e. Abarámbuzi tíbareeb' ézi ngagi
a-ba-rambuz-i ti-ba-a-reeb-a *e-zi* n-gagi
AUG-2-tourist-NMLZ NEG-2.SBJ-N.PST-see-FV DEM-10.PROX 10-gorilla
'(The) tourists have not seen these gorillas.'

The example given in (7) below is from Morris & Kirwan (1972: 151), and it shows that the initial morpheme of the demonstrative can be omitted if the demonstrative comes immediately after a negative verb.[7] However, in the discourse studied, there are no cases where the demonstrative following a negative verb loses its initial element. Hence, the effects of the negation rule are likely to be offset by not placing the demonstrative immediately after the negative verb and this is the strategy that speakers mostly use (8a–8b), as was also observed in Taylor (1985: 137)[8]. Besides, in the spoken register, due to phonetic factors, it appears that some speakers may omit the initial morpheme of the demonstrative perhaps to aid production.

(7) Asiimwe (2014: 124)
Tindíkwenda ki kitabo
ti-n-riku-end-a *ki* ki-tabo
NEG-1.SBJ-IPFV-want-FV 7.PROX 7-book
'I do not want this book.' (Morris & Kirwan 1972: 151, glosses added)

[7] Glosses are mine.
[8] (8b) is a typical Rukiga sentence.

(8) Asiimwe (2014: 181)

a. Tindíkwend' ékitab' éki
ti-n-riku-end-a e-ki-tabo *e-ki*
NEG-1.SBJ-IPFV-want-FV AUG-7-book DEM-7.PROX
'I do not want this book.'

b. Ekitab' eki tíndakyenda
e-ki-tabo *e-ki* ti-n-ra-ki-end-a
AUG-7-book DEM-7.PROX NEG-1SBJ-IPFV-7-want-FV
'This book, I do not want it.'

According to speakers who I worked with for this study, the structures shown in (8a–8b) are the most natural forms and are the most widely used by speakers.[9] Although some of the speakers I consulted claim that it is acceptable to use a demonstrative without the initial morpheme in spoken discourse, others indicate that without the initial element of the demonstrative (as in (7)), the construction sounds odd.

Note further that, when a noun or its modifier follows the locative element *omu* 'in' or *aha* 'at/on', the nominal obligatorily loses the augment (Taylor 1985: 88), as shown in (9a). [10] However, when a demonstrative immediately follows either *omu* or *aha* (as in (9b)), the initial element of the demonstrative (which manifests as *e-* in (9b)) is retained. In addition, an obligatory suffix *-ri* is added to the locative element.

(9) a. Engagi nizituur' ómu kibira
E-n-gagi ni-zi-tuur-a o-mu ki-bira
AUG-10-gorilla IPFV-10.SBJ-live-FV AUG-18.in 7-forest
'(The) gorillas live in a/the forest.'

b. Engagi nizituur' ómury' éki kibira
E-n-gagi ni-zi-tuur-a *o-mu-*(ri)* *e-ki* ki-bira
AUG-10-gorilla IPFV-10.SBJ-live-FV AUG-18.in-SUFF DEM-7.PROX 7-forest
'(The) gorillas live in this forest.'

In addition to the suffix *-ri* being obligatorily added onto *omu/aha* before a demonstrative (9b), the suffix is required when *omu/aha* precedes nominal elements, which inherently bear no augment such as proper names (10a), pronouns

[9] I worked with four native speakers of Runyankore and two speakers of Rukiga.

[10] Taylor argues that *omu* and *aha* are prepositions. The idea of a nominal element is adopted from Asiimwe (2014) who argues that these elements have nominal properties (see Asiimwe 2014: 143–145 for a discussion).

(10b), the invariant *buri* 'every' (10c) and numerals (10d) (cf. Taylor 1985: 88–89). Since the demonstrative requires an obligatory *-ri* suffix on *aha/omu*, it can, therefore, be argued to belong to the same category of nominals which possess no augment.[11,12]

(10) a. aha*(ri) Kábagaráme
a-ha-ri Kabagarame
AUG-16.at-SUFF 23.Kabagarame
'at Kabagarame'

b. omu*(rí)iwe
o-mu-ri iwe
AUG-18-SUFF you
'in you'

c. omu*(ri) burí kyaro
o-mu-ri buri ki-aro
AUG-18-SUFF every 7-village
'in every village'

d. aha*(ri) mukáaga
a-ha-ri mu-kaaga
AUG-16-SUFF 6-six
'at six/out of six'

Further evidence to support the claim that the initial element of the demonstrative *a-* is the historical core of the demonstrative, responsible for definiteness meaning, comes from examining the referential (definite) morpheme *-a*, which appears to be the grammaticalised form of this core demonstrative (cf. Asiimwe 2014, 2016). An appropriate agreement prefix is attached as shown in (11) (but see Asiimwe 2016: 68 for a list of the forms for noun classes 1–18).

[11]The status of the suffix *-ri* remains unclear, and further research is needed to establish its full morphosyntactic properties. The element *-ri* is added when *omu/aha* is immediately followed by a demonstrative, proper name, pronoun, the invariant *buri* 'every' and numeral. It may be that it performs a similar function to the augment since nouns belonging to these categories cannot take an augment. See Beermann & Asiimwe (forthcoming) for further discussion of this.

[12]The locative class 17 *-ku-* is less productive in Runyankore-Rukiga (see Beermann & Asiimwe forthcoming).

(11) a. gwá muhánda
gu-a mu-hánda
3-REF 3-path
'that (other) path'

b. byá bitabo
bi-a bi-tabo
8-REF 8-book
'those (other) books'

c. rwá rutindo
ru-a ru-tindo
11-REF 11-bridge
'that (other) bridge'

Following Diessel's (1999) characterization of the grammaticalisation process associated with demonstratives, and additional evidence (cf. Asiimwe 2014, 2016), the demonstrative and the functional element *-a* can be seen to share semantic features, meaning that (i) they can be used interchangeably as shown in (11). Although the medial demonstrative form can be used in the place of *-a* without altering the semantics, it is mostly the distal demonstrative that can replace *-a* to locate a referent that is assumed to be accessible in the mind of the hearer. (ii) Like the demonstrative, the morpheme *-a* can be used as a tracking device for a referent already mentioned in the previous discourse. A medial demonstrative (*ogwo*) can replace *wa* as in (12) without altering the semantics of the discourse (this function of the demonstrative is discussed in §4.1). (iii) The morpheme *-a* shares with the demonstrative the feature of denoting a referent about which both the speaker and the hearer have common knowledge (as in (13)).

The context for both (13a) and (13b) is that speaker (A) asks the hearer (B) whether B got the book which A asked for, a referent that is assumed to have been talked about by A and B previously.[13]

(12) Asiimwe (2014: 215)
Enkundi ku erikumara kuragara omuzaire naaruga aha kiriri, *omwana* nibamushohoza aheeru…Nibaronda omwojo n'omwishiki b'omuka endiijo, reero omwishiki naaheeka *wa mwereere* naagyenda n'ogwo mwojo ou baija hamwe nibaza omu kishaka …(Karwemera 1994: 109)[14]

[13]Note that where Runyankore speakers use the distal demonstrative e.g., *ka-ri(ya)* (see Table 1) Rukiga speakers use *-a* (e.g., *kariya katabo* versus *ka katabo*).

[14]In the interest of space, long excerpts are not glossed.

'… After the umbilical cord has fallen off, the days for the mother to remain in the house with the child (after the child has been born) are over, and the *child* is taken outside…they get a boy and a girl from another family and the girl carries *that baby* on her back and goes with that boy, whom she has come with, to the bush …'

(13) a. Tindíkumanya yáába ká katabo waakantúngîire
ti-n-riku-many-a yaaba *k-a* ka-tabo
NEG-1.SBJ-IPFV-know-FV whether 12-REF 12-book
w-aa-ka-n-túngi-ire
2SG-PST-12-1SG-find-PFV
'I do not know whether you got for me that other (small) book.'

b. Asiimwe (2014: 205)
Tindíkumanya yáába *kárí(ya)* (á)katabo waakantúngîire
ti-n-riku-many-a yaaba *ka-ri(ya)* (a)-ka-tabo
NEG-1.SBJ-IPFV-know-FV whether 12-DIST AUG-12-book
w-aa-ka-n-tungi-ire
2SG-PST-12-1SG-find-PFV
'I do not know whether you got for me that other (small) book.'

In the phrase *wa mwereere*, 'that baby', the distal demonstrative *oriya* can be used in the place of *wa*: *oriya mwereere* 'that baby'. The two phonological words *wa* and *oriya*, in this case, are in complementary distribution. This shows that they play the same role and, perhaps, the morpheme *-a* evolved from the demonstrative core morpheme *a-*.

2.2 Proximal, medial and distal demonstratives

Many Bantu languages have a three, or four-way system of demonstratives, expressing the distance of the speaker or hearer (Nicolle 2012, Van de Velde 2019) from the deictic center. Runyankore-Rukiga typically follows a three-way demonstrative system marking the referent close to speaker (proximal), close to the hearer (medial) and far from both the speaker and hearer (distal). Also, most Bantu languages have retained the proximal suffix *-no* or its cognate *-nu* from Proto-Bantu (Ashton et al. 1954, Nicolle 2012, Ahn & van der Wal 2019). Runyankore-Rukiga maintains the suffix *-nu* commonly used as an anaphoric demonstrative in narratives (§2.3). The proximal demonstrative, which typically reflects the position of a referent close to the speaker, has no overt morphological marker for distance as shown in (14a) (also see Table 1). The medial demonstrative, also known as the demonstrative of reference, which refers to an entity

near the hearer has the suffix *-o* (14b). This suffix is also found in many other Bantu languages (cf. Ashton et al. 1954, Nicolle 2012, 2014, Ahn & van der Wal 2019, among others).

(14) a. eki kibira
e-ki ki-bira
DEM-7.PROX 7-forest
'this forest.'

b. ekyo ki-bira
e-ki-o ki-bira
DEM-7-MED 7-forest
'that forest.'

According to Taylor (1985: 135), the distal demonstrative is further divided between reference to visible and invisible entities. For distant objects, which are nevertheless visible to the speaker and hearer, Taylor observes that the suffix *-riya* is used, while for distant and invisible entities, he argues that *-ri* is used. However, in contrast to Taylor's (1985) claim, *-ri* and *-riya* are here considered to be variants of the same form (compare (15) and (16) but see also Table 1). They are used interchangeably for both visible and invisible referents and the use of either of the forms depends mainly on an individual's choice (Asiimwe 2014).[15] Example (15) shows that a speaker can choose either the *-ri* or its variant *-riya* for both visible and invisible referents. It should also be noted that referents in noun classes 1 and 9 only occur with the long form (*-riya*) as exemplified in (17a) with a class 9 noun. Note further that the same noun in the plural form (class 10) permits either the short or long form of the distal demonstrative (17b). Nouns in classes 1 and 9 have only a vowel as their demonstrative concord while the rest use a consonant plus a vowel, which might be the reason why the two noun classes behave differently with the distal demonstrative forms.

(15) Bárí/riya báâna nibaróndʼ oha?
∅-ba-ri/riya ba-ana ni-ba-rond-a o-ha?
DEM-2-DIST 2-child PROG-2.SBJ-look-FV 1-who
'Who are those children looking for?'

(16) Asiimwe (2014: 184)
Omwishiki oú yaashabíre akamugira ati "Shaná wááza kunshwéra obanze óité báríya báána baawe." (Karwemera 1975: 20)

[15]Note that with referents which are far from the speaker and the hearer, the core morpheme of the demonstrative is morphologically unmarked.

o-mu-ishiki o-u y-aa-shab-ire a-ka-mu-gir-a a-ti
AUG-1-girl AUG-REL.PRO 3.SBJ-PST-ask-PFV 3.SBJ-REM-1OM-say-FV 1-that
"Shana w-aa-za ku-n-shwer-a o-banz-e o-it-e
May.be 1-PRS-go INF-1.SBJ-marry-FV 2SG.SBJ-first-SBJV SG-kill-SBJV
∅-*ba-riya* ba-ana ba-awe
DEM-2-DIST 2-child 2-CONN.yours
'The girl whom he asked for a hand in marriage told him 'If you are to marry me, you will first kill those children of yours'.

(17) a. Embúz' érí*(ya)
e-n-buzi ∅-*e-ri*(ya)*
AUG-9-goat DEM-9-DIST
'that goat (over there)'

b. Embúzi zíri(ya)
e-n-buzi ∅-*zi-ri(ya)*
AUG-10-goat DEM-10-DIST
'those goats (over there)'

For referents that are far from the speaker but visible to both the speaker and the hearer, especially in spoken discourse, the final *-a* may be lengthened and pronounced with a high tone for emphasis, as in (18)–(19).

(18) enté ériyááá
e-n-te ∅-*e-riya* (Taylor 1985: 136–137)
AUG-9-COW DEM-9-DIST
'that cow right over there'

(19) Asiimwe (2014: 184)
Réébá ékintu kiríkugamba nkírí múríyááá.
∅-reeb-a e-ki-ntu ki-riku-gamb-a n-ki-ri
2PL-see-FV AUG-7-thing 7SBJ-PROG-talk-FV LD.COP-7-DIST
∅-*mu-riya* (Mubangizi 1966: 28)
DEM-16-DIST
'Look, the creature that is making some sound is there, in there (far but visible).'

In contrast, among the speakers of Rukiga, it is the vowel /i/ of the demonstrative suffix that may be lengthened and also pronounced with a high tone to refer to an object that is far but visible: e.g. *múrííya* 'in there'. The lengthening of the vowel is at times accompanied by pouted lips (cf. Morris & Kirwan 1972: 59) or a

pointing gesture for extra emphasis. Table 1 gives a summary of the forms of the demonstratives indicating the distance of the speaker or hearer from the deictic centre.

Table 1: Demonstratives in Runyankore-Rukiga adapted from Taylor (1985: 136). †: Visible and invisible referents.

Noun class		Proximal (this)	Medial (that)	Distal (that) †
1	-mu-	ogu	ogwo	oriya
2	-ba-	aba	abo	bari/bariya
3	-mu-	ogu	ogwo	guri/guriya
4	-mi-	egi	egyo	giri/giriya
5	-ri-	eri	eryo	riri/ririya
6	-ma-	aga	ago	gari/gariya
7	-ki-	eki	ekyo	kiri/kiriya
8	-bi-	ebi	ebyo	biri/biriya
9	-n-	egi	egyo	eriya
10	-n-	ezi	ezo	ziri/ziriya
11	-ru-	oru	orwo	ruri/ruriya
12	-ka-	aka	ako	kari/kariya
13	-tu-	otu	otyo	turi/turiya
14	-bu-	obu	obwo	buri/buriya
15	-ku-	oku	okwo	kuri/kuriya
16	-ha-	aha	aho	hari/hariya
17	-ku-	oku	okwo	kuri/kuriya
18	-mu-	omu	omwo	muri/muriya
20	-gu-	ogu	ogwo	guri/guriya
21	-ga-	aga	ago	gari/gariya

2.3 The demonstrative suffix *-nu*

The demonstrative may also take the suffix form *-nu* to which an appropriate noun class prefix is attached. The *-nu* form is commonly used in storytelling to refer to a referent that has already been mentioned (cf. Morris & Kirwan 1972: 59; also see Nicolle 2012, 2014 for Digo). It is usually used in narratives with third-person human referents. Also, *-nu* is used with personified nouns in folktales. The literal translation of *-nu* is 'this one', hence a proximal form of the demonstrative. This can be seen in example (20).

(20) Asiimwe (2014: 186)
Omukáma ayeta ómwigárire Nyakwégira. Kú áija ámubuuza atí "Ógu níwe shó ékiti?" *Onú* ati "Buzima niwé tatá ékiti" …(Mubangizi 1966: 31)
o-mu-kama a-et-a o-mu-igarire Nyakwegira. Ku
AUG-1-king 1.SBJ-call-FV AUG-1-princess 1.Nyakwegire When
a-ij-a a-mu-buuz-a a-ti "o-gu ni-w-e
1SBJ-come-FV 3.SBJ-2OM-ask-FV 1-that DEM-1.PROX COP-1-REL.PRO
sho ekiti?" *o-nu* a-ti "Buzima ni-w-e tata ekiti
father.your real 1-PROX 1-that true COP-1-REL.PRO 1.father real
'The king called the princess. And when she came, he asked her "Is this your real father?" Then this one (the princess) answered: "Yes he is my real father".'

2.4 Identificational demonstrative *n-*

Another demonstrative in Runyankore-Rukiga is formed by the nasal *n-*[16] which appears before the noun class prefix. Taylor (1985: 138) refers to this form as an emphatic demonstrative which means 'here she/he/it is'. I call this morpheme a locative demonstrative copulative because it marks the location of a referent with a copulative sense. This form is also found in Runyoro-Rutooro (JE.11-JE.12), a language cluster that is closely related Runyankore-Rukiga. For Runyoro-Rutooro, Rubongoya (1999) identifies this form as a nasal morpheme that is used for things that are visible and specified. IsiXhosa (S.41), a Bantu language of South Africa, also has this form of demonstrative (Du Plessis et al. 1992), performing the same function as in Runyankore-Rukiga. In Runyankore-Rukiga, the locative demonstrative copulative morpheme *n-* occurs in all the three demonstrative divisions (21a–21c).

(21) a. Amíízi ngága
A-ma-izi *n-ga-ga*
AUG-6-water LD.COP-6-6
'Here is the water.'

b. Amíízi ngágo
A-ma-izi *n-ga-ga-o*
AUG-6-water LD.COP-6-6-MED
'The water, there it is.'

[16]This form of the demonstrative should not be confused with the locative demonstrative discussed in §2.5.

c. Amíízi ngárí(ya)
a-ma-izi *n-ga-ri(ya)*
AUG-6-water LD.COP-6-DIST
'There is the water (far from speaker and hearer but visible).'

d. Amíízi ngagári
a-ma-izi *n-ga-ga-ri*
AUG-6-water LD.COP-6-6-DIST
'There is the water (far from speaker and hearer but visible).'

Observe that with demonstrative locative copulatives, the noun class concord is doubled in the proximal and medial demonstrative types (see Table 2 below) and typically not in the distal although in some varieties of Rukiga, the class prefix is duplicated in the distal demonstrative (21d) as well. However, it remains unclear why exactly the noun class concord is duplicated. It might also be argued that one of the two identical morphemes performs a function other than noun class concord. However, this is a question that needs to be investigated further. The prefix *n-* has a locative meaning and as such directs the addressee to a more specific location of a referent and an appropriate gesture can accompany this form.

The locative demonstrative copulative *n-* is used to locate entities which may be seen or have been referred to previously in a more specific place. Concerning the locative noun classes, the identificational morpheme *n-* can only combine with *ha* (class 16), e.g., *mpaha (n-ha-ha)* 'Here it (place) is' but not *ku* (class 17), **Nkuku* or *mu* (class 18): **mmumu (n-mu-mu).*

2.5 Locative demonstratives

Locative constructions in Bantu have generally attracted a lot of interest due to their morpho-syntactic versatility (see for example, Bresnan & Kanerva 1989, Cocchi 2000, Marten 2012, Bloom Ström 2015, Zeller 2013, 2017). Runyankore-Rukiga has three locative noun classes: *ha* (class 16), *ku* (class 17) and *mu* (class 18). The locative prefix *ha* marks a specific place, *ku* indicates a general location, while *mu* signals an internal location (Taylor 1985, Asiimwe 2014, Beermann & Asiimwe forthcoming). Generally, the class 16 locative noun prefix *ha-* is more productive than the other two classes, as it is the only prefix that can be attached to verbs and to most nominal modifiers (cf. Asiimwe 2014). Nevertheless, locative demonstratives can be formed with all three classes. There are also different realisations of the proximal demonstrative which are used for more specific versus

Table 2: The locative demonstrative copulative (adapted from Asiimwe 2014: 187)

Noun class	Proximal	Medial	Distal
1 -mu-	ngugu	ngugwo	nguri(ya)
2 -ba-	mbaba (n-ba-ba)	nbabo	mbari(ya)
3-mu-	ngugu	ngugwo	nguri(ya)
4 -mi-	ngigi	ngigyo	ngiri(ya)
5 -i/ri-	ndiri (n-ri-ri)	ndiryo (n-ri-ri-o)	ndiri(ya) (n-ri-ri(ya))
6 -ma-	ngaga	ngago	ngari(ya)
7 -ki-	nkiki	nkikyo (n-ki-ki-o)	nkiri(ya)
8 -bi-	mbibi (n-bi-bi)	mbibyo (n-bi-bi-o	mbiri(ya)
9 -n-	ngigi	ngigyo (n-gi-gi-o)	ngiri(ya)
10 -n-	nzizi	nzizo (n-zi-zi-o)	nziri(ya)
11 -ru-	nduru (n-ru-ru)	nduryo (n-ru-ru-o)	nduri(ya)
12 -ka-	nkaka	nkako	nkari(ya)
13 -tu-	ntutu	ntutyo (n-tu-tu-o)	nturi(ya)
14 -bu-	mbubu (n-bu-bu)	mbubwo (n-bu-bu-o	mburi(ya)
15 -ku-	nkuku	nkukwo (n-ku-ku-o)	nkuri(ya)
16 -ha-	mpaha (n-ha-ha)	mpaho (n-ha-ha-o)	Mpariya (n-ha-ri(ya))
17 -ku-	-	-	-
18 -mu-	-	-	-

non-specific locations while there are no variants for the medial locative demonstrative. As for the distal demonstrative, a more specific location is marked by lengthening the vowel /i/ of the suffix especially in the Rukiga dialects. The different forms of locative demonstratives are given in Table 3.

Among the forms given in Table 3, there are those which mark a place that is specific or very close to the speaker. These include the following: *hanu*/*hanuuya*, *kunu*/*kunuuya* and *munu*/*munuuya* (22), and these contrast with the regular forms *aha, oku,* and *omu*. Locative demonstratives that denote a more specific location are formed by removing the augment and attaching the suffix *-nu(uya)* to the locative prefix with either a long /u/ or /i/ depending on the dialect. Speakers of Rukiga mostly lengthen the vowel of the suffix *-nuuya or -riiya* for the proximal or distal demonstrative respectively (compare (22a) and (22b)).

Table 3: Locative demonstratives in Runyankore-Rukiga

Proximal		Medial		Distal	
aha/ hanu/ hanuuya	'here'	aho	'there'	hari(ya)/ hariiya	'there ('at' place not close to both speaker and hearer'
oku/ kunu/ kunuuya	'this side/ place'	okwo/ okwe[a]	'that side/ place'	kuri(ya)/ kuriiya	'there (wider (unspecified place)'
omu/ munu/ munuuya	'in here'	omwo/ omwe	'in there'	muriya/ muriiya	'in there (an 'in-location' far from speaker and hearer)'

[a]There are dialectal differences; the variants *okwe/omwe/muriiya/kuriiya/hariiya* are commonly used in Rukiga.

(22) a. Amíízi gashuké múnúúya/múríiya Rukiga
a-ma-izi ga-shuk-e *mu-nuuya/mu-riiya*
AUG-6-water 6SBJ-pour-SBJV 18-here/18-there
'(You) Pour the water (exactly) in here/there.'

b. Amíízi gashuké múnu/múríya Runyankore
a-ma-izi ga-shuk-e *mu-nu/mu-riya*
AUG-6-water 6SBJ-pour-SBJV 18-here/18-there
'(You) Pour the water (exactly) in here/there.'

Locative demonstratives are sometimes categorized as demonstrative adverbs (Dixon 2003) (also see §3.5). As adverbs, they can appear after the verb indicating a location described by the verb (23). They can also appear as locative demonstrative pronouns as shown in (24). For future research, a detailed analysis of the categorial status of locative demonstratives is recommended.

(23) Engagi zaarabá áha
e-n-gagi z-aa-rab-a *a-ha*
AUG-10-gorilla 10.SBJ-N.PST-FV AUG-here
'(The) Gorillas have passed here.'

(24) Waagira ngu nookorerá Mbarara? Mbarará ókwo toríkubáása kúmpangirayó ákarimo kúnu kumburíírwe?
w-aa-gir-a ngu ni-o-kor-er-a Mbarara?
2SG-N.PST-say-FV that IPFV-2SG.SBJ-work-APPL-FV 23.Mbarara.
Mbarara o-ku-o ti-o-riku-baas-a
23.Mbarara DEM-17-MED NEG-2SG.SBJ-IPFV-can-FV
ku-n-hang-ir-a=yo a-ka-rimo *ku-nu* ku
INF-1SG-fodge-APL-FV=23 AUG-12-job 17-this.side but
n-buri-ir-w-e?
1SG-fail-APPL-PASS-PFV
'You have said that you work from Mbarara, right? Is it possible for you to find a small job for me there in Mbarara as I have failed to find one this side?'

2.6 The manner demonstrative *-ti*

Another form of the demonstrative found in Runyankore-Rukiga is *-ti* which expresses the manner in which something is done or perceived. This category is close to what Dixon (2003) identifies as verbal demonstrative and it has both nominal and verbal properties. As a nominal demonstrative, it can be used pronominally or adnominally for either proximal or medial location of a referent. It does not refer to entities far from both speaker and hearer. As a verbal demonstrative, it combines with personal pronouns as shown in (25). This form expresses manner in a similar way to that discussed by Guérin (2015), also in relation to the meaning 'do like this' suggested in Dixon (2003: 72). In terms of morphology, it appears as a suffix to which an appropriate noun class concord is attached. As mentioned already, the manner demonstrative modifies nominals (26a) and predicates (26b)

(25)

	Proximal (*like this*)	medial (*like that*)
1st pers.sg.	n-ti	n-ty-o/n-sy-o[17]
1st pers. pl.	tu-ti	tu-ty-o/tu-sy-o
2nd pers.sg.	o-ti	o-ty-o/o-sy-o
2nd pers. pl.	mu-ti	mu-ty-o/mu-syo
3rd pers. sg. human[18]	a-ti	a-ty-o/a-sy-o
3rd pers. pl. human	ba-ti	ba-ty-o/ba-sy-o

[18]The form *-syo* is commonly used in Rukiga.

[18]For non-human entities, the shape of the prefix takes the concordial form of the modified head noun.

(26) a. Context: In a visit to a zoo, one sees a three-horned chameleon for the first time.
Enyarujw' etí tínkagíreebahóga!
e-nyaruju *e-ti* ti-n-ka-gi-reeb-a=ho=ga
AUG-9.chameleon 9-PROX NEG-1SG.SBJ-ASP-9.OM-see-FV=PART=never
'I have never seen a chameleon like this one.'

b. Enyarujú neetambúra eti
e-nyaruju ni-e-tambur-a e-ti
AUG-9.chameleon IPFV-9-walk-FV 9-PROX
'A chameleon moves likes this (while demonstrating the way a chameleon moves)'

Unlike the basic forms of the demonstrative discussed in §2.2, which can either be pre- or postposed, the demonstrative *-ti* strictly follows the noun (27) or verb (28) it modifies. In addition, as an adnominal or pronominal demonstrative, it can be replaced by a basic demonstrative as shown (29).

(27) a. Ondééteré ékitabo kítí
o-n-reet-er-e e-ki-tabo ki-tí
2SG.OM-1SG-bring-APPL-SBJV AUG-7-book 7-PROX
'(You) bring for me a book like this (one).'

b. * Ondeetere ki-ti e-ki-tabo
o-n-reet-er-e ki-ti e-ki-tabo
2SG.SBJ-1SG-bring-APPL-SBJV 7-PROX AUG-7-book
'(You) bring for me a book like this (one).'

c. Ondéétere kítí
o-n-reet-er-e ki-tí
2SG.SBJ-1SG-bring-APPL-SBJV 7-PROX
'(You) bring for me like this (one).'

(28) a. Abarungí bakora bátyo
a-ba-rungi ba-kor-a ba-ty-o
AUG-2-good/nice 2SBJ-do-FV 2-like.that-MED
'Lit. 'Good people do like that.' 'Good people behave like that'.'

b. * A-ba-rungi ba-ty-o ba-kor-a
AUG-2-good/nice 2-like.that-MED 2SBJ-do-FV

(29) Ondééteré ékitabo nk'éki
o-n-reet-er-e e-ki-tabo nk'-e-ki
2SG.SBJ-1SG-bring-APPL-SBJV AUG-7-book like-DEM-7.PROX
'(You) bring for me a book like this (one).'

Generally, the manner demonstrative has both endophoric and exophoric uses. It can be used for referents in the physical environment and its anchor is an entity, an action, a situation or an activity in a conversation or in a previous text.

So far, we have seen different forms of demonstratives found in Runyankore-Rukiga. The next section discusses the syntax of demonstratives, focusing on the position of the demonstrative in relation to the noun it modifies and its distribution in the verbal domain.

3 Syntax of Runyankore-Rukiga demonstratives

3.1 Position of the demonstrative in the NP

A demonstrative, like most other nominal modifiers in Bantu languages, generally shows concord with the noun it occurs with either as an adnominal or predicative demonstrative. In terms of position in the nominal domain, nominal modifiers in African languages generally follow the noun (c.f., Zeller 2020). However, there is considerable flexibility among some nominal modifiers including demonstratives. Van de Velde (2005) categorises Bantu languages into three on the basis of the position of the adnominal demonstrative: (i) Bantu languages in which the adnominal demonstrative is always preposed, (ii) languages where the demonstrative is postposed and (iii) languages that present a free word order between the noun and the demonstrative. Some of the Bantu languages of East Africa from Van de Velde's list which allow only postnominal demonstratives include Hema (JE.10), Ganda (JE.15), and Haya (JE.22).[19] However, (Lu)Ganda (JE.15), which is spoken in Uganda and neighbours Runyankore-Rukiga, allows a preposed demonstrative as well (Kawalya p.c.) as shown in (30). Van de Velde (2005) describes Rundi (JE.62) as a language which seems to allow only the prenominal demonstrative. Outside the Van de Velde's sample, Gunnink (2018) shows that the demonstrative in Fwe (K.402) is preferred in the prenominal position. In Chiyao, the demonstrative is canonically postnominal. In this language, according to Taji (2024 [this volume]) a prenominal demonstrative can only be allowed if it cooccurs with a reduced form of the demonstrative (that occurs without the initial

[19] Makhuwa (P.31) also allows only postnominal demonstratives (van der Wal 2010).

element) in the postnominal position as illustrated in (31). Runyankore-Rukiga falls in the category of languages which allow free word order between the noun and the demonstrative (Taylor 1985, de Blois 1970, Van de Velde 2005, Asiimwe 2014). Other Bantu languages spoken in Uganda which allow the demonstrative in the prenominal and postnominal position include Lugwere (Ahn & van der Wal 2019) and Runyoro-Rutooro (Rubongoya 1999).

(30) Luganda (Kawalya p.c.)
Bano *(á)bávúbuka abámánsámánsá ssénte baziggya wa?
ba-no *(a)-ba-vubuka a-ba-mansa-mansa ssente
2-PROX AUG-2-youth AUG-2.REL-scatter-scatter 10.money
ba-zi-ggy-a wa
2SBJ-10OM-get-FV where
'These youths who spend money lavishly, where do they get it from?

(31) Chiyao (Taji (2024 [this volume]))

a. *aú* m-gundá *u*
3.DEM.PROX. 3-farm 3.DEM.PROX
'this farm'

b. * *aú* m-gundá
3.DEM.PROX 3-farm
'this farm'

There is no consensus in terms of the canonical position of the adnominal demonstrative in Runyankore-Rukiga. According to Wald (1973), the demonstrative in Runyankore-Rukiga was historically prenominal. In contrast, Taylor (1985) considers the demonstrative in Runyankore-Rukiga to be canonically a postnominal modifier, which can, however, precede the noun for emphasis (Taylor 1985: 89). Synchronically, nominal modifiers in Runyankore-Rukiga apart from the quantifier *buri* 'every' typically occur in the postnominal position (confirming Zeller's (2020) observation) although can freely move to the prenominal position. The demonstrative may either precede or follow the head noun as the examples in (32) show. Different factors may influence the position of the demonstrative in the nominal domain, including individuals' preference, dialectal variation, genre or register, and spoken versus written discourse (cf. Asiimwe 2014 for further discussion of this). The position of the demonstrative may not influence the role the demonstrative plays in the nominal domain. However, this needs to be investigated, putting into consideration that information structure, for example, might influence word order in the nominal domain.

(32) a. Ndéétera ébiníka égyo!
N-reet-er-a e-binika *e-gi-o*
1SG.SBJ-bring-APPL-FV AUG-9.kettle DEM-9-PROX
'Bring me that kettle!'

b. Ndéétera égyo biníka!
N-reet-er-a *e-gi-o* binika
1SG.SBJ-bring-APPL-FV DEM-9-PROX 9.kettle
'Bring me that kettle!'

The example in (33) is extracted from a weekly local Runyankore-Rukiga newspaper called *Orumuri*.[20] In this newspaper, discourse demonstratives are consistently postnominal. In literary works, the demonstrative occurs either prenominally or postnominally (cf. examples (34)). In spoken discourse, it has been observed that the demonstrative is mostly preposed and the augment on the modified noun retained regardless of its function (see §3.2).

(33) Asiimwe (2014: 197)
Bwanyima y'eka y'abantu 4 kwitwa oburwaire butamanyirwe, abashaho bagyerizeho kukyebera nikwo kushanga ngu n'oburwairwe bwa MARBURG [...] *Endwara egi* ebarukireho omuri Kabale [...] *oburwaire obu* nibukwata nka Ebola [...]
Orumuri newspaper (October 22–28, 2012)
'After four members of one family had died of an unknown disease, doctors carried out tests, and found out the disease to be Marburg [...] *This disease* broke out in Kabale [district][...] *this disease* has signs like those of Ebola [hemorrhagic fever]'

(34) Asiimwe (2014: 207)
Eshaaha y'okwetebeekanisiza okutemba endegye ekarindwa yaahika; ntyo naasiibuurana n'abo abaabaire banshendekyereize. Emigugu yangye naagyehisya haihi, naaza omu runyiriri. Ntyo naahika ahi barikushwijumira tikiti, paasipoota hamwe n'emigugu. Tikiti naagiha *empagare* nungi, nayo ndeeba yaagirabyamu amaisho kandi yaateeraho sitampu. *Emigugu* yangye bagirabya omu kyoma, bangira ngu tiinyine nshonga yoona. *Egyo mpangare* engira ngu *egyo migugu* niinyija kugishanga Gatiwick. (Mugumya 2010: 1)
'Time to prepare for boarding the plane came. I bid farewell to those who had accompanied me. I got my baggage closer and joined the queue. I then

[20]Publication of the newspaper was suspended in May 2020.

approached the passport, air-ticket and baggage checking desk. I handed my ticket to a *nice looking lady*. She checked it and then stamped it. My *baggage* was sent on a conveyor belt, and I was told that there was no problem. *That lady* told me that I will find *that baggage* at Gatwick.'

3.2 Demonstratives and the head noun augment

According to Taylor (1985), if the noun is preceded by a demonstrative, the augment of the noun is omitted. This is also reported in Dewees (1971) and de Blois (1970). However, there is evidence indicating that synchronically, the augment can be retained on the head noun with a prenominal demonstrative (35). Hence, since the augment and the demonstrative can co-occur, they are not in complementary distribution.

(35) Asiimwe (2014: 199)
[…] biriyoni 15 ezi baihire ahari difensi omukama we naagira ngu timurikuziihaho. Kandi nabo abantu bagira ngu Your Excellency kasita eki twabaire nituteeraho esente nyingi aha rutaro tukaba twine orutaro omu north, tukaba twine orutaro nkahi [gap]. Mbwenu hati obu rutakiriho katuzite omu kurwanisa *aba abakazi* 16 abarikufa buriizooba, *aba abantu* 300 abarikwitwa omushwija buriizooba, *aba abaana* 435 abarikufa ahabw'endwara ezi twakubaire nitutamba, tube nikyo twaza kukora kandi nyowe nindeeba tikyakubaire kiri ekizibu.[21]
'[…] 15 billion which was to be deducted from the defense budget, the chief [the president] as for him, he says that they cannot deduct it. And as for them, the people say that 'Your Excellency' we were allocating more money to war [defense] because there was war in the north, we had a war… where [gap]? Now that the war is no more, let us use this money in the struggle to minimize the level of death of *these 16 women* who die every day [of maternal health related complications], *these 300 people* who die of fever [malaria] every day, [and] *these 435 children* who die as a result of diseases we could prevent. That is what we should now do and to me, I see that would not be a problem.'

The use of an augment on a noun with a preposed demonstrative appears to be a recent occurrence. While it is commonly found in spoken discourse, it is hardly found in written works. Languages are not static; retaining an augment on the

[21]Recorded from a radio program, *Katuhurirane*, loosely translated as 'Let us hear from one another' on *Radio West* on 21.09.2012: 9.00pm, EAT.

noun with a prenominal demonstrative can be attributed to language change and language contact. It may be the case that Runyankore-Rukiga speakers are influenced by Luganda[22], since in Luganda the augment of the noun modified by a prenominal demonstrative is obligatory as in *bano *(a)bavubuka*[23] 'these youths' in (30) above. Bantu languages spoken in Uganda that are in close proximity with Luganda geographically are borrowing heavily from Luganda. Borrowing is witnessed almost at all levels of linguistics, especially at the lexical level.

Regarding the role of the augment on a noun preceded by a demonstrative, some speakers do not attach any meaning to it. However, when it is present on a noun, it appears to play the pragmatic role of adding emphasis to the noun (cf. Asiimwe 2014). In the radio recording given in (35) above, the speaker consistently uses the augment on nouns preceded by demonstratives and the augment appears to encode an additional feature of emphasis on these nouns. In the written discourse examined, there is only one case that has been identified where the augment is retained on the head noun with a prenominal adnominal demonstrative, given in (36). The extract is from a book written by a Rukiga speaker. One can, perhaps, predict that Rukiga speakers may have already introduced the augment on a noun preceded by a demonstrative in the formal discourse, which may be an indication that Rukiga is at a further stage in the process of language change.

(36) Asiimwe (2014: 199)
Ku baabaire bahikaho, babashangisa aha irembo enjugano zibaikiriziine. Bakaba babanza kubooreka ente kaasha (tikirikumanyisa ngu *egi ente* eine akaasha omu buso, kureka nikimanyisa ente nungi erikuhita ezindi)[…] Rukiga (Karwemera 1994: 86)
'After arriving there, they would meet at the gate and they [the girl's relatives] would look at the bride price which had been agreed upon. They would first show them the cow *kasha* (*kasha* does not mean that *that (specific) cow* has a white spot on its forehead, instead it means a nice looking healthy cow amongst all the cows brought)[…]'

3.3 Double demonstratives in one NP

As reported in Asiimwe (2014), two demonstratives of the same form can co-occur in Runyankore-Rukiga. One demonstrative may appear preposed and the other

[22]Luganda appears to have a great influence on many indigenous languages in Uganda.
[23]When the augment is removed from the noun on the Luganda phrase, *Bano bavúbúka* becomes a copula clause 'These are youths.'

one postposed, as in (37a), or both may appear in a sequence in the same position, either prenominally (37b) or postnominally with no difference in interpretation. Two demonstratives of different forms can also co-occur in the same NP. For example, the locative copulative demonstrative can co-occur with a non-locative demonstrative (38) and add a level of emphasis (see also §4.3). It is also possible for pronominal forms to co-occur. The use of two adnominal demonstratives adds emphasis but may also be used for confirmatory reasons: *egi ngigi* 'this very one' as in (38).

(37) a. Egyo njú égyó éióríkureeba kúri niharááramú omugabe
e-gi-o n-ju *e-gi-o* e-i
DEM-9-MED 9-house DEM-9-MED AUG-9.REL.PRO
o-riku-reeb-a ku-ri ni-ha-raar-a=mu o-mu-gabe
2SG.SBJ.REL-PROG-see-FV 17-DIST IPFV-16OM-sleep-FV=18 AUG-1-king
'That house that you see over there, a King lives there.'

b. Egyó égyó énju éí oríkureeba kúri niharááramú omugabe
e-gi-o *e-gi-o* e-n-ju e-i
DEM-9-MED DEM-9-MED AUG-9-house AUG-9.REL.PRO
o-riku-reeb-a ku-ri ni-ha-raar-a=mu o-mu-gabe
2SG.SBJ.REL-PROG-see-FV 17-DIST IPFV=16OM-sleep-FV=18 AUG-1-king
'That house that you see over there, a King lives there.'

(38) Asiimwe (2014: 210)
Aha murúndí ógu nkataayaayira ényanja ya Lomond na Ness. Egi ngígí ekabá neegámbwahó kúkye
a-ha mu-rundi o-gu n-ka-taayaay-ir-a e-n-yanja
AUG-16 3-time DEM-3.PROX 1SG.SBJ-REM-visit-appl-FV AUG-9-lake
y-a Lomond na Ness. *e-gi* *n-gi-gi*
9-conn 23.Lomond and 23.Ness DEM-9.PROX LDCOP-9.PROX-9.PROX
e-ka-ba ni-e-gamb-w-a=ho ku-kye
9-REM-be IPFV-9-talk-PASS-FV=16.PART 15-little
(Mugumya 2010: 59)

'This time around, I visited Lake Lomond and Ness. For this very one, no one talked about it much.'

3.4 Pronominal demonstratives

Demonstratives can play a pronominal role. A pronominal demonstrative replaces a noun phrase in an argument position of a verb. A pronominal demon-

strative identifies a referent that has been previously established in the discourse, or a referent that can be identified in the physical environment (39) or generally accessible by the hearer. The example in (39) shows the use of the demonstrative pronoun in an adverbial position (but see next Section). In (40), *ekyo* is used as an endophoric demonstrative pronoun. Note that the demonstrative *-nu* (see also discussion in §2.3 and §4.2) can only be used pronominally, as shown in (41) and it is mostly used in narratives.

(39) Omukázi naakund' éki
o-mu-kazi ni-a-kund-a *e-ki*
AUG-1-woman IFPV-1.SBJ-like-FV DEM-7.PROX
'The woman likes this (one).'

(40) Ekyo nookimányahó ki?
e-ki-o ni-o-ki-mány-a=ho ki
DEM-7-MED IPFV-2.SBJ-7OM-know-FV=16.PART what
'What do you know about that?'

(41) Ónu ebishumuuruzo akabitíína tarábiréesire.
o-nu e-bi-shumuurur-o a-ka-bi-tiin-a
1-PROX AUG-8OM-key-NMLZ 3.SBJ-REM-8OM-fear-FV
ti-a-ra-bi-reet-ire (Mubangizi 1997: 46)
NEG-3.SBJ-REM-8OM-bring-PFV
'This one feared to bring the keys.'

3.5 Demonstrative adverbs

Demonstratives can be used as verbal modifiers (Diessel 1999). The same forms of adnominal demonstratives occur as demonstrative adverbs (42a–42c). We noted that the noun class 17 locative prefix is not available as an adnominal demonstrative prefix but can be used when the demonstrative is used pronominally (as shown in Table 3 for locative demonstratives) and adverbially (42c). Generally speaking, demonstrative adverbs are commonly used pronominally. This tendency has also been observed in Makhuwa (van der Wal 2010).

(42) a. Engagi zaaraba aha/hánu
e-n-gagi z-aa-rab-a *a-ha/há-nu*
AUG-10-gorilla 10.SBJ-N.PST-pass-FV DEM-16/16-PROX
'Gorillas passed here.'

b. Engagi niziraará ómu
e-n-gagi ni-zi-raar-a *o-mu*
AUG-10-gorilla IPFV-10.SBJ-sleep-FV DEM-18.PROX
'Gorillas sleep in here.'

c. Engagi nizirééba kúri
e-n-gagi ni-zi-reeb-a *ku-ri*
AUG-10-gorilla PROG-10-see-FV 17-there
'Gorillas are facing the other side.'

The same forms of demonstratives can be used for temporal deixis. Any of the three demonstrative types can be used to refer to time (43a–43b). The proximal demonstrative specifically refers to the current time (43c) while either the proximal or the distal form can be used to refer to time in the past (43a–43b). The locative demonstrative for the current time may also take the *-nu* form as shown in (43d) in a non-verbal clause. The form *-nu*, which is not the locative copulative demonstrative commonly used in narratives (§2.3) is presumably borrowed from Runyoro-Rutooro where it is used when reference is made to the present time e.g., *obusumi bunu* 'these days'.

(43) a. Obwo tukaba nitusháárura ómugúsha
o-bw-o tu-ka-ba ni-tu-shaarur-a o-mu-gusha
DEM-14-MED 1PL.SBJ-REM-BE IPFV-1PL-harvest-FV AUG-3-sorghum
'At that time, we were harvesting sorghum.'

b. Búríya tukabá nitusháárura ómugúsha
∅-*bu-riya* tu-ka-ba ni-tu-shaarur-a o-mu-gusha
DEM-14-DIST 1PL.SBJ-REM-be IPFV-1PL-harvest-FV AUG-3-sorghum
'At that time, we were harvesting sorghum.'

c. Obu turimú nitusháárura ómugúsha
o-bu tu-ri=mu ni-tu-shaarur-a
DEM-14.PROX 1PL.SBJ-be=18.EXPLET IPFV-1PL-harvest-FV
o-mu-gusha
AUG-3-sorghum
'We are now harvesting sorghum.'

d. Búri tibwó búnu
∅-bu-ri ti-bu-o bu-nu
DEM-14-DIST NEG-14-PRON 14-PROX
'The current time is not the same as that (past) time.'

Not all forms of the demonstrative discussed in §2 can modify verbs. The pronominal demonstrative *-nu* (§2.3), for instance, does not occur in the verb phrase (44a). An identificational demonstrative also cannot immediately follow a verb in the main clause. It is only felicitous in a relative clause construction (44b) where it has a pronominal but not adverbial role. In the main clause, an identificational demonstrative may follow the ordinary demonstrative *aha.* When it immediately precedes or follows the verb, it functions as a nominal and requires a (locative) object agreement marker to correspond with (44c–44d).[24] Moreover, the identificational form *n-* only combines with noun class 16 but not 17 and 18 locatives (§2.4).

(44) a. Engagi zaabyama *zinu

b. Mpaha ahú zaaba zíbyami
n-ha-ha a-hu z-aa-ba zi-byami
LD.COP-16-16 AUG-16.RM 10.SBJ-N.PST-be 10.SBJ-sleep
'It is (exactly) here where they were sleeping.'

c. Wááharénga mpáho
w-aa-ha-reng-a *n-ha-ha-o*
2SG.SBJ-N.PST-16OM-pass-FV LD.COP-16-16-MED
It (the place) is right there, you are about to bypass it.'

d. Mpaho wááharenga
n-ha-ha-o w-aa-ha-reng-a
LD.COP-16-16-MED 2SG.SBJ-N.PST-16OM-pass-FV
'It (the place) is right there, you are about to bypass it'

The goal of this section was to discuss the syntax of demonstratives in Runyankore-Rukiga. The section has highlighted the fact that the demonstrative can either occur in the prenominal or postnominal position. We also note that the augment, especially in the spoken discourse, can appear on the noun preceded by a demonstrative. The same forms of demonstratives can be used adnominally, pronominally, as adverbs and as temporal deictic markers although not all forms

[24] *Mpaho* can be used expletively as a confirmatory pragmatic marker. As a pragmatic marker, it also functions to express surprise or noteworthiness:

(i) Mpáho wáákimanya
n-ha-ha-o w-aa-ki-many-a
DEM-16-16-MED 1SG.SBJ-N.PST-7OM-know-FV
'(I confirm that) you have understood it.'

of demonstratives discussed in §2 can be used adverbially. Next, I turn to the basic roles demonstratives play in the grammar of Runyankore-Rukiga.

4 Functions of demonstratives

This section discusses the pragmatic roles demonstratives play both in discourse and in the physical context. For purposes of this analysis, I follow Diessel's (1999: 7) categorization of demonstratives into exophoric (§4.1) and endophoric demonstratives (§4.2). Other functions not covered under these two broad categories, such as emphasis and specificity, are discussed in §4.3.

4.1 Exophoric uses

The exophoric category symbolizes the basic use from which all non-anaphoric uses of the demonstrative derive. The exophoric demonstratives accompany referents which are mostly visible and accessible in the spatial environment (45)–(46). For a referent that is available in the physical environment, the demonstrative can be further accompanied by a pointing gesture or a specific eye gaze or pointing lips (cf. Fillmore 1997, Diessel 1999, 2012, Lyons 1999, Levinson 2004) to guide the hearer further to the intended referent.

(45) Abarámbuzi tíbaareebá ézo ngagi
A-ba-rambuzi ti-ba-a-reeb-a *e-z-o* n-gagi
AUG-2-tourist-NMLZ NEG-2-N.PST-see-FV DEM-10-DIST 10-gorilla
'The tourists have not seen those gorillas.'

(46) Asiimwe (2014: 201)
Obú wíízire nsígarira n'ogú mwána nzé kwéreeter' ótwizi
obu w-a-iz-ire n-sigar-ir-a na *o-gu* mu-ana
since 2SG-PRS-come-PFV 1SG.SBJ-stay-APPL-FV with DEM-1.PROX 1-child
n-z-e ku-e-reete-er-a o-tu-izi
1SG.SBJ-go-FV INF-REFL-bring-APPL-FV AUG-12-water
'Now that you have come, stay with this child while I go to fetch for myself some water.'

An exophoric demonstrative may also be used to refer to an entity not visible to either the speaker or the hearer but assumed to be present in the physical environment (47). The medial demonstrative form is used for this purpose.

(47) Context: At the sound of a loud bang outside as heard from inside a house.
Eky' ékyángwa n'énki?
e-ki-o e-ky-a-ngw-a ni enki?
DEM-7-PROX AUG-7.REL-N.PST-fall-FV COP what
'What is that that has fallen?'

Relatedly, the pronominal proximal demonstrative can be used to refer to an invisible and unknown referent. Imagine a situation where the interlocutors are moving in a car; one looks outside and sees a vacuum flask thrown by the roadside and marvels as in (48): the person who threw the flask by the roadside cannot be identified nor be seen.

(48) Ogu shí furásika yaaginagira ki?
o-gu shí furásika y-aa-gi-nag-ir-a ki
DEM-1.PROX DM 9.flask 3SG.SBJ-N.PST-9OM-throw-APPL-FV why
'Why has this one thrown the vacuum flask?'

Exophoric demonstratives further take on a symbolic role (Fillmore 1997, Levinson 2004). According to Diessel (1999: 94), the symbolic demonstrative draws on knowledge about a larger situational context, which involves more than what is immediately visible in the surrounding situation. The symbolic use of the demonstrative is exemplified in (49).

(49) Asiimwe (2014: 204)
Abantu b'ómury' éky'(e)kyaro n'ábahíngi
a-ba-ntu b-a o-mu-ri *e-ki* (e)-ki-aro ni
AUG-2-person 2-CONN AUG-18-SUFF DEM-7.PROX AUG-7-village COP
a-ba-hingi
AUG-2-farmer
'People of this village are farmers.'
'People who live in this village are farmers.'

It is the proximal form of the demonstrative that is used for symbolic reference. Hence, the demonstrative use of *eki* in (49) is based on common knowledge about the larger situational context or the symbolic use of *ekyaro* 'village'. The symbolic demonstrative is a form of deictic demonstrative which does not take any form of gesture because it involves activating knowledge about the communicative event and the referent (Diessel 1999: 94).

4.2 Endophoric uses

The second major category consists of demonstratives whose role is to identify participants in an ongoing discourse. Diessel (1999) considers such to be typically anaphoric (also see Himmelmann 1996, Lyons 1999, Levinson 2004, Guillemin 2011 among others). The endophoric category also includes demonstratives with a recognitional role, a term that is attributed to Himmelmann (1996). The speaker draws the hearer to locate a referent in an on-going discourse. A demonstrative takes an anaphoric role if the referent exists in previous discourse. The antecedent of a demonstrative may be a noun phrase (noun phrase anaphora) or a piece of text – a clause, a paragraph, or even a full story (this is called "textual anaphora" according to Dixon (2003: 64)). Example (50) below illustrates noun phrase anaphora. Proximal demonstratives are not commonly used in Runyankore-Rukiga texts for phrase anaphora in contrast to some other Bantu languages such as Digo (Nicolle 2007, 2014). However, they can be used to indicate topic continuation in an ongoing discourse. Medial demonstratives are widespread in discourse especially, to signal a shift in topic (51) back to the major topic introduced previously and mostly used in a full NP. [25] Distal demonstrative forms and the proximal narrative demonstrative are used in turn to indicate shifts and turns between two participants in a narrative. This is possible when there is no intervening NP as exemplified with an excerpt from a folktale by Mubangizi (1966: 56–57) in (52). After a number of turns taken using only demonstratives (cf. (52)), full NPs are resumed.

(50) *Ekitóngore* éki Jáne yaabaire naakorá nakyó kikaba kitári mu mateeka. *Ekitongore ékyo* kikakingwa.
E-ki-tongore e-ki Jane y-aa-ba-ire
AUG-7-organisation AUG-7.REL.PRO 1.Jane 3SM.SBJ-PST-be-PFV
ni-a-kor-a na=ki-o ki-ka-ba ki-ta-ri mu ma-teeka.
IPFV-3SG-work-FV with=7-PRO 7.SBJ-REM-be 7-NEG-be in 6-law
e-ki-tongore e-ki-o ki-ka-king-w-a
AUG-7-organisation DEM-7-MED 7.SBJ-REM-close- PASS-FV
'The company that Jane worked with was illegally established. That company was closed.'

(51) Asiimwe (2014: 207)
Eshaaha y'okwetebeekanisiza okutemba endegye ekarindwa yaahika; ntyo naasiibuurana n'abo abaabaire banshendekyereize. Emigugu yangye

[25] For convenience, the passage in (33) is repeated in (50).

naagyehisya haihi, naaza omu runyiriri. Ntyo naahika ahi barikushwijumira tikiti, paasipoota hamwe n'*emigugu*. Tikiti naagiha *empagare* nungi, nayo ndeeba yaagirabyamu amaisho kandi yaateeraho sitampu. *Emigugu* yangye bagirabya omu kyoma, bangira ngu tiinyine nshonga yoona. *Egyo mpangare* engira ngu *egyo migugu* niinyija kugishanga Gatiwick. (Mugumya 2010: 1).
'Time to prepare for boarding the plane came. I bid farewell to those who had accompanied me. I got my baggage closer, and joined the queue. I then approached the passport, air-ticket and baggage checking desk. I handed my ticket to a nice looking lady. She checked it and then stamped it. My baggage was sent on a conveyor belt, and I was told that there was no problem. *That lady* told me that I will find *that baggage* at Gatwick.'

(52) A conversation between Kaaremeera and Rwamunyoro extracted from a folktale by Mubangizi (1966: 56–57).
Amugira ati "Kááremeera!"
a-mu-gira a-ti Kaaremeera
3SG.SBJ-1OM-say 1-that 1.Kaaremeera
"He said "Kareemeera!"
Ónu ati "Éé!"
o-nú ati ee
1-PROX 1-that yes
"This one said "Yes!"
Oríya ati "Ente kú ébaagwá ni biihá ébifá busha bitariirwe
o-ríya a-ti "E-n-te kú é-baag-w-á ni bi-ihá
1-DIST 1-that AUG-9-cow when 9.SBJ-slaughter-PASS-FV is 8-what
é-bi-fá busha bi-ta-ri-ir-w-e
AUG-8-waste nothing 8-NEG-eat-APPL-PASS-FV
'The other one said, "When a cow is slaughtered, which parts of it are not eaten?"'
Ónu abanzá áteekáteeka
ó-nu a-banzá á-teekáteek-a…
1-PROX 3SG-first-FV 3SG-think-FV
'This one had to first think…'
Oríya ati f'ókugamba óku óríkubímanya."
o-riya a-ti fa o-ku-gamb-a o-ku o-riku-bi-many-a."
1-DIST 1-that just AUG-INF-speak-FV AUG-how 1.SBJ-IPFV-8OM-know-FV
'The other said, "Just mention those that you know."

Ónu ati "Ndeeb-a nibanaga ámíîsho, nibanaga óbwongo, nibanaga óruhango, nibanagá endurwe
o-nu a-ti "N-reeb-a ni-ba-nag-a a-ma-isho, ni-ba-nag-a
1-PROX 1-that 1SG-see-FV IPFV-2-throw-FV AUG-6-eye IPFV-2-throw-FV
o-bu-ongo, ni-ba-nag-a o-ru-hango, ni-ba-nag-a
AUG-14-brain IPFV-2-throw-FV AUG-11-gallbladder IPFV-2-throw-FV
e-n-rurwe…
AUG-9-bile duct
"This one said, "I see eyes being thrown, the brain being thrown, I see the bile duct being thrown, I see the gallbladder being thrown…"
Oríya ati ébyo nibyó óríkumanyá byónka?
o-riya a-ti e-bi-o ni-bi-o o-riku-many-a bi-onka?"
1-DIST 1-that DEM-8-MED COP-8-REL.PRO 1-IPFV-know-FV 8-only
"The other said, "Is that all you know?"
Rwamunyóro ati "Amahémbe n'ómukíra orábiríire
Rwamunyoro a-ti "A-ma-hembe na o-mu-kira
1.Rwamunyoro 1-that AUG-6-horn and AUG-3-tail
o-ra-bi-ri-ire?
2SG.SBJ-ever-8OM-eat-PFV
'Rwamunyoro said, "Have you ever eaten the horns and the tail?"'
Karemeera ati "Shana óyenzire kugira ébyokushánzirana nangwá kutabiiré kunjúma
Karemeera a-ti "Shana o-yend-ire ku-gira e-bi-a
1.Karemeera 1-that maybe 2SG-want-PFV INF-say AUG-8-conn
o-ku-shanz-ir-ana nangwá ku-ta-b-iré ku-n-júm-a."
AUG-joke-APPL-RECP but INF-NEG-be-PFV INF-1SG-abuse-FV
'Karemera said "Unless you just want to make fun of me, if not abusing me."'

A pronominal form in a narrative, as *ogwo* in (53) is used for the referent that forms the most important topic in the previous discourse. The demonstrative *ogwo* is associated with a referent that is already established. A pronominal form is felicitous because the antecedent appears in the recent discourse and *ogwo* can be recognised to refer to a human referent for example, not *ente* 'cow' which is non-human and belongs to a different noun class, yet both can be tracked from the immediate preceding discourse.

(53) Asiimwe (2014: 208)
Ku baahikire omu kyaro ky'owaabo, baateekyerereza Omukama waabo

eby'omuhiigo, n'eby'oburungi bwa munyaanya wa Muyanda, n'oburungi bw'ente zi yaabaire atungire. *Omukama ku yaabihuriire yaagira ati "Ntashwire* ***ogwo*** *ndyashwera oha?"* Omukama ahabwokwenda ngu ashwere ogwo mwishiki kandi anyagye n'ente za Muyanda, akateekateeka eihe ry'okuza kurwanisa. (Karwemera 1975: 21).
'When they returned to their village, they told their King about hunting and the beauty of Muyanda's sister, and about the beauty of the cows which Muyanda reared. *When the King heard all that he said 'If I don't marry that one, whom shall I marry?'* Because the King wanted to marry that girl and to rustle Muyanda's cows, he organised a militia group to go and fight with.'

The boldfaced anaphoric pronominal demonstrative *ogwo* in (53) also selects the most salient referent from the previous discourse which forms the main topic for the subsequent discourse. This contrasts with the medial demonstrative in a full NP which provides additional information about a major participant in the foregoing discourse. This is especially if there are other NPs introduced in between, such as *egyo mpangare* in (51) which is reactivated after a series of other intervening NPs (see also Nicolle 2014 for Digo).

Demonstratives not only refer to NPs but also to whole sentences, paragraphs or even a full story. In text anaphora, the antecedent of the demonstrative must refer to the immediately preceding discourse (Himmelmann 1996: 224). *Ekyo*, a medial demonstrative in (54), refers to the whole preceding proposition.

(54) Asiimwe (2014: 209)
Tukabá tuteera órunyiriri rw'ókuzá omu kinaabiro-kihorónyo kwékoraho. Abándi ab'émicwe etagunjúkire bakabá bamarayó ebyanda, *ékyo* kireeterá abantu kwéshanyá ógwó otaríkwenda kuhéérezá ábandi omugisha gw'ókwéshemeza. (Mugumya 2010: 3)
'We would queue to go to the bathroom to clean ourselves. Other people with no good manners would spend there a long time, and *that* would make people angry at that person who does not want to give others a chance to clean themselves.'

This section has underscored the diversity of roles that demonstratives perform in Runyankore-Rukiga both in the physical world and in discourse. The exophoric demonstratives refer to events and objects in the physical environment but also serve to activate shared knowledge between interlocutors. In discourse, demonstratives activate and reactivate participants but also highlight major discourse participants. Demonstratives also may indicate turns and shifts between

discourse participants in a narrative as indicated in (52), for instance. A detailed study of discourse roles of demonstratives in spontaneous speech and natural contexts with the help of a corpus can reveal more specific roles.

4.3 Other functions of the demonstrative: emphasis and particularisation

Emphasis and particularisation are other roles played by the Runyankore-Rukiga demonstrative. These functions are commonly realised when two demonstratives co-occur. Similar to Runyankore-Rukiga, Taji (2024 [this volume]) notes that double demonstratives in Chiyao realise emphasis. By particularisation, this means that an object is singled out from other objects. The demonstratives may be of the same form (55) or of a different form. For example, a locative demonstrative copula can co-occur with a locative demonstrative to particularise or emphasise a more specific location as in (56). In addition, when an augment on the head noun modified by a prenominal demonstrative is retained, it is said to add emphasis to the noun (see §3.2)

(55) Egyó égyó ésímu niyo ndíkwenda
e-gy-o *e-gy-o* e-simu ni-yo n-riku-end-a
DEM-9-MED DEM-9-MED AUG-9.phone COP-9.REL.PRO 1SG-IPFV-want-FV
'It is that phone (particularly that one) that I want [may be accompanied by a gesture].'

(56) Aha mpáha bakáhabáha?
a-ha *n-ha-ha* ba-ka-ha-ba-h-a
DEM-16.PROX-here LDCOP-16-16 3PL-REM-OM16-OM2-give-FV
'Were you given this (exact) place?'

Other demonstrative forms used to encode additional emphasis on the location of an entity include: *okwe nkukwe* 'exactly there' for class 17 and *omwe mpaho* 'exactly in there' for class 18, depending on the relative distance of location of an entity from both the speaker and the hearer (refer to the various locatives forms in Table 3). The use of these latter forms depends on the dialect of the speaker and age as these forms seem to be common among the young Rukiga speakers. It is observed that Runyankore speakers simply double the demonstrative locative to denote a more specific location of a referent: *aha (a)ha* 'exactly here'; *okwo okwe* 'exactly there' and *omwo omwe* 'exactly in there'. Moreover, doubling of demonstratives is a common strategy for expressing emphasis in Bantu languages (c.f., discussion in Malinga 1980 for isiXhosa and Gunnink 2018 for Fwe).

5 Conclusion

This chapter has offered an overview of the morphology, the syntax and the functions of demonstratives in Runyankore-Rukiga. Various forms of the demonstrative and usage have been discussed. Evidence for the claim that the initial element of the demonstrative is not an augment but a core morpheme, hence an indispensable element has been presented. This morpheme responsible for deictic and anaphoric meanings, as discussed, is prevalent in many Bantu languages. Another key aspect of the current analysis is the manner demonstrative *-ti* which seems to have not received much attention in previous work. A detailed follow up study of this form can establish its connection with the other demonstrative classes, its functions, historical origin and also the study might be extended to other Bantu languages to find out how far it is spread in the Bantu language zones. An additional key issue for further investigation that emerges from this chapter concerns the role of the augment of the noun appearing with a prenominal demonstrative. An augment retained on the noun preceded by an adnominal demonstrative seems to be more pronounced in spoken discourse and is hardly found in written works. This phenomenon shows how flexible the spoken register is, compared to the written form. This occurrence could also be attributed to language contact with the neighboring Luganda where the augment is grammatically required when the noun appears with a pre-modifying demonstrative and many Runyankore-Rukiga speakers are also bilingual or trilingual in Luganda.

Another pertinent question to explore further is the extent to which information structure interacts with syntax. Such a study might provide answers to what determines the position of the adnominal demonstrative in the Runyankore-Rukiga language cluster. An in-depth study of discourse demonstratives, their distribution and the roles they play, is also still pending. This work would benefit greatly from the presence of an extensive Runyankore-Rukiga corpus.

The present chapter is a contribution to descriptive and comparative studies of Bantu languages. Insights can be drawn from the current study to conduct further specific and detailed studies about the morpho-syntax and functions of demonstratives in Runyankore-Rukiga and other related Bantu languages.

Abbreviations

APPL	applicative
ASP	aspect
AUG	augment
COP	copula
DEM	demonstrative morpheme
DIST	distal
DM	discourse marker
EXPLET	expletive

FV	final vowel
CONN	connective
INF	infinitive
IPFV	imperfective
LDCOP	locative demonstrative copulative
MED	medial
NEG	negation
NMLZ	nominalizer
N.PST	immediate past
OM	object marker
PART	partitive
PERS.	person
PL	plural
PROX	proximal
PROG	progressive
PRS	present tense
REF	referential
REM	remote past
REFL	reflexive
REL	relative
REL.PRO	relative pronoun
SBJ	subject marker
SBJV	subjunctive
SG	singular

References

Ahn, Dorothy & Jenneke van der Wal. 2019. What does that Lugwere demonstrative refer to? A semantic analysis of proximity and exteriority. *Studies in African Linguistics* 48(1). 1–24.

Ashton, Ethel O., Enoch M. K. Mulira, E. G. M. Ndawula & Archibald Norman Tucker. 1954. *A Luganda grammar*. London, New York & Toronto: Longmans, Green & Co.

Asiimwe, Allen. 2014. *Definiteness and specificity in Runyankore-Rukiga*. Stellenbosch: Stellenbosch University of South Africa. (Doctoral dissertation). http://hdl.handle.net/10019.1/95926 (16 March, 2015).

Asiimwe, Allen. 2016. Investigating the connection between the demonstrative and the definite morpheme *-a* in Runyankore-Rukiga. *South African Journal of African Languages* 36(1). 65–73.

Beermann, Dorothee & Allen Asiimwe. Forthcoming. Locatives in Runyankore-Rukiga. In Eva-Marie Bloom Ström, Rozenn Guérois, Hannah Gibson & Lutz Marten (eds.), *Approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.

Bernsten, Jan. 1998. Runyakitara: Uganda's "new" language. *Journal of Multilingual and Multicultural Development* 19(2). 93–107. DOI: 10.1080/01434639808666345.

Bloom Ström, Eva-Marie. 2015. The word order in Swahili adnominal constructions with locative demonstratives. *Nordic Journal of African Studies* 24(2). 118–145.

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.

Cocchi, Gloria. 2000. Locative constructions in Bantu. *Quaderni del Dipartimento di Linguistica dell'Università di Firenze* 10. 43–54.

de Blois, Kornelis Frans. 1970. The augment in the Bantu languages. *Africana Linguistica* 4. 85–165.

Dewees, John William. 1971. *The role of syntax in the occurrence of the initial vowel in Luganda and some other Bantu languages*. Madison, WI: University of Wisconsin. (Doctoral dissertation).

Diessel, Holger. 1999. *Demonstratives: Form function and grammaticalization*. Amsterdam: John Benjamins.

Diessel, Holger. 2012. Bühler's two-field theory of pointing and naming and the deictic origins of grammatical morphemes. In Kristin Davidse, Tine Breban, Lieselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and language change: New reflections*, 37–50. Amsterdam: John Benjamins.

Dixon, Robert Malcolm Ward. 2003. Demonstratives: A cross-linguistic typology. *Studies in language* 27(1). 61–112.

Du Plessis, Jacobus Albertus, Jan Adriaan Du Plessis & Eli Golden Mputa. 1978. *IsiXhosa 4*. Bellville: Via Afrika.

Du Plessis, Jacobus Albertus, Jan Adriaan Du Plessis & Marianna W. Visser. 1992. *Xhosa syntax*. Pretoria: Via Africa.

Fillmore, Charles J. 1997. *Lectures on deixis*. Stanford, CA: CSLI Publications.

Guérin, Valerie. 2015. Demonstrative verbs: A typology of verbal manner deixis. *Linguistic Typology* 19(2). 141–199.

Guillemin, Diana. 2011. *The syntax and semantics of a determiner system: A case study of Mauritian Creole*. Amsterdam, Philadelphia: John Benjamins.

Gunnink, Hilde. 2018. *A grammar of Fwe: A Bantu language of Zambia and Namibia*. Ghent University. (Doctoral dissertation).

Himmelmann, Nikolaus P. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.), *Studies in anaphora* (Typological Studies in Language 33), 203–252. Amsterdam: John Benjamins.

Karwemera, Festo. 1975. *Shutama nkutekyerereze*. Nairobi, Kampala & Dar es Salaam: East African Literature Bureau.

Karwemera, Festo. 1994. *Emicwe n'emigyenzo y'abakiga*. Kampala: Fountain Publishers.

Ladefoged, Peter, Ruth Glick & Clive Cripper. 1972. *Language in Uganda*. Nairobi: Oxford University Press.

Levinson, Stephen C. 2004. Deixis. In Laurence R. Horn & Gergory Ward (eds.), *The handbook of pragmatics*, 97–121. Oxford: Blackwell.

Lewis, Paul M., Gary Simons & Charles Fennig. 2013. *Ethnologue: Languages of the world*. 17th edn. Dallas, TX: SIL International.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Maho, Jouni Filip. 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 (20 March, 2013).

Malinga, R. M. M. 1980. Differences between the ordinary demonstrative and the emphatic demonstrative in Xhosa. *Stellenbosch Studies in African Languages* 2. 1–11.

Marten, Lutz. 2012. Agreement in locative phrases in Luganda. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics (Cologne 2009)*, 433–443. Cologne: Köppe.

Morris, Henry Francis & Brian Edmond Renshaw Kirwan. 1972. *Runyankore grammar*. Kampala: East African Literature Bureau.

Mubangizi, K. Benedicto. 1966. *Nkuganire i*. Kisubi: Marianum Press.

Mubangizi, K. Benedicto. 1997. *Abagyenda bareeba*. Memorial Single Volume. Kisubi: Marianum Press.

Mugumya, Levis. 2010. *Omuteizi omuri bungyereza*. Kabale: Mwesigwa Mugabi Publications.

Nicolle, Steve. 2007. Textual functions of Chidigo demonstratives. *SOAS Working Papers in Linguistics* 15. 159–171.

Nicolle, Steve. 2012. Semantic-pragmatic change in Bantu *-no* demonstrative forms. *Africana Linguistica* 18. 193–234.

Nicolle, Steve. 2014. Discourse functions of demonstratives in eastern Bantu narrative texts. *Studies in African Linguistics* 43(2). 125–144.

Rubongoya, L. T. 1999. *A modern Runyoro-Runtooro grammar*. Köln: Rüdiger Köppe Verlag.

Simons, Gary F. & Charles D. Fennig (eds.). 2017. *Ethnologue: Languages of the world*. 20th edn. Dallas, TX: SIL International. http://www.ethnologue.com (4 April, 2017).

Taji, Julius. 2024. Demonstratives in Chiyao: An analysis of their form, distribution and functions. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 17–41. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663763.

Taylor, Charles. 1985. *Nkore-Kiga*. London: Croon-Helm.

Valinande, Nzama K. 1984. *The structure of Kinande*. Washington DC: Georgetown University. (Doctoral dissertation).

Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics with special focus on Bantu and Mande*, 425–441. Tervuren: Royal Museum for Central Africa.

Van de Velde, Mark. 2019. Nominal morphology and syntax. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 237–269. London & New York: Routledge.

van der Wal, Jenneke. 2010. Functions of demonstratives in Makhuwa narratives. *Africana Linguistica* 16. 183–213.

Visser, Marianna. 2008. Definiteness and specificity in the isiXhosa determiner phrase. *South African Journal of African Languages* 28(1). 11–29.

Wald, Benji. 1973. Syntactic change in the lake languages of Northeast Bantu. *Studies in African Linguistics* 4(3). 237–268.

Zeller, Jochen. 2013. Locative inversion in Bantu and predication. *Linguistics* 51(6). 1107–1146.

Zeller, Jochen. 2017. *Locatives in Bantu*. Unpublished ms. http://www.jzeller.de/pdf/Loc.pdf (24 July, 2017).

Zeller, Jochen. 2020. Syntax. In Rainer Vossen & Gerrit Dimmendaal (eds.), *The Oxford handbook of African languages*, 66–87. Oxford: Oxford University Press.

Chapter 4

Concord and agreement in Eastern Bantu: The augment and noun classes in Nyakyusa

Amani Lusekelo
University of Dar es Salaam

The recent discussion of parameters of morphosyntactic variation motivated further scrutiny of the properties of augment and object markers available on bare nouns and complex noun phrases in Nyakyusa (M31). The focus of this chapter is on the (non-)occurrence of the V-augment and CV-particle, the role of demonstratives, and the word-order within the noun phrases. The CV-particle appears to derive from the proximal demonstrative. This is confirmed by its complementary distribution with both the proximal demonstrative and the V-augment. The main role of the CV-particle is to indicate contrastive focus of the referent. In addition, the anaphoric demonstrative *-la* 'that/those' occurs in complementary distribution with the augment, as both are related to the indication of definiteness. With regard to the role of object prefixes, Nyakyusa reveals that object marking may provide definite readings with verbs which take optional object markers (e.g. *piija* 'cook' and *bɪɪka* 'put'). However, definite interpretations are mandatory with verbs which require obligatory object prefixes, e.g. *bona* 'see' and *bʊʊla* 'inform'. Therefore, object marking is not associated with the realisation of the augment.

1 Introduction

The contribution of this chapter is two-fold. Firstly, the parameters for morphosyntactic variation in Bantu languages, as articulated by Guérois et al. (2017), opened another avenue to re-examine the morphosyntactic properties of nouns and noun phrases provided in previous studies for Nyakyusa (de Blois 1970, Lusekelo 2009, Mbope 2016, Persohn 2017). For instance, de Blois (1970) suggested that

Amani Lusekelo. 2024. Concord and agreement in Eastern Bantu: The augment and noun classes in Nyakyusa. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 85–108. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663767

both V-augments and CV-augments occur in Nyakyusa. Data offered in sections 2 and 3 substantiates that both the V-augment, CV-augment (termed CV-particle herein) and prenominal demonstrative realise definiteness in Nyakyusa. Based on the theory of definiteness (Lyons 1999) and contrast focus (Repp 2010), the proper functions of the augment are provided in this chapter. I establish the role of the augment as a marker of (in)definiteness.

Secondly, the interpretation of (in)definite sentence(s) involves both the presence and/or absence of the augment and object markers in Bantu languages with augments. For instance, Visser (2010) claims that definite readings in Xhosa are obtained once an object marker is cliticized on the verb and the object noun is marked with an augment. This claim is confirmed in Bantu languages without augments. For instance, with regard to object marking, Marten & Kula (2012) argue that the use of the object marker with non-animate NPs is associated with definiteness in Swahili. However, Riedel (2009) argues that some sentences provide definite readings without prefixing the object marker in Swahili. In §4, I argue that the distinction between definiteness and indefiniteness in Nyakyusa is partly associated with object marking with verbs which take an object prefix optionally. In verbs which require a mandatory object prefix, the object prefix does not indicate definiteness.

2 The shape and concord of bare nouns

2.1 The shape and distribution of the augment

The first parameter of Guérois et al. (2017) requires an investigation of the morphology of the augment. Data shows that bare nouns[1] in Nyakyusa[2] consist of an augment (always a V-augment), a noun class prefix and a stem, as exemplified in (1). Notice that more examples are provided for noun classes 5/6 due to variations in the shape of the noun class 5.

(1) cl. 1 *ʊ-mʊ-ndʊ* ‘person’
cl. 2 *a-ba-ndʊ* ‘persons’
cl. 3 *ʊ-m-piki* ‘tree’

[1]The notion “bare noun” is used to refer to nouns composed of an augment, noun class prefix and a stem without any modification or quantification. A bare noun is a noun in isolation. This is opposed to complex noun phrases which are comprised of a head noun and at least one modifier or quantifier.

[2]Most examples of bare nouns for Nyakyusa come from Felberg (1996). Some sentences for Nyakyusa come from Lusekelo (2012) and Persohn (2017). Other examples were constructed by the author who is a native speaker of the language.

	ʊ-n-kota	‘medicine’
cl. 4	*ɪ-mɪ-piki*	‘trees’
	ɪ-mɪ-kota	‘medicines’
cl. 5	*ɪ-lɪ-lasi*	‘bamboo tree’
	ɪ-lasi	‘bamboo tree’
	ɪ-lɪ-bwe	‘stone’
	ɪ-bwe	‘stone’
	ɪ-lɪ-fumbi	‘egg’
	ɪ-fumbi	‘egg’
	ɪ-ly-abi	‘underpant’
	ɪ-ly-osi	‘smoke’
cl. 6	*a-ma-boko*	‘arms, hands’
	a-ma-lasi	‘bamboo trees’
	a-ma-isi (amiisi)	‘water’
	a-ma-bwe	‘stones’
	a-ma-abi	‘underpants, underwear’
cl. 7	*ɪ-kɪ-lundi*	‘leg’
	ɪ-kɪ-boko	‘arm, a hand’
	ɪ-kɪ-paale	‘calabash’
cl. 8	*ɪ-fɪ-lundi*	‘leg’
	ɪ-fɪ-boko	‘arms, hands’
	ɪ-fɪ-paale	‘calabashes’
cl. 9	*ɪ-nguku*	‘fowl’
	ɪ-mbwa	‘dog’
cl. 10	*ɪ-nguku*	‘fowls, chickens’
	ɪ-mbwa	‘dogs’
	ɪ-mbabu	‘firewood (PL)’
cl. 11	*ʊ-lʊ-kama*	‘milk’
	ʊ-lʊ-babu	‘firewood (SG)’
cl. 12	*a-ka-kuku*	‘chick’
	a-ka-lasi	‘small bamboo tree’
cl. 13	*ʊ-tʊ-kuku*	‘chicks’
	ʊ-tʊ-lasi	‘small bamboo trees’
cl. 14	*ʊ-bʊ-ndʊ*	‘humanity’
cl. 15	*ʊ-kʊ-lima*	‘to farm, to hoe’
	ʊ-kʊ-seka	‘to laugh’
cl. 16	*pa-kaja*	‘at home/homestead’
cl. 17	*ku-kaja*	‘to the home/homestead’
cl. 18	*mu-ngaja*	‘in the grove’

The data in (1) substantiate five points regarding Nyakyusa. The first point concerns the distribution of the augment across noun classes. The underived nouns are primarily characterized by the word-structure AUG-NCP-root because an augment is found before noun class prefix. The shape of the augment is regularly a vowel-augment (V-augment) in bare nouns. The V-augment is realised as *ɪ-*, *a-* and *ʊ-*, as discussed in previous works (de Blois 1970, Felberg 1996, Lusekelo 2009, 2013, Mbope 2016, Persohn 2017). In fact, the vowel quality of the augment is a copy of the noun class prefix vowel. Locative classes 16–18 do not contain the augment.

The *ɪ*-augment prolifically occurs in six noun classes: 4, 5, 7, 8, 9 and 10. In addition, the loanwords which are integrated into noun classes 9/10 take the *ɪ*-augment irrespective of the presence or absence of the nasal consonant noun prefix, as exemplified in (2).

(2) cl. 9/10 *ɪ-bulausɪ* [<*blouse*: English] ‘blouse’
ɪ-hela [<*Heller*: German] ‘money’
ɪ-kaabatɪ [<*kabati*: Swahili] ‘cupboard’
ɪ-kaabɪkɪ [<*cabbage*: English] ‘cabbage’
ɪ-katani [<*katani*: Swahili] ‘sisal’
ɪ-naulɪ [<*nauli*: Swahili] ‘bus fare’
ɪ-ndalama [<*dirham*: Arabic] ‘money’
ɪ-ndobo [<*ndoo*: Swahili] ‘bucket’
ɪ-sisala [<*scissor*: English] ‘a pair of scissor’

The *ɪ*-augment occurs in other nouns in classes 9/10 which do not bear the nasal. The *ɪ*-augment is for both singularity and plurality, as shown in (3).

(3) cl. 9/10 *ɪ-sekema* ‘fever’
ɪ-fula ‘rainfall’
ɪ-sanu ‘weed(s)’

The *ʊ*-augment is also productive, as it occurs in classes 1, 3, 11 and 13. The *a*-augment occurs in noun classes 2, 6 and 12. Both augments occur in loans, e.g. *alumasɪ* [<Swahili: *almasi*] ‘diamond’, noun class 6, *amafʊta* [<Swahili: *mafuta*] ‘oil, fat’, noun class 1, *ʊnnesi* [<Swahili: *nesi*] ‘nurse’ and *ʊnsinjala* [<English: *messenger*] ‘office clerk’.

The borrowed nouns in Nyakyusa obtain the shape AUG-NCP-root when they are incorporated into the lexicon, as illustrated in (4).

(4) cl. 1 *ʊ-n-sikali* [<*askari*: Swahili] 'a police officer, a soldier'
cl. 2 *a-ba-sikali* [<*askari*: Swahili] 'police officers, soldiers'
cl. 3 *ʊ-m-papaju* [<*mpapai*: Swahili] 'a paw paw tree'
cl. 4 *ɪ-mɪ-papaju* [<*mipapai*: Swahili] 'paw paw trees'
cl. 5 *ɪ-lɪ-galasi* [<*glass*: English] 'a glass, spectacles'
cl. 7 *ɪ-kɪ-kombe* [<*kikombe*: Swahili] 'a cup'
cl. 8 *ɪ-fɪ-kombe* [<*vikombe*: Swahili] 'cups'

The augment is absent in some noun classes in Nyakyusa. De Blois (1970: 119) states correctly that Nyakyusa has no V-augments in locative classes 16, 17 and 18. The data in (1) confirm that the augment does not surface in any of the locative noun classes 16, 17 and 18.

The second point concerns the status of the *ɪ*-augment in noun class 5. Data in (1) above show the optionality of the noun class *-lɪ-* (cl. 5). Nouns in class 5 may occur or may not occur with a noun prefix, as exemplified by *ɪlɪbwe/ɪbwe* 'stone' and *ɪlɪlasi/ɪlasi* 'bamboo tree'. More data with the optional singular class prefix, together with their corresponding plural pairings, are provided in (5) below.

(5)	cl. 5 *ɪ-(lɪ)-jabʊ*	'cassava'	cl. 6 *a-ma-jabʊ*	'cassava (PL)'
	cl. 5 *ɪ-(lɪ)-kina*	'machine'	cl. 6 *a-ma-kina*	'machines'
	cl. 5 *ɪ-(lɪ)-jɪko*	'kitchen'	cl. 6 *a-ma-jɪko*	'kitchens'
	cl. 5 *ɪ-(lɪ)-koonda*	'wasp'	cl. 6 *a-ma-koonda*	'wasps'
	cl. 5 *ɪ-(lɪ)-kumbuulu*	'a hoe'	cl. 6 *a-ma-kumbuulu*	'hoes'
	cl. 5 *ɪ-(lɪ)-lopa*	'blood'	cl. 6 *a-ma-lopa*	'blood'
	cl. 5 *ɪ-(lɪ)-sosɪ*	'tear'	cl. 6 *a-ma-sosɪ*	'tears'

Some nouns in class 5 occur with the noun prefix, as demonstrated by *ɪlyabi* 'underpant' and *ɪlyosi* 'smoke' in (1). Here the removal of the noun class prefix *-lɪ-* is unacceptable. Therefore, the forms **ɪabi* and **ɪosi* are unacceptable.

In nouns with optional *-lɪ-*, the status of the augment *ɪ-* in class 5 is worth mentioning here. First and foremost, do we treat the noun without a class prefix as being preceded by an augment or a noun class prefix? The answer to this question will be that the augment does not function as the class prefix, as found by Legère (2005) for Kwanyama (spoken in Namibia). In Nyakyusa, nouns in class 5 may contain a silent nominal prefix. The augment is always overt.

The third point substantiated by data in (1) above surrounds claims made in previous studies for Nyakyusa. De Blois (1970: 114) characterised Nyakyusa as lacking an augment in kinship terms, proper names, titles, and their plural forms (i.e. nouns in classes 1a and 2a). Persohn (2017: 41) suggested that the *ʊ*-augment

is optional for kinship terms, proper names, some living beings, and some loans. This use of the augment is related to specific grammatical conditions, which have not been articulated by De Blois and Persohn.

Nyakyusa allows the use of the *ʊ*-augment and the *a*-augment in kinship terms, some proper names and titles, as illustrated in (6).

(6)

cl. 1a	*ʊ-nyoko*	'your mother'	cl. 2a *a-ba-nyoko*	'your mothers'
cl. 1a	*ʊ-taata*	'my father'	cl. 2a *a-ba-taata*	'my fathers'
cl. 1a	*ʊ-maama*	'mother'	cl. 2a *a-ba-maama*	'mothers'
cl. 1a	*ʊ-baaba*	'father'	cl. 2a *a-ba-baaba*	'fathers'
cl. 1a	*ʊ-juuba*	'mother'	cl. 2a *a-ba-juuba*	'mothers'
cl. 1a	*ʊ-Tuntufye*	'Tuntufye'	cl. 2a *a-ba-Tuntufye*	'the Tuntufyes'
cl. 1a	*ʊ-Mwasekage*	'Mwasekage'	cl. 2a *a-ba-Mwasekage*	'the Mwakaseges'
cl. 1a	*ʊ-Malija*	'Mary, Maria'	cl. 2a *a-ba-Malija*	'the Marys'
cl. 1a	*ʊ-m-puuti*	'priest'	cl. 2a *a-ba-puuti*	'priests'
cl. 1a	*ʊ-malafyale*	'chief'	cl. 2a *a-ba-malalafyale*	'chiefs'
cl. 1a	*ʊ-n-nabo*	'their mother'	cl. 2a *a-ba-nabo*	'their mothers'

Some proper names that contain the prefix for class 11 do not contain the *ʊ*-augment. Nouns which contain the augment belong to class 11, as illustrated in (7).

(7)

cl. 11	*ʊ-lusajo*	'blessing'	cl. 1a *lusajo*	'proper name'
cl. 11	*ʊ-lusekelo*	'happiness'	cl. 1a *lusekelo*	'proper name'
cl. 11	*ʊ-lupakisyo*	'fear'	cl. 1a *lupakisyo*	'proper name'
cl. 11	*ʊ-lugano*	'love'	cl. 1a *lugano*	'proper name'
cl. 11	*ʊ-lusubilo*	'hope'	cl. 1a *lusubilo*	'proper name'
cl. 11	*ʊ-lutufyo*	'praise'	cl. 1a *lutufyo*	'proper name'

The fourth point concerns the V-augments *ɪ-*, *a-* and *ʊ-* which remain present in derived nouns in Nyakyusa. Therefore, derived nouns are characterised by the configuration AUG-NCP-stem-NOML, as exemplified in (8). In Nyakyusa, nomilising suffixes (marked as NOML) are V-shaped, realised mainly as *-e*, *-i* and *-o*. The applicative is often used to derive deverbatives in combination with these nominalising suffixes, such as *ɪ-simb-il-o* 'a pen' (< *samba* 'write').

(8)

bina	'become ill'	cl. 1	*ʊ-m-bin-e*	'patient'
paapa	'give birth'	cl. 2	*a-ba-paap-i*	'parents'
jweega	'make noise'	cl. 3	*ʊ-n-jweeg-o*	'noise'
simba	'write'	cl. 9	*ɪ-simb-il-o*	'pen'
koma	'hit'	cl. 5	*ɪ-lɪ-kom-el-o*	'playing/dancing tail'
sambula	'scoop'	cl. 8	*ɪ-fɪ-sambul-il-o*	'ladles, scoopers'
seka	'laugh'	cl. 11	*ʊ-lʊ-sekel-o*	'happiness'

The fifth and last point concerns the diminutive class. Felberg (1996: 3–5) lists inherent nouns in classes 12/13 (*aka-/ʊtʊ-*), which are few under primary classification in Nyakyusa. Typical nouns in this class include *akabalilo-ʊtʊbalilo* 'time', *akaaja-ʊtwaja* 'homestead(s)', *akalʊʊlʊ-ʊtʊlʊʊlʊ* 'ululation(s)' and *akasumo-ʊtʊsumo* 'tale(s)'.

Most class 12/13 nouns are derived diminutive nouns. Persohn (2017: 43) uses the examples *akapango* 'story' and *ʊtʊpango* 'stories', which are examples of derived nouns which come from the verb *panga* 'narrate, tell a story'. Diminutive nouns (and sometimes pejorative nouns) are derived from classes 1 to 11, as illustrated in (9).

(9)

cl. 1	*ʊ-mw-ana*	'child'	cl. 12	*a-ka-ana*	'small child'
cl. 2	*a-ba-ana*	'children'	cl. 13	*ʊ-tw-aana*	'small children'
cl. 3	*ʊ-m-piki*	'tree'	cl. 12	*a-ka-piki*	'small tree'
cl. 4	*ɪ-mɪ-piki*	'trees'	cl. 12	*ʊ-tʊ-piki*	'small trees'
cl. 5	*ɪ-(lɪ)-lasi*	'bamboo'	cl. 12	*a-ka-lasi*	'small bamboo'
cl. 6	*a-ma-lasi*	'bamboos'	cl. 13	*ʊ-tʊ-lasi*	'small bamboo'
cl. 7	*ɪ-kɪ-kota*	'chair'	cl. 12	*a-ka-kota*	'small chair'
cl. 8	*ɪ-fɪ-kota*	'chairs'	cl. 13	*ʊ-tʊ-kota*	'small chairs'
cl. 9	*ɪ-mbene*	'goat'	cl. 12	*a-ka-pene*	'small goat'
cl. 10	*ɪ-mbene*	'goats'	cl. 13	*ʊ-tʊ-pene*	'small goats'
cl. 11	*ʊ-lʊ-goje*	'rope'	cl. 12	*a-ka-goje*	'small rope'
cl. 11	*ʊ-lʊ-kama*	'milk'	cl. 12	*ʊ-tʊ-kama*	'little milk'

2.2 The function of the augment in Nyakyusa in relation to its role in Eastern Bantu

Previous studies suggested that the augment functions to mark definiteness in Bantu languages (cf. de Blois 1970, Bokamba 1971, Hyman & Katamba 1993, Legère 2005, Visser 2010, Goodness 2013, Asiimwe 2014, Petzell & Kühl 2017). They suggest that the augment functions like articles in English, Hungarian or French. Three exemplary cases are provided below.

Bokamba (1971) found in Dzamba (C322, spoken in the DRC) that the augment functions like articles in Western European languages (mainly the English definite article *the*). It is argued that Dzamba permits augments on definite nouns only, as shown in (10–11).

(10) Dzamba (Bokamba 1971: 220)
mɔ-ibi (mɔɔ) anyɔlɔki o-ndaku
1-thief (one) entered AU-9.house
'A thief entered the house.'

(11) o-mɔ-ibi (*mɔɔ) anyɔlɔki o-ndaku
AUG-1-thief (*one) entered AU-9.house
'The thief entered the house.'

In both examples the object noun phrases carry the augment and have definite interpretation. However, the subject noun phrases differ. In example (10), *mɔibi* 'a thief' is not preceded by an augment, hence we obtain an indefinite reading of the subject noun phrase. Example (11) *omɔibi* 'the thief' is preceded by an augment, hence we get a definite interpretation. These examples also reveal that the augment and the numeral *mɔɔ* 'one' cannot co-occur (Bokamba 1971). Based on such data, it is argued that an NP without an augment indicates an indefinite reading. The presence of the augment warrants a definite interpretation in Dzamba.

In support of the claim that the augment functions as an (in)definite marker, Asiimwe (2014: 71) shows the following examples from Luganda (JE15, spoken in Uganda) in which the presence and absence of the augments signal (in)definiteness. The presence of the augment in (12) marks definiteness while its absence in (13) signals indefiniteness.

(12) Luganda (Asiimwe 2014: 71)
U-mw-ana a-a-fwaaya i-ci-tabu
AUG-1-child 3PL-PST-want AUG-7-book
'The child wanted the book.'

(13) U-mw-ana a-a-fwaaya ci-tabu
AUG-1-child 3PL-PST-want 7-book
'The child wanted a book.'

Irrespective of the suggestion above, the use of an augment is often difficult to describe in Luganda and Nyankore-Kiga. Most speakers of Nyankore-Rukiga are not certain of the proper choices for the use of deletion of the augment (Asiimwe 2014: 183–184). This kind of uncertainty is also mentioned by Hyman & Katamba (1993) and Goodness (2013). In fact, the use of the augment to mark definiteness is not a straightforward mechanism in other Bantu languages, as confirmed by Hyman & Katamba (1993) for Luganda, Goodness (2013) for Shinyiha and Petzell & Kühl (2017) for Luguru (both spoken in Tanzania). Petzell & Kühl (2017) show that the use of demonstratives plays an important role for providing definite readings in Luguru. This may point to the suggestion that the occurrence or non-occurrence of the augment is not related in a straightforward way with the use of the definite article *the* in English.

The augment is used in a variety of functions in Nyakyusa. Its distribution is robust, but it has limited connection with the expression of (in)definiteness. For instance, Persohn (2017: 343) provides some texts whose nouns bear no augment. Text (14) shows that only the noun *iisikʊ* 'day' bears a V-augment.

(14) a. po tʊ-kʊ-tɪ **kalʊlʊ** na **nsyɪsyɪ** ba-a-lɪ ba-manyaani fiijo
then 1PL-PRS-say hare COM skunk 2-PST-COP 2-friend intense
'We say, Hare and Skunk were good friends.'

b. b-end-aga b-oosa kʊkʊtɪ kʊ-no bi-kʊ-bʊʊk-a
2-travel-IPFV 2-all every 17-prox 2-PRS-go-FV
'They went together wherever they went.'

c. po **ii-sikʊ** lɪ-mo **kalʊlʊ** na **nsyɪsyɪ** ba-a-bʊʊkile n-kʊ-fwɪm-a
then 5-day 5-one hare COM skunk 2-PST-go-PFV 18-15-hunt-FV
'So one day Hare and Skunk went to hunt.'
(Persohn 2017: 343)

Persohn (2017: 352) provides other texts whose nouns bear a V-augment. He suggests that the definiteness and indefiniteness distinction is realised using the augment. Text (15) shows that all nouns bear the V-augment, e.g. *abatasi* 'ancestors' and *ɪndingala* 'drum'.

(15) a. ba-a-li=po **a-ba-ndʊ** b-a ijolo **a-ba-tasi**
2-PST-COP AUG-2-person 2-assoc old times AUG-2-ancestor
'There were the people of old times, the ancestors.'

b. ba-a-li n=**ii-penenga** **ɪ-ndingala** ɪ-jɪ
2-PST-COP COM=5-type.of.drum AUG-9-drum AUG-prox
j-aa-job-igw-aga
9-PST-speak-PASS-IPFV
'They had the drum.' (Persohn 2017: 352)

Irrespective of the presence or absence of the V-augment, the texts in Persohn (2017) are provided with definite and indefinite readings. This entails that the definite and indefinite interpretations of the sentences require some context to be used. For example, the sentence in (16) can be interpreted as indefinite or definite.

(16) ʊ-malafyale a-fw-ele ɪ-ngiga
AUG-1.chief SM1-wear-PFV AUG-9.crown
'a/the chief wears a/the crown.'

The sentence in (16) will be interpreted as indefinite if interlocutors are unaware of the chief and crown. However, if the chief is known by interlocutors, a definite reading obtains. Likewise, once the crown is known by interlocutors, a definite reading obtains.

The use of demonstratives is paramount in drawing the distinction between definite and indefinite readings. Lyons (1999) pointed out that the anaphoric use of demonstratives provides a definite reading, while the referential use of demonstratives gives a deictic interpretation of nouns. In Bantu languages, scholars have shown that the anaphoric use of demonstratives is linked to definiteness and focus (Visser 2010, Asiimwe 2014, Ndomba 2017, Kimambo 2018b). The Nyakyusa text in (17) shows the use of demonstratives to indicate definiteness.

(17) ʊ-n-nyambala jʊ-mo a-fik-ile ku-buhesya. a-ba-ag-ile
AUG-1-man 1-one SM1-arrive-PFV 17-new-land SM1-OM2-find-PFV
a-ba-fwimi. a-ba-fwimi ba-la ba-li-mw-amb-ile
AUG-2-hunter AUG-2-hunter 2-those SM2-PST-OM1-invite-PFV
ʊ-n-nyambala jʊ-la.
AUG-1-man 1-that
'A certain man arrived at a new country. He found hunters. The hunters invited the man.'

The data in (17) above shows that the augment occurs in the indefinite noun *ʊnnyambala jʊmo* 'certain man'. Also, it occurs in definite nouns *ʊnnyambala jʊla* 'that man'. However, it is the use of anaphoric demonstratives *bala* 'those' and *jʊla* 'that' which provides definite readings. Lyons (1999) states that anaphoric demonstratives tend to focus the referent. In this case, we obtain definite and focused nouns in Nyakyusa.

The prenominal demonstratives are also anaphoric in Bantu languages (Visser 2008, Kimambo 2018b). Nyakyusa allows prenominal demonstratives, as shown in (18). Usually, the prenominal demonstrative occurs in complementary distribution with the augment (Lusekelo 2009). Thus, either the V-augment or prenominal demonstratives indicate definiteness.

(18) jʊ-la n-nyambala a-ba-buul-ile ba-la ba-fwimi
1-that 1-man SM1-OM2-tell-PFV 2-those 2-hunter
'The man informed the hunters.'

The role of the augment in Dzamba, Luganda, Nyankore-Kiga and Xhosa is directly related to the function of articles in English, French and Hungarian. Its

presence entails a definite reading. This is not the case for Nyakyusa. Rather, the prenominal demonstrative and V-augment occur in complementary distribution, which is an indication that focused nouns would not require the V-augment.

In the context that provides an indefinite reading, the V-augment is not required, as exemplified in (19a). However, in a context that provides a definite reading, the V-augment is required, as shown in (19b).

(19) a. u-pimb-ile ma-toki ma-ki?
SM2SG-carry-PFV 6-banana 6-what
'What kind of bananas are you carrying?'

b. u-pimb-ile a-ma-toki ga-liku?
SM2SG-carry-PFV AUG-6-banana 6-what
'Which bananas are you carrying?'

c. m-bimb-ile ma-toki ma-bifwe
SM1SG-carry-PFV 6-banana 6-ripe
'I am carrying ripe bananas.'

d. m-bimb-ile a-ma-toki a-ma-bifwe
SM1SG-carry-PFV AUG-6-banana AUG-6-ripe
'I am carrying the ripe bananas.'

The V-augment is not required in interrogatives (19a), although it is attested in this context as well (19b). Both are indefinite contexts. In the case of an indefinite answer, no V-augment appears as in (19c). In contrast, the V-augment is typically (although not obligatorily) required in definite contexts (19d). The context of speech also provides clues to definiteness.

2.3 CV-particle in the nominal domain in Nyakyusa

The discussion about the presence of the CV-augment in Nyakyusa begins with de Blois (1970: 93). He reported that the reconstruction of the augment in Proto-Bantu included a CV-shape, mainly **ga-*, **ba-*, **jɨ-* and **jʉ-* as reported in Meeussen (1967). The interpretation of nouns preceded by a CV-augment is "with definite meaning" (Meeussen 1967: 99). The reconstructed proto-words include **jʉ muntu* 'the person', **ba bantu* 'the persons', **ji mbúa* 'the dog' and **sɨ mbúa* 'the dogs' (1967: 99). In many Bantu languages, the CV-augment was reduced to a V-augment, except a few languages (including Nyakyusa) whose CV-augment is restricted to emphasized nouns (de Blois 1970: 92–93).

De Blois (1970) shows that Nyakyusa possesses the CV-augment, and that it is still in use alongside the normal V-augment. He notes that "the CV-type is

used instead of the V-type to give more prominence or emphasis to the noun" (de Blois 1970: 98). This means that when the CV-augment is used, it shows a special meaning of emphasis, as illustrated by *jʊ-mʊ-ndʊ* 'the very man' and *ga-ma-heelu* 'only terms of abuse'.

Persohn (2017: 44) separated the CV-augment from the V-augment. With regard to the former, he suggested that "nouns carrying a pronominal prefix instead of an augment, express an emphatic notion translatable along the lines of 'the very X' as in *lʊ-lw-ala* 'the very grindstone' and *gʊ-n-tʊ* 'the very head'" (2017: 44).

The data in (20) show that the CV-augment is present in all 18-noun classes in Nyakyusa. Even the locative classes 16, 17 and 18 bear the CV-augment. Note that the morphology of the CV-augment is represented in Table 1.

(20)

cl. 1	*jʊ-m-puuti*	'only the priest'
cl. 2	*ba-ba-ndʊ*	'the very persons'
cl. 3	*gʊ-m-papaju*	'the very paw paw tree'
cl. 4	*gɪ-mɪ-papaju*	'the very paw paw trees'
cl. 5	*lɪ-lɪ-lasi*	'only the bamboo tree'
cl. 6	*ga-ma-lasi*	'only the bamboo trees'
cl. 7	*kɪ-kɪ-kota*	'the very chair'
cl. 8	*fɪ-fɪ-kota*	'the very chairs'
cl. 9	*jɪ-mbwa*	'only the dog'
cl. 10	*sɪ-mbwa*	'only the dogs'
cl. 11	*lʊ-lʊ-kama*	'only the milk'
cl. 12	*ka-ka-kʊkʊ*	'only the chick'
cl. 13	*tʊ-tʊ-kʊkʊ*	'only the chicks'
cl. 14	*bʊ-bʊ-ndʊ*	'only the humanity'
cl. 15	*kʊ-kʊ-kuuta*	'only to cry'
cl. 16	*pa-pa-sʊkʊʊlʊ*	'at the very school'
cl. 17	*kʊ-kʊ-kaja*	'to the very home'
cl. 18	*mʊ-mʊ-supa*	'in the very bottle'

However, what is called the CV-augment is not attached to the lexical noun; rather it occurs as a separate element, as correctly suggested by Persohn (2017). This CV-particle derives from proximal demonstratives, as shown in Table 1.

The prenominal demonstrative and the CV-particle occur in complementary distribution. Example (21) shows the use of the V-augment, while the prenominal demonstrative restricts its occurrence, as shown in (22). The CV-augment restricts the occurrence of the V-augment (23), as well as the prenominal demonstrative (24).

Table 1: Proximal demonstrative and CV-particle

Cl.	Proximal demonstrative			CV-augment		
1	*ʊmʊndʊ*	*ʊjʊ*	this person	*jʊ*	*mʊndʊ*	the person
2	*abandʊ*	*aba*	these persons	*ba*	*bahesya*	the persons
3	*ʊmwenda*	*ʊgʊ*	this dress	*gʊ*	*mpiki*	the tree
4	*ɪmɪpiki*	*ɪgɪ*	these trees	*gɪ*	*mipiki*	the trees
5	*ɪlyabi*	*ɪlɪ*	this pant	*lɪ*	*lyabi*	the pant
6	*amaabi*	*aga*	these pants	*ga*	*maabi*	the pants
7	*ɪkɪtili*	*ɪkɪ*	this cap	*kɪ*	*kitili*	the cap
8	*ɪfɪtili*	*ɪfɪ*	these caps	*fɪ*	*fitili*	the caps
10	*ɪmbwa*	*ɪjɪ*	this dog	*jɪ*	*mbwa*	the dog
10	*ɪmbwa*	*ɪsɪ*	these dogs	*sɪ*	*mbwa*	the dogs
11	*ʊlʊkili*	*ʊlʊ*	this stick	*lʊ*	*lʊkili*	the stick
12	*akakuku*	*aka*	this chick	*ka*	*kakuku*	the chick
13	*ʊtʊkuku*	*ʊtʊ*	these chicks	*tu*	*tukuku*	the chicks
14	*ʊbʊtatu*	*ubu*	this trinity	*bʊ*	*bʊtatu*	the trinity
15	*ukulima*	*uku*	this farming	*kɪ*	*kulima*	the farming
16	*pakaja*	*apa*	this homestead	*pa*	*pakaja*	at the homestead
17	*kukaja*	*aka*	that homestead	*ka*	*kukaja*	to the homestead
18	*ntwaja*	*ʊmo*	in homestead	*mʊ*	*ntwaja*	in the homestead

(21) ɪ-mbeba sɪ-tafwine ɪ-ndefu
AUG-10.rat SM10-tear.PFV AUG-10.mat
'Rats damaged mats.'

(22) ɪsɪ mbeba sɪ-ta-fwine ɪ-ndefu
DEM10 10.rat SM10-tear.PFV AUG-10.mat
'(Specifically) the rats damaged mats.'

(23) sɪ mbeba sɪ-ta-fwine ɪ-ndefu
CV-10 10.rat SM10-tear.PFV AUG-10.mat
'(Specifically) the rats damaged mats.'

(24) * sɪ ɪsɪ mbeba sɪ-ta-fwine ɪ-ndefu
CV-10 DEM10 10.rat SM10-tear.PFV AUG-10.mat

Two theories can account for the definiteness marking. On the one hand, based on the theory of definiteness and specificity (Lyons 1999), the examples above

demonstrate that the CV-augment indicates typically a specific and focused referent in Nyakyusa.

The theory of information structure, mainly contrastive focus (Aboh 2004, Repp 2010), can account for the role of the CV-augment and prenominal demonstrative. In example (25), the CV-particle is used to focus the contrast of *bwalwa* 'alcohol' from other foodstuff. It means that the alcoholics did not vomit anything else but alcohol. Likewise, the contrastive focus in example (26) indicates that Tuntufye wears nothing but *kɪtili* 'cap'. It means that Tuntufye, who is perhaps a child, had not put on any other articles of clothing.

(25) A-ba-leefi ba-teek-ile bʊ bwalwa
AUG-2-alcoholic SM2-vomit-PFV AUG.14 14.alcohol
'The alcoholics vomited nothing but alcohol.'

(26) Tuntufye a-fw-ele kɪ kɪ-tili
1.Tuntufye SM1-wear-PFV AUG.7 7-cap
'Tuntufye put on only a cap.'

The tools of information structure help to analyse properly the emphatic purpose of the CV-augment highlighted by de Blois (1970) and Persohn (2017). It means that the CV-augment is used to provide contrastive focus in which of the many alternative referents, it foregrounds one.

3 Configuration and concord in complex noun phrases

Rijkhoff (2002) highlights the fact that complex noun phrases contain the lexical noun with its dependents. The dependents of the lexical noun include articles, demonstratives, adjectives, numerals etc. (Alexiadou et al. 2007). Both articles and demonstratives are used to indicate definiteness (Lyons 1999).

Given this backdrop, two points are discussed in this section. Firstly, I investigate whether the (non-)occurrence of the augment on adjectives has implications for the (in)definite interpretation in Nyakyusa. In Nyankore-Kiga and Rutooro, the occurrence of the augment on the adjectives indicates a definite reading, and its absence indicates indefinite interpretation (Kaji 2009, Asiimwe 2014).

In Nyankore-Kiga, example (27) shows an adjective *murungi* 'good' without an augment, indicating indefiniteness, while example (28) shows that an adjective *umurungi* 'good' has an augment which provides definiteness (Asiimwe 2014). Similarly, Rutooro provides an indefinite interpretation for the adjective without the augment (29), and a definite reading for the adjective with the augment (30) (Kaji 2009).

(27) Nyankore-Kiga (Asiimwe 2014: 120)
o-mu-shaija mu-rungi
AUG-1-man 1-good
'a good man'

(28) o-mu-shaija u-mu-rungi
AUG-1-man AUG-1-good
'the good man'

(29) Rutooro (Kaji 2009: 246)
e-ki-tábu ki-ríngi
AUG-7-book 7-good
'a good book'

(30) e-ki-tábu e-ki-ríngi
AUG-7-book AUG-7-good
'the good book'

Secondly, I investigate further whether the co-occurrence of the augment and demonstratives overlap in the marking of definiteness in Nyakyusa. In some Bantu languages, the prenominal demonstrative, which provides anaphoric reference to the noun (Rugemalira 2007, Petzell & Kühl 2017, Kimambo 2018b), restricts the augment occurring on the noun (Visser 2002, Van de Velde 2005).

In Xhosa, the canonical order is demonstrative > noun, which provides a deictic reading. The augment does not occur in the nouns in (31), while example (32) shows the occurrence of the augment and provides a definite reading (Visser 2002).

(31) Xhosa (Visser 2002: 287)
lo m-fana m-de
this 1-young man 1-tall
'this tall young man'

(32) o-m-fana o-m-de
AUG-1-young man AUG-1-tall
'the tall young man'

Lusekelo (2009) showed that the prenominal demonstrative restricts the augment in Nyakyusa. In addition, as highlighted in section 2.2, both the augment and prenominal demonstrative have implications for the marking of definiteness in Nyakyusa. In this section, I discuss the role of the demonstrative and adjective in marking definiteness.

Nyakyusa reveals this pattern of noun phrase order: (DEM) > *noun* > DEM > ADJ (Lusekelo 2009). The example in (33) shows the word-order N > DEM > NUM, while example (34) shows the word-order N > DEM > ADJ. These examples confirm the canonical pattern of deictic referent nouns in Nyakyusa.

(33) Tu-ku-songol-a **ɪ-fɪ-kota ɪfɪ fɪ-hano**
SM1PL-PRES-curve-FV AU-8-chair DEM8 8-five
'We curve these five chairs.'

(34) **a-ba-ana-ngu aba a-ba-tali** ba-fik-ile ʊlʊ
AUG-2-child-POSS1SG 2.DEM AUG-2-tall SM2-arrive-PFV now
'These tall children of mine have just come now.'

The word order DEM > *noun* > ADJ provides a definite interpretation. In fact, the prenominal demonstrative, which indicates contrastive focus, restricts the augment from occurring on the noun, as shown in (35–36). Given this pattern, both the augment and demonstrative function to indicate the definite determiner in the nominal domain of Nyakyusa.

(35) aba ba-ana-ngu a-ba-tali ba-fik-ile ʊlʊ
2.DEM 2-child-POSS1PL AUG-2-tall SM2-arrive-PFV now
'The two tall children of mine have just come now.'

(36) aga ma-boko ma-bili ga-nya-lile
6.DEM 6-hand 6-two SM6-be.dirty-PFV
'The two arms have become unclean.'

In the complex noun phrase, the occurrence of the augment on adjectives is not restricted, as shown in (35) above. However, the numeral does not host the augment (36).

The anaphoric use of the demonstrative is also determined by the shape in Bantu languages. Mwamzandi (2014: 60) highlights that the demonstrative *-le* 'that/those' provides deictic and anaphoric references when introducing the topic referent in Swahili. The word order *noun* > DEM introduces an activated and recent referent (37), while the word order DEM > *noun* reintroduces the inactive referent into current discourse (38).

(37) M-sichana yule a-li-ingi-a
1-girl 1-DEM SM1-PST-inter-FV
'That girl entered.'

(38) Yu-le m-sichana a-li-ingi-a
1-DEM 1-girl SM1-PST-enter-FV
'That girl entered.'

The anaphoric demonstraitve in Nyakyusa is *-la* 'that/those'. It occurs in all noun classes, as exemplified below. In example (39), the anaphoric demonstrative is *gʊla* 'that' for class 3, while example (40) contains the anaphoric *gala* 'those' for class 6. In example (41), the anaphoric demonstrative is for noun class 2.

(39) pa-kyinja ba-a-gʊ-bwene ʊ-mw-esi gʊ-la
16-year SM2-PST-OM6-see.PFV AUG-6-moon 3-that
'Last year, they saw that moon.'

(40) a-ma-gali ga-la go-onangike lılımo
AUG-6-car 6-those SM6-wreck today
'Those cars have wrecked today.'

(41) a-ba-ndu ba-la ba-pimb-ile i-ly-afuli
AUG-2-person 2-DEM SM2-carry-PFV AUG-5-umbrella
'Those people carry umbrellas.'

The canonical word order of noun > DEM is also attested for the anaphoric demonstratives, as illustrated above. However, it also occurs in prenominal position, as shown in (42–43). The NPs *bala bandʊ* 'the very people' and *lila lılasi* 'the very bamboo tree' have the word order DEM > noun.

(42) **bala ba-ndʊ** ba-ø-fik-ile
2.those 2-person SM2-PST-arrive-PFV
'The (expected/known) people have arrived.'

(43) Atu a-tem-ile **lila lı-lasi**
1.Atu SM1-cut-PFV 5.that 5-bamboo
'Atu has cut the (exact) bamboo tree.'

At this juncture, two points are underscored here. On one hand, the presence of the augment may indicate both the indefinite and definite referents in Nyakyusa. And the absence of the augment indicates clefting, which is associated with focus. Any referent in focus is foregrounded; therefore, the augment cannot occur on the noun. This is confirmed by its complementary distribution with the prenominal demonstrative, which indicates definite and focused nouns (Lyons 1999). Therefore, the augment in Nyakyusa behaves like a typical determiner.

Word order is not the primary mechanism to indicate the definite interpretation of referents in Nyakyusa. The anaphoric demonstrative *–la* 'that/those' provides two kinds of definiteness: (i) the word order noun > DEM is associated with a deictic referent which is active in the discourse; (ii) the pattern DEM > noun is related to the introduction of an inactive referent which is referred to in the current situation.

Based on the observations above, the occurrence or non-occurrence of the augment cannot yield definiteness independent of the context. The proper identification of indefiniteness has to be determined by the context of speech, whether signalling active or inactive referents, as also discussed by Mwamzandi (2014).

Since the position of the determiner is now identified, I suggest the structure of the Nyakyusa nominal domain in (44). The augment, the prenominal demonstrative, and the postnominal demonstrative are all indicators of definiteness.

(44) (determiner) *noun* (determiner) (modifiers)
(DEM) (AUG) (DEM) (ADJ) (NUM)

However, the postnominal demonstrative occurs in complementary distribution with the possessive, which is another category that indicates definiteness (Lyons 1999). In Nyakyusa, the possessives occur postnominally and do not host the augment, as illustrated in (45–46). In these sentences, the possessives provide a specificity interpretation of the referents.

(45) ʊ-m-piki gʊ-ake gʊ-mel-ile
AUG-3-tree 3-his SM3-grow-PFV
'His tree has grown.'

(46) ɪ-mɪ-piki gɪ-ake gɪ-mel-ile
AUG-4-tree 4-his SM4-grow-PFV
'His trees have grown.'

The co-occurrence of possessives and demonstratives is permitted in Nyakyusa (Lusekelo 2009). In example (47), the demonstrative precedes the possessive, while in example (48), the possessive precedes the demonstrative. These examples substantiate the fact that both possessives and demonstratives occur in the determiner slot in the nominal domain.

(47) a-ba-ana aba ba-angu ba-a-fik-ile
AUG-2-child 2.these 2-mine SM2-PST-arrive-PFV
'These children of mine arrive.'

(48) a-ba-ana ba-angu aba ba-a-fik-ile
AUG-2-child 2-mine 2.these SM2-PST-arrive-PFV
'These children of mine arrive.'

Given the fact that both demonstratives and possessives occur in postnominal position, I propose the structure of the Nyakyusa nominal domain in (49). Since the augment is introduced, this structure differs from the one suggested in Lusekelo (2009).

(49) (determiner) *noun* (determiner) (modifiers)
(DEM) (AUG) (DEM) (POSS) (ADJ)(NUM)

4 Object marking and the status of the augment in bare nouns

Object marking relates to the marking of (in)definiteness and (non-)specificity in Bantu languages (Marten et al. 2007, Riedel 2009, Marten & Kula 2012, Visser 2010, Kimambo 2018a). For instance, Marten & Kula (2012: 242) argue that the use of the object marker with non-animate NPs is associated with definiteness or specificity in Swahili. The example in (50) demonstrates an indefinite/non-specific reading with no object marker in the verb, while the example in (51) demonstrates a definite/specific reading because of the object prefix.

(50) Swahlii, Marten & Kula (2012: 242)
Ni-li-on-a ki-tabu
SM1-PST-see-FV 7-book
'I saw a/the book.'

(51) Ni-li-**ki**-on-a ki-tabu
SM1-PST-OM7-see-FV 7-book
'I saw the book.'

Visser (2010) argues that object marking and occurrence of the V-augment yield definite and specific readings. The example (52) shows a non-specific reading of *ngubo* 'blanket' because there is no object prefix and no V-augment on the noun. But the example in (53) shows that the object prefix and V-augment indicate a specific reading of the object noun *ingubo* 'blanket'.

(52) Xhosa (Visser 2010: 302)
I-intombi a-zi-hlambi ngubo
AU-10.girl NEG-SM10-PRES-wash 9.blanket
'(The) girls do not wash any (not specific) blanket.'

(53) **I**-intombi a-zi-yi-hlambi **i**-ngubo
AU-10.girl NEG-SM10-OM9-wash AUG-9.blanket
'The girls do not wash the (specific) blanket.'

However, the (in)definite interpretation is not always associated with object marking for some Bantu languages without augments. Riedel (2009) offers cases from Shambala and Swahili. The example (54) confirms that in Shambala "object marking with an inherently definite noun phrase is optional" and example (55) shows that in Swahili "definite readings are available without object marking" (Riedel 2009: 49).

(54) Shambala (Riedel 2009: 49)
N-za-(mw)-ona ng'wanae
SM1-PFV.DJ-OM1-see 1child.POSS.3S
'I saw his child.'

(55) Swahili (Riedel 2009: 49)
Ni-li-penda sana ki-tabu chake cha kwanza
SM1-PST-like much 7-book 7.her 7ASS. first
'I liked her first book a lot.'

Object marking in Nyakyusa is determined by the lexical semantics of the verb (Lusekelo 2012), similar to Chiyao and Luguru (Marten & Ramadhani 2001, Taji 2020). Some verbs take compulsory object prefixes, while other kinds of verbs take optional object markers.

Some of the verbs taking compulsory object prefixes include *bona* 'see' (56–57) and *bʊʊla* 'inform' (58). Notice also that the augment occurs on the lexical object nouns *abahesya* 'guests', *ɪkɪkota* 'chair' and *abakipanga* 'congregation'.

(56) ʊ-lʊgano a-a-*(ba)-bwene a-ba-hesya mmajolo
AUG-1.Lugano SM3-PST-OM2-see.PFV AUG-2-guest yesterday
'Lugano saw (the) guests yesterday.'

(57) ʊ-lʊgano **a-*(kɪ)**-bwene (ɪ-kɪ-kota)
AUG-1.Lugano SM1-OM7-see.PFV AU-7-chair
'Lugano saw a/the chair.'

(58) ʊ-m-pʊʊtɪ a-*(**ba**)-bʊʊl-ile a-ba-kipanga.
PPX-1-priest SM-OM2-tell-PFV PPX-2-congregation
'A/the priest told (the) congregation.'

Other verbs do not require an object prefix, as illustrated by the verb *piija* 'cook' (59) and *bɪɪka* 'put' (60). Notice also that the lexical nouns *ɪfɪndu* 'food' and *ɪndeko* 'pot' host the augment. Object marking in this type of verbs would mark definite referents.

(59) ʊ-lʊgano a-piij-ile ɪ-fɪ-ndu
AUG-1.Lugano SM1-cook-PFV AUG-8-food
'Lugano has cooked some food.'

(60) ba-a-bɪɪk-ile ɪ-ndeko kʊ-sofu
SM2-PST-put-PFV AUG-10.pot 17-inner-room
'They put pots in the inner room.'

In examples (61–62), object marking is associated with a definite interpretation. The object prefix is therefore used to signal definite referents, and the augment still occurs in both lexical object nouns *abaana* 'children' and *ɪndeko* 'pots'.

(61) ʊ-lʊgano a-ba-piij-il-e a-ba-ana ɪ-fɪ-ndu
AUG-1.Lugano SM1-OM2-cook-APPL-PFV AUG-2-child AUG-8-food
'Lugano has cooked for the children the food.'

(62) ba-a-si-bɪɪk-ile ɪ-ndeko kʊ-sofu
SM2-PST-OM10-put.PFV AUG-10.pot 17-inner-room
'They put the pots in the inner room.' (Lusekelo 2012)

5 Conclusions

This chapter described the morphosyntactic properties of the augment in Nyakyusa. The bare nouns in Nyakyusa contain this morphology: AUG+NCP+stem. The V-augment is prolific in the language. I showed that the augment cannot co-occur with the prenominal demonstrative because both have the same function of indication of definiteness.

The nominal domain in Nyakyusa reveals the use of the CV-particle, previously called CV-augment (de Blois 1970, Mbope 2016, Persohn 2017). The CV-particle functions to indicate contrastive focus of the referent. As a result, it is in complementary distribution with both the proximal demonstrative and the V-augment.

The occurrence of the anaphoric demonstrative, realised as *-la* 'that/those', yields the word orders noun > DEM and DEM > noun. The former word order introduces the deictic and active referent, while the latter introduces the inactive referent in discourse. The prenominal anaphoric demonstrative occurs in complementary distribution with the V-augment and CV-augment. This substantiates the fact that they realise one single function of indication of definiteness.

The adnominal demonstrative which occurs in the postnominal position has similar functions as the possessives. Lyons (1999) shows that possessives indicate definiteness. In the foregoing discussion, I showed that the nominal domain in Nyakyusa hosts the elements captured in (63).

(63) (determiner) *noun* (determiner) (modifiers)
(DEM) (AUG) (DEM) (POSS) (ADJ) (NUM)

Morever, the object prefix is required in some verbs such as *bona* 'see' and optional in other verbs such as *piija* 'cook'. As a result, it has no influence on the occurrence or non-occurrence of the augment in Nyakyusa. However, when it occurs optionally (in verbs like *piija* 'cook'), it yields a definite reading.

Abbreviations

1, 2, 3 etc.	noun class number	PFV	perfective aspect marker
ADJ	adjective	PL	plural
AUG	augment	POSS	possessive
DEM	demonstrative	PROX	proximal
FV	(default) final vowel	PST	past tense marker
OM	object marker	SG	singular
NCP	noun class prefix	SM	subject marker
NUM	numeral		

References

Aboh, Enoch O. 2004. Topic and focus within D. *Linguistics in the Netherlands* 21. 1–12.

Alexiadou, Artemis, Liliane Haegeman & Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin: Mouton de Gruyter.

Asiimwe, Allen. 2014. *Definiteness and specificity in Runyankore-Rukiga*. Stellenbosch: Stellenbosch University of South Africa. (Doctoral dissertation). http://hdl.handle.net/10019.1/95926 (16 March, 2015).

Bokamba, Georges D. 1971. Specificity and definiteness in Dzamba. *Studies in African Linguistics* 2. 217–38.

de Blois, Kornelis Frans. 1970. The augment in the Bantu languages. *Africana Linguistica* 4. 85–165.

Felberg, Knut. 1996. *Nyakyusa-English-Swahili and English-Nyakyusa dictionary*. Dar es Salaam: Mkuki na Nyota Publishers.

Goodness, Devet. 2013. What triggers the augment in Shinyiha. *Journal of Education, Humanities and Sciences* 2(1). 33–46.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list*. Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.

Hyman, Larry M. & Francis X. Katamba. 1993. The augment in Luganda: Syntax or pragmatics? In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 209–256. Stanford: CSLI, Stanford University.

Kaji, Shigeki. 2009. Tone and syntax in Rutooro, a toneless Bantu language of Western Uganda. *Language Science* 31(1/2). 239–247.

Kimambo, Gerald E. 2018a. Object marking in Swahili: Definiteness, specificity, or both? *South African Journal of African Languages* 38(1). 27–35.

Kimambo, Gerald E. 2018b. The morpho-syntactic and semantic-pragmatic realisation of definiteness and specificity in Swahili. *Ghana Journal of Linguistics* 7(1). 65–83.

Legère, Karsten. 2005. Preprefix or not – that is the question. In Erhard F. K. Voeltz (ed.), *Studies in African linguistic typology*, 251–262. Amsterdam, Philadelphia: John Benjamins.

Lusekelo, Amani. 2009. The structure of the Nyakyusa noun phrase. *Nordic Journal of African Studies* 18(4). 305–331.

Lusekelo, Amani. 2012. *Inflectional and derivational morphology in optimality theory: Multiple object-nouns and co-occurrence of verbal extensions in Kinyakyusa*. Gaborone: University of Botswana. (Doctoral dissertation).

Lusekelo, Amani. 2013. DP-internal and V-external agreement patterns in Bantu: Re-statement of the facts on Eastern Bantu languages. *Journal of Linguistics and Language in Education* 7(1). 19–47.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Marten, Lutz & Nancy Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30(2). 237–253.

Marten, Lutz, Nancy Kula & Nhlanhla Thwala. 2007. Parameters of morpho-syntactic variation in Bantu. *Transactions of the Philological Society* 105. 253–338.

Marten, Lutz & Deograsia Ramadhani. 2001. An overview of object marking in Kiluguru. *SOAS Working Papers in Linguistics* 11. 259–275.

Mbope, Ngusubila. 2016. *The pre-prefix in Kinyakyusa*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

Mwamzandi, Mohamed Y. 2014. *Swahili word order choices: Insights from information structure*. Arlington, TX: University of Texas at Arlington. (Doctoral dissertation).

Ndomba, Rodrick G. 2017. *The structure and derivation of the determiner phrase in Kiswahili*. Dublin: University College Dublin. (Doctoral dissertation).

Persohn, Bastian. 2017. *The verb in Nyakyusa: A focus on tense, aspect and modality*. Berlin: Language Science Press.

Petzell, Malin & Karoline Kühl. 2017. The influence of non-linguistic factors on the usage of the pre-prefix in Luguru. *Linguistic Discovery* 15(1). 35–48.

Repp, Sophie. 2010. Defining 'contrast' as an information-structural notion in grammar. *Lingua* 120. 1333–1345.

Riedel, Kristina. 2009. *The syntax of object marking in Sambaa: A comparative Bantu perspective*. Leiden: LOT Publications.

Rijkhoff, Jan. 2002. *The noun phrase*. London: Oxford University Press.

Rugemalira, Josephat M. 2007. The structure of the Bantu noun phrase. *SOAS Working Papers in Linguistics* 15. 135–148.

Taji, Julius. 2020. Verb stem influence on object marking in Chiyao. *Southern African Linguistics and Applied Language Studies* 38(4). 351–365. DOI: 10.2989/16073614.2020.1859942.

Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics with special focus on Bantu and Mande*, 425–441. Tervuren: Royal Museum for Central Africa.

Visser, Marianna. 2002. The category DP in Xhosa and Northern Sotho. *South African Journal of African Languages* 22(4). 280–293.

Visser, Marianna. 2008. Definiteness and specificity in the isiXhosa determiner phrase. *South African Journal of African Languages* 28(1). 11–29.

Visser, Marianna. 2010. Definiteness and specificity in Xhosa determiner phrase. In Karstens Legère & Christina Thornell (eds.), *Bantu languages: Analysis, description and theory*, 295–314. Koeln: Ruediger Koeppe Verlag.

Chapter 5

The morphosyntax of locative expressions in Kiwoso

Aurelia Mallya
University of Dar es Salaam

This paper discusses the morphological and syntactic properties of locative expressions in Kiwoso. It provides an account of the locative forms and their properties in relation to nominal and verbal morphology. The findings show that locative nouns in Kiwoso are formed by means of a locative suffix *-(i)n*. It also shows that the traditional Bantu locative class prefixes (*ku-*, *pa-*, *mu-*) are unproductive in Kiwoso. However, the locative class 17 prefix *ku-* triggers agreement on all nominal and verbal modifiers, indicating that locative meanings are still part of the noun class system in the language. The data show that Kiwoso exhibits two post-final locative enclitics – *=ho* and *=u*. Both particles are used to indicate locative objects, albeit with different interpretations. The post-final *=ho* relates directly to the semantics of the locative noun *kundo*, while *=u* corresponds to the interpretation of the locative noun *ando*. This paper contributes to the understanding of locatives within the Bantu language family in general, and offers new insights about locatives in Kiwoso, an area which has not received extensive treatment in the previous literature.

1 Introduction

Locative constructions have received extensive attention in the previous literature on Bantu languages. Descriptive accounts suggest that locative expressions are marked differently both within and between Bantu languages (Marten et al. 2007, Persohn & Devos 2017). In a number of Bantu languages, locative expressions are derived by attaching the class 16, 17 and 18 prefixes to a noun (see Rugemalira 2004, Petzell 2008, Riedel & Marten 2012, Guérois 2016, Van de Velde

Aurelia Mallya. 2024. The morphosyntax of locative expressions in Kiwoso. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 109–128. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663769

2019, among others). However, while some languages such as Kagulu (Petzell 2008), Bemba (Marten 2012) and Chichewa (Bresnan & Kanerva 1989) have maintained all of the three locative prefixes, others like Kivunjo-Chaga (Moshi 1995) and Sesotho (Demuth & Mmusi 1997) exhibit only two productive prefixes. Moreover, languages such as Haya and Zulu exhibit only one productive locative prefix (Riedel & Marten 2012).

In addition to the prefixation strategy, locative nouns in Bantu languages may also be derived by means of suffixation (cf. Grégoire 1975, Guérois 2016). This strategy is predominantly attested in Eastern and Southern Bantu languages and most of the languages that employ a locative suffix lack locative prefixes. Scholars have also noted that there are Bantu languages that employ both prefixes and suffixes in marking locatives (Marten 2010, 2012). This is the case in Nguni languages of Southern Africa, for example, in which locative noun class 25 (*e-*) and the suffix *-(i)ni* are used jointly to derive locative nouns (van der Spuy 2014). It has also been noted that in languages in zone P30 (spoken in Mozambique), the traditionally recognized locative prefixes (those of classes 16-18) can be used in combination with the locative suffix *-ni* to derive locative expressions (Guérois 2016).

The variation in locative constructions has attracted the attention of a wide range of scholars who are interested in investigating the nature of locative expressions in individual Bantu languages, particularly in relation to the domain of morphosyntax. The present paper contributes to the on-going description and discussion of the morphosyntax of locative nouns in Bantu, using data from the Tanzanian Bantu language Kiwoso. The chapter aims to address issues regarding the morphosyntax of Kiwoso locative expressions, with reference to Guérois et al.'s (2017) parameters. Guérois et al. (2017) propose 142 descriptive parameters aimed at examining morphosyntactic variation in Bantu languages. For the purposes of the present study, I have selected four parameters to address key issues pertaining to Bantu locative constructions: i) the forms of locative expressions in Kiwoso, ii) agreement patterns, iii) locative subject and object marking, and iv) the presence or absence of locative postverbal enclitics.

The rest of this paper is organized as follows: §2 provides a brief linguistic profile of Kiwoso, while an overview of the noun class system of the language is presented in §3. Locative nouns, their forms and the associated agreement system are discussed in §4. §5 summarizes and concludes the discussion offered in this chapter.

The Kiwoso data presented in this work are based on the intuition of the author as a native speaker, complemented by acceptability judgements provided by two other native speakers of Kiwoso. The primary data are supplemented

by secondary data obtained from existing written documents, particularly the dissertations by Mallya (2011, 2016) and Mushi (2005). Examples from other languages used in this chapter are taken from various sources and are acknowledged accordingly.

2 Linguistic profile of Kiwoso

Kiwoso is an eastern Bantu language spoken predominantly in the Kilimanjaro region of Tanzania. In the survey carried out by the Languages of Tanzania Project, it was reported that Kiwoso is spoken by approximately 81,000 people who are scattered across different districts of the Kilimanjaro region (LOT 2009). Native speakers of Kiwoso are mainly found in the administrative areas of Moshi Rural, Hai, Siha, and Moshi Town Districts. Maho (2009) classifies Kiwoso as one of the Zone E languages belonging to the Chagga group (E60) and Kiwoso specifically is coded as E621D (Maho 2009).

Kiwoso is one of a large number of under-studied and under-described languages of Tanzania. The only available literature on Kiwoso is a dictionary (Kagaya & Olomi 2009), two unpublished MA dissertations (Mallya 2011, Mushi 2005) and a PhD thesis (Mallya 2016). Although the present paper is not intended to provide a full linguistic description of Kiwoso, some background information on the noun class system is presented before embarking on the more specific discussion of the morphology and the syntax of locatives, the primary focus of this paper.

3 The Kiwoso noun class system

Kiwoso displays the typical Bantu noun class system and exhibits 14 noun classes, as illustrated in Table 1. For each noun class presented in the table, the nominal prefix, an example word, the subject and object agreement morphemes, adjective and possessive prefixes, and the three forms of demonstrative are also shown.

Most of the noun classes in classes 1–10 appear in a singular-plural pairing. More specifically, classes 1, 3, 5, 7, 9, and 11 contain singular nouns, while classes 2, 4, 6, 8, and 10 contain their plural counterparts. However, not all classes conform to this pairing system. For example, class 11 nouns form their plural counterparts in class 6, and class 14 nouns lack plural counterparts. The singular-plural pairing system of noun classes found in Kiwoso is illustrated in Figure 1 below. The class 11/6 plural pairing is exemplified by the examples in (1).

Table 1: The Kiwoso noun class system

Class	Nominal Prefix	Example	Gloss	SM	OM	Adj Prefix	Poss Prefix	Dem1	Dem2	Dem3
1	*mu-*	*mu-na*	child	*a-*	*n-*	*n-*	*o-*	*e-tu*	*e-to*	*u-lya*
2	*wa-*	*wa-na*	children	*wa-*	*wa-*	*wa-*	*wa-*	*e-wa*	*e-wo*	*wa-lya*
3	*n-*	*n-ji*	tree	*u-*	*i-*	*n-*	*o-*	*e-tu*	*e-to*	*u-lya*
4	*mi-*	*mi-ji*	trees	*i-*	*i-*	*mi-*	*ta-*	*e-ti*	*e-to*	*tya*
5	*i-*	*i-du*	ear	*lyi-*	*lyi-*	*lyi-*	*lya-*	*e-lyi*	*e-lyo*	*lya*
6	*ma-*	*ma-du*	ears	*a-*	*wa-*	*ma-*	*a-*	*e-wa*	*e-wo*	*alya*
7	*ki-*	*ki-andu*	knife	*ki-*	*ki-*	*ki-*	*ki-*	*e-kyi*	*e-kyo*	*kya*
8	*shi-*	*shi-andu*	knives	*shi-*	*shi-*	*shi-*	*shi-*	*e-shi*	*e-sho*	*shya*
9	*N-*	*mburu*	goat	*i-*	*i-*	*ngi-*	*a-*	*e-yi*	*e-yo*	*iya*
10	*N-*	*mburu*	goats	*ti-*	*ti-*	*ngi-*	*ta-*	*e-ti*	*e-to*	*tya*
11	*u-*	*u-dende*	leg	*lu-*	*lu-*	*lu-*	*lo-*	*e-lu*	*e-lo*	*lou*
14	*u-*	*u-doko*	laziness	*lu-*	*lu-*	*lu-*	*lo-*	*e-lu*	*e-lo*	*lou*
16	*a-*	*a-ndo*	place	*ku-*	*ku-*	*ku-*	*ko-*	*yaa*	*yoo*	*alya*
17	*ku*	*ku-ndo*	place	*ku-*	*ku-*	*ku-*	*ko-*	*kunu*	*kulya*	*kulya*

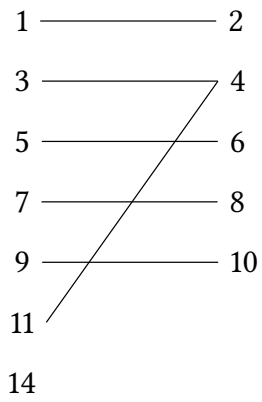


Figure 1: Singular/plural noun class pairings in Kiwoso

(1) a. Lelo ni-a-le-many-a u-dende na kyaara
Lelo INIT-2SM-PST-cut-FV 11-leg by 7.axe
'Lelo cut (his) leg by (means of) an axe'

b. Lelo ni-a-le-many-a ma-dende na kyaara
Lelo INIT-2SM-PST-cut-FV 6-leg by 7.axe
'Lelo cut (his) legs by (means of) an axe'

In many Bantu languages, the class 15 prefix *ku-* is the prefix for infinitival nouns (Katamba 2003, Van de Velde 2019). However, Kiwoso differs from the majority of Bantu languages in relation to the infinitive marker. In Kiwoso, infinitives are marked with the class 5 prefix *i-* which also triggers class 5 subject and object agreement similarly to other class 5 nouns. The infinitive morpheme in Kiwoso can be illustrated using infinitives such as *ikora* 'to cook', *idema* 'to cultivate', *isoma* 'to read', and *iseka* 'to laugh'. Interestingly, the Tanzanian Bantu language Rangi also employs some class 5 infinitives. However, in Rangi, the class 5 infinitive no longer appears to be the productive (nor dominant) noun class for the formation of infinitives. Rather, it is used in addition to the more widespread class 15 infinitive marking (Gibson 2012).

Note also that while many Bantu languages form diminutives by assigning nouns to classes 12 and 13, which are amongst the classes reconstructed for diminutives in Proto-Bantu (Meeussen 1967), diminutives in Kiwoso are expressed by a shift into classes 7/8 and the associated prefixes *ki-*/*shi-*. For example, *iwee* 'stone', *kiwee* 'small stone' *shiwee* 'small stones' and *uwoko* 'hand' *kiwoko* 'small hand' *shiwoko* 'small hands'.

The following section discusses the locative noun classes 16 (*a-*) and 17 (*ku-*) which are of particular relevance in this paper.

4 Locative nouns in Kiwoso

4.1 Unproductive locative prefixes

In Kiwoso, there are two locative nouns, namely *ando* and *kundo,* both signifying 'place'. However, the nouns *ando* and *kundo* are pragmatically different in that the former (*ando)* can be interpreted as a place which is definite, specific, known, and near to both the speaker and the hearer, whereas the latter (*kundo)* refers to a place which is indefinite, unspecific, unknown, and far from both the speaker and the hearer.

The two shades of meaning associated with the locative nouns *ando* and *kundo* can be seen in examples (2) and (3), respectively. It is important to note from the outset that the only grammatically active locative classes in the language are class17 marked by *ku-* and class 16 marked by *a-*. However, in contrast, locative agreement in Kiwoso is regularly marked with class 17 (see also §4.3.1).

(2) a. a-lya a-ndo ku-cha
16-DEM3 16-place 17-nice
'There is a nice place' [definite].

b. lya wa-na wa-le-ch-a a-le-end-a a-ndo ka-woiya-u
REL 2-child 2-PST-come-FV 3SG-PST-go-FV 16-place CONSC-keep-there
sau
silent
'When the children arrived, s/he went to that place and kept silent' [definite]

(3) a. ku-lya ku-ndo ku-cha
17-DEM3 17-place 17-nice
'There is a nice place' [indefinite].

b. wa-ka wa-le-fik-a ku-ndo ku-lya umbe ti-lekumb-o
2-woman 2-PST-arrive-FV place 17-DEM3 10.cow 10-sell-PASS
'Women reached the place where cows were sold'. [indefinite]

Examples (2) and (3) suggest that, in common with many other Bantu languages, locative prefixes in Kiwoso can function as noun class markers in the sense that they can be attached to nominal stems yielding locative meanings. Note also that the locative noun class reconstructed as class 18 **mʊ-* in Proto-Bantu and found synchronically as a variant of *mu-* in a number of other Bantu languages, does not exist in Kiwoso.

In addition to specific place names such as *Dar es Salaam*, *Arusha*, *Tanga*, and *Kampala,* there are general place names in Kiwoso such as *kinaange* 'market', *mmba* 'house', *shuule* 'school' *kai* 'attic' and *bo* 'home'. These names are inherently locative in nature and have to be unmarked for locative, as ungrammatical construction in (4b) illustrates. However, similarly, to derived locative nouns, inherent locative nouns take class 17 agreement, as exemplified in the locative inversion construction in (4c), based on the sentence in (4a). (See also §4.3.1).

(4) a. wa-ka wa-le-koon-a kinaange
2-woman 2SM-PST-meet-FV market
'The women met at the market (place).'

b. * wa-ka wa-le-koon-a kinaangen
2-woman 2SM-PST-meet-FV market
'The women met at the market (place).'

c. kinaange ku-le-koon-a wa-ka
market 17-PST-meet-FV 2-woman
'The Market is a place where women used to meet.'

Locative constructions such as in (4) are interpreted differently in terms of discourse-pragmatics. In (4a), the locative noun *kinaange* 'market' serves as a focus while in (4c) the noun encodes a topic. (See Marten & Gibson 2016, Marten & van der Wal 2014 and Mallya 2020 for further details on locative inversion constructions).

4.2 Locative suffixation

The present section provides a brief introduction to the locative suffix *-(i)ni* in Bantu languages, before discussing the morphology of locative nouns in Kiwoso. As shown in the introduction, apart from the commonly established pattern of locative marking which involves the three locative prefixes from classes 16, 17 and 18, some Bantu languages derive locative nouns by means of the locative suffix *-(i)ni* (or variants thereof).

Although the suffix *-(i)ni* is widely attested in eastern and southern Bantu languages, there is currently no consensus on its origins. Different scholars have put forth different proposals on the source of this suffix. For example, Meinhof (1941/42) as cited in Samson & Schadeberg (1994: 128) proposes that the locative suffix is derived from the locative class prefix 18 (*mu-*). Meinhof's proposal is further supported by Güldemann (1999) who argues that the suffix *-(i)ni* was originally a marker of inessive relations which later developed into a general

locative. However, Samson & Schadeberg (1994) have convincingly demonstrated that the locative suffix is the result of grammaticalization of the word *-*ini* 'liver'.

Some Bantu languages use double locative marking, combining both prefixation and suffixation. For example, this pattern is found in the P30 languages spoken in Mozambique (Guérois 2016) and southern Bantu Nguni languages (Fleisch 2005, van der Spuy 2014). The P30 languages use the prefixes of classes 16, 17 and 18 in addition to the locative suffix *-ni*, whereas the Southern Bantu languages use a combination of the class 25 prefix *e-* and the locative suffix (*-i*)*ni*. In contrast, locative marking in Kiwoso is solely achieved through suffixation, as will be further shown in the following section.

Locative nouns in Kiwoso are derived by attaching a locative suffix -(*i*)*n* to the noun. This contrasts with Bantu languages in which locative expressions are achieved by means of locative prefixes such as Bemba (Marten 2010, 2012), Kagulu (Petzell 2008), and Chichewa (Bresnan & Kanerva 1989). Examples of the use of the locative suffix in Kiwoso are shown in Table 2 below. The data presented in this paper indicate that there is an instance of vowel coalescence in Kiwoso, when the locative suffix (*-i*)*n* is attached to nouns that end with the vowel *-a*. In such instances, the vowel changes into *-e*, as the examples in Table 2 illustrate.

Table 2: Locativised nouns in Kiwoso

ordinary nouns	gloss	nouns with -(*i*)*n*	gloss
ndubhi	'calabash'	*ndubhin*	'in the calabash'
nlango	'door'	*nlangon*	'at the door'
nlima	'mountain'	*nlimen*	'on/at/ the mountain'
nungu	'pot'	*nungun*	'in the pot'
muda	'water'	*muden*	'in the water'
kitara	'bed'	*kitaren*	'on bed'
umbe	'cow'	*umben*	'at/among cows'
irike	'warmth'	*iriken*	'in the warmth'

Depending on the context, the locative suffix -(*i*)*n* in Kiwoso demonstrates all shades of meanings expressed by the traditionally recognized locative prefixes *pa-*, *ku-*, and *mu-*. Suffixation as a means of deriving locative nouns has been attested in other East African Bantu languages such as Kikuyu (Mugane 1997), Kiswahili (Grégoire 1975), Kamba (Kioko 2005), and in southern Bantu languages such as Tswana (Creissels 2011) and Swati (Marten 2010).

Prototypically, in many Bantu languages, class 16 expresses nearness, specific and definite location. Class 17 denotes remoteness, unspecific and indefinite location, while class 18 indicates interiority, inside or location within (see Grégoire 1975, Maho 1999, Fleisch 2005, Marten et al. 2007, Guérois 2016). Although the specific meaning expressed by these prefixes differs across languages, the aforesaid are the general meanings associated with the locative classes. Illustrative examples are provided in (5).

(5) a. wa-ndu wa-le-id-a ruko-n
2-people 2SM-PST-enter-FV 9.kitchen-LOC
'People entered in (i.e., inside) the kitchen.'

b. wa-ka wa-le-lal-a ki-tare-n
2-woman 2SM-PST-sleep-FV 7-bed-LOC
'Women slept on the bed.'

c. wa-na wa-le-shaam-a n-lime-n
2-child 2SM-PST-climb-FV 3-mountain-LOC
'Children went to the mountain.'

d. duke-n ku-le-ch-a wa-ndu
9.shop-LOC 17-PST-come-FV 2-people
'At the shop there came people.'

The locative expressions *ruko-n* 'in the kitchen' in (5a) denotes an inside or interior location, *kitare-n* 'on the bed' in (5b) and *nlime-n* 'to the mountain' in (5c) indicate general and non-specific locations, whereas *duken* 'at the shop' in (5d) expresses a specific, definite location. These examples show that the locative suffix -(*i*)*n* in Kiwoso can be used to express a range of nuances of meaning which are associated the locative classes 16, 17 and 18 cross-Bantu. As mentioned in §4.1, the locative suffix *-*ini* and related forms has been considered to be the grammaticalized form of the lexeme meaning 'liver', and it is thought to have originally been used to denote interior location before it expanded further to denote other locative relations in languages such as Kiwoso. (For further details about the suffix across Bantu languages see Grégoire (1975: 185–204) and Güldemann (1999: 51–52)).

The available evidence suggests that semantically, locative suffixes cannot occur with animate nouns in some languages. For example, in Kiswahili, nouns such as *mtu* 'person' and *nguruwe* 'pig' and *paka* 'cat' cannot be locativised by means of the suffix -*ni* (Rugemalira 2004). This means that the constructions such as *mtu-ni*, *nguruwe-ni*, and *paka-ni* are unacceptable. In contrast, in Kiwoso, the

locative suffix -(*i*)*n* can be affixed to animate nouns to form locative nouns. The examples in (6) demonstrate this process.

(6)

Ordinary nouns	gloss	locativised nouns	gloss
wandu	'people'	*wandun*	'at/in/with/by the people'
mburu	'goat'	*mburun*	'by the goat'
umbe	'cow'	*umben*	'by the cow'
baka	'cat'	*baken*	'by the cat'
kite	'dog'	*kiten*	'by the dog'

Apart from the nouns exemplified in (6), the locative suffix -(*i*)*n* in Kiwoso is also used to mark abstract locations. The suffix can be attached to abstract nouns, such as *reema* 'darkness', *mmbari* 'sun', and *ngoo* 'heart' to form locative nouns, as the forms in (7) illustrate.

(7)

Ordinary nouns	gloss	locativised nouns	gloss
reema	'darkness'	*reemen*	'in the darkness'
mmbari	'sun'	*mmbarin*	'in the sun'
ngoo	'heart'	*ngoon*	'in/from/the heart'

The data presented in (6) and (7) suggest that in Kiwoso, animate and inanimate nouns, as well as abstract entities can express places or locations by simply adding the locative suffix -(*i*)*n*. Note also that agreement on the dependents of the nouns is marked by the invariant locative class 17 prefix *ku-*. Agreement patterns are explained further in §4.3.1 below.

4.3 Locative agreement patterns

4.3.1 Locative marking within NPs

Locative expressions are also realized differently in terms of morphology. In Bantu languages, locative nouns are often associated with different types of agreement markers. Usually, in languages where locative classes 16, 17, and 18 are productive, a series of concordial class prefixes are associated with the derived nouns. In Cuwabo and Makhuwa (Guérois 2016), Bemba (Marten 2012) and Chichewa (Bresnan & Kanerva 1989, Carstens 1997), for example, all three locative prefixes exhibit full agreement with other elements in a construction. This is demonstrated in the examples in (8) from Chichewa (Carstens 1997: 362).

(8) a. pa-nyumba pa-ku-on-ek-a ngati pa-ku-psy-a
16-9house 16-ASP-see-STAT-FV like 16-ASP-burn-FV
'The house and surrounding yard look like they are burning.'

b. ku-nyumba ku-ndi ku-tali
17-9house 17-DEM 17-far
'That house and its environs are far away.'

c. mu-nyumba mu-ku-nunkh-a
18-9house 18-ASP-stink-FV
'Inside the house stinks.'

In contrast to languages such as Chichewa, some Bantu languages such as Herero and Lozi (Marten et al. 2007) distinguish three locative noun classes, but only one or two of the classes are reflected in the agreement pattern of these languages. Lozi and Herero for example exhibit a three-way distinction in the class prefix of locative nouns, but subject agreement is exclusively marked by the class 17 prefix. Similarly, in Kinyarwanda, locative agreement on predicates is invariably marked by the class 16 prefix (Zeller & Ngoboka 2018).

This observation suggests that the absence of a three-way locative class prefix distinction on derived nouns does not preclude a three-way locative noun class prefix system on nominal modifiers and verb agreement. Grégoire (1975) has pointed out that locative nouns in languages such as Kiswahili, Shambala, and Bondei are consistently achieved by means of the locative suffix *-(i)ni*, but agreement markers on dependents reflect the three-way class distinction, as the examples in (9) from Kiswahili show (Carstens 1997: 402).

(9) a. nyumba-ni pa-ngu pa-zuri
9.house-LOC 16-my 16-good

b. nyumba-ni kw-angu ku-zuri
9.house-LOC 17-my 17-good

c. nyumba-ni mw-angu m-zuri
9.house-LOC 18-my 18-good
'at/in my good house'

Indeed, this is the system that is seen in Kiwoso. In common with other languages that express location by means of the suffix *-(i)ni*, and in which locative prefixes are unproductive, Kiwoso exhibits agreement markers on different locative nominal modifiers. However, in Kiwoso, agreement on dependents is marked by the invariant locative class 17 prefix *ku-*. Examples in (10) are illustrative of this pattern.

(10) a. ruko-n ko-ke ku-cha
9.kitchen-LOC 17-POSS 17-nice
'at/in his/her nice kitchen'

b. ** ruko-n lya-ke lyi-cha
9.kitchen-LOC 9-POSS 9-nice

c. ki-tare-n ko-ke ku-cha
7-bed-LOC 17-POSS 17-nice
'on his/her nice bed'

d. ** ki-tare-n ki-ake ki-cha
7-bed-LOC 7-POSS 7-nice

The examples in (10) show that it is the locative suffix *n-* which controls agreement on other modifiers, such as possessives and adjectives and not the prefix of the inherent noun. Marten (2012) describes such an agreement system as 'outer' agreement. Like many other Bantu languages, the noun *kundo* 'place' in Kiwoso reflects a remnant of a locative class 17 prefix. The data presented above further indicate that unlike Zigua and Kamba (Marten 2012) which show agreement with the original noun class of the locative noun, Kiwoso does not license inner agreement, as the unacceptability of the examples in (10b) and (10d) show. In Kiwoso, when the locative class prefix *ku-* triggers agreement on the modifiers, as in (10a) and (10c), the emphasis is on the location (i.e., the modifier provides information about the location).

4.3.2 Locative verbal marking

Locative nouns in Bantu languages such as Kagulu (Petzell 2008), Chichewa (Bresnan & Kanerva 1989) and Haya (Riedel 2010) exhibit subject agreement on the verb, as examples in (11) and (12) from Chichewa and Kagulu, respectively, demonstrate.

(11) (Bresnan & Kanerva 1989: 3)
m-mi-tengo mw-a-khal-a a-nyani
18-4-tree 18-PERF-sit-FV 2-baboon
'In the trees are sitting baboons'.

(12) (Petzell 2008: 75)
ku-m-lomo ku-fimb-a ku-gati
17-3-mouth 17-swell-FV 17-inside
'The mouth has swollen inside'.

Chichewa and Kagulu are examples of Bantu languages which distinguish three locative noun classes, viz. 16-18 (Marten et al. 2007). In Chichewa, locative noun class agreement on the verb is reflected by the presence of subject markers for classes 16, 17, and 18, as shown in (13).

(13) a. pa-msika-pa pa-badw-a nkhonya
16-3market-6DEM 16SM-be_born-FV 10fist
'At this village the fight is going to break out.'

b. ku-mu-dzi ku-na-bwer-a a-lendo
17-3-village 17SM-PST-come-FV 2-visitor
'To the village came visitors.'

c. m-nkhalango mw-a-khal-a mi-kango
18-9forest 18SM-PRF-remain-FV 4-lion
'In the forest have remained lions.'

However, not all Bantu languages reflect the full three-way locative noun class distinctions. In Kinyarwanda, Subwa and Sukuma, for example, locative agreement on the verb is restricted to class 16 regardless of the class of the locative noun (cf. Maho 1999). Similarly, in Lozi locative subjects are invariably marked by class 17 *ku-* (Marten et al. 2007).

Subject agreement with locative nouns is not only attested in languages with a locative prefix. The agreement is also exhibited in the languages which mark locative nouns with the locative suffix *-(i)ni*. Example (14) from Kiswahili shows that agreement on the verb is marked by locative classes 16-18, regardless of the fact that the language does not mark locative nouns through locative class prefixes (Carstens 1997: 402).

(14) nyumba-ni pa-/ku-/m-na wa-tu wengi
9.house-LOC 16SM-/17SM-/18SM-has 2-people 2.many
'In/at the house has many people.'

As has been shown in §4.2, Kiwoso derives locative expressions by means of the suffix *-(i)n*. However, subject agreement is consistently with locative class 17 prefix for all locative nouns. Examples in (15) are illustrative of this.

(15) a. duke-n ku-le-ch-a wa-ndu wa-fye
9.shop-LOC 17SM-PST-come-FV 2-people 2-many
'In/at the shop came many people'.

b. ku-le-ch-a wa-ndu wa-fye
17-PST-come-FV 2-person 2-many
'There came many people'.

c. mmba ku-le-id-a mbefu
9.house 17-PST-enter-FV 10.ants
'In the house entered ants'.

Note that the locative prefix *ku-* in (15) indicates a location or a place. The prefix is a grammatical locative subject marker; thus, constructions with the locative prefix *ku-* cannot be interpreted as impersonal constructions in Kiwoso. This interpretation (locative) holds even when the lexical locative subject is not mentioned, as (15b) exemplifies. Example (15c) shows that inherent locative nouns, i.e., locative nouns without locative morphology, also trigger class 17 subject prefix on the verb (see also Mallya 2016).

In addition to locative subject, locative expressions in a number of Bantu languages trigger locative object agreement. Examples from Kivunjo-Chaga (Moshi 1995: 138) and Haya (Riedel & Marten 2012: 282) in (16) and (17), respectively, demonstrate this.

(16) wa-fee wa-ku-ichi (kayi)
2-parent 2-OM17-know (9.attic)
'The parents know there (the attic place)'.

(17) n-ka-ha-gul-a
1SM-PST-OM16-buy-FV
'I bought it (the place)'.

Examples in (16) and (17) illustrate locative object marking in two Bantu languages. However, not all Bantu languages can license locative object markers. Studies indicate that languages such as Lozi, Chasu, Yeyi and the languages of the Nguni group do not realize locative object markers (Marten et al. 2007). Languages of zone P30 such as Cuwabo and Makhuwa also lack locative object markers (Guérois 2016), making locative object marking another area of variation amongst Bantu languages.

However, Kiwoso can realise locative objects on the verb. The locative object is marked by the locative class 17 prefix *ku-* only, as illustrated by the examples in (18).

(18) a. mmba wa-le-me-**ku**-loly-a
9.house 2SM-PST-PERF-OM17-see-FV
'In the house they have seen (it) there.'

b. wa-ndu wa-le-**ku**-many-a (Muchi)
2-people 2SM-PST-OM17-know-FV Moshi
'People knew (recognized) (it) there (Moshi).'

4.4 Locative verbal enclitics

A locative enclitic, as commonly found in many Bantu languages, is a morpheme that can be attached to the verb to license locative expressions. A large number of Bantu languages exhibit a locative enclitic which establishes the location in which a particular event takes place (see Persohn & Devos 2017 for further discussion and examples of this). Kiwoso exhibits two postverbal locative enclitics, namely =*ho* and =*u*. These markers are considered to be enclitics since they occur after all other suffixes, including the final vowel (see example 19b). These locative enclitics can only be attached to the verb to contribute the locative semantics, as exemplified in (19).

(19) a. wa-na wa-le-bhik-a ki-tabu i-kari-n
2-child 2-PST-put-FV 7-banana 5-car-LOC
'The children put a book in the car'

b. wa-na wa-le-bhik-a=**ho/u** ki-tabu
2-child 2-PST-put-FV=LOC 7-book
'The children put there a book'

c. wa-na wa-le-bhik-a ki-tabu
2-child 2-PST-put-FV 7-book
'The children put the book'

The examples in (19) show that a locative enclitic =*ho/u* is an obligatory part of the verb *bhika* 'put' when a full locative noun is omitted, as the unacceptability of the sentence in (19c) demonstrates. The obligatory locative enclitic =*ho/u* in example (19b) refers to an object argument.

The data from Kiwoso show that although the two elements (=*ho* and =*u*) function as true locative objects, their interpretation is slightly different from each other. On the one hand, =*ho* is used to indicate a place or a location which is indefinite, non-specific and which is far from both the speaker and the hearer. On the other hand, =*u* is used when both the speaker and the hearer are certain about the place or the location, and such a location or a place is specific and closer to both the speaker and the hearer. For example, in (20a) the locative noun *nnda* 'land/field' is assumed to be far from both the speaker and the hearer. This contrasts with example (20b). The use of the demonstratives *kulya* 'there' (afar) and *alya* 'there' (near) serve to confirm the difference between the enclitics =*ho* and =*u*. In other words, *kulya* cannot co-occur with =*u* and *alya* cannot co-occur with =*ho*.

(20) a. wa-ka wa-le-ur-a nnda (kulya) wa-ka-dema=ho
2-woman 2SM-PST-buy-FV land (DEM3) 2SM-PERF-cultivate=LOC
soko
beans
'Women bought land (there) and planted there beans'.

b. wa-ka wa-le-ur-a nnda (alya) wa-ka-dema=u
2-woman 2SM-PST-buy-FV land (DEM2) 2SM-PERF-cultivate=LOC
soko
beans
'Women bought land (there) and planted there beans'.

In terms of interpretation, the clitic =*ho* mirrors the meaning assigned to the locative noun *kundo* 'place', while the semantics of the clitic =*u* matches the one associated with the locative noun *ando* 'place', as also shown in (20) (cf. §4.1 for details on the semantic differences of the nouns *ando* and *kundo*).

More examples of the use of the post-final locative enclitics as objects in Kiwoso is exemplified in (21).

(21) a. duke-n ku-le-ch-a=**ho** wa-ndu
9.shop-LOC 17-PST-come-FV=LOC 2-people
'At the shop came (there) people.'

b. ku-le-ch-a-=**ho** wa-ndu (duke-n)
17-PST-come-FV=LOC 2-people (9.shop-LOC)
'There came (there) people (at the shop).'

c. wa-ndu wa-le-many-a=**ho**
2SM-people 2SM-PST-know-textscFV=LOC
'People knew (recognized) (it) there (the place).'

d. ** wa-ndu wa-le-ku-many-a=**ho**
2SM-people 2-PST-17-know-FV=LOC
'People knew (recognized) there (the place).'

The examples in (21) indicate that locative enclitics in Kiwoso can optionally co-occur with the corresponding lexical object noun *duken* 'at the shop' (21b), but not with the locative object agreement prefix *ku-* (21d). This implies that the prefix *ku-* is an object agreement marker and the post-verbal locative enclitics =*ho/=u* in Kiwoso are in complementary distribution. However, both enclitics, =*ho* and =*u* can co-occur with the lexical locative subject as well as locative subject agreement marker, as evidenced in (21a) and (21c). In (21a), the enclitic =*ho* is

an anaphoric locative agreement marker, whereas in (21c), it functions as a true locative object.

In the majority of Bantu languages, locative enclitics, when present, correspond to the three locative noun classes, 16, 17 and 18. For example, the three locative enclitics *=vo*, *=wo*, and *=mo* in Cuwabo originate from the three locative noun prefixes *va-*, *o-*, and *mu-*, respectively (see Guérois 2017: 5). Additionally, Gunnink (2017: 3) reports that in Fwe, the verbal locative enclitics *=ho*, *=ko*, and *=mo* correspond to the locative noun classes 16, 17 and 18, respectively. However, unlike Cuwabo and Fwe where enclitics are derived from demonstrative forms of different locative noun classes, it is not easy to ascertain the origin of the two locative enclitics (*=ho* and *=u*) in Kiwoso because the language lacks the locative prefixes and the demonstrative forms of the two available locative noun classes (16 and 17) do not correspond the locative enclitics identified in Kiwoso. As mentioned earlier, locative nouns in Kiwoso are derived through a suffix *-(i)n* and class 17 prefix *ku-* is only productive in agreement (see §4.2).

5 Summary and conclusion

This chapter has examined the morphosyntax of locative expressions in Kiwoso. It has shown that locative expressions in Kiwoso are achieved by means of a locative suffix -(*i*)*n*. The data presented show that although the three traditionally recognized locative noun class prefixes are not productive in Kiwoso, the locative agreement prefix *ku-* is consistently used with all nominal and verbal modifiers. Although the locative noun class prefixes in Kiwoso are unproductive, the language has maintained some features of the locative system common to Bantu languages, as evidenced in both the nominal and verbal morphology. The analysis offered in this chapter indicates that Kiwoso further has two post-verbal locative enclitics which function as locative arguments.

The use of the locative suffix -(*i*)*n* can be viewed as an innovation to compensate for the lost locative prefixes in the language. The chapter has also looked at the forms of locative expressions in Kiwoso and some of their syntactic properties. It would be interesting to conduct further research on the post-final locative enclitics so as to establish their different forms, origin, and their broader functions other than as locative arguments.

Abbreviations

ADJ	adjective marker	PERF	perfective
ASP	aspect	POSS	possessive marker
CONSC	consecutive	PRN	pronoun
DEM1	demonstrative of the first series	PRO	pronominal
DEM2	demonstrative of the second series	PST	past tense
DEM3	demonstrative to the second series	SM	subject marker
FV	final vowel	STAT	stative
INIT	Initial element	1, 2, 3…	noun classes 1, 2, 3…
LOC	locative	*	Proto-Bantu
OM	object marker	(**…)	unacceptable sentence

References

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.

Carstens, Vicki. 1997. Empty nouns in Bantu locatives. *Linguistic Review* 14. 361–410.

Creissels, Denis. 2011. Tswana locatives and their status in inversion construction. *Africana Linguistica* 17. 33–52.

Demuth, Katherine & Sheila Mmusi. 1997. Presentational focus and thematic structure in comparative Bantu. *Journal of African Languages and Linguistics* 18(1). 1–19. DOI: 10.1515/jall.1997.18.1.1.

Fleisch, Axel. 2005. A cognitive semantic approach to the linguistic construal of UPPER SPACE in Southern Ndebele. *Southern African Linguistics and Applied Language Studies* 23(2). 139–154.

Gibson, Hannah. 2012. *Auxiliary placement in Rangi: A Dynamic Syntax perspective*. London: SOAS, University of London. (Doctoral dissertation).

Grégoire, Claire. 1975. *Les locatifs en bantou*. Tervuren: Musée Royal de l'Afrique Centrale.

Guérois, Rozenn. 2016. The locative system in Cuwabo and Makhuwa (P30 Bantu languages). *Linguistique et Langues Africaines* 2. 43–75.

Guérois, Rozenn. 2017. Locative enclitics in Cuwabo. *Africana Linguistica* 23. 86–116.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list*. Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.

Güldemann, Tom. 1999. Toward a grammaticalization and typological account of the ka- possessive in southern Nguni. *Journal of African Languages and Linguistics* 2(2). 157–184.

Gunnink, Hilde. 2017. Locative clitics in Fwe. *Africana Linguistica* 23. 120–135.

Kagaya, Ryohei & Rogati Olomi. 2009. *A Kiwoso vocabulary*. Tokyo: Research Institute for Languages, Culture of Asia & Africa (ILCAA), Tokyo University of Foreign Studies.

Katamba, Francis X. 2003. Bantu nominal morphology. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 103–120. London: Routledge.

Kioko, Angelina Nduku. 2005. *Theoretical issues in the grammar of Kikamba*. München: Lincom Europa.

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

Maho, Jouni Filip. 1999. *A comparative study of Bantu noun classes*. Gothenburg: Acta Universitatis Gothoburgensis.

Maho, Jouni Filip. 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 (20 March, 2013).

Mallya, Aurelia. 2011. *Locative expressions in Kiwoso*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Mallya, Aurelia. 2016. *Argument realization, causation and event semantics in Kiwoso*. Sellenbosch: Stellenbosch University. (Doctoral dissertation).

Mallya, Aurelia. 2020. Locative-subject alternation constructions in Kiwoso. *Ghana Journal of Linguistics* 9(2). 1–21.

Marten, Lutz. 2010. The great siSwati locative shift. In Anne Breitbarth, Christopher Lucas, Sheila Watts & David Willis (eds.), *Continuity and change in grammar*, 249–267. Amsterdam: Benjamins.

Marten, Lutz. 2012. Agreement in locative phrases in Luganda. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics (Cologne 2009)*, 433–443. Cologne: Köppe.

Marten, Lutz & Hannah Gibson. 2016. Structure building and thematic constraints in Bantu inversion constructions. *Journal of Linguistics* 52(3). 565–607.

Marten, Lutz, Nancy Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105. 253–338.

Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318–368.

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.
Meinhof, Carl. 1941/42. Entstehung und Gebrauch der Lokativendung in Bantusprachen. *Zeitschrift für Eingeborenen-Sprachen* 32. 161–164.
Moshi, Lioba. 1995. Locatives in Kivunjo-Chaga. In Akinbiyi M. Akinlabí (ed.), *Theoretical approaches to African linguistics*, 129–145. Trenton NJ: Africa World Press.
Mugane, John M. 1997. *A paradigmatic grammar of Gĩkũyũ*. Stanford, CA: CSLI (Center for the Study of Language & Information) publications.
Mushi, C. 2005. *Verbal derivational in Kikibosho*. Dar es Salaam: University of Dar es Salaam. (MA thesis).
Persohn, Bastian & Maud Devos. 2017. Post-final locatives in Bantu: Axes of variation and non-locative functions. *Africana Linguistics* 23. 3–46.
Petzell, Malin. 2008. *The Kagulu language of Tanzania: Grammar, texts and vocabulary*. Köln: Rüdiger Köppe.
Riedel, Kristina. 2010. Relative clauses in Haya. *ZAS Papers in Linguistics* 53. 211–225. DOI: 10.21248/zaspil.53.2010.399.
Riedel, Kristina & Lutz Marten. 2012. Locative object marking and the argument-adjunct distinction. *South African Linguistics and Applied Language Studies* 30(2). 277–292.
Rugemalira, Josephat M. 2004. Locative arguments in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics*, 285–295. Köln: Rüdger Köppe.
Samson, Ridder & Thilo C. Schadeberg. 1994. Kiinimacho cha mahali: Kiambishi tamati cha mahali -ni. *Afrikanistische Arbeitspapiere* 37. 127–138.
Van de Velde, Mark. 2019. Nominal morphology and syntax. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 237–269. London & New York: Routledge.
van der Spuy, Andrew. 2014. The morphology of the Zulu locative. *Transactions of the Philological Society* 112(1). 61–79.
Zeller, Jochen & Jean Paul Ngoboka. 2018. Agreement with locatives in Kinyarwanda: A comparative analysis. *African Languages and Linguistics* 39(1). 65–106.

Chapter 6

A comparative study of the locative system in South-Tanzanian Bantu languages

Gastor Mapunda[a] & Fabiola Hassan[b]
[a]University of Dar es Salaam [b]University of Dodoma

The paper presents a comparative analysis of locative expressions in four South-Tanzanian Bantu languages, namely Bena (G60), Ngoni (N12), Yao (P21), and Makhuwa (P31). We more particularly explore the locative marking strategies within noun phrases, the issue of locative agreement, and locative inversion constructions. The article pursues two objectives: (1) To describe the form of locative affixes in each of the four languages, and (2) to establish resemblances and dissimilarities between four neighbouring languages spoken in the south of Tanzania. The findings show that, although the locative systems of the four sampled languages are overall very similar, Makhuwa still exhibits a few divergent features.

1 Introduction

In Eastern Bantu Languages, locative expressions have received enormous attention from various scholars (cf., among others, Harries 1965, Rugemalira 2004, Buell 2007, Marten 2012, Barlew 2013, Marten & van der Wal 2014, Guérois 2016, Zeller forthcoming). These different studies give insight into the high degree of variation of Bantu locatives.

The present article aims to show how Bena, Ngoni, Yao, and Makhuwa, four Eastern Bantu languages, vary in the expression of their locative noun phrases and locative clauses. The languages mentioned above have been selected because they represent different language groups that are found in the Eastern Bantu area, and are geographically close. Additionally, these languages are familiar to

Gastor Mapunda & Fabiola Hassan. 2024. A comparative study of the locative system in South-Tanzanian Bantu languages. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 129–152. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663771

the authors of this chapter. Bena is an Eastern Bantu language spoken in the Southern Highlands of Tanzania, mostly in Njombe District. It is also spoken in the north-western part of Songea District, the north-eastern part of Mbeya District, the southern part of Mufindi District, and the south-western part of Ulanga District. Guthrie (1971) classifies Bena under zone (G60) together with Ki-Hehe, Shi-Sango, Ki-Kinga, Ki-Kisi, and Ki-Wanji. Chaula (1989: 115) identifies seven main dialects of Bena, namely Lupembe, Masakati, Sovi, Maswamu, Mavemba, Ilembula, and Ulanga. Makhuwa (P31) is spoken in the north of Mozambique (Cabo Deldago, Nampula, Niassa, and Zambézia provinces), in Malawi (Mulanje and Tyholo), and in the southern part of Tanzania (Kröger 2005). In Tanzania, the principal regions where Makhuwa speakers live are Mtwara, Lindi, Morogoro, and the Coast. Ismail (2000) lists no less than twelve dialects,[1] most of which are located in Mozambique. This article focuses on the Imithupi dialect spoken in Tanzania, next to the three other languages analyzed in this chapter. Yao (P21) is spoken in Malawi, Mozambique, Tanzania, Zambia, and Zimbabwe. The current article uses data from Yao spoken in Tanzania, specifically in Masasi and Tunduru Districts. Finally, the Tanzanian Ngoni (N12) has four dialects, namely Maposeni-Peramiho, Likonde-Kigonserat, Matimira, and Rwanda. The data for the current study are based on the Maposeni-Peramiho dialect, which is also the best known (Mapunda 2015, Ngonyani 2003). These four variants analyzed in this chapter are spoken in Tanzania, as shown in Map 1 below. The four variants are in close geographical proximity, and the speakers of these dialects understand each other well.

Specifically, the article describes and is structured along the following lines, which resume some of the morphosyntactic parameters proposed by Guérois et al. (2017): i) what are the formal strategies of locative marking on nouns? (§2); ii) how does locative agreement operate within NPs and VPs? (§3); iii) are locative inversion constructions attested? (§4). As a general result, the paper shows how Makhuwa tends to behave differently from the other three languages.

The primary data used in this study were obtained through interviews with two adult consultants from each language. More specifically, consultants were prompted to translate Swahili sentences into their language. Then, the translations were cross-checked for consistence. Follow-up questions were also asked when additional information was required.

[1] As we have not engaged with a comparative study of these twelve dialects, we cannot comment on their similarities and differences.

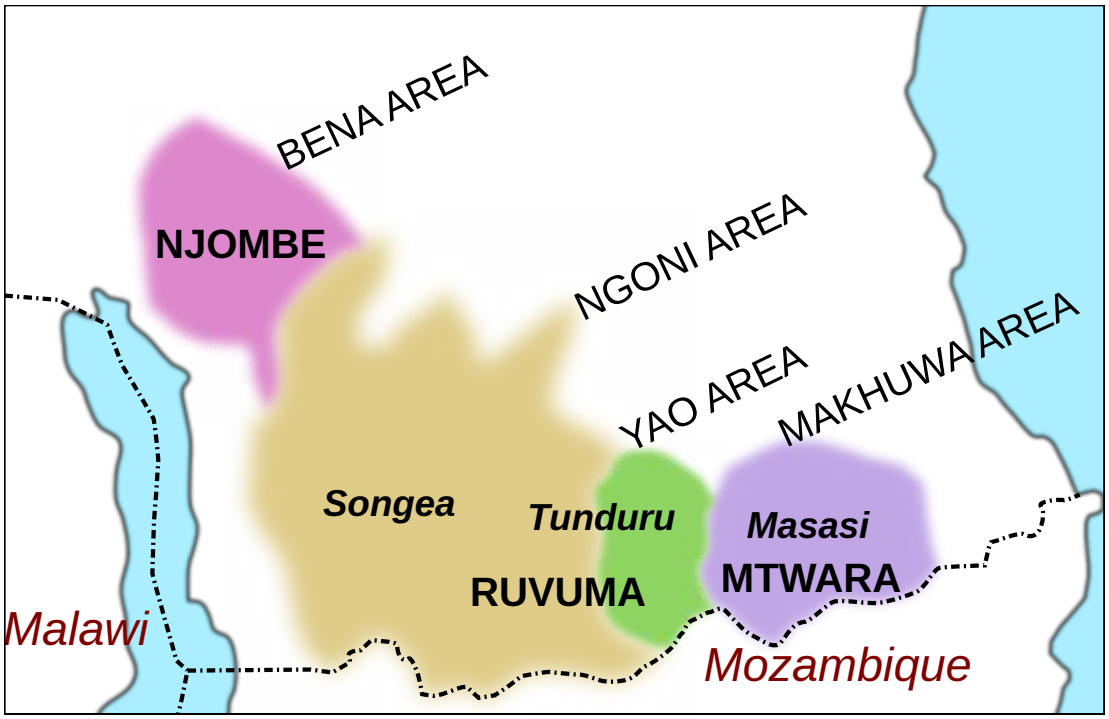


Figure 1: Approximate locations of the Bena, Ngoni, Yao and Makhuwa speaking areas in Tanzania (map produced by Sebastian Nordhoff based on work by G. Mapunda and F. Hassan)

2 Locative marking strategies

In many Bantu languages location is marked by nominal prefixation (Rugemalira 2004). Four locative prefixes were reconstructed to Proto-Bantu: class 16 *pa-, class 17 *ku-, class 18 *mu-, and class 23/25 *i/e-[2] (Bleek 1862–1869, Guthrie 1948, 1967–1971, Meeussen 1967). In semantics terms, **pa-* means nearness, adjacency, definiteness, specificity, limitedness or known location; **ku-* implies remoteness, farness, unspecificity, generalness, unlimitedness, not necessarily known or direction/towardness; and **mu-* denotes withinness, interiority or enclosed location. Whilst many Bantu languages have retained the first three historical locative prefixes, e.g. Shona [S10] illustrated in (1), other languages have retained two affixes, e.g. Vunjo [E62] (Mcha 1979) and yet others have retained only one class, e.g. Kiwoso [E621d] (Mallya 2011).

[2]This locative affix is attested in very few Bantu languages. Ganda (JE15 Uganda) is an example, e.g. *e-Kampala* 'in Kampala'.

(1) class 16 ***pa***-*imba* 'at the house'
class 17 ***ku***-*imba* 'to the house'
class 18 ***mu***-*imba* 'in the house'

A second strategy to expression locative consists in suffixing *-(i)ni ~ -n ~ -eng* to the end of the noun stem. Examples in (2) illustrate this strategy in several Eastern languages.

(2) a. Swahili [G42] *nyumba-**ni*** 'at/to/in the house'
b. Sesotho [S33] *thab-**eng*** 'to/on the mountain' (Machobane 1995: 120)
c. Chaga [E30] *ruko-**nyi*** 'at/to/in the kitchen' (Moshi 1995: 131)

Note that the loss of locative morphology in these languages is restricted to noun class prefixes. As will be made clear in §2, noun modifiers and verb forms controlled by a locative head noun necessarily host locative agreement markers (Mpiranya 2015).

Double affixation, which involves both a locative noun class prefix and a locative suffix, is a third strategy very rarely attested across Bantu. As far as we know, only P30 languages productively exhibit double affixation,[3] as illustrated in (3) with Cuwabo [P34] and Makhuwa-Enahara [P31] which both combine class 17 prefix *o-* and locative suffix *-ni*.

(3) a. Cuwabo ***o***-*ma-básá-**ni*** (cl.17) 'at work'
b. Makhuwa-Enahara ***o***-*n-tékô-**ni*** (cl.17) 'at work' (Guérois 2016: 51)

Guérois (2016) suggests that whilst locative prefixes were inherited, the suffixation of *-ni* is a later innovation resulting from a contact situation with Swahili.

Finally, it should be noted that names of places or cities do not commonly host locative marking. In Swahili, for instance, cities like Tokyo, London, or Paris are not modified when used locatively (Mkude 2005: 153). Cities from Tanzania also do not host locative markers (e.g. Arusha, Dar es Salaam, Dodoma, Mbeya).

Two of the strategies above are attested in our sample of languages, namely prefixation (for Bena, Ngoni, and Yao) and double affixation (for Makhuwa), as shown in (4).[4]

[3] In Swati [S43], some locative nouns necessarily combine the class 25 locative prefix *e-* and the locative suffix *-ini*, e.g. *e-ndl-ini* 'at/to/in the house' (Marten 2010: 254), but this double affixation is restricted to a few nouns only.

[4] N is a homorganic nasal, i.e. its surface realization depends on its phonetic environment, such as N > [m] / _ bilabial C and N > [ŋ] / _ velar C.

(4)

	Bena	Ngoni	Yao	Makhuwa-I.	
cl.16	***pa****-kaye*	***pa****-nyumba*	***pa****-nyumba*	***va****-nupa-****ni***	‘at the house’
cl.17	***ku****-kaye*	***ku****-nyumba*	***ku****-nyumba*	***u****-nupa-****ni***	‘to the house’
cl.18	***mu****-kaye*	***mu****-nyumba*	***n****-nyumba*	***n****-nupa-****ni***	‘in the house’

In each sampled language, locative prefixes are additive, i.e. they are added to the stem which has an inherent noun class prefix. For instance, in Makhuwa, the locative prefixes *va-*, *u-* and *N-* can be added to the noun stem *mwiri* ‘tree’ which has an inherent noun class 3 prefix *mw-* as seen in example (5).

(5) cl.16 ***va****-mw-iri-****ni*** ‘at the tree’ [Makhuwa-I.]
cl.17 ***u****-mw-iri-****ni*** ‘to the tree’
cl.18 ***m****-mw-iri-****ni*** ‘in the tree’

Looking back at Table 2, we see that Makhuwa-Imithupi behaves as Makhuwa-Enahara illustrated in (3), i.e. locative nouns are by default built upon the combination of a locative prefix (class 16 *va-*, class 17 *u-*, and class 18 *N-*) and the locative suffix *-ni*. Only a few nouns deviate from this pattern by not admiting the locative suffix. These are lexicalized locatives such as *vachula* ‘at the top’ and *uchulu* ‘to the top’, and proper locative nouns such as names of towns, countries and continents, as in (6).

(6) Makhuwa

a. i. ***u****-Dar es salaam*
17-Dar es Salaam
‘to Dar es Salaam’

ii. * ***u****-Dar es salaam-****ni***
17-Dar es Salaam-LOC

b. i. ***u****-Tanzania*
17-Tanzania
‘to Tanzania’

ii. * ***u****-Tanzania-****ni***
17-Tanzania-LOC

On the other hand, mere locative prefixation with common nouns seems to be strictly prohibited in Makhuwa-Imithupi, as shown in (7). Both nouns *patsári* ‘market’ and *matta* ‘field’ are made locative by double affixation, i.e. by one of the three locative prefixes and by the locative suffix *-ni* (*u-patsári-ni* ‘to the market’ and *m-matta-ni* ‘in the field’). In this respect, it is worth noting a dialectal difference with Makhuwa-Enahara, whereby certain common nouns may be marked

for locative uniquely through prefixation (8). Some others freely add the locative suffix *-ni* (8).

(7) Makhuwa-Imithupi

a. * ***u****-patsári*
17-9.market
'to the market'

b. * ***m****-matta*
18-field
'in the field'

(8) Makhuwa-Enahara (Guérois 2016: 53–54)

a. ***o****-patsári*
17-market
'at/to the market'

b. ***m****-mátta(-****ni****)*
18-field-LOC
'in the field'

If the locative suffix *-ni* is present in certain locative expressions in Bena, Ngoni, and Yao, its use is not productive at all. In these three languages, locative NPs are expressed through prefixation only. In Ngoni, for instance, *-ni* is present in two specific contexts, i.e. in lexicalized locative NPs (9)[5] and in borrowed locative NPs (10). Since *-ni* is not segmentable in these words, it does not convey any locative meaning, hence locative prefixation is still needed.

(9) *mfuleni* 'well' *pa-mfuleni* 'at the well' [Ngoni]
bomani 'town' *pa-bomani* 'to/in (?) town'

(10) *m-ji****ni*** 'town' < Swahili *mjini* [Ngoni]

pa*-m-ji****ni*** (cl.16)	'at the town'
ku*-m-ji****ni*** (cl.17)	'to the town'
mu*-m-ji****ni*** (cl.18)	'in the town'

In Yao, the locative suffix *-ni* may optionally be added on loan words (11), reminiscent of equivalent Makhuwa double affixed locative NPs *umsikitini* 'to the mosque', *ukanisani* 'to the church', *ushuleni* 'to school' and *umahakamani* 'to the court'. Different from Yao, in Bena the locative marker *-ni* cannot be added (12).

[5]However, it cannot be excluded that these lexicalized locative NPs are originally loans from Swahili where locative is marked by suffixation.

(11) Yao

msikiti	'mosque' (cl.3)	→	***pa**-m-sikiti(-**ni**)*	'at the mosque' (cl.16)
kanisa	'church' (cl.5)	→	***ku**-kanisa(-**ni**)*	'to the church' (cl.17)
shule	'school' (cl.9)	→	***ku**-shule(-**ni**)*	'to the school' (cl.18)
mahakama	'court' (cl.9)	→	***mu**-mahakama(-**ni**)*	'in the court' (cl.18)

(12) Bena

msikiti	'mosque' (cl.3)	→	***pa**-m-sikiti(*-ni)*	'at the mosque' (cl.16)
kanisa	'church' (cl.5)	→	***ku**-kanisa(*-ni)*	'to the church' (cl.17)
sule	'school' (cl.9)	→	***ku**-sule(*-ni)*	'to the school' (cl.18)
mahakama	'court' (cl.9)	→	***mu**-mahakama(*-ni)*	'in the court' (cl.18)

Table 1 shows a summary of locative marking in Bena, Ngoni, Yao, and Makhuwa.

Table 1: Locative marking in Bena, Ngoni, Yao, and Makhuwa

locative marking strategies	prefixation only	suffixation only	prefixation + suffixation
Bena Ngoni Yao	yes	no	yes in (Swahili) loans
Makhuwa-I.	restricted to lexicalized locative nouns + proper geographical names	no	yes by default

As seen in Table 1, the locative marking strategies attested in the selected languages are prefixation and double affixation. The sampled languages differ from other Eastern Bantu languages spoken in Tanzania such as Swahili and Chagga, whose nouns become locative via suffixation only.

3 Locative agreement

Several linguists have discussed locative agreement systems in Bantu languages (e.g., among others, Stucky 1976, Harford 1983, Kahigi 2005, Marten 2012, Ngunga & Mpofu-Hamadziripi 2013). Agreement occurs: i) within locative NPs, between the locative head noun and its modifiers; ii) within clauses, between the locative head noun and its dependent verb. The two types of agreement are discussed in the two following subsections.

3.1 Agreement within NPs

Locative agreement is a morphosyntactic process whereby the dependent elements in the locative NP agree with the locative. Noun dependents here involve possessives, associatives, adjectives, and demonstratives. They are commonly referred to as modifiers. Agreement-wise, languages show a three-way distinction (e.g. Marten 2012, Machobane 1995). Firstly, there are languages with an inner agreement system, whereby the inherent noun class prefix of a noun controls the agreement between the locative head and its dependents. This is shown in (13) with Runyambo [JE21], where the first person singular possessive stem *nje* agrees in noun class with *citabo* 'book', i.e. class 7. Secondly, there are languages with an outer agreement system, whereby noun modifiers receive locative agreement prefixes. In (14), the Swahili first person singular possessive stem *angu* takes class 18 agreement to express withinness. Thirdly, there are languages which exhibit both outer and inner agreement systems. In these languages, the inherent noun class prefix of a noun or the locative prefix controls the agreement between the locative head and its dependents. In Tshiluba [L31], demonstratives modifying locative nouns may agree either with the leftward locative prefix (15a) or with the inherent noun prefix (15b).

(13) o-mu-**ci**-tabo **ca**-nje
AUG-18-7-book 7-my
'In my book' (Runyambo, Rugemalira 2004: 6)

(14) chumba-ni **mw**-angu
7.room-LOC 18-POSS.1SG
'In my room' (Swahili, Mkude 2005: 154)

(15) a. **mu**-di-kopu e-**mu** mu-di mu-tooke
18-5-cup DEM-18 SM18-be 18-clean
'This cup is clean inside'

b. mu-**di**-kopu e-**di** mu-di mu-tooke
18-5-cup DEM-5 SM18-be 18-clean
'The space inside this cup is clean' (Tshiluba, Stucky 1976: 180)

Based on this typology, our data show that Bena and Yao both have outer and inner types of agreement on all types of modifiers, namely adjectives, connectives, demonstratives, and possessives, just like Tshiluba in (15). Bena data are provided in Table 2 and Yao data in Table 3. The difference in meaning is not entirely clear-cut, but it seems that outer agreement gives more emphasis on

the locative aspect of the event, i.e. it relates to a place and not somewhere else. On the other hand, inner agreement gives more importance to the modifier as such. For instance, *pakaye inofu* 'to a good house' in Bena, provided in Table 2, underlies the fact that the house is good (and not bad).

Table 2: Outer and inner agreement in Bena

	Outer AGR	Inner AGR
ahele ... 'he has gone'	**pa**-kaye **pa**-nofu 16-9.house 16-good 'to a good house'	pa-**kaye** **i**-nofu 16-9.house 9-good 'to the place where the house is good'
	pa-kaye **pa** vaanu 16-9.house 16.CON 2.people 'to the people's house'	pa-**kaye** **ja** vaanu 16-9.house 9.CON 2.people 'to the place where the house is of the people'
	pa-kaye **pa**-la 16-9.house 16.DEM.II 'to that house'	pa-**kaye** **i**-la 'to the 16-9.house 9.DEM.II place of that house'
	pa-kaye **pa**-angu 16-9.house 16-POSS.1SG 'to my house'	pa-**kaye** **ya**-angu 16-9.house 9-POSS.1SG 'to my house'

Ngoni resembles Runyambo (illustrated in (13) above): outer and inner types of agreement are only attested between the locative head noun and demonstratives. The other modifiers (adjectives, connectives, possessives) may only receive inner agreement, whereas outer agreement is ungrammatical, as seen in Table 4.

From our sample, Makhuwa differs the most, as it only displays outer agreement (as in Swahili in (14) above). This is illustrated in Table 5.

Table 3: Outer and inner agreement in Yao

	Outer AGR	Inner AGR
ajawile... 'he has gone'	**pa**-nyumba **pa**-ambone 16-9.house 16-good 'to the good house'	pa-**nyumba ja**-ambone 16-9.house 9-good 'to the place where the house is good'
	pa-nyumba **pa** vandu 16-9.house 16.CON 2.people 'to the people's house'	pa-**nyumba ja** vandu 16-9.house 9.CON 2.people 'to the place where the house is of the people'
	pa-nyumba a-**pa**-la 16-9.house AUG-16-DEM.III 'to that house'	pa-**nyumba** a-**ja**-la 16-9.house AUG-9-DEM.III 'to the place of that house'
	pa-nyumba **pa**-angu 16-9.house 16-POSS.1SG 'to my house'	pa-**nyumba ja**-angu 16-9.house 9-POSS.1SG 'to the place of my house'

Table 4: Inner agreement in Ngoni

ahambi... 'he has gone'	Outer AGR	Inner AGR
'to a good house'	***pa**-nyumba **pa**-bwina 16-9.house 16-good	pa-**nyumba** ya-bwina 16-9.house 9-good
'to the house of the people'	***pa**-nyumba 16-9.house **pa**-vanu 16.CON-2.people	pa-**nyumba ya** vanu 16-9.house 9.CON 2.people
'to that house'	**pa**-nyumba **pa**-la 16-9.house 9-DEM.III	pa-**nyumba** y**i**-la 16-9.house 9-DEM.III
'to my house'	***pa**-nyumba ***pa**-angu 16-9.house 16-POSS.1SG	pa-**nyumba ya**-angu 16-9.house 9-POSS.1SG

Table 5: Inner agreement in Makhuwa-Imithupi

ahorwa... 'he has gone'	Outer AGR	Inner AGR
to a good house	**va**-i-nupa-ni **va**-orera 16-9-house-loc 16-good	*va-i-nupa-ni *y-orera 16-9-house-loc 9-good
to the house of the people	**va**-i-nupa-ni 16-9-house-loc **va**-atu 16.CON-people	*va-i-nupa-ni *y-atu 16-9.house-loc 9.CON-people
to that house	**va**-i-nupa-ni **va**-le 16-9-house 9-DEM.III	*va-**i**-nupa-ni *i-le 16-9-house 9-DEM.III
to my house	**va**-i-nupa-ni **va**-aka 16-9-house 16-POSS.1SG	*va-**i**-nupa-ni *y-aka 16-9-house 9-POSS.1SG

3.2 Agremeent within VPs

Within VPs, locative indexation on the verb usually involves subject, object and relative prefixation as well as locative cliticization. Locative verbal enclitics are not attested in the selected languages. Therefore, in this chapter, we only discuss locative subject and object prefixation. In the four sampled languages, locative subject prefixes exist for the three historical locative classes. Class 16 *pa-* (or variant *va-* in Makhuwa-I.) is illustrated in (16). More examples of locative verbal agreement are described in §4 on locative inversion construction.

(16) Class 16 locative subject prefixes

Bena	**pa**-i-nung-a	a-ma-futa	**pa-kaye**
	SM16-PRS-smell-FV	AUG-6-oil	16-9.house
Ngoni	**pa**-gi-nung'-a	ma-huta	**pa-nyumba**
	SM16-smell-FV	6-oil	16-9.house
Yao	**pa**-ku-nung-a	ma-huta	**pa-musi**
	SM16-PRS-smell-FV	6-oil	16-9.house
Makhuwa	**va**-no-nukh-a	ma-khura	**va-nupa-ni**
	SM16-PRS-smell-FV	6-oil	16-9.house- LOC

'It smells oil at the house.'

Locative object marking is also attested in our sample, except in Makhuwa where object marking is restricted to classes 1 and 2 (17).

(17) Class 16 locative object prefixes

Bena	u-mw-ana	a-ku-**pa**-nogw-a	**pa-sule**
	AUG-1-child	SM1-PRS-OM16-like-FV	16-school
Ngoni	mw-ana	a-**pa**-gan-i	**pa-shuli**
	1-child	SM1-OM16-like-FV	16-school
Yao	mw-anache	a-ku-**pa**-sak-a	**pa-shule**
	1-child	SM1-PRS-OM16-like-FV	16-school
Makhuwa	*mw-ana	a-no-**va**-tun-a	**va-shule-ni**
	1-child	SM1-PRS-OM16-like-FV	16-9.school-LOC

'The child likes school.' (lit. 'The child likes there at the school.')

3.3 Summary

Table 6 summarizes the locative agreement system as found in each sampled language. As can be seen, Bena and Yao behave alike: both languages allow locative inner and outer agreement within NPs and both have locative subject and object verbal markers. Ngoni is very similar, except for outer agreement which is restricted to the demonstratives, whereas it is observed with all modifiers in Bena and Yao. Makhuwa, in turn, differs from the other three languages in two respects: first it prohibits inner agreement, second it does not have locative object markers.

4 Locative inversion constructions

Locative inversion (LI) is part of those inversion constructions whereby a logical subject, i.e. the highest thematic role selected by the verb, occupies a postverbal position and the locative phrase is raised to the preverbal position where it grammatically behaves like a regular subject, i.e. it controls agreement on the verb. This change in word order is often motivated by information-structural considerations (Marten & van der Wal 2014, Hamlaoui 2014). Two types of LI are traditionally distinguished (Buell 2007): formal agreeing LI and semantic agreeing LI. The former relies on locative morphology and implies that languages have maintained a productive locative system. This is the case in Chewa as shown in (18) where the verb *li* 'be' agrees with the preverbal locative phrase *kumudzi* 'to the village'. Semantic agreeing LI, in turn, involves nouns which are inherently locative without any additional locative marking. This is illustrated in Zulu (19)

Table 6: Overview of locative agreement systems

Agreement system	Bena	Ngoni	Yao	Makhuwa
within NPs				
inner agreement				
with adjectives	✓	✓	✓	✗
with connectives	✓	✓	✓	✗
with demonstratives	✓	✓	✓	✗
with possessives	✓	✓	✓	✗
outer agreement				
with adjectives	✓	✗	✓	✓
with connectives	✓	✗	✓	✓
with demonstratives	✓	✓	✓	✓
with possessives	✓	✗	✓	✓
within VPs				
locative subject marker	✓	✓	✓	✓
locative object marker	✓	✓	✓	✗

with *lezi zindlu* '(in) these houses', which triggers subject agreement on *hlala* 'live'.

(18) a. Chi-tsime chi-li **ku-mu-dzi**. [Chewa]
7-well SM7-be 17-3-vilage
'The well is to the village.'

b. **Ku-mu-dzi ku**-li chi-tsime.
17-3-village SM17-be 7-well
'To the village there is a well.' (Salzmann 2005: 5)

(19) a. **Aba-ntu** aba-dala **ba**-hlala **lezi** **zi-ndlu**. [Zulu]
2-people 2-old SM2-live 10.DEM.I 10-house
'Old people live in these houses.'

b. **Lezi** **zi-ndlu** **zi**-hlala aba-ntu aba-dala.
10-DEM.I 10-house SM10-live 2-people 2-old
'(In) these houses live old people.' (Buell 2007: 107–108)

More recently, Guérois (2014) shows that both locative LI and semantic LI exist in Cuwabo [P34]. Other languages such as Olutsootso [JE32b] and Swahili

(Marten & van der Wal 2014), and Kinyarwanda [JD61] (Ngoboka 2016) show the same feature. Our collected data show no evidence of semantic LI constructions; only formal LI is attested in the 4 sampled languages, in accordance with the most common Bantu pattern (Marten & van der Wal 2014). An example of each language is provided below.

(20) a. a-ma-futa ma-gi-nung'-a mu-shumba [Bena]
AUG-6-oil SM6-PRS-smell-FV 18-7.room
'Oil is smelling in the room'

b. **mu-shumba mu**-gi-nung-a a-ma-futa
18-7.room SM18-PRS-smell-FV AUG-6-oil
'In the room is smelling oil'

(21) a. ma-huta ma-gi-nung'-a mu-chumba [Ngoni]
6-oil SM6-PRS-smell-FV 18-7.room
'Oil is smelling in the room'

b. **mu-chumba mu**-gi-nung'-a ma-huta
18-7.room SM17-PRS-smell-FV 6-oil
'In the room is smelling oil'

(22) a. ma-huta ma-ku-nung-a mu-ch-umba [Yao]
6-oil 6-PRS-smell-FV 18-7-room
'Oil is smelling in the room'

b. **mu-chumba mu**-ku-nung-a ma-huta
18-7.room-LOC SM18-PRS-smell-FV 6-oil
'In the room is smelling oil'

(23) a. ma-khura a-no-nukh-a n-ch-umba-ni [Makhuwa]
6-oil 6SM-PRS-smell-FV 18-7-room-loc
'Oil is smelling in the room'

b. **n-chumba-ni n**-no-nukh-a ma-khura
18-7.room-LOC SM18-PRS-smell-FV 6-oil
'In the room is smelling oil'

Other examples with the copula verb *li ~ ri* 'be' are provided in (24)–(27).

(24) **mu-shumba mu**-li mw-ana [Bena]
18-7.room SM18-be 1-child
'In the room there is a child.'

(25) **mu-chumba mu**-wi (na) mw-ana [Ngoni]
18-7.room SM18-be (with) 1-child
'In the room there is (with) a child.'

(26) **mu-nyumba mu**-li mw-anache [Yao]
18-9.house SM18-be 1-child
'In the house there is a child.'

(27) **m-nupa-ni m**-ri mw-ana [Makhuwa]
18-9.house-LOC SM18-be 1-child
'In the house there is a child.'

The preverbal locative phrase behaves, in many ways, just like a regular subject. Like in most Bantu languages, finite verbs in the four sampled languages have an obligatory subject prefix that agrees with the subject NP in noun class. In LI constructions, the subject prefix of the verb obligatorily agrees with the preverbal locative phrase, in one of the three locative noun classes. Such agreement is a clear indicator of the subject status of the fronted locative phrase.

As a grammatical subject and discourse topic, the fronted locative NP may be dropped or may be postponed clause-finally. In both cases, it keeps licensing subject agreement on the verb. This is shown below.

(28) **mu**-li mw-ana (**mu-shumba**) [Bena]
SM18-be 1-child (18-7.room)
'There is a child (in the room).'

(29) **mu**-wi (na) mw-ana (**mu-chumba**) [Ngoni]
SM18-be (with) 1-child (18-7.room)
'There is (with) a child (in the room).'

(30) **mu-li** mw-anache (**mu-nyumba**) [Yao]
SM18-be 1-child (18-9.house)
'There is a child (in the room).'

(31) **m-ri** mw-ana (**m-nupa-ni**) [Makhuwa]
SM18-be 1-child (18-9.house-LOC)
'There is a child (in the room).'

On the other hand, the inverted subject appears immediately after the verb, i.e. the object position, but maintains a thematic role of subject. Its presence is mandatory. Omitting the inverted subject would make the sentence ungrammatical, as seen in the examples below.

(32) * mu-shumba mu-li [Bena]
18-7.room SM18-be
lit. 'In the room there is.'

(33) * mu-chumba mu-wi(na) [Ngoni]
18-7.room SM18-be
lit. 'In the room there is.'

(34) * mu-nyumba mu-li [Yao]
18-9.house SM18-be
lit. 'In the house there is.'

(35) * m-nupa-ni m-ri [Makhuwa]
18-9.house-LOC SM18-be
lit. 'In the house there is.'

Despite its postverbal object position, the inverted subject does not really behave as an object. First, it cannot be object-marked on the verb as seen in examples (36)–(39).

(36) * mu-sh-umba mu-i-**ma**-nung-a **a-ma-futa** [Bena]
18-7-room SM18-PRS-OM6-smell-FV AUG-6-oil
'In the room is smelling it, oil'

(37) * mu-ch-umba mw-i-**mu**-nung'-a **ma-huta** [Ngoni]
18-7-room SM18-PRS-OM6-smell-FV 6-oil
'In the room is smelling it, oil'

(38) * mu-chumba mu-**ma**-kunung-a **ma-huta** [Yao]
18-7.room SM18-OM6-smell-FV 6-oil
'In the room is smelling it, oil'

(39) * m-chumba-ni m-no-**mw**-unl-a **mw-ana** [Makhuwa]
18-7.room-LOC SM18-PRS-OM6-smell-FV 1-child
'In the room is crying him, the child'

Second, the logical subject cannot be passivized, as seen in (40)–(43).

(40) * **ma-futa** ma-i-nung-**w**-a (ni mu-ki-yumba) [Bena]
6-oil SM6-PRS-smell-PASS-FV (by 18-7-room)
'Oil is smelled (in the room)'

(41) * **ma-huta** ma-inung'-iw-a (ni mu-chumba) [Ngoni]
6-oil SM6-smell-PASS-FV (by 18-7.room)
'Oil is smelled (in the room)'

(42) * **ma-huta** ma-kungung-**w**-a (ni mu-chumba) [Yao]
6-oil SM6-smell-PASS-FV (by 18-7.room)
'Oil is smelled (in the room)'

(43) * **ma-khura** ma-no-nukh-**iy**-a (ni m-chumba-ni) [Makhuwa]
6-oil SM6-PRS-smell-PASS-FV (by 8-7.room-LOC)
'Oil is smelling (in the room)'

Third, the logical subject cannot be extracted by relativization, as seen in (44)–(47).

(44) * ani ye mu-kaye i-vemb-a? [Bena]
who 1.DEM 18-9.house SM1.PRS-cry-FV.REL
'Who is it that is crying in the house?'

(45) * yani mwe mu-nyumba i-vemb-a? [Ngoni]
who 1.DEM 18-9house SM1.PRS-cry-FV.REL
'Who is it that is crying in the house?'

(46) * nduni jwelejo m-nyumba a-ku-lil-a? [Yao]
who 1.DEM 18-9.house SM1.PRS-cry-FV.REL
'Who is it that is crying in the house?'

(47) * mpani yo m-nupa-ni a-no-unl-a? [Makhuwa]
who 1.DEM 18-9.house-LOC SM1-PRS-cry-FV.REL
'Who is it that is crying in the house?'

As noted by Bresnan & Kanerva (1989), the impossibility to object-mark, to passivize and to relativize the postverbal logical subject of a LI construction, suggests that it is not a typical object complement of the verb. Yet, its inflexible immediate-after-the-verb position and its obligatory presence still liken it to a core argument rather than an adjunct.

Argument structures involved in LI may differ. For example, Demuth & Mmusi (1997) argue that in Tswana, LI is possible with active transitive verbs. In contrast, in Chewa, Bresnan & Kanerva (1989) observe that those verbs do not allow LI. In the four sampled languages, LI is possible with unaccusative verbs, i.e. intransitive verbs which take one argument with the semantic role of theme. The verb may in most cases also take a locative argument. Examples of these verbs are

'smell', 'be full', 'spread', and 'germinate'. Examples in (20)–(23) above illustrate the point with the verb 'smell'. However, LI is no longer possible when unaccusative verbs are used in the passive voice. Examples in (48)–(51) illustrate this point.

(48) * ku-sh-umba ku-i-nung'-w-a a-ma-futa (na va-ana) [Bena]
17-7-room SM17-PRS-smell-PASS-FV AUG-6-oil (by 2-child)
'To the room is being smelled the oil (by the children)'

(49) * ku-ch-umba ku-i-nung'-iw-a ma-huta (na va-ana) [Ngoni]
17-7-room SM17-PRS-smell-PASS-FV 6-oil (by 2-child)
'To the room is being smelled the oil (by the children)'

(50) * mu-ch-umba mu-ku-nung-w-a ma-huta (ni va-ana) [Yao]
17-7-room-LOC SM17-PRS-smell-PASS-FV 6-oil (by 2-child)
'In the room is being smelled the oil (by the children)'

(51) * n-ch-umba-ni n-no-nukh-w-a ma-khura (na ashana)
18-7-room-LOC SM18-PRS-smell-PASS-FV 6-oil (by 2.child)
[Makhuwa]
'In the room is being smelled the oil (by the children)'

On the other hand, unergative verbs do not allow LI. Unergative verbs are intransitive verbs that are semantically distinguished by having an agent argument. Examples of these verbs are 'vomit', 'defecate', 'run', and 'cry'. Bena examples in (52) illustrate the point with the verb *vemba* 'cry'.

(52) a. * mu-shumba mu-vemb-a mw-ana [Bena]
18-7.room SM18-cry-FV 1-child
'The child is crying in the room.'

b. * mu-chumba mu-vemb-a mw-ana [Ngoni]
18-7.room SM18-cry-FV 1-child
'The child is crying in the room.'

c. * mu-ki-yumba mu-vemb-a mw-ana [Yao]
18-7-room SM18-cry-FV 1-child
'The child is crying in the room.'

d. * n-chumba-ni n-no-unl-a mw-ana [Makhuwa]
18-7.room-loc SM18-prs-cry-FV 1-child
'The child is crying in the room.'

In the same way, passivised unergative verbs cannot appear in LI. Example in (53) illustrates the point with the verb *vembwa* 'cried by'.

(53) a. * mu-sh-umba mu-vemb-w-a (ni mw-ana) [Bena]
18-7-room SM18-cry-PASS-FV (by 1-child)
'It is being cried in the room (by the child).'

b. * mu-ch-umba mu-vemb-w-a (ni mw-ana) [Ngoni]
18-7-room SM18-cry-PASS-FV (by 1-child)
'It is being cried in the room (by the child).'

c. * mu-ki-yumba mu-vemb-w-a (ni mw-ana) [Yao]
18-7-room SM18-cry-PASS-FV (by 1-child)
'It is being cried in the room (by the child).'

d. * n-ch-umba-ni n-no-unl-w-a (ni mw-ana) [Makhuwa]
18-7-room SM18-prs-cry-PASS-FV (by 1-child)
'It is being cried in the room (by the child).'

Transitive verbs, which add a thematic object to the argument structure, fail to undergo LI. This is expected when the thematic object precedes the inverted subject, as the latter necessarily follows the verb. The order inverted subject-theme is nevertheless just as ungrammatical. Infelicitous examples are provided in (54)–(57) with the verbs 'cultivate' and 'put'.

(54) a. * a-pa-ono pa-limil-e i-ki-tu kuku [Bena]
AUG-16-place SM16-cultivate-PRF AUG-7-thing 1.grandfather
'Grandfather has cultivated something on the place.'

b. * a-pa-ono pa-limil-e kuku i-ki-tu
AUG-16-place SM16-cultivate-PRF 1.grandfather AUG-7-thing
'Grandfather has cultivated something on the place.'

(55) a. * ap-a naha pa-limil-e chi-tu gogu [Ngoni]
16-DEM int SM16-cultivate-PRF 7-thing 1.grandfather
'Grandfather has cultivated something on the place.'

b. * ap-a naha pa-limil-e gogu chi-tu
16-DEM int SM16-cultivate-PRF 1.grandfather 7-thing
'Grandfather has cultivated something on the place.'

(56) a. * pa-m-keka pa-vichil-e chi-ndu baba [Yao]
16-3-mat SM16-put-PRF 7-thing 1.father
'Father has put something on the mat.'

b. * pa-m-keka pa-vichil-e baba chi-ndu
16-3-mat SM16-put-PRF 1.father 7-thing
'Father has put something on the mat.'

(57) a. * va-m-pasa-ni va-ho-wesh-a i-tu athatha [Makhuwa]
16-3-mat-LOC SM16-PRF-put-FV 7-thing 1.father
'Father has put something on the mat.'

b. * va-m-pasa-ni va-ho-wesh-a athatha i-tu
16-3-mat-LOC SM16-PRF-put-FV 1.father 7-thing
'Father has put something on the mat.'

On the other hand, passivized transitive verbs do allow LI. Examples in (58)–(61) illustrate the point with the verb 'being put'.

(58) a-pa-ono pa-limil-w-e i-ki-tu (ni kuku) [Bena]
AUG-16-place SM16-put-PASS-PRF AUG-7-thing (by 1.grandfather)
'Something has been cultivated on the place (by grandfather).'

(59) ap-a naha pa-lim-iw-e i-ki-tu (na gogu) [Ngoni]
16-place int SM16-put-PASS-PRF AUG-7-thing (by 1.grandfather)
'Something has been cultivated on the place (by grandfather).'

(60) pa-m-keka pa-vichil-w-e chi-ndu (ni baba) [Yao]
16-3-mat SM16-put-PASS-PRF 7-thing (by 1.father)
'Something has been put on the mat (by father).'

(61) va-m-pasa-ni va-ho-wesh-iy-a i-tu (ni athatha) [Makhuwa]
16-3-mat-LOC SM16-PRF-put-FV 7-thing (by 1.father)
'Something has been put on the mat (by father).'

Table 7 summarizes the findings for LI. As argued above, the four sampled languages behave alike, both in terms of types of LI allowed (formal versus semantic) and the interaction between LI and the argument structure.

5 Conclusion

This paper has provided a comparative description of the locative system of four South-Tanzanian Bantu languages, namely Bena, Ngoni, Yao and Makhuwa. The study shows that these languages overall exhibit similar locative constructions with similar properties. This is particularly clear with LI constructions, which

Table 7: Locative inversion in Bena, Ngoni, Yao, and Makhuwa

Parameters	Bena	Ngoni	Yao	Makhuwa
Types of LI				
formal agreeing LI	✓	✓	✓	✓
semantic agreeing LI	✗	✗	✗	✗
LI and argument structure				
active unaccusative verb	✓	✓	✓	✓
passive unaccusative verb	✗	✗	✗	✗
active unergative verb	✗	✗	✗	✗
passive unergative verb	✗	✗	✗	✗
active transitive verb	✗	✗	✗	✗
passive transitive verb	✓	✓	✓	✓

show identical properties. Furthermore, the four languages make a productive use of the three historical locative prefixes of class 16, 17 and 18 in both nominal and verbal domains. While Bena and Yao are strictly identical for all properties discussed in this paper, Ngoni differs from the three others in that it does not allow outer agreement within NPs (except with demonstrative modifiers). The most notable differences come from Makhuwa. In this language, in addition to locative prefixation, locative nouns are further marked with a locative suffix *-ni*. The only cases of exception are lexicalized locatives and nouns refering to administrative-geographical entities, such as names of towns or countries. What looks like double affixation in Bena, Ngoni or Yao is attested in loanwords only, especially from Swahili. Furthermore, Makhuwa is the only sampled language which does not allow inner agreement within NPs. Only outer agreement is attested. One last major difference observed in Makhuwa is the absence of a full paradigm of object prefixes. The system eroded to such a point that only classes 1/2 have object agreement markers in the language. In contrast, Bena, Ngoni and Yao have full object markers paradigms, which includes locative object markers.

Bena and Yao, in spite of sharing identical locative features, are geographically not proximal. In fact, and as already shown in Map 1, Bena and Yao areas are separated by the Ngoni linguistic group. Influence from Swahili, as a lingua franca across north-eastern Bantu, is perceptible in all four languages, with lexical borrowing of words such as *mafuta/mahuta* ‘oil’, or in Bena and Ngoni, *chumba* ‘room’, and *lima* ‘cultivate’. As far as the locative system is concerned, however, only Makhuwa seems to have been more directly affected by Swahili through the

suffixation of *-ni* on locativized nouns. Beyond Swahili influence, the few examples retrieved in this paper may not warrant any conclusion on mutual influence within the sampled languages.

Avenues for future research would at least involve extending the study to include locative verbal enclitics which have been excluded from this paper because of a lack of clear data in the selected languages and the difficulty to further investigate on them ex situ. As explained in the introduction, the languages surveyed here represent a convenience sample. Further light could be shed on the microvariation of locative systems in Eastern Bantu through a broader comparative work covering a certain number of Eastern Bantu languages to see how our four sampled languages fit in a wider geographical area.

Abbreviations

FV	Final Vowel
LOC	Locative
OM	Object Marker
PRF	Perfective
PRS	Present
PASS	Passive
REL	Relative
SM	Subject Marker
TAM	Tense Aspect Mood
DEM	demonstrative
INT	intensifier

References

Barlew, Jefferson. 2013. Point of view in Mushunguli locatives. In Ọlanikẹ Ọla Orie & Karen W. Sanders (eds.), *Selected Proceedings of the 43rd Annual Conference on African Linguistics*, 115–129. Somerville, MA: Casdadilla proceedings Project.

Bleek, Wilhelm H. I. 1862–1869. *A comparative grammar of South African languages*. London: Trübner.

Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50.

Buell, Leston. 2007. Semantic and formal locatives: Implications for the Bantu locative inversion typology. *SOAS Working Papers in Linguistics* 15. 105–120.

Chaula, Edigio Hangaika Yamligile. 1989. *Aspects of Ki-Bena phonology: The case of Ki-Mavemba variety*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Demuth, Katherine & Sheila Mmusi. 1997. Presentational focus and thematic structure in comparative Bantu. *Journal of African Languages and Linguistics* 18(1). 1–19. DOI: 10.1515/jall.1997.18.1.1.

Guérois, Rozenn. 2014. Locative inversion in Cuwabo. *ZAS Papers in Linguistics* 57. 49–71.

Guérois, Rozenn. 2016. The locative system in Cuwabo and Makhuwa (P30 Bantu languages). *Linguistique et Langues Africaines* 2. 43–75.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list*. Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.

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

Guthrie, Malcolm. 1971. *Comparative Bantu*, vol. 2. London: Gregg International Publishers.

Guthrie, Malcolm. 1967–1971. *Comparative Bantu: An introduction to the comparative and prehistory of the Bantu languages*, vol. 1–4. London: Gregg International.

Hamlaoui, Fatima. 2014. A note on bare-passives in (selected) Bantu and Western Nilotic languages. *ZAS Papers in Linguistics* 57. 160–182.

Harford, Carolyn. 1983. Locative pseudo-subject in Shona. *Journal of African languages and linguistics* 5. 131–155.

Harries, Lyndon. 1965. Locative agreement in Swahili. *Swahili* 35(2). 70–73.

Ismail, Joseph Hokororo. 2000. *Aspects of Makua phonology: The case of Meto variety*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Kahigi, Kulikoyela. 2005. The Sisumbwa noun: Its classes and derivation. In David P. B. Massamba (ed.), *LOT occasional papers in linguistics*, vol. 1, 117–154. Dar es Salaam: University of Dar es Salaam.

Kröger, Oliver. 2005. Report of a survey of coastal Makhuwa dialects. *Journal of Language Survey Reports* 2. 9108.

Machobane, 'Malillo 'Matsepo. 1995. The Sesotho locative constuctions. *Journal of African Languages and Linguistics* 16(2). 115–136. DOI: 10.1515/jall.1995.16.2.115.

Mallya, Aurelia. 2011. *Locative expressions in Kiwoso*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Mapunda, Gastor. 2015. An analysis of the vitality of the intangible cultural heritage of the Ngoni people of Tanzania: Lessons for other ethnolinguistic groups. *Nordic Journal of African Studies* 24(2). 169–185.

Marten, Lutz. 2010. The great siSwati locative shift. In Anne Breitbarth, Christopher Lucas, Sheila Watts & David Willis (eds.), *Continuity and change in grammar*, 249–267. Amsterdam: Benjamins.

Marten, Lutz. 2012. Agreement in locative phrases in Luganda. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics (Cologne 2009)*, 433–443. Cologne: Köppe.

Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318–368.

Mcha, Yohana. 1979. *The locative in Vunjo*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

Mkude, Daniel J. 2005. *The passive construction in Swahili*. Tokyo: Institute for Languages, Cultures of Asia & Africa.

Moshi, Lioba. 1995. Locatives in Kivunjo-Chaga. In Akinbiyi M. Akinlabí (ed.), *Theoretical approaches to African linguistics*, 129–145. Trenton NJ: Africa World Press.

Mpiranya, Fidèle. 2015. *Swahili grammar and workbook*. London & New York: Routledge.

Ngoboka, Jean Paul. 2016. *Locatives in Kinyarwanda*. Durban: University of KwaZulu-Natal. (Doctoral dissertation).

Ngonyani, Deo. 2003. *A grammar of Chingoni*. München: Lincom.

Ngunga, Armindo & Nomalanga Mpofu-Hamadziripi. 2013. A study of Bantu locatives: The case of Shona and Yao. *Journal of LASU* 1(1). 44–55.

Rugemalira, Josephat M. 2004. Locative arguments in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics*, 285–295. Köln: Rüdger Köppe.

Salzmann, Martin David. 2005. *Theoretical approaches to locative inversion*. Zurich: University of Zurich. (MA thesis).

Stucky, Susan U. 1976. Locatives as objects in Tshiluba: A function of transitivity. *Studies in Linguistic Sciences* 6(2). 174–202.

Zeller, Jochen. Forthcoming. Locatives. In Lutz Marten, Nancy Kula, Ellen Hurst & Jochen Zeller (eds.), *The Oxford guide to the Bantu languages*. Oxford: Oxford University Press.

Part II

The verbal domain

Chapter 7

Multiple-object constructions in Ganda

Nobuko Yoneda & Judith Nakayiza

The possible number of pre-stem object markers (OMs) and the symmetrical/asymmetrical nature of objects in double-object constructions in Bantu languages have been widely discussed (Bresnan & Moshi 1993, Marlo 2015, Marten & Kula 2012, Zeller 2014, among others). While both objects display the same syntactic properties in symmetrical languages, only one object has the syntactic properties of the primary object in asymmetrical languages. Ganda (JE15), spoken in Uganda, is considered a language that allows two OMs (Ssekiryango 2006, Marlo 2015), and also a symmetrical object language, according to the criteria of Bresnan & Moshi (1993). However, according to our observations, three OMs are possible in certain situations. Although Ganda predominantly shows the behaviour of a symmetrical object language, some asymmetrical behaviour can be also observed. This paper demonstrates how object NPs and OMs behave in multiple-object constructions in Ganda. It also shows the asymmetrical tendency of the language and where three OMs are possible.

1 Introduction

Multiple pre-stem object markers (OMs) and the symmetrical/asymmetrical nature of objects in Bantu languages have received much attention in Bantu research (Bresnan & Moshi 1993, Marlo 2015, Marten & Kula 2012, Zeller 2014, among many others). In double-object constructions in symmetrical languages, both objects display the same syntactic properties, whereas in asymmetrical languages, only one object has the syntactic properties of the primary object (see (4)), and the other is restricted in its syntactic behaviour (Bresnan & Moshi 1993).

Ganda, spoken in Uganda, is considered a language that allows two OMs (Ssekiryango 2006, Marlo 2015, Ranero 2019, van der Wal 2020), and is considered

Nobuko Yoneda & Judith Nakayiza. 2024. Multiple-object constructions in Ganda. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 155–189. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663773

a symmetrical language (Ssekiryango 2006). Our observations reveal, however, that three OMs are possible in certain situations. Moreover, although Ganda predominantly shows the behaviour of a symmetrical language, some asymmetrical behaviour can also be observed.

This chapter examines how object NPs and OMs behave in multiple-object constructions and in which morpho-syntactic contexts asymmetry may emerge in Ganda. In §2, we lay out the background of the research on multiple-object constructions. In §3, we show the behaviour of object NPs and OMs in double-object constructions, while in §4 we focus on triple-object constructions. In §5 we discuss the characteristics of multiple-object constructions in Ganda regarding the behaviour and constraints that are observed in §3 and §4. The data are examined against the parameters of object marking proposed by Guérois et al. (2017), in particular parameters 75, 76, 78, 109, and 110.

2 Background to the research

Ganda is spoken in Uganda where it is used both as a first language and as a language of wider communication. In Maho's (2009) classification, Ganda is classified as JE15 in the major group of Nyoro-Ganda. Ganda has been relatively well researched and a reference grammar (Ashton et al. 1954) and numerous textbooks have been written.

Like many other Bantu languages, the Ganda verb is agglutinative, and constructed with a root and different kinds of affixes: subject and object agreement prefixes (SM, OM), affixes that determine the tense, aspect and mood (PreSM, TAM, and Final), and derivational suffixes (DER), as follows:[1]

(1) tebáátúzímbira

PreSM-	SM-	TAM-	OM-	ROOT	-DER	-Final
te-	bá-	á-	tú-	zímb	-ir	-a
NEG-	SM3PL-	PST-	OM1PL-	build	-APPL	-FV

'they did not build for us'

There are some Bantu languages in which the object NP and the corresponding OM cannot co-occur in a clause, others in which the existence of the OM affects the finiteness, and still others in which the presence/absence of the OM depends on the animacy of the corresponding object. Ganda is a language that allows the co-occurrence of the object NP and the corresponding OM. Moreover, the OM is

[1]All Ganda data come from the second author.

optional regardless of the animacy of the object NP, and the existence of an OM does not affect the definiteness of the object NP, as shown in (2) and (3).

(2) a. N-á-lábyé omu-lénzí.
SM1SG-PST-meet.PRF 1-boy

b. N-á-**mú**-lábyé omu-lénzí.
SM1SG-PST-OM1-meet.PRF 1-boy
'I saw a/the boy.'

(3) a. N-a-gúl-a eki-tábo.
SM1SG-PST-buy-FV 7-book

b. N-a-**kí**-gúl-a eki-tábo.
SM1SG-PST-OM7-buy-FV 7-book
'I bought a/the book.'

In the literature, the following are generally assumed to be the main syntactic properties of the primary object in Bantu languages.

(4) Hyman & Duranti (1982: 220)
a. has access to the position immediately following the verb
b. is capable of assuming the subject role through passivization
c. can be expressed by an object marker within the verbal complex

Bresnan & Moshi (1993) divide Bantu languages into two types according to the syntactic behaviour of these objects, namely symmetrical and asymmetrical languages. In symmetrical languages, both (or all) objects can exhibit the syntactic properties of the primary object shown in (4), whereas in asymmetrical languages, only one object can exhibit the syntactic properties of the primary object (Bresnan & Moshi 1993: 47). Example (5) is from Tswana, a symmetrical language, and (6) is an example of Herero, an asymmetrical language.

(5) Tswana (symmetrical type) (Marten et al. 2007: 269)
a. ke ape-ets-e ngwana kuku
SM1.PRS cook-APPL-PRF 1.child 9.chicken
'I cooked the child the chicken'

b. ke ape-ets-e kuku ngwana
SM1.PRS cook-APPL-PRF 9.chicken 1.child
'I cooked the chicken for the child'

(6) Herero (asymmetrical type) (Marten et al. 2007: 269)

a. mávé tjàng-ér-é òvà-nâtjé ò-mbàpírà
PRG.SM2 write-APPL-FV 2-children 9-letter
'They are writing the children a letter'

b. * mave tjang-er-e o-mbapira ova-natje
PRG.SM2 write-APPL-FV 9-letter 2-children
(Intended meaning: 'They are writing a letter to the children')

In Tswana, a symmetrical language, both orders of objects are acceptable in a double-object construction. In Herero, an asymmetrical language, only the benefactive/goal can access the position immediately following the verb. This asymmetry correlates with various factors, such as the semantic role and animacy hierarchy, depending on the language.

Following this idea, in this chapter we discuss the behaviour of the objects in multiple-object constructions in Ganda in terms of symmetry/asymmetry. The criteria to determine the primary object are as follows: (a) if it can be placed immediately after the verb, (b) if it can be the subject in a passive, and (c) if it can be pronominalized inside the verbal complex.

Ganda has some ditransitive verbs such as *-gamba* 'tell' or *-wa* 'give' just like English. They are underived ditransitive verbs, as shown in (7).

(7) a. N-á-gámbyé emi-kwánó ama-úlire.
SM1SG-PST-tell.PRF 4-friends 6-news
'I told friends the news.'

b. N-á-wáddé emi-kwánó amá-tooke.
SM1SG-PST-give.PRF 4-friends 6-bananas
'I gave friends bananas.'

In addition, there are verbs that become ditransitive by derivation, such as in the applicative or causative verb forms of transitive verbs, as shown in (8b) and (8c), respectively.

(8) a. N-a-fúmbyé amá-tooke.
SM1SG-PST-cook.PRF 6-bananas
'I cooked bananas.'

b. N-á-fúmb-íddé aba-ana amá-tooke.
SM1SG-PST-cook-APPL.PRF 2-children 6-bananas
'I cooked bananas for children.'

c. N-á-fúmb-ísízzá aba-ana amá-tooke.
SM1SG-PST-cook-CAUS.PRF 2-children 6-bananas
'I have made children cook bananas.'

In §3 and §4, we will examine the behaviour of the objects of underived ditransitive, applicative, and causative verbs.

3 Double-object constructions

This section demonstrates the behaviour of each object in double-object constructions, including applicative and causative constructions besides underived ditransitive verbs. For each, we will discuss the three criteria, word order in §3.1, passivization in §3.2, and object marking in §3.3. Relative animacy of the objects is indicated with A > B (A is higher than B), A < B (A is lower than B), or A = B (equal animacy).

3.1 Word order of object NPs

As already shown in (4a), the NP which can be placed immediately after the verb (hereafter IAV) is considered the primary object. Here we examine, using the ditransitive, applicative, and causative verbs with double objects, which of the two objects can be placed IAV, and whether the animacy or semantic role of the objects plays a role.

3.1.1 Ditransitive verbs

A typical ditransitive verb is *-wa* 'give'. The semantic roles associated with the objects of this verb are those of recipient and theme. The recipient is *emikwano* 'friends' in (9) and (10), and the themes are *amatooke* 'bananas' in (9) and *embwa* 'dog' in (10). In both examples, the recipient is higher in terms of animacy than the theme.

(9) [RECIPIENT: friends (human), THEME: bananas (entity)] RECIPIENT > THEME

a. Máamá a-wáddé emi-kwánó amá-tooke.
1.mother SM1-give.PRF 4-friends 6-bananas
'Mother has given the friends bananas.'

b. Máamá a-wádde amá-tooke emi-kwánó.
1.mother SM1-give.PRF 6-bananas 4-friends
'Mother has given bananas to the friends.'

(10) [RECIPIENT: friends (human), THEME: dog (animal)] RECIPIENT > THEME

a. Máamá a-wádde emi-kwánó embwá.
1.mother SM1-give.PRF 4-friends 9.dog
'Mother has given the friends a dog.'

b. Máamá a-wádde embwá emi-kwánó.
1.mother SM1-give.PRF 9.dog 4-friends
'Mother has given a dog to the friends.'

When the recipient is higher than the theme in animacy, either object of the ditransitive verb *-wa* 'give' can be placed IAV as shown in (9) and (10).

Ssekiryango (2006: 70) shows examples in which only the recipient can be placed IAV when both objects are of equal animacy. Ranero (2019: 599) also states that the order of object NPs in ditransitive sentences is strictly fixed in that the 'recipient (goal)/benefactive' is placed before the 'theme'. In our data, however, both orders are acceptable. Although the order in which the recipient appears in the IAV position is more natural than the other order as shown in (11). However, still the order of these objects seems not to be as strict as Ranero (2019) claims.[2] The presence of '?' in front of the sentence indicates that the utterance is less natural than the other, but is still grammatical and acceptable.

(11) [RECIPIENT: daughter (human), THEME: hunter (human)] RECIPIENT = THEME

a. Kabáka ya-wá mu-walá we omu-yízzi.
1.king SM1.PST-give.FV 1-daughter 1.his 1-hunter
Interpretation 1: 'The king gave his daughter the hunter.'
?Interpretation 2: 'The king gave his daughter to the hunter.'

b. Kabáka ya-wá omu-yízzi mu-walá we.
1.king SM1.PST-give.FV 1-hunter 1-dughter 1.his
?Interpretation 1: 'The king gave the hunter to his daughter.'
Interpretation 2: 'The king gave the hunter his daughter.'

[2] Some speakers say that interpretation 2 in (11a), (13a) and interpretation 1 in (11b), (13b) are not acceptable. However, these orders in which the theme precedes the recipient or benefactive is often used, especially among older generations. Such differences as the disagreement between our data and data of Ssekiryango (2006) and Ranero (2019) are also seen in other properties. There might be generational or/and areal variation.

In (11), both objects are human and there is no difference in animacy. Although (11a), in which the recipient NP is placed IAV, is more natural and is preferred for the translation of 'the king gave his daughter the hunter', (11b) is still possible, and thus, both (11a) and (11b) are ambiguous with the reading 'the king gave his daughter to the hunter'.

With ditransitive verbs, the objects are symmetrical with respect to word order, although there is a preference for placing the recipient in the IAV position.

3.1.2 Applicative verbs

The applicative verb form in Ganda is formed by adding the derivational suffix *-ir*.[3] When a verb appears in the applicative form, a new object is introduced, the applied object (AO), which contrasts with the base object (BO), which is the original object of the base verb without derivation. A typical semantic role associated with the applied object is the benefactive,[4] and the base object is the theme.

When the applied object is higher than the base object in animacy, either object can be placed IAV as shown in (12).

(12) [AO (BEN): friends (human), BO (THEME): bananas (entity)] AO > THEME

a. Máama a-fúmb-ír-á mi-kwánó gyange amá-tooke.
1.mother SM1-cook-APPL-FV 4-friends 4.my 6-bananas

b. Máama a-fúmb-ír-á amá-tooke mi-kwánó gyange.
1.mother SM1-cook-APPL-FV 6-bananas 4-friends 4.my
'Mother is cooking bananas for my friends.'

When both objects are of equal animacy, the order in which the applied object precedes the base object (13a) is more natural than the other order, and usually the object IAV is interpreted as the applied object. However, the opposite order is still possible, resulting in two possible interpretations, as shown in (13). Only the context allows the hearer to make a choice between the two possible interpretations.

[3] When this suffix appears with the perfect final *-ili*, it appears as *-dde* as a result of the application of morpho-phonological rules.

[4] In Ganda, the applicative verb with its applied object X can be interpreted as 'for X' or 'on behalf of X' depending on the context. We therefore use the standard semantic role 'benefactive (BEN)' for an applied object with either of these interpretations.

(13) [AO (BEN): daughters (human), BO (THEME): man (human)] AO= THEME

a. Máama a-kúb-íddé ba-walá be omu-sájja.
1.mother SM1-beat-APPL.PRF 2-daughters 2.her 1-man
Interpretation 1: 'Mother has beaten a man for/because of her daughters.'
?Interpretation 2: 'Mother has beaten her daughters for/because of the man.'

b. Máama a-kúb-íddé omu-sájja ba-walá be.
1.mother SM1-beat-APPL.PRF 1-man 2-daughters 2.her
?Interpretation 1: 'Mother has beaten a man for/because of her daughters.'
Interpretation 2: 'Mother has beaten her daughters for/because of the man.'

In cases where the semantic role of the applied object is the benefactive, its animacy is rarely lower than that of the base object. However, when the semantic role of the applied object is the reason or motivation, the animacy of the applied object can be lower. The semantic role of the applied object in (14) is a reason, and its animacy is indeed lower than that of the base object. When the base object is higher up on the animacy hierarchy than the applied object, the order in which the base object precedes the applied object (14a) is more natural than the other order (14b),[5] although both orders are possible.

(14) [AO (REASON): party (entity), BO (THEME): friends (human)] AO < THEME

a. Tú-jjá ku-yít-ir-a mi-kwánó gyaffe embága.
SM1PL-FUT INF-call-APPL-FV 4-friends 4.our 9.party

b. ? Tú-jjá ku-yít-ir-a embága mi-kwánó gyaffe.
SM1PL-FUT IND-call-APPL-FV 9.party 4-friends 4.our
'We will call our friends for a party.'

With the applicative verb, as well as with the ditransitive verb, the objects are symmetrical in terms of word order, and neither the semantic role nor animacy determines the ordering of the object NPs, although there seems to be a moderate tendency that either the benefactive or one which is higher in animacy is preferably placed IAV.

[5]This might not be because of the animacy feature, but because of the semantic role. It is obvious that showing an example with benefactive is ideal, but it is not easy to find a good example with benefactive in lower animacy than theme.

3.1.3 Causative verbs

The causative verb form in Ganda is formed by adding the suffix *-is*[6] to the verb root. When the verb appears in the causative form, the causee appears as an object.

When the causee is a human and the theme an inanimate entity, placing the causee IAV is more natural, but both orders are still grammatical as shown in (15). Therefore, based on these examples, we can conclude that both the causee and theme can be placed IAV.

(15) [CAUSEE: friends (human), THEME: banana (entity)] CAUSEE > THEME

a. Máama a-fúmb-ísízzá mi-kwánó gyange amá-tooke.
1.mother SM1-cook-CAUS.PRF 4-friends 4.my 6-bananas

b. ? Máama a-fúmb-ísízzá amá-tooke mi-kwánó gyange.
1.mother SM1-cook-CAUS.PRF 6-bananas 4-friends 4.my
'Mother has made my friends cook bananas.'

However, in cases where the causee and theme have equal animacy, only the causee can be placed IAV. In other words, only the NP which is placed IAV is interpreted as the causee as shown in (16). The sign # indicates that the utterance is grammatical but does not have the intended meaning in which the causee is 'girls' and the theme is 'boy'. Therefore the sentence cannot be used for the interpretation prefixed with *.

(16) [CAUSEE: girls (human), THEME: boy (human)] CAUSEE = THEME

a. Máama a-gób-ésézzá aba-wála omu-lénzí.
1.mother SM1-chase-CAUS.PRF 2-girls 1-boy
'Mother has made the girls chase away the boy.'

b. # Máama a-gób-ésézzá omu-lénzí aba-wála.
1.mother SM1-chase-CAUS.PRF 1-boy 2-girls
Interpretation 1: 'Mother has made the boy chase away the girls.'
*Interpretation 2: 'Mother has made the girls chase away the boy.'

In Ganda, an instrument can be expressed as the causee. In (17), the instrument *omúggo* 'stick' appears as the causee. The causee is an inanimate entity and the theme is an animal; the animacy of the causee is thus lower than that of the theme. Both the causee *omúggo* 'stick' and the theme *embwá* 'dog' can be placed IAV as shown in (17).

[6]When this suffix appears with the perfect final *-ili*, it appears as *-isizza* as the result of the application of some morpho-phonological rules.

(17) [CAUSEE: stick (entity), THEME: dog (animal)] CAUSEE < THEME

a. Maama a-kúb-íssá omú-ggo embwa.
1.mother SM1.PST-beat-CAUSE.PRF 3-stick 9.dog

b. Maama a-kúb-íssá embwa omú-ggo.
1.mother SM1.PST-beat-CAUS.PRF 9.dog 3-stick
'Mother beat a dog with a stick.' (literally meaning: 'mother caused a stick to beat a dog.')

From examples (15) and (17), it seems that with the causative verb as well, objects are symmetrical in terms of word order, except when both objects are of equal animacy. Indeed, (16) shows an asymmetrical behaviour in that the causee must be placed in IAV position when both objects are equal in their animacy.

3.1.4 Summary of word order of object NPs

When the animacy of the theme is lower than that of the other object, these objects behave symmetrically (although there is a moderate preference for a non-theme to be placed IAV) in all ditransitive, applicative, and causative verbs, with two exceptions. One is with causatives. When both objects of a causative verb are of equal animacy, only the causee can be placed IAV and must precede the other object. It is unlike the case of ditransitive or applicative, in which there may be ambiguity in interpretation. Another exception is seen in applicative. When the animacy of the applied object is lower than that of the theme, placing the theme IAV is preferable.

Ssekiryango (2006: 69) claims that objects with higher animacy appear IAV as the primary object, and Ranero (2019: 599) claims that the order is fixed according to the semantic roles. However, our data show that in most cases both objects can appear IAV regardless of their semantic role or animacy.

These facts are summarized in Table 1. We conclude that, in terms of word order, objects are predominantly symmetrical in Ganda.

3.2 Passivization

The passive sentence in Ganda is constructed by adding the derivational suffix *-w*[7] after the verb root. No overt marker is used to introduce the agent noun phrase, as is shown in (18a) (cf. (18b)):

[7]When this suffix appears with the perfect final *-ili*, it appears as *-íddwa* or *-éddwá* as a result of the application of some morpho-phonological rules.

Table 1: IAV positioning of ditransitive objects. *: not acceptable, ?: less natural but acceptable

Recipient > Theme	Rec.	Yes	Rec. = Theme	Rec.	Yes	Rec. < Theme	Rec.	–
	Theme	Yes		Theme	?		Theme	–
Benefactive > Theme	Ben.	Yes	Ben. = Theme	Ben.	Yes	Reason < Theme	reason	?
	Theme	Yes		Theme	?		Theme	Yes
Causee > Theme	Causee	Yes	Cau. = Theme	Causee	Yes	Cau. < Theme	Causee	Yes
	Theme	?		Theme	*		Theme	Yes

(18) a. Eki-tabo ki-som-eddwa aba-ntu bangi.
7-book SM7-read-PASS.PRF 2-people 2.many
'The book has been read by many people.'

b. Aba-ntu bangi ba-somye eki-tabo.
2-people 2.many SM2-read.PRF 7-book
'Many people have read the book.'

The ability to become the subject of the passive sentence is one of the key syntactic properties of the primary object in Bantu languages as shown earlier in (4b). We will see which object can be the subject of a passive sentence in double-object constructions with ditransitive, applicative, and causative verbs.

3.2.1 Ditransitive

(19a) is a passive example in which the recipient is the subject, and in (19b) the theme is the subject of the passive verb.[8] (19c) is the corresponding active sentence.

[8]No preposition is used to introduce the actor in passive sentences in Ganda. Therefore, the order of recipient/theme and actor is also an interesting issue. However, in this paper, we concentrate on their ability to appear as the subject of a passive sentence. In actual use the actor ('maama' in (19)) is often deleted (see Ssekiryango's analysis about Ganda not allowing the presence of the agent in a passive sentence derived from a double-object construction (2006: 72)).

(19) [RECIPIENT: friends (human), THEME: banana (entity)] RECIPIENT > THEME

a. Mi-kwánó gyange gí-wéréddwá amá-tooke (máama).
4-friends 4.my SM4-give.PASS.PRF 6-bananas 1.mother
'My friends have been given bananas (by my mother).'

b. Ama-tóóké gá-wéréddwá mi-kwánó gyange (máama)
6-bananas SM6-give.PASS.PRF 4-friends 4.my 1.mother
'Bananas have been given to my friends (by my mother).'

c. Máamá a-wáddé mi-kwánó gyange amá-tooke.
1.mother SM1-give.PRF 4-friends 4.my 6-bananas
'Mother has given my friends bananas.'

With ditransitive verbs, both objects can be the subject in a passivized sentence when the recipient is higher than the theme in animacy, as shown in (19).

(20) exemplifies cases where both objects are of equal animacy. (20a) is an example in which the recipient *muwalá wa kabaka* 'king's daughter' is the subject, and in (20b) the theme *omuyízzi* 'a hunter' is the subject of the passive verb. (20c) is the corresponding active sentence.

(20) [RECIPIENT: daughter (human), THEME: hunter (human)] RECIPIENT = THEME

a. Mu-walá wa kabaka a-wereddwa omu-yízzi
1-daughter GEN 1.king SM1-give.PASS.PRF 1-hunter
Interpretation 1: 'The daughter of king has been given the hunter.'
?Interpretation 2: 'The daughter of king has been given to the hunter.'

b. Omu-yízzi a-wereddwa mu-walá wa kabaka.
1-hunter SM1-give.PASS.PRF 1-daughter GEN 1.king
?Interpretation 1: 'The hunter has been given to king's daughter.'
Interpretation 2: 'The hunter has been given the king's daughter.'

c. Kabáka a-wadde mu-walá we omu-yízzi.
1.king SM1-give.PRF 1-daughter 1.his 1-hunter
'The king has given his daughter the hunter.'

Both objects can be the subject of the passivized sentence, though the preference is given for the recipient to be the subject. As a result, (20a) and (20b) are both ambiguous in their interpretation.

3.2.2 Applicative

The examples in (21) show passive sentences with an applicative verb. The applied object (benefactive) is the subject in (21a), and the theme is the subject in (21b). (21c) is the corresponding active sentence.

(21) [AO (BEN): friends (human), BO (THEME): banana (entity)] AO > THEME

a. Mi-kwánó gyange gi-fúmb-ír-íddwá amá-tooke.
4-friends 4.my SM4-cook-APPL-PASS.PRF 6-bananas
'My friends have been cooked bananas (by mother).'

b. Amá-tooke gá-fúmb-ír-íddwá mi-kwánó gyange.
6-bananas SM6-cook-APPL-PASS.PRF 4-friends 4.my
'Bananas have been cooked for my friends (by mother).'

c. Máama a-fúmb-íddé mi-kwánó gyange amá-tooke.
1.mother SM1-cook-APPL.PRF 4-friends 4.my 6-bananas
'Mother has cooked bananas for my friends.'

With applicative verbs, both objects can be the subject of a passivized sentence as was also shown with the ditransitive verbs above. Therefore, both objects behave symmetrically in terms of their ability to appear as the subject of a passive sentence when the benefactive is higher in animacy than the theme.

The examples in (22) are cases in which both objects are equal in animacy.

(22) [AO (BEN): daughters (human), BO (THEME): man (human)] AO = THEME
aba-walá ba-kúb-ir-íddwá omu-sájja.
2-daughters SM2-beat-APPL-PASS.PRF 1-man
Interpretation 1: 'Someone beat a man for daughters.'
Interpretation 2: 'Someone beat daughters for a man.'

(22) can have both interpretations 1 and 2. *Abawala* 'daughters' is the benefactive in Interpretation 1, and is the theme in Interpretation 2. That is, both benefactive and theme can be the subject of the passive. Therefore (22) is ambiguous.

(23) is an example in which the theme is higher than the applied object in animacy. The semantic role of the applied object is reason. In this case, only the theme can be the passive subject as shown in (23b). (23a) in which 'party' is the subject is grammatical but has a different meaning.

(23) [AO (REASON): party (entity), BO (THEME): friends (human)] AO < THEME

a. # Embága e-jja ku-yit-ir-w-a mi-kwánó gyaffe.
9.party SM9-FUT INF-call-APPL-PASS-FV 4-friends 4.our
'The party will be held for our friends.'

b. Mi-kwano gyaffe gi-jja ku-yit-ir-w-a embaga.
4-friends 4.our SM4-FUT INF-call-APPL-PASS-FV 9.party
'Our friends will be called for a party.'

c. Tú-jjá ku-yít-ir-a mi-kwánó gyaffe embága.
SM1PL-FUT INF-call-APPL-FV 4-friends 4.our 9.party
'We will call our friends for a party.'

Therefore, in applicatives, when the animacy of the applied object (benefactive) is equal to the theme or higher, objects behave symmetrically, while objects behave asymmetrically when the animacy of the applied object (a reason) is lower than the theme. However, cases like (23) in which the semantic role of the applied object is a reason/motivation may need to be treated separately based on a number of other factors.

3.2.3 Causative

The examples in (24) show passive sentences with a causative verb. The causee is the subject in (24a) and the theme is the subject in (24b).

(24) [CAUSEE: friends (human), THEME: bananas (entity)] CAUSEE > THEME

a. Mi-kwánó gyange gí-fúmb-ísiddwa amá-tooke.
4-friends 4.my SM4-cook-CAUS.PASS.PRF 6-bananas
'My friends have been caused to cook bananas.'

b. Amá-tooke gá-fúmb-ísíddwá mi-kwano gyange.
6-bananas SM6-cook-CAUS.PASS.PRF 4-friends 4.my
'Bananas have been caused to be cooked by my friends.'

As shown, with causative verbs as well, both objects can be the subject of a passivized sentence. However, when the causee and theme are of equal animacy, only the causee and not the theme can be the passive subject, as shown in (25).

(25) [CAUSEE: girls (human), THEME: boy (human)] CAUSEE = THEME

a. Aba-wála ba-gób-és-éddwá omu-lénzí máama
2-girls SM2-chase-CAUS-PASS.PRF 1-boy 1.mother
'The girls were caused to chase away a boy by mother.'

b. # Omu-lénzí a-gób-és-éddwá aba-wála máama.
1-boy SM1-chase-CAUS-PASS.PRF 2-girls 1.mother
Interpretation 1: 'The boy was caused to chase away the girls by mother.'
*Interpretation 2: 'The boy was chased away by the girls caused by mother.'

The same holds for cases in which the causee is lower than the theme in animacy, only the causee can be the subject of the passivized sentence, as shown in (26).

(26) [CAUSEE: stick (entity), THEME: dog (animal)] CAUSEE < THEME

a. Omú-ggo gwa-kúb-ísíddwá embwa máamá.
3-stick SM3.PST-beat-CAUS.PASS.PRF 9.dog 1.mother
'A stick was used by mother to beat the dog.' (literally meaning: 'A stick was caused to beat the dog by mother.')

b. * Embwá ya-kúb-ísíddwá omú-ggo máamá.
9.dog SM9.PST-beat-CAUS.PASS.PRF 3-stick 1.mother
(Intended meaning: 'The dog was beaten with a stick by mother.')

3.2.4 Summary of properties of passivization

The data presented thus far can be summarized as in Table 2. Concerning the ability of being the subject of a passivized sentence, objects are generally symmetrical with the exception of causative verbs. For causative verbs, the theme can be the subject of the passivized sentence only when its animacy is lower than the causee. The semantic role "reason" is again an exception here, and exhibits asymmetry.

3.3 Object marking

The third syntactic property of the primary object in Bantu languages is pronominalization, namely, which object can be expressed by an OM within the verbal complex.

3.3.1 Ditransitive

Example (27) shows a ditransitive verb. Both objects appear as noun phrases in (27c). The recipient *emikwánó* 'friends' is pronominalized and appears as an

Table 2: Object ability to be the subject of a passivized sentence

Recipient > Theme	Rec.	Yes	Rec. = Theme	Rec.	Yes	Rec. < Theme	Rec.	–
	Theme	Yes		Theme	?		Theme	–
Benefactive > Theme	Ben.	Yes	Ben. = Theme	Ben.	Yes	Reason < Theme	reason	*
	Theme	Yes		Theme	Yes		Theme	Yes
Causee > Theme	Causee	Yes	Cau. = Theme	Causee	Yes	Cau. < Theme	Causee	Yes
	Theme	Yes		Theme	*		Theme	*

OM in (27b), and the theme *amátooke* 'bananas' is pronominalized and appears as an OM in (27c). As shown in (27), either object in ditransitive verbs can be pronominalized and expressed as an OM when the recipient is higher than the theme in animacy.

(27) [RECIPIENT: friends (human), THEME: banana (entity)] RECIPIENT > THEME

a. Máama a-wáddé emi-kwánó amá-tooke.
1.mother SM1-give.PRF 4-friends 6-bananas
'Mother has given friends bananas.'

b. Máama a-ba-wádde amá-tooke.
1.mother SM1-OM2-give.PRF 6-bananas
'Mother has given them (friends) bananas.'

c. Máama a-ga-wádde emi-kwánó.
1.mother SM1-OM6-give.PRF 4-friends
'Mother has given them (bananas) to friends.'

Both objects can appear as OMs in an utterance as in (28). In (28a), the recipient OM is placed immediately before the stem (IBS, hereafter). This is the natural order, and the other order in (28b) is odd, although not completely ungrammatical (as is marked by "??").

(28) a. Máama a-ga-bá-wadde.
1.mother SM1-OM6-OM2-give.PRF

b. ?? Máama a-ba-gá-wadde.
1.mother SM1-OM2-OM6-give.PRF
'Mother has given them (bananas) to them (friends).'

According to Ranero (2019: 599), the OM that agrees with the theme must precede the one that agrees with the recipient or benefactive. Our data also show that the other order is odd, but it is still not ungrammatical.[9]

This preference for the order of the OMs can be seen more clearly when the OM agrees with a 1st person singular subject as shown in (29). (29a), in which the 1SG OM referring to the recipient appears IBS, is fine, while (29b), in which the theme OM appears IBS, is ungrammatical.

(29) a. Máama a-gá-n-wadde. (> agampadde)
1.mother SM1-OM6-OM1SG-give.PRF

b. * Maama a-n-ga-wadde
1.mother SM1-OM1SG-OM6-give.PRF
'Mother has given them (bananas) to me.'

In fact, not all Ganda speakers accept (28b). Even then, (28b) is still not as bad as (29b) for these speakers. This suggests that there is a subtle but clear difference in the acceptability between (28b) and (29b). We will return to this issue later in §4.3.

The examples in (30) show cases in which both objects are equal in animacy.

(30) [RECIPIENT: daughters (human), THEME: hunter (human)] RECIPIENT = THEME

a. Kabáka ya-ba-wá omu-yízzi.
1.king SM1.PST-OM2-give.FV 1-hunter
Interpretation 1: 'The king gave the hunter to them (his daughters).'
?Interpretation 2: 'The king gave them (his daughters) to the hunter.'

b. Kabáka ya-mu-wa ba-wala be.
1.king SM1.PST-OM1-give.FV 2-daughters his
?Interpretation 1: 'The king gave him (the hunter) to his daughters.'
Interpretation 2: 'The king gave his daughters to him (the hunter).'

c. Kabáka ya-mu-ba-wá.
1.king SM1.PRF-OM1-OM2-give.FV
'The king gave him (the hunter) to them (his daughters).'

[9] (28b) is odd but not ungrammatical. For example, the second author's grandmother used such utterances. However, it is not common and is only marginally acceptable today it seems.

d. # Kabáka ya-ba-mu-wá.
1.king SM1.PRF-OM2-OM1-give.FV
Interpretation 1: 'The king gave them (his daughters) to him (the hunter).'
*Interpretation 2: 'The king gave him (the hunter) to them (his daughters).'

Both objects can be pronominalized and appear as OM as shown in (30a) and (30b). Data from Ssekiryango (2006) and Ranero (2019) show that only the recipient can be pronominalized when one of the objects is pronominalized. Our data also show that pronominalization of a recipient is more natural. However, the pronominalization of a theme is also possible, and hence there is ambiguity. When both objects appear as OMs simultaneously, the OM that agrees with the recipient must appear closer to the verb stem, as shown in (30c) and (30d). This agrees with the data from Ssekiryango (2006) and Ranero (2019).

3.3.2 Applicative

(3a) is an example of applicative verb. Both objects appear as noun phrases in (31a), the benefactive *mikwánó gyange* 'my friends' appears as an OM in (31b), and the theme *amátóóké* 'bananas' appears as an OM in (31c). Either object can appear as an OM when the benefactive is higher than the theme in animacy, as is also the case with ditransitive verbs.

(31) [AO (BEN): friends (human), BO (THEME): banana (entity)] AO > THEME

a. Máama a-fúmb-íddé mi-kwánó gyange amá-tóóké.
1.mother SM1-cook-APPL.PRF 4-friends 4.my 6-bananas
'Mother has cooked bananas for my friend.'

b. Máama ya-ba-fúmb-idde amá-tóóké.
1mother SM1-OM2-cook-APPL.PRF 6-bananas
'Mother has cooked bananas for them.'

c. Máama ya-ga-fúmb-íddé mi-kwánó gyange.
1.mother SM1-OM6-cook-APPL.PRF 4-friends 4.my
'Mother has cooked it for my friends.'

It is also possible that both objects appear as OMs at the same time, as shown in (32). Placing the one that agrees with the benefactive (or that of higher animacy) IBS is much more natural than the other order. (32b), in which the theme appears as an OM, is very unnatural, although it is not completely ungrammatical.

(32) a. Máama a-ga-bá-fumb-idde.
1.mother SM1-OM6-OM2-cook-APPL.PRF

b. ?? Máama a-ba-gá-fúmb-idde.
1.mother SM1-OM2-OM6-cook-APPL.PRF
'Mother has cooked it for them.'

The examples in (33) show cases in which both objects are equal in animacy. The benefactive (AO) *bawalá be* 'her daughters' appears as an OM in (33a), the theme (BO) *omusájja* 'man' appears as an OM in (33b), and both objects appear as OMs in (33c) and (33d).

(33) [AO (BEN): daughters (human), BO (THEME): man (human)] AO = THEME

a. Máama a-ba-kúb-íddé omu-sájja.
1.mother SM1-OM2-beat-APPL.PRF 1-man
'Mother has beaten a man for/on behalf of them (her daughters).'

b. # Máama a-mu-kúb-íddé ba-walá bwe.
1.mother SM1-OM1-beat-APPL.PRF 2-daughters 2.her
'Mother has beaten her daughters for/on behalf of him (a man).'

c. Máama a-mu-ba-kúb-íddé.
1.mother SM1-OM1-OM2-beat-APPL.PRF
'Mother has beaten him (a man) for/on behalf of them (her daughters).'

d. # Máama a-ba-mu-kúb-íddé.
1.mother SM1-OM2-OM1-beat-APPL.PRF
'Mother has beaten them (her daughters) for/on behalf of him (a man).'

It is only the benefactive that can be pronominalized and appear as an OM when both objects are equal in animacy, as shown in (33a) and (33b). Also when both objects appear as OMs, the one that agrees with the benefactive must appear IBS, as shown in (33c).

(34) is an example in which the animacy of the theme *mikwánó gyaffe* 'our friends' is higher than that of the applied object *embága* 'party'. Both objects can appear as OMs, just like in (31), where the animacy relation of the two objects is the other way around.

(34) [AO (REASON): party (entity), BO (THEME): friends (human)] AO < THEME

a. Tú-jjá ku-yít-ir-a mi-kwánó gyaffe embága.
SM1PL-FUT INF-call-APPL-FV 4-friends 4.our 9.party
'We will call our friends for a party.'

b. Tú-jjá ku-gi-yít-ir-a mi-kwánó gyaffe.
SM1PL-FUT INF-OM9-call-APPL-FV 4-friends 4.our
'We will call our friends for it (a party).'

c. Tú-jjá ku-ba-yít-ir-a embága.
SM1PL-FUT INF-OM2-call-APPL-FV 9.party
'We will call them (our friends) for a party.'

However, when both objects appear as OMs, placing the OM that agrees with the theme IBS is much more natural than the other order, as shown in (35b).

(35) a. ?? Tú-jjá ku-ba-gi-yít-ir-a.
SM1PL-FUT INF-OM2-OM9-call-APPL-FV

b. Tú-jjá ku-gi-ba-yít-ir-a.
SM1PL-FUT INF-OM9-OM2-call-APPL-FV
'We will call them (our friends) for it (a party).'

Therefore, both the symmetrical behaviour (with respect to pronominalization (34)) and asymmetrical behaviour (with respect to the order of OMs (35)) are observed here.

3.3.3 Causative

(36) is an example of a causative verb. Both objects appear as noun phrases in (36a), the causee *mikwánó gyange* 'my friends' appears as an OM in (36b), and the theme *amátóóké* 'bananas' appears as an OM in (36c). In the case of a causative verb, either of the objects can be pronominalized and appear as an OM, as was also the case with ditransitive and applicative verbs.

(36) [CAUSEE: friends (human), THEME: bananas (entity)] CAUSEE > THEME

a. Máama a-fúmb-ísízzá mi-kwánó gyange amá-tooke.
1.mother SM1-cook-CAUS.PRF 4-friends 4.my 6-bananas
'Mother has made my friends cook bananas.'

b. Máama a-ba-fúmb-ísízzá amá-tooke.
1.mother SM1-OM2-cook-CAUS.PRF 6-bananas
'Mother has made them (my friends) cook bananas.'

c. Máama a-ga-fúmbísízzá mi-kwánó gyange.
1.mother SM1-OM6-CAUS.PRF 4-friends 4.my
'Mother has made my friends cook them (bananas).'

In the examples in (37), in which both objects appear as OMs, the OM that agrees with the causee *mikwánó gyange* 'my friends' is placed IBS in (37a), and the opposite order is shown in (37b). The former order is acceptable, but not the latter. (37b) is not ungrammatical, but is very odd.

(37) a. Máama a-ga-bá-fúmb-ísízzá.
1.mother SM1-OM6-OM2-cook-CAUS.PRF
'Mother has caused them to cook it.'

b. ?? Máama a-ba-gá-fúmb-ísízzá.
1.mother SM1-OM2-OM6-cook-CAUS.PRF
'Mother has caused them to cook it.'

When the causee is higher than the theme in animacy, either object can be pronominalized as we have seen in (37); however, when both objects are of equal animacy, only the causee can be pronominalized and expressed as the OM, as shown in (38).

(38) [CAUSEE: girls (human), THEME: boy (human)] CAUSEE = THEME

a. Máama a-gob-ésézzá aba-wála omu-lénzi.
1.mother SM1-chase-CAUS.PRF 2-girls 1-boy
'Mother has made the girls chase away the boy.'

b. Máama a-ba-gób-ésézzá omu-lénzi.
1.mother SM1-OM2-chase-CAUS.PRF 1-boy
'Mother has made them (the girls) chase away the boy.'

c. # Máama a-mú-gób-ésézzá aba-wála.
1.mother SM1-OM1-chase-CAUS.PRF 2-girls
'Mother has made him (the boy) chase away the girls.'

In (38), the causee *abawála* 'the girls' and the theme *omulénzi* 'boy' are at the same level of animacy. In this case, only the causee can appear as an OM. Therefore, in (38c), in which *omulénzi* 'the boy' appears as an OM, *abawála* 'the girls' cannot be interpreted as the causee. Here we can see some clear asymmetrical characteristics determined by the semantic role.

Both objects can appear as OMs at the same time in the order shown in (39), which is also the case when the two objects are of equal animacy. In this case too, the OM that agrees with the causee must be placed IBS.

(39) a. Máama a-mú-ba-gób-ésézzá.
1.mother SM1-OM1-OM2-chase-CAUS.PRF
'Mother has made them (the girls) chase him (the boy) away.'

b. # Máama a-bá-mu-gób-ésézzá.
1.mother SM1-OM2-OM1-chase-CAUS.PRF
'Mother has made him (the boy) chase them (the girls) away.'

(40) and (41) are examples in which the causee is lower than the theme in animacy. The causee is pronominalized in (40a) and the theme is pronominalized in (40b). The behaviour is thus symmetrical and both objects can be pronominalized as shown in (40a) and (40b). However, when both objects are pronominalized and appear as OMs at the same time, their order is asymmetrical. The OM that is placed IBS is the one that corresponds to the causee, and the other order is odd as shown in (41).

(40) [CAUSEE: stick (entity), THEME: dog (animal)] CAUSEE < THEME

a. Máama a-gu-kúb-íssá embwa.
1.mother SM1.PST-OM3-beat-CAUS.PRF 9.dog
'Mother beat a dog with it (stick).' (literally meaning: 'Mother caused it (stick) to beat dog.')

b. Máama a-gi-kúb-íssá omú-ggo.
1.mother SM1.PST-OM9-beat-CAUS.PRF 3-stick
'Mother beat it (dog) with a stick.'

(41) a. Máama a-gi-gu-kúb-íssá.
1.mother SM1.PST-OM9-OM3-beat-CAUS.PRF

b. ?? Máama a-gu-gi-kúb-íssá.
1.mother SM1.PST-OM3-OM9-beat-CAUS.PRF
'Mother beat it (a dog) with it (a stick).'

Likewise with causative verbs, both the symmetrical and asymmetrical behaviours are observed.

3.3.4 Summary of object marking

The data presented in this section regarding pronominalization are summarized in Table 3. Table 4 summarizes the facts regarding the possible ordering of OMs when both objects appear at the same time. As Table 3 shows, the language is

Table 3: Pronominalization (appearing as an OM)

Recipient > Theme	Rec.	Yes	Rec. = Theme	Rec.	Yes	Rec. < Theme	Rec.	–
	Theme	Yes		Theme	?		Theme	–
Benefactive > Theme	Ben.	Yes	Ben. = Theme	Ben.	Yes	Reason < Theme	reason	Yes
	Theme	Yes		Theme	*		Theme	Yes
Causee > Theme	Causee	Yes	Cau. = Theme	Causee	Yes	Cau. < Theme	Causee	Yes
	Theme	Yes		Theme	*		Theme	Yes

Table 4: Ability of the OM to be placed immediately before the stem. *: not acceptable, ?: less unnatural but acceptable, ??: very odd but not ungrammatical

Recipient > Theme	Rec.	Yes	Rec. = Theme	Rec.	Yes	Rec. < Theme	Rec.	–
	Theme	??		Theme	*		Theme	–
Benefactive > Theme	Ben.	Yes	Ben. = Theme	Ben.	Yes	Reason < Theme	reason	??
	Theme	??		Theme	*		Theme	Yes
Causee > Theme	Causee	Yes	Cau. = Theme	Causee	Yes	Cau. < Theme	Causee	Yes
	Theme	??		Theme	*		Theme	??

symmetrical except when both objects are of equal animacy. On the other hand, the order of OMs is clearly asymmetrical as Table 4 shows.

Compared to the other two properties (object order and passivization), pronominalization seems to most clearly highlight the asymmetry in the language especially with respect to the order of OMs. Ssekiryango (2006) and Ranero (2019) also report that although both objects can be pronominalized symmetrically, the order of OMs is rigidly fixed. Here again, the behaviour of the "reason" semantic role is an exception to the rule.

3.4 Findings and summary of double-object constructions in Ganda

In double-object constructions in Ganda, both objects can be (i) placed IAV, (ii) the subject of a passive sentence, and (iii) pronominalized. These facts show that Ganda is a symmetrical object language (Bresnan & Moshi 1993). They show that even in cases where the interpretation of the semantic roles of the two objects becomes ambiguous, the grammar allows both objects to equally assume the primary object position. At the same time, however, some asymmetrical characteristics are also observed, such as the preference for the primary object to be a non-theme (recipient, benefactive or causee). Another noticeable asymmetrical feature is the order of OMs. These asymmetrical characteristics seem particularly prominent with causative verbs. The order of OMs is not included in the main syntactic properties of the primary object in Bantu languages shown earlier in (4). However, it is still an important characteristic observed in Ganda. van der Wal (2020: 216–217) points out that although the ordering of OMs does not necessarily follow the thematic roles in other Bantu languages with multiple object markers, the order of OMs is determined by their semantic role in Ganda.

When the semantic role of the applied object is not the benefactive but the reason, it behaves differently from other cases. It is not clear at this point if this difference is due to the animacy hierarchy or the semantic role of the "benefactive" and "reason". This deserves further investigation. Another important finding is the restriction on the appearance of the OM that agrees with the 1st person singular. We will discuss this in §4.3 below.

4 Triple-object constructions

Verbs in Ganda do not allow three object NPs; however, triple-object constructions are possible, albeit restricted. The conditions of triple-objects in Ganda are that (i) they appear with the applicative forms of ditransitive verbs, such as *-wa* 'give', *-soba* 'ask', and *-gamba* 'tell', and (ii) the applied object (benefactive) has to be indicated by an OM. Therefore, the semantic roles of the objects in the triple-object constructions must be the recipient, theme, and benefactive. The different ways in which these can therefore be expressed are as follows:

(42) a. with an OM (benefactive) + two object NPs (recipient and theme)
b. with two OMs (benefactive, and recipient or theme) + an object NP (recipient or theme)
c. with three OMs (benefactive, recipient, and theme)

Table 5: Symmetrical/asymmetrical nature of double objects

Relative animacy	Semantic role of objects	IAV	Passive	OM	IBS
RECIPIENT > THEME	Recipient	Yes	Yes	Yes	Yes
	Theme	Yes	Yes	Yes	??
BENEFACTIVE > THEME	Benefactive	Yes	Yes	Yes	Yes
	Theme	Yes	Yes	Yes	??
CAUSEE > THEME	Causee	Yes	Yes	Yes	Yes
	Theme	?	Yes	Yes	??
RECIPIENT = THEME	Recipient	Yes	Yes	Yes	Yes
	Theme	?	?	?	*
BENEFACTIVE = THEME	Benefactive	Yes	Yes	Yes	Yes
	Theme	?	Yes	*	*
CAUSEE = THEME	Causee	Yes	Yes	Yes	Yes
	Theme	*	*	*	*
REASON < THEME	Reason	?	*	Yes	??
	Theme	Yes	Yes	Yes	Yes
CAUSEE < THEME	Causee	Yes	Yes	Yes	Yes
	Theme	Yes	*	Yes	??

In this section, we will show the possible orders of the object NPs as schematized in (42a), and the possible orders of OMs as in (42b) and (42c).

4.1 Order of object NPs

As mentioned above, placing three NPs following the verb is not allowed (43a) and the applied object has to be indicated by an OM in triple-object constructions in Ganda (43b). Alternatively, the benefactive has to appear with a preposition as shown in (43c), in which case it is no longer a triple-object construction.

(43) a. * N-gámb-idde máamá emi-kwano ama-wulire.
SM1SG-tell-APPL.PRF 1.mother 4-friends 6-news
(Intended meaning: 'I have told the news to friends for/on behalf of my mother.')

b. N-mu-gámb-idde emi-kwáno ama-wúlire.
SM1SG-OM1-tell-APPL.PRF 4-friends 6-news
'I have told the news to friends for/on behalf of her (mother).'

c. N-gámbye emi-kwáno ama-wúlire kulwá máama.
SM1SG-tell.PRF 4-friends 6-news for 1.mother
'I have told friends the news for/on behalf of my mother.'

Since the applied object must be pronominalized and expressed as the OM, the question relates to the order of the other objects, namely, the order of theme object NP and the recipient object NP. The examples in (44) show that these two object NPs can appear in either order.

(44) a. N-mu-gámb-idde emi-káno ama-wúlire.
SM1SG-OM1-tell-APPL.PRF 4-friends 6-news
b. N-mu-gámb-idde ama-wúlire emi-káno.
SM1SG-OM1-tell-APPL.PRF 6-news 4-friends
'I have told the news to friends for her (mother).'

4.2 The order of OMs

Other possible forms of triple-object constructions are two OMs + one NP (42b) and three OMs (42c). Here, we will show the ordering of the OMs.

4.2.1 Two OMs + one NP

As mentioned above, the applied object (benefactive) must appear as an OM; therefore, one of the object NPs must be either a recipient or theme. The possible combinations of OMs are {recipient OM+ benefactive OM} and {theme OM+ benefactive OM}, as exemplified in (45) and (46) respectively. The benefactive OM is at IBS in (45a) and (46a), and the other OM is at IBS in (45b) and (46b). As (45) and (46) show, the order is determined by the semantic roles. The OM that refers to the applied object appears IBS. The other order is very odd, but not ungrammatical.

(45) a. A-ba-mú-gámb-idde ama-úlire.
SM1-OM2-OM1-tell-APPL.PRF 6-news
'He has told them (friends) the news for her (mother).'
b. ?? A-mu-bá-gámb-idde ama-úlire.
SM1-OM1-OM2-tell-APPL.PRF 6-news

(46) a. A-ga-mú-gámb-idde emi-káno.
SM1-OM6-OM1-tell-APPL.PRF 4-friends
'He has told it (the news) to friends for her (mother).'

b. ?? A-mu-gá-gámb-idde emi-káno
SM1-OM1-OM6-tell-APPL.PRF 4-friends

When the OM is 1st person singular, it must be placed IBS, and any other order is ungrammatical. In (47), the applied object is 1st person singular, and its OM must appear IBS. The other order is ungrammatical as shown in (47b). Compare this with (45b) or (46b), which are very unnatural but not ungrammatical.

(47) a. A-bá-n-gámb-idde ama-wúlire.
SM1-OM2-OM1SG-tell-APPL.PRF 6-news
Interpretation 1: 'He has told them the news for me.'
*Interpretation 2: 'He has told me the news for them'

b. * A-n-bá-gámb-iddé ama-wúlire.
SM1-OM1SG-OM2-tell-APPL.PRF 6-news

If the OM of 1st person singular is placed IBS, it is only interpreted as the benefactive, never as the recipient, as shown in (47a). If the benefactive OM of 3rd person plural (class 2) is IBS, it is ungrammatical as shown in (47b).

The benefactive OM must appear IBS; however, the OM of 1st person singular must appear IBS as well. There is a conflict, then, when the applied object is not 1st person singular. In the case when the 1st person singular OM does not refer to the benefactive, the applicative verb form cannot be used, and the benefactive must be expressed in a prepositional phrase as shown in (48a,b). In this case, the position IBS is not for the benefactive but for the 1st person singular. Therefore, restriction of 1st person singular OM has priority over the benefactive restriction for IBS.

(48) a. A-ga-n-gámbye kulwábwe.
SM1-OM6-OM1SG-tell.PRF for.them
'He has told it (the news) to me for them'

b. A-n-gámbye ama-wúlire kulwábwe.
SM1-OM1SG-tell.PRF 6-news for.them
'He has told me the news for them'

c. * A-n-ba-gámbye ama-wúlire.
SM1-OM1SG-OM2-tell.PRF 6-news
(Intended meaning: 'He has told me the news for them')

The place restriction of the 1st person singular OM is very strict, and seems to be more than just a "tendency" or "preference" which was observed for the other hierarchies. We will further discuss this in §4.3.

4.2.2 Three OMs

Ganda is considered a language that allows two OMs (Ssekiryango 2006, Marlo 2015). However, according to our observation, in fact, up to three OMs are possible. Expressions such as the following are therefore widespread.

(49) a. O-ki-bá-n-gámb-idde.
SM2SG-OM7-OM3PL-OM1SG-tell-APPL.PRF
'You have told it to them for me.'

b. O-ki-ba-tú-gámb-idde.
SM2SG-OM7-OM3PL-OM1PL-tell-APPL.PRF
'You have told it to them for us.'

All three objects, namely the benefactive, the recipient, and the theme, can be expressed by OMs at the same time. Here, we show the possible orders of these three OMs.

The applied (benefactive) OM must appear IBS, as with the double-object construction (§3.3.2). The examples in (50) show that the 2nd person singular benefactive OM *kú-* is necessarily placed IBS. As long as this condition is met, the order of both theme and recipient OMs is interchangeable, as shown in (50a) and (50b).

(50) a. THEME-REC-BEN
N-ki-ba-kú-gámb-idde.
SM1SG-OM7-OM3PL-OM2SG-tell-APPL.PRF

b. REC-THEME-BEN
N-ba-kí-kú-gámb-idde.
SM1SG-OM3PL-OM7-OM2SG-tell-APPL.PRF
'I have told it to them for/on behalf of you(sg).'

(51) exemplifies a situation in which the 3rd person plural OM *bá-* placed IBS may not be interpreted as the recipient but rather as the benefactive. The 2nd person singular OM *kú-* thus receives the recipient interpretation, as shown in (51a) and (51b) respectively.

(51) a. THEME-BEN-REC
N-ki-ku-bá-gámb-idde.
SM1SG-OM7-OM2SG-OM3PL-tell-APPL.PRF
*Interpretation 1: 'I have told it to them for/on behalf of you(sg).'
Interpretation 2: 'I have told it to you(sg) for/on behalf of them.'

b. BEN-THEME-REC
N-ku-ki-bá-gámb-idde.
SM1SG-OM2SG-OM7-OM3PL-tell-APPL.PRF
*Interpretation 1: 'I have told it to them for/on behalf of you(sg).'
Interpretation 2: 'I have told it to you(sg) for/on behalf of them.'

(52) shows examples in which the theme OM *ki-* is placed IBS. This is very unnatural regardless of the ordering of the other OMs, as shown in (52a) and (52b) (note again the ??).

(52) a. REC-BEN-THEME
?? N-ba-kú-kí-gámb-idde.
SM1SG-OM3PL-OM2SG-OM7-tell-APPL.PRF
??Interpretation 1: 'I have told it to them for/on behalf of you (sg).'
*Interpretation 2: 'I have told it to you (sg) for/on behalf of them.'

b. BEN-REC-THEME
?? N-ku-ba-kí-gámb-idde.
SM1SG-OM2SG-OM3PL-OM7-tell-APPL.PRF
*Interpretation 1: 'I have told it to them for/on behalf of you (sg).'
??Interpretation 2: 'I have told it to you (sg) for/on behalf of them.'

In (52a), the 2nd person singular OM *ku-* can still be interpreted as the benefactive. The same holds with the 3rd person plural OM *ba-* in (52b). This suggests that the closer the OM is to IBS, the more likely it is to be interpreted as the benefactive, although these sentences are still very odd.

However, here again, the 1st person singular OM *n-* behaves differently. This must be placed IBS as shown in (53a) and (53b), and other orders are all ungrammatical as shown in (53c)–(53f).

(53) a. A-ki-bá-**n**-gámb-idde.
SM3SG-OM7-OM3PL-OM1SG-tell-APPL.PRF

b. A-ba-kí-**n**-gámb-idde.
SM3SG-OM3PL-OM7-OM1SG-tell-APPL.PRF

c. * A-ki-**n**-bá-gámb-idde.
SM3SG-OM7-OM1SG-OM3PL-tell-APPL.PRF

d. * A-**n**-kí-bá-gámb-idde.
SM3SG-OM1SG-OM7-OM3PL-tell-APPL.PRF

e. * A-**n**-ba-kí-gámb-idde.
SM3SG-OM1SG-OM3PL-OM7-tell-APPL.PRF

f. * A-ba-**n**-kí-gámb-idde.
SM3SG-OM3PL-OM1SG-OM7-tell-APPL.PRF
'He has told it to them for/on behalf of me.'

Therefore, when the benefactive is not 1st person singular, it cannot appear as an OM as shown in (54). This is the same restriction we saw earlier in (47) in §4.2.1 above.

(54) a. * A-**n**-kí-ba-gámb-idde.
SM3SG-OM1SG-OM7-OM3PL-tell-APPL.PRF

b. A-ki-n-gámbye kulwábwe.
SM3SG-OM7-OM1SG-tell.PRF for.them
'He has told it to me for them.'

4.3 Constraint of the 1st person singular OM *n-*

Marlo (2014) discusses a unique behavior of the 1st person singular OM *n-* when it appears alongside the reflexive as shown in (55).

(55) Unique properties of 1SG and reflexive OP[10] (Marlo 2014: 5)

a. The 1st person singular OP and the reflexive are generally required to surface closest to the verb stem (Polak 1983: 297) and may therefore be in different morphological or syntactic positions from other OPs (Buell 2005, Muriungi 2008).

b. The 1st person singular and reflexive are the highest on animacy-topicality and person-number hierarchies, which are known to play a role in object marking (Alsina 1994, Contini-Morava 1983, Duranti 1979, Rugemalira 1993).

c. Most OPs have a CV- shape, but 1SG and reflexive are generally unique in having monophone N- and V-.

As we have seen in §3.3 and §4.2, this is true for Ganda as well; the 1st person singular OM *n-* must always be placed IBS regardless of its semantic role. This is not a preference or tendency, but rather is obligatory.

This constraint only holds for 1st person singular, not for 1st person plural. Unlike the case of 1st person singular OM, the 1st person plural OM can appear even when it is not the benefactive as shown in (56).

[10] OP = Object Prefix, referred to as the OM in this chapter.

(56) A-kí-tú-ba-gámb-idde.
SM3SG-OM7-OM1PL-OM3PL-tell-APPL.PRF
'He has told it to us for/on behalf of them.'

The restriction that the 1st person singular OM must be placed at IBS does not seem to be due to the hierarchy, since it only holds for 1st person singular and not for 1st person plural. Of course, it is possible that there is a difference in the hierarchy between singular and plural. As mentioned in (55b), the 1st person singular OM is the highest on the animacy-topicality and person-number hierarchies. However, all the asymmetrical characteristics that can be observed, due to the hierarchy, are not as rigid as the rule regarding the placement of the 1st person singular OM. Therefore, we should think of this constraint on the 1st person singular OM independently of the restrictions on other objects.

The most likely explanation is a morpho-phonological one. All OMs in Ganda take the form CV except for the 1st person singular OM. Only the 1st person singular OM *n-* does not itself constitute a syllable (see (55c)). It must therefore appear either alongside the vowel of the TAM marker, or merge with the initial consonant of the stem in order to form a syllable. This means that it cannot appear in front of other OMs. However, it is still "the most likely explanation", but remains a topic that requires further research.

4.4 Findings and summary of triple-object constructions in Ganda

Based on what we have discussed in §4.1–§4.3, the following generalizations can be identified (57) regarding the triple-object construction in Ganda:

(57) a. The triple-object construction occurs with the applicative form of ditransitive verbs.
b. Placing three object NPs after the verb is not accepted; therefore, at least one of the three objects must appear as an OM.
c. The applied object must always appear as the OM and cannot appear as an NP.
d. Up to three OMs can appear at the same time.
e. The OM that appears IBS necessarily has the benefactive interpretation.
f. When the verb has an OM which agrees with the 1st person singular object, it must be placed IBS and must be the benefactive.

As summarized in (57), although verbs in Ganda cannot be followed by three object NPs, the triple-object construction itself is possible.

In the triple-object construction, the applied object must always be expressed as an OM, and this is a further restriction unique to the triple-object construction. In the double-object construction, the ordering of OMs is asymmetrically determined by semantic role, but pronominalization is symmetric in that any of the two objects can be freely pronominalized. In the triple-object construction, then, the semantic role hierarchy also figures crucially in pronominalization. As a result, the triple-object construction turns out to be more rigidly asymmetrical.

5 Conclusion

We have shown how object NPs and OMs behave in multiple-object constructions in Ganda.

Considering the symmetrical versus asymmetrical nature of objects in Ganda, both objects in a double-object construction can be (i) placed IAV,[11] (ii) pronominalized, and (iii) passivized. According to the criteria proposed by Bresnan & Moshi (1993), Ganda seems to be a "perfect" symmetrical object language. Ssekiryango (2006) claims that Ganda is a symmetrical language although he also shows asymmetrical data. The data we presented in this chapter are perhaps more supportive of Ssekiryango (2006) than his data in support of the idea that Ganda is a symmetrical language. Even if the resulting interpretation of the sentence is ambiguous, objects still behave symmetrically. This can be seen as strong evidence that Ganda is a symmetrical object language.

However, this language also shows some characteristics of an asymmetrical object language. There is a preference to place the recipient, benefactive, or causee at the IAV position, to pronominalize these elements and to passivize them as opposed to the theme. A preference to treat the recipient, benefactive, and causee as the primary objects can be seen with respect to all three criteria. In addition to these criteria, the order of OMs is fixed asymmetrically. These asymmetrical features are affected by the semantic role, rather than the animacy hierarchy.

It is not surprising to find some asymmetrical features in the languages whose objects are considered as symmetrical, and this seems to be very natural. It is unnecessary to reconsider such languages as asymmetrical languages according

[11] As we mentioned in §3.1.1, Ranero (2019) claims that the order of postverbal objects in the ditransitive is strictly "goal/ben – theme". However, according to our data, both "goal/ben – theme" and "theme – goal/ben" orders are acceptable.

to such asymmetrical features. Presumably, it must be common that some asymmetrical characteristics exist even in a language which is considered predominantly symmetrical, and languages in which objects behave completely symmetrical must be very rare. However, it is still meaningful, especially for a microvariation study, to find out where or based on which semantic (or other grammatical) features such asymmetrical characteristics can be observed. Interestingly, in Ganda, the semantic role affects the preference or naturalness of the order of OMs, but the factor which most strongly affects the order of the OMs is a morpho-phonological condition, particularly for the 1st person singular OM *n-*. Therefore, there seems to be different kinds of restrictions in Ganda, one based on semantic roles, and the other, presumably a morpho-phonological condition.

Summarizing the symmetry/asymmetry in Ganda, the objects show a symmetrical nature with respect to ordering and passivization, with some exceptions in cases where the animacy of both objects are the same. They are symmetrical with respect to pronominalization, but not in the ordering of OMs. The asymmetry emerges as a result of the hierarchical principles of semantic roles rather than animacy. Regarding the possible number of OMs, although Ganda has previously been considered a language that allows two OMs (Ssekiryango 2006, Marlo 2015), we have shown that three OMs are possible, and the conditions where this is possible.

The emergence of the asymmetrical characteristics in Ganda can be seen as part of a more general concept of the "emergence of the unmarked" (cf. Bresnan 1997, 2001), whereby a grammatical property or restriction that is usually not observed, or "hidden", suddenly surfaces or comes in effect in a corner of a language under a certain condition. Many "exceptions" in language might be then understood as "emergent properties" once we recognize the relevant, crucial conditioning factors.

Acknowledgements

This paper is based on our presentation at the 9th World Congress of African Linguistics held in 2018 in Rabat, Morocco. We would like to thank all those who gave us helpful comments during the conference. We would also like to thank the reviewers, who provided very helpful comments and suggestions for the previous versions of this paper. Of course any mistakes are our own responsibility. This research is supported by JSPS Grants-in-Aid for Scientific Research (Grant Number 19K00550, 19H01254) and the ILCAA Joint Research Project "Typological Study of Microvariation in Bantu (2)".

References

Alsina, Alex. 1994. Bantu multiple objects: Analyses and fallacies. *Linguistic Analysis* 24. 153–174.

Ashton, Ethel O., Enoch M. K. Mulira, E. G. M. Ndawula & Archibald Norman Tucker. 1954. *A Luganda grammar*. London, New York & Toronto: Longmans, Green & Co.

Bresnan, Joan. 1997. The emergence of the unmarked pronoun: Chichewa pronominals in Optimality Theory. In Ashlee C. Bailey, Kevin E. Moore & Jeri L. Moxley (eds.), *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society: Special Session on Syntax and Semantics in Africa*, 26–46. Berkeley, CA: Berkeley Linguistics Society.

Bresnan, Joan. 2001. Emergence of the unmarked pronoun. In Géraldine Legendre, Sten Vikner & Jane Grimshaw (eds.), *Optimality theoretic syntax*, 114–142. Cambridge, MA: MIT Press.

Bresnan, Joan & Lioba Moshi. 1993. Object asymmetry in comparative Bantu syntax. In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 47–91. Stanford, CA: CSLI Publications.

Buell, Leston. 2005. *Issues in Zulu verbal morpho-syntax*. Los Angeles, CA: University of California. (Doctoral dissertation).

Contini-Morava, Ellen. 1983. Ranking of participants in Kinyarwanda: The limitations of arbitrariness in language. *Anthropological Linguistics* 25. 425–435.

Duranti, Alessandro. 1979. Object clitic pronouns in Bantu and the topicality hierarchy. *Studies in African Linguistics* 10. 31–45.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list*. Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.

Hyman, Larry M. & Alessandro Duranti. 1982. On the object relation in Bantu. In Paul J. Hopper & Sandra A. Thompson (eds.), *Syntax and semantics: Studies in transitivity*, 217–239. New York: Academic Press.

Maho, Jouni Filip. 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 (20 March, 2013).

Marlo, Michael R. 2014. Exceptional patterns of object marking in Bantu I. *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). 1–65. DOI: 10.1515/jall-2015-0001.

Marten, Lutz & Nancy Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30(2). 237–253.

Marten, Lutz, Nancy Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105. 253–338.

Muriungi, Peter. 2008. *Phrasal movement inside Bantu verbs: Affix scope and order in Kĩĩtharaka*. Tromsø: University of Tromsø. (Doctoral dissertation).

Polak, Louise. 1983. Le réfléchi en bantou. *Africana Linguistica* 9(1). 271–304.

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 papers from the 47th Annual Conference on African Linguistics*, 595–622. Berlin: Language Science Press.

Rugemalira, Josephat M. 1993. Bantu multiple "object" constructions. *Linguistic Analysis* 23. 226–252.

Ssekiryango, Jackson. 2006. Observations on double object construction in Ganda. In Olaoba F. Arasanyin & Michael A. Pemberton (eds.), *Selected Proceedings of the 36th Annual Conference on African Linguistics*, 66–74. Somerville, MA: Cascadilla Proceedings Project.

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 Generative Syntax 6), 199–234. Berlin: Language Science Press. DOI: 10.5281/zenodo.3541755.

Zeller, Jochen. 2014. Three types of object marking in Bantu. *Linguistische Berichte* 239. 347–367.

Chapter 8

Object marking in four Mozambican Bantu languages

Armindo Ngunga[a] & Crisófia Langa da Câmara[a]
[a]Centro de Estudos Africanos Universidade Eduardo Mondlane

Object marking in Bantu is an area which has received substantial attention (e.g. Marten & Kula 2012, Riedel 2009, Marten et al. 2012, Zeller 2014). In many languages of this group, a morpheme which is co-referential with the object can also be incorporated into the verb structure. The present chapter, which looks at data from four Mozambican languages, aims to document and describe the situation in these languages with respect to object marking. The data show that the languages analyzed in this paper can be divided into three groups: Group 1, composed of Cinyungwe and Ciwutee, in which object marking is not obligatory with particular object NPs; Group 2 comprising of Citshwa, in which the object marker and the object argument cannot co-occur, object marking is not obligatory with particular object NPs and there is no locative object markers; and Group 3, Ciyaawo, in which object markers are not obligatory with particular object NPs and only the benefactive object can be expressed by an object marker in double object constructions. Taking into account the data from the four languages, we suggest that the obligatory requirement for an object marker [+OM] associated with some transitive verbs and structures should be added as a seventh parameter to the six parameters of variation in object marking in Bantu put forward by Marten & Kula (2012). Due to the existence of transitive verbs subcategorized as [+OM], we further encourage scholars to examine these parameters of variation in other Bantu languages in light of these features of variation.

1 Introduction

The Bantu languages are known for the systematic way in which grammatical relations are morphologically marked in the verbal structure (Ngunga (2014)). In

Armindo Ngunga & Crisófia Langa da Câmara. 2024. Object marking in four Mozambican Bantu languages. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 191–227. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663775

the case of object marking, many of these languages exhibit agreement with both the subject and the object. Agreement with the subject is usually grammatically obligatory, while the status of object marking is often more pronominal (Marten & Ramadhani 2001) and may be optional. Object markers are affixes or clitics that identify and cross-reference an object argument. Thus, the phenomenon of object marking on verbs in Bantu languages is a mechanism for referring to discourse-familiar entities, similar to pronominalization (Sikuku et al. 2018). In this paper, we use the term object marking to refer to the way a lexical object is reflected within the verb structure through a co-referential marker (Baker 1988, Corbett 2006, Deen 2006, amongst others).

The ambiguity of the status of object markers in Bantu has received considerable attention in the literature (Marten & Kula 2012, Riedel 2009, Marten et al. 2012, Zeller 2014 among others). In Bantu languages, a wide range of prefixes (subject, object, tense, aspect, mood, negation, and other markers) and suffixes (derivational and inflectional) can be attached to the lexical verb root. In many languages, the object markers (OMs) are attached directly to the verb stem. See the examples in (1) and (2) presented below where the objects are cross-referenced by the forms wa- and ci- respectively.

(1) Kiswahili (G42) (Riedel 2009: 46)
A-li-**wa**-won-a
SM1-PAST-OM2-see-FV
'he saw them'

(2) Cinyungwe (N43)
Iye a-da-**ci**-mog-a.
he SM1SG-PFV-OM7-jump-FV
'he jumped it'

The slot immediately before the verb root has largely been identified as the OM slot in the Bantu verb structure (Ngunga (2014)) . However, there is no such consensus about the grammatical status of OMs as pronominal or agreement markers (Bresnan & Mchombo 1987, Deen 2006, Riedel 2009), since in individual languages the OMs behave differently.

The object argument may or may not co-occur with the object marker depending on a series of syntactic, pragmatic and semantic factors. This means that there is a difference between Bantu languages concerning the possibility of the co-occurrence of the object marker and the corresponding object argument (Ngunga 2014, Zeller 2014). The example in (3) illustrates that the co-occurrence

of the object marker and the object argument is possible in IsiZulu, but it is not possible in Kinyarwanda, as example (4) illustrates.

(3) Isizulu (S42) (Zeller 2012: 219)
A-ba-ntwana ba-ya-si-thand-a lesi si-kole
AUG-2-child SM2-DIS-OM7-like-FV DEM7 7-school
'the children like this school.'

(4) Kinyarwanda (L12) (Riedel 2009: 76)
* A-ba-aana ba-a-ra-bi-ri-ye i-bi-ryo ejo.
AUG-2-child SM2-REM-DIS-OM8-eat-PFV AUG-8-food yesterday
Intended: 'the children ate the food yesterday'

According to Iorio (2015), the co-occurrence of the object marker and the co-referring object argument is only possible if the latter is right dislocated. This dislocation results in a definiteness and specificity effect on the object arguments with which they co-occur. See the Bembe examples below:

(5) Bembe (D54) (Iorio 2015: 196)

a. mwana a-a-yak-a ngyoʔa. [VO]
SM1-child SM1.SG-PFV-kill-FV SM9.snake
'the child has killed a/*the snake.'

b. mwana a-a-**ya**-yak-a. [OM-V]
SM1-child SM.SG1-PFV-OM9-kill-FV
'the child has killed it.'

c. * mwana a-a-**ya**$_i$-yak-a ngyoʔa$_i$ *[OM-V O]
SM1-child SM1.SG-PFV-OM9-kill-FV SM9.snake
Intd.: 'the child has killed a/he snake'

d. mwana a-a-**ya**$_i$-yak-a, ngyoʔa$_i$ [OM-V] [O]
SM1-child SM1.SG-PFV-OM9-kill-FV SM9.snake
'the child has killed it, the/*a snake (that is)'

The co-occurrence of the object marker and the nominal object is not the only variation that is found in object marking in Bantu languages. According to van der Wal (2015), in languages which allow the occurrence of the object marker and the co-referring, there is a great deal of variation as to which objects are marked by an object marker. In Nyarutu, for example, it is usually the animate, definite and/or given objects that are doubled by an object marker (van der Wal 2015, 2016). Therefore, the example in (6b) is ungrammatical because animate objects must be doubled by an object marker.

(6) a. Nyarutu (F32) (van der Wal 2015: 6 [via Hualde 1989: 182])
n-a-**mʊ**-onaa Maria.
SM1SG-PST-1OM-see 1.Maria
'I saw Maria.'

b. * n-a-onaa Maria.
SM1SG-PST-OM1-see 1.Maria
Intd: 'I saw Maria.'

At this point we notice that although OMs occur in the verbal structure of several Bantu languages, their occurrence and/or co-occurrence with the object argument is determined by several factors. According to Marten & Ramadhani (2001), in Kiluguru (G30), a language spoken in Tanzania, the distribution of object marking in simple transitive predicates is partly motivated by the syntactic context and semantic aspects. Moreover, in some cases, the object is interpreted as being more definite when the object agreement is present. The other, probably more important, dimension to the analysis of object marking in Kiluguru is pragmatic, since the use of object agreement is more related to the anaphoric structure of the discourse and to evaluations by speakers in relation to what they hear. Marten & Ramadhani (2001) also claim that in this language, in contrast to simple predicates, verbs such as *-ona* 'see', *-phika* 'find' and *-ing'a* 'give' obligatorily require the use of the OM with the lexical object. This is a point to which we will return also for the languages under examination here in (§3) also for the languages under examination.

The high degree of diversity in the morphological marking of object arguments in Bantu languages is reflected in the diversity of proposals made by several authors using data from different languages. Marten & Kula (2012), for instance, present six parameters for the investigation of variation in object marking based in 16 Bantu languages.

A different approach, adopted by Zeller (2014), divides Bantu languages into three types, namely: Type 1, where object markers are agreement markers; Type 2, where object markers are pronominal clitics; and Type 3, where object marking is a reflex of A-bar movement of the corresponding object.

Aissen (2002) claims that object marking is based on semantic and pragmatic grounds. In some languages, it is the pragmatic character of the object that determines whether it is obligatory or optional, or if it is excluded. In pragmatic and morphological object marking, the objects that most resemble subjects are overtly case-marked, whereas syntactic objects are obligatorily case-marked if they stand in a position that is more marked for an object. This is the case for

languages like Hungarian and Malayalam (see Aissen (2002)) for further details. For Aissen (2002: 437), the prominence scale for direct objects is as follows:

Animacy: Human > Animate > Inanimate

Definiteness: Personal pronoun > Proper name > Definite NP > Indefinite specific NP > Non-specific NP

This variation in object marking reflects the tension between two principles: iconicity, which prefers semantic markedness to be expressed by morphology, and economy, which would rather be devoid of structure whenever possible (Aissen 2002).

The current paper seeks to contribute to the understanding of object marking in Bantu by applying Marten & Kula's (2012) six parameters to four Mozambican Bantu languages. We also develop a seventh parameter for the four languages analyzed in this paper, which is related to the obligatoriness of object markers with specific transitive verbs and specific structure as we shall see in §3.

This paper is organized as follows. Following this introduction, we present Marten & Kula's (2012) six parameters (§2). We then apply these parameters to four Mozambican Bantu languages (§3), before presenting some conclusions (§4).

2 A parametric approach to object marking in Bantu

Studies of object marking have shown differences regarding the realization of object markers in Bantu languages. One of these studies is that of Marten & Kula (2012), who identified a number of micro-parameters that determine cross-Bantu variation. Marten & Kula (2012) present six parameters relating to the investigation of the variation in object marking in 16 Bantu languages (7). The languages of their study are: Bemba (M42), Chaga (Kivunjo) (E62b), Chichewa (N31), Ha (D66), Haya (E22), Kinyarwanda (D60), Lozi (K21), Makhuwa (P31), ciNsenga (N41), Otjiherero (R31), Ruwund (L53), Sambaa (G23), siSwati (S43), Kiswahili (G42), Setswana (S31), and Yeyi (R41).

(7) Morphosyntactic parameters of object marking in Bantu (Marten & Kula 2012: 5).
   (i) Can the object marker and the object argument co-occur?
   (ii) Is an object marker obligatory with particular object NPs?
   (iii) Are there locative object markers?
   (iv) Is object marking restricted to one object marker per verb?

(v) Can either benefactive or theme objects be expressed by an object marker in double object constructions?
(vi) Is an object marker required/optional/disallowed in object relatives?

2.1 (i) The co-occurrence of object markers and lexical objects

In some Bantu languages, such as Kiswahili (8), there are no restrictions of co-occurrence of an object marker and a co-referential overt NP. This means that the object marker can be used together with an overt NP. However, in other languages like Otjiherero, the object marker cannot co-occur with an overt NP (9).

(8) Kiswahili (G42) (Marten & Kula 2012: 240)
ni-li-mw_i-on-a Juma$_i$.
SM1SG-PAST-OM1-see-FV Juma
'I saw Juma'

(9) Otjiherero (R31)(Marten & Kula 2012: 240)
* mb-é vé mún-ù òvá-nátjè.
SM1SG-PAST OM2 see-FV 2-children
Intd.: 'I saw the children'

The examples in (8) and (9) illustrate that Swahili behaves differently from Otjiherero. In Swahili (8), the co-occurrence of the object marker (-*mw*-) and the overt object NP (Juma) yields a grammatical result which is not possible in Otjiherero (9).

2.2 (ii) The obligatoriness of object markers with specific classes of objects

This parameter of variation relates to cases where the co-occurrence of object markers and co-referential NPs is obligatory with specific NPs. This can be found in Swahili for example, where object marking is obligatory with animate objects, particularly nouns which refer to humans, as shown in (10a, b) below:

(10) Kiswahili (G42) (Riedel 2009: 46)

a. ni-li-*(**mw$_i$**)-on-a m-toto$_i$
SM-PFV-(OM1)-see-FV SM1-child
'I saw his child'

b. * ni-li-on-a m-toto
SM-PFV-see-FV SM1-child
Intd: 'I saw the child'

In contrast, the use of the object marker in Kiswahili is structurally optional with inanimate NPs, as exemplified in (11) from Marten & Kula (2012: 241):

(11) ni-li-(ki)-on-a ki-tabu
SM-PFV-(OM7)-see-FV 7-book
'I saw a/the book'

There are also languages where the thematic role of the object determines whether it can co-occur with an object marker. For instance, in Ruwund, the benefactive object can co-occur with the object marker (12a), but cannot occur with a theme object (10b).

(12) Ruwund (D62) (Marten & Kula 2012: 241)

a. ka-ma-mu-tum-in mwâan.
INF-OM6-OM1-send-APPL 1.child
'to send the child them'

b. * ka-ma-mu-tum-in mwâan ma-long.
INF-OM6-OM1-send-APPL 1.child 6-plates
'to send the child the plates'

2.3 (iii) The presence of locative object markers

In languages like Cinsenga and Setswana where locative objects can be expressed by locative object markers, locative nouns and locative object markers can co-occur. This is shown in the examples in (13) and (14):

(13) Cinsenga (N41) (Marten & Kula 2012: 243)
ku-Lilongwe n-a-ku-ziw-a.
17-Lilongwe SM1-PRES-OM17-know-FV
'Lilongwe I know it (there)'

(14) Setswana (S31a)
ke a gó itsé.
SM1 PRES OM17 know
'I know it (there)'

However, some other languages do not have locative object markers.

2.4 (iv) The multiple object markers

The fourth parameter of object marking variation proposed by Marten & Kula (2012) is related to the number of object markers allowed per inflected verb form. There are languages that do not allow multiple object markers in the same verb. This is what happens in Bemba which allows one object marker, as shown in (15) below:

(15) Kiswahili

a. ni-li-m-p-a.
SM1-PST-OM1-give-FV
'I gave him (it)'

b. * ni-li-i-m-p-a.
SM1-PST-OM9-OM1-give-FV
Intd: 'I gave him it'

In contrast to the Bemba examples present abov, each lexical object in Sambaa may have its corresponding OM in the verb structure. Consider the example in (16):

(16) Sambaa (L12) (Riedel 2009: 72)
n-za-**ha**-**ci**-**m**-nka Stella ki-tabu.
SM-PFV-OM16-OM7-OM1-give Stella 7-book
'I gave Stella a book there'

This fact has led scholars like Henderson (2006) and Zeller (2014) to argue that in Bantu, OMs can function as agreement markers and pronominal clitics.

It is important to note that in languages with multiple object marking, there is variation as to which objects are marked. In Bemba it is possible to mark more than one object if both object markers are animate (17a) or if the object marker closest to the verb is the first person singular *n-* (17b).

(17) Bemba (M42) (Marten & Kula 2012: 245)

a. mù-kà-**bá**-**mú**-éb-él-á-kó.
SM1-FUT-OM2-OM1-tell-APPL-FV-PRO17
'you will tell them for him.'

b. mú-ká-**cí**-**mù**-**n**-twààl-íj-é-kó.
SM2-FUT-OM7-OM1-OM1-return-APPL-FV-PRO17
'you should return it to him/her for me.'

However, in other languages object markers can co-occur in an unrestricted manner.

2.5 (v) The object marking in double object constructions

According to Marten & Kula (2012), another well-known parameter of variation relates to object marking in double object constructions. There are languages in which only the benefactive object in a double object construction can be object marked and those in which either the benefactive or the theme can be marked.

The ungrammaticality of (18b) illustrates that in Chichewa only the benefactive object can be object-marked on the verb in a double object construction. In contrast, in Otjiherero the theme can also be object-marked (19):

(18) Chichewa (N31a) (Marten & Kula 2012: 247)

a. a-lenje a-ku-wá-phík-ir-á zí-túmbúwa (a-nyani).
2-hunters SM2-PRES-OM2-cook-APPL-FV 8-pancakes 2-baboons
'the hunters are cooking (for) them (the baboons) some pancakes'

b. * a-lenje a-ku-wá-phík-ir-á a-nyani (zí-túmbúwa).
2-hunters SM2-PRES-OM8-cook-APPL-FV 2-baboons 8-pancakes

(19) Otjiherero (R31) (Marten & Kula 2012: 247)
Má-yé ì tjángér-é òvà-nâtjé.
PRES.SM OM9 write-APPL-FV 2-children
'they are writing the children it'

2.6 (vi) The object marking in relative clauses

The last parameter proposed by Marten & Kula (2012) pertains to the use of object markers in object relative clauses. In descriptive terms, three groups of language types can be distinguished: (i) those where object markers are required in object relatives (e.g. Setswana); (ii) those where object markers are optional (e.g. Swahili) and (iii) those where object markers are not allowed in object relative clauses (e.g. Lozi). These three types are illustrated by examples (20), (21) and (22) below.

(20) Setswana (Marten & Kula 2012: 248)

a. di-kwelo tse ke di bone-ng …
10-books REL10 SM1.PAST OM10 see-REL
'the books which I saw them…'

b. * di-kwelo tse ke bone-ng …
10-books REL10 SM1.PAST see-REL
'the books which I saw them…'

The example in (20b) is ungrammatical because the object relative construction does not have an object marker.

The example of the second type of language is exemplified by Swahili, where object marking in object relatives is possible, but not required (21):

(21) Swahili
ki-tabu amba-cho ni-li-(ki-)som-a …
7-books REL-7 SM1-PAST-(OM7)-read-FV
'the book which I read (it)'

The third pattern is shown with the example from Lozi, where object markers in object relatives are not allowed.

(22) Lozi (K21) (Marten & Kula 2012: 248)
* buka ye-ne-ba-(ye)-bon-i ba-nana fa-tafule ki-ye-tuna.
9.book 9.REL-PAST-SM2-(OM9)-see-FV 2-children 16-table COP-SM9-big
Intd: 'the book which the children saw it on the table is big.'

Moreover, example (22) is important because it illustrates that objects are not required in relative constructions.

3 Object marking in four Mozambican Bantu languages

In the present section, we examine properties of object marking in four Mozambican Bantu languages, namely, Cinyungwe, Citshwa, Ciwutee, and Ciyaawo. We analyse these languages using six parameters of variation from Marten & Kula (2012) The languages analyzed in this paper were chosen on the basis of available information and our own knowledge as native speakers of Ciyaawo (first author) and Cinyungwe (second author). In §3.1 we start our discussion by analyzing data from Cinyungwe.

3.1 Object marking in Cinyungwe

Some Bantu languages show restrictions on the co-occurrence of an object marker and the co-referential object argument. The first parameter presented by Marten & Kula (2012) identifies the conditions under which an object marker can co-occur with a corresponding object argument after the verb.

3.1.1 (i) Can the object marker and the object argument co-occur?

Cinyungwe is a Mozambican Bantu language spoken in Tete Province by 457,290 speakers (Ngunga & Faquir 2011: 108). In Cinyungwe, the co-occurrence of the object argument and the OM within the same sentence is possible only if the object is a dislocated adjunct (i.e. it is not *in situ*), as illustrated in (23) below:

(23) Cinyungwe (N43)

a. baba a-da-nyamul-a m-wana.
1.dad SM1-PFV-hold-FV 1-child
'dad held a child'

b. baba a-da-**mu**-nyamul-a.
1.dad SM1-PFV-OM1-hold-FV
'dad held (her/him) the child'

c. * baba a-da-**mu$_i$**-nyamul-a m-wana$_i$
1.dad SM1-PFV-OM1-hold-FV 1-child
Intd: 'dad held (her/him), the child'

d. baba a-da-**mu$_i$**-nyamul-a, (m-wana)$_i$
1.dad SM1-PFV-OM-hold-FV 1-child
'dad has held (her/him), the child'

In example (23a) *mwana* 'child' is non-specific. The presence of the OM *-mu-* in (23b) means that this is an appropriate response to a question such as "What did dad do to the child?". The example in (23c) is ungrammatical because the object marker and the NP co-occur, which is prohibited in Cinyungwe. The example in (23c) shows that in Cinyungwe doubling an object marker with an *in situ* object is unacceptable in neutral discourse contexts. Note however that this sentence is acceptable in a context in which the speaker wants to convince the hearer that the action happened and s/he even saw father holding the child, i.e. for emphatic purposes or for certainty. The pause after the verb in example (23d) is obligatory and indicates that the NP is dislocated, and represents the only way such a sentence is acceptable in this context.

3.1.2 (ii) Is an object marker obligatory with particular object NPs?

The other aspect of variation with respect to the co-occurrence of the OM and the object argument found in Bantu languages relates to whether an object marker is obligatory with a specific object argument. In Cinyungwe, object marking is

not obligatory with specific object arguments of any type. See the examples presented below:

(24) Cinyungwe

a. mw-ana a-da-won-a ng'ombe.
1-child SM1SG-PFV-see-FV 9.cow
'the child saw the cow'

b. mw-ana a-da-**yi**-won-a, (ng'ombe).
1-child 1SM-PFV-OM9-see-FV 9.cow
'the child saw it, (the cow)'

(25) a. mw-ana a-da-won-a mu-ti.
1-child 1SM-PFV-see-FV 3-tree
'the child saw the tree'

b. mw-ana a-da-**wu**-won-a, (mu-ti).
1-child 1SM-PFV-OM9-see-FV 3-tree
'the child saw it, (the tree)'

The examples presented above illustrate that in Cinyungwe the OM is not obligatory with a specific object argument because as can be seen (24) the object is an animate and in (25), the object is an inanimate. Nonetheless, the co-occurrence of non-animate NPs and object marker is related to definiteness or specificity. In (24), the object argument is animate while in (25), the object argument is non-animate.

However, in contrast to what we described in (24) and (25) above, object marking with the verb *-wona* 'to see' is obligatory. See the examples in (26) and (27) below.

(26) a. a-da-??(**mu**)$_i$-won-a iye$_i$ dzulo.
1SM-PFV-OM1-see-FV he yesterday
'he saw him yesterday'

b. a-da-??(**wa**)-won-a iwo dzulo.
1SM-PFV-OM2-see-FV they yesterday
'they saw them yesterday'

c. a-da-??(**wa**)-won-a yavu dzulo.
1SM-PFV-OM2-see-FV grandma yesterday
'they saw her (the grandma) yesterday'

(27) a. a-da-**mu**$_i$-pas-a iye$_i$ ci-mbamba.
1SM-PFV-OM1-give-FV he 7-beans
'he gave him beans'

b. a-da-**wa**$_i$-pas-a iwo$_i$ ci-mbamba.
1SM-PFV-OM2-give-FV they 7-beans
'he gave them beans'

c. a-da-**mu**$_i$-pas-a ci-mbamba mayi.
1SM-PFV-OM1-give-FV 7-beans 1.mother
'he gave her (the mother) beans'

The data presented in (27) illustrate that in Cinyungwe object marking is obligatory with pronominal objects with the verb *-won-* 'to see'. Moreover, examples (26c) and (27c) illustrate that the obligatoriness of the object marker in the verb may relate to the verb and not the pronominal object per se.

3.1.3 (iii) Are there locative object markers?

In Cinyungwe, locative objects can be expressed by locative object markers and they can co-occur with their overt locative nouns but not in neutral context. Consider the examples in (28):

(28) a. pa-xikola$_i$, nd-a-(**pa$_i$**)-yend-a.
16-school SM1SG-PFV-(OM16)-go-FV
'to school, I (really) went to (there)'

b. ku-muyi$_i$, u-ndza-(ku$_i$)-pit-a
17-home SM2SG-FUT-(OM17)-pass-FV
'home, you will (really) pass by (it)'

c. * mu-nyumba u-da-**mu**-pit-a
18-house SM2SG-PFV-OM18-pass-FV
Intd: 'inside the house, you will pass by'

d. pa-xikola$_i$ nd-a-*(**pa$_i$**)-won-a.
16-school SM1SG-PFV-(OM16)-see-FV
'school I saw (there)'

In (28a) and (28b), we see that only class 16 and 17 locative objects can be expressed by locative object markers on the verb and that locative object markers cannot co-occur with locative objects in the same clause. It is important to note that this co-occurrence happens when the speaker wants to expresses his or her

knowledge concerning an issue. In example (28c), class 18 cannot be expressed by a locative object marker in the verb structure while in (28d), omission of the object marker renders the sentence ungrammatical. This means that the verb *-wona* 'to see' requires an object marker.

3.1.4 (iv) Is object marking restricted to one object marker per verb?

Another parameter discussed in Marten & Kula (2012) that we focus on here concerns the number of object markers that can occur in an inflected verb structure. In Cinyungwe, only one object marker per inflected verb is permitted. See example (29):

(29) a. mw-ana a-da-won-es-a Siriza mu-ti.
1-child SM1-PFV-see-CAUS-FV Siriza 3-tree
'the child made Siriza see the tree'

b. mw-ana a-da-(*__mu)-wu__-won-es-a, Siriza mu-ti.
1-child SM1-PFV-(OM9)-OM3-see-CAUS-FV Siriza 3-tree
'the child saw it, (the tree)'

c. mw-ana a-da-(*__wu)-mu__-won-es-a, Siriza mu-ti.
1-child SM1-PFV-(OM3)-OM1-see-CAUS-FV Siriza 3-tree
'the child saw it, (the tree)'

Example (29b) illustrates that only one object marker is permitted. In this sentence the class 3 object marker occurs immediately before the verb root and (29c) shows that changing the order of the object markers does not alter the ungrammaticality of the sentence.

3.1.5 (v) Can either benefactive or theme objects be expressed by an object marker in double object constructions?

In Cinyungwe, either benefactive or theme objects can be expressed by an object marker in double object constructions. This is illustrated in the examples in (30).

(30) a. Mayi a-da-__mu__-phik-ir-a ci-manga, Siriza.
Mayi SM1-PFV-OM1-cook-APPL-FV 7-maize Siriza
'the mother cooked her (Siriza) maize'

b. Mayi a-da-__ci__-phik-ir-a Siriza, ci-manga.
Mayi SM1-PFV-OM7-cook-APPL-FV Siriza 7-maize
'the mother cooked Siriza it (the maize)'

3.1.6 (vi) Is an object marker required/optional/disallowed in object relatives?

In Cinyungwe, object markers are generally optional in object relative clauses (31), although again it is not allowed to mark the object argument with the verb *-won-* 'to see' (31).

(31) ma-bvembe y-omwe mayi a-ndza-(ma)-bweres-a yanitapira …
6-watermelon REL-6 1.mother SM1-FUT-(OM6)-bring-FV sweet
'the watermelons that mum shall bring (them) are sweet'

(32) ma-bvembe y-omwe mayi a-ndza-*(ma)-won-a yanitapira …
6-watermelon REL-6 1.mother SM1-FUT-(OM6)-see-FV sweet
'the watermelons that mum shall see (them) are sweet'

The difference between the examples in (31) and (32) reflects the different object marking properties associated with different verb types in Cinyungwe. We do not explore the impact of the verb types on object marking properties in any further detail here although this would be a good avenue for future research.

It is important to note that verb types are not part of the Marten & Kula (2012) parameters and in this paper, we add verb types as seventh parameter. In terms of the parameters under examination here, the answers for Cinyungwe are "yes" for the five parameters (ii), (iii) and (iv), (v) and (vi) and "no" for (i) and (ii) (Table 1).

Table 1: Parametric variation in object marking in Cinyungwe

(i)	Can the object marker and the object argument co-occur?	✗
(ii)	Is an object marker obligatory with particular object NPs?	✗
(iii)	Are there locative object markers?	✓
(iv)	Is object marking restricted to one object marker per verb?	✓
(v)	Can either benefactive or theme objects be expressed by an object marker in double object constructions?	✓
(vi)	Is an object marker required/optional/disallowed in object relatives?	✓
(vii)	Is an object marker obligatory with particular verb?	✗

After presenting data of object marking in Cinyungwe, in the next section we look at object marking in Citshwa.

3.2 Object marking in Citshwa

Citshwa is a Mozambican Bantu language with 693,386 speakers. Speakers are found in the three southern provinces Inhambane, Gaza and Maputo and in two central provinces Manica and Sofala (Ngunga & Faquir 2011). Citshwa has six dialects: Xikhambani, spoken in Panda District; Xirhonga, spoken in Massinga; Xihlengwe, spoken in Morrumbene, Massinga and Funhalouro Districts; Ximhandla, spoken in Vilankulo District; Xidzhonge (or Xidonge), spoken in Inharrime District; Xidzivi, spoken in Morrumbene and Homoine Districts. The data analyzed in this paper were provided by a speaker of the Ximhandla dialect via elicitation.

3.2.1 (i) Can the object marker and the object argument co-occur?

Ngunga (2014) shows that an object marker and the object argument can co-occur in Citshwa, and provides the examples in (33) to support this observation.

(33) Tshwa (S51) (Ngunga 2014: 187)

a. Polina a-nyik-ile pawu ci-n'wanana
Polina SM1-give-PFV 5.bread 7-child(a.small.one)
'Polina gave the child some bread'

b. Polina a-**ci**$_i$-nyik-ile ci$_i$-n'wanana$_i$ pawu
Polina SM1-OM7-give-PFV 7.child(a.small.one) 5.bread
'Polina gave the child some bread'

c. Polina a-**ci**-nyik-ile pawu
Polina SM1-OM7-give-PFV 5.bread
'Polina gave her bread'

d. Polina a-**gi**-nyik-ile
Polina SM1-OM7-give-PFV
'Polina gave her (it)'

In (33a), there is no OM present in the verb structure. In (33b), the class 7 OM prefix is co-referential with the indirect lexical object NP *cin'wanana* 'child'. These examples show that when there are two objects, a direct and an indirect object, that it is the indirect object with which the OM in the verb structure agrees. It is also worth noting that the word order changes in such cases. While in (33a) the word order is S-V-DO-IO (subject, verb, direct object, indirect object), in (33b) the word order is S-V-IO-DO, which seems to suggest a locality (adjacency) principle in the agreement between the OM and the indirect lexical object. In

(33c), the indirect object noun is not realized but the construction is acceptable if it is part of a conversation where the referent can be recovered from context. This is also what happens in (33d) where the OM cross-references a class 7 noun.

In (34), we present another example which shows that there are important different pragmatic interpretations to be considered when the object marker and the object argument co-occur in an intransitive verb in Citshwa. According to our consultant, in example (34b), the co-occurrence of the object marker and the object argument does not appear out of the context. For him, any Citshwa speaker hearing this sentence out of the context can ask, "Which meat are you talking about?", "Why are you telling me that?". Thus, it seems like in (34b), we are talking about a specific meat. In Citshwa, OM-doubling brings this specificity and giveness reading of the object. That is why we propose that in Citshwa, the object and the co-referring direct object cannot co-occur out of the blue.

(34) a. mu-fana w-a-g-a nyama
1-boy SM1.PRS-eat-FV 9.meat
'The boy eats the meat'

b. mu-fana w-a-**yi**-g-a nyama
1-boy SM1.PRS-OM9-eat-VF 9.meat
'The boy eats the meat'

In Citshwa, there are cases where the co-occurrence of the object argument and the OM within the same sentence has a different meaning to the one described in (34b) above. Thus, if the speaker avoids the co-occurrence of the object argument and the object marker by dislocating the object argument, this results in emphasis on how the boy loves eating meat. An example of the co-occurrence of the object argument and the object marker and the resulting interpretation is shown in (35) below:

(35) mu-fana w-a-**yi**-g-a, nyama
1-boy SM1.PRS-OM9-eat-VF 9.meat
'the boy eats a lot of meat'

The example in (35) can also have a totality interpretation when the speaker is telling the hearer not to be afraid thinking that the boy shall not finish the meat the hearer is giving him because as he knows, the boy loves meat and he can eat it with the bones.

3.2.2 (ii) Is an object marker obligatory with particular object NPs in Citshwa?

In Citshwa, animate as well as inanimate objects can appear with the object marker, although its presence is optional in both cases. See the examples in (36a–d), presented below.

(36) a. Zabhela a-won-ile mbzana
Zabhela SM1-see-PFV 9.dog
'Zabhela saw a dog'

b. Zabhela a-(**yi**)-won-ile mbzana
Zabhela SM1-(OM9)-see-PFV 9.dog
'Zabhela saw it'

c. Tereza a-tsal-ile papilu
Tereza SM1-write-PFV 5.letter
'Teresa wrote the letter'

d. Tereza a-(**gi**)-tsal-ile papilu
Tereza SM1-(OM5)-write-PFV 5.letter
'Tereza wrote it'

The examples presented in (36) above illustrate that object marking is not obligatory. That is, sentences in (36a–b) are still grammatical even if the object marker is not present. In (36c–d), where the object argument is an inanimate argument, we can see that the occurrence of the object marker in the verb structure is still not obligatory. The difference between these examples is that (36a) and (36c) are statements and (36b) and (36d) are context-based sentence structures. They are used to clarify what was not previously understood in the first statement. There is also another interpretation statement that can be added in the interpretation of (36b) and (36d). For our Citshwa speaker, the example in (36b) and (36d) can also be used for emphatic purposes where the speaker is trying to make clear how beautiful the words written in the paper were.

Thus, the OM may not be obligatory but the presence or absence of the OM changes the interpretation of each sentence.

3.2.3 (iii) Are there locative object markers?

In Citshwa, there are no locative prefixes of the form similar to the ones we described in Cinyungwe (cf. §3.1). Therefore, locativization is expressed by the suffix *-eni* attached to the NP. The examples in (37) illustrate that the locative object marker is only recovered from the verb for class 17.

(37) a. ci-kolw-eni, u-ta-famb-a
7-school-LOC SM2SG-FUT-go-FV
'to school, you will go'

b. ci-kolw-eni, u-ta-**ku**-famb-a
7-school-LOC SM2SG-FUT-OM17-go-FV
'to school, you will go there (wanting or not)'

c. ndlw-ini, u-ta-nghen-a
10-house-LOC SM2SG-FUT-enter-FV
'in the house, you will get in'

d. ndlw-ini, u-ta-**(ku)**-nghen-a
10-house-LOC SM2SG-FUT-OM17-enter-FV
'in the house, you will get in'

The absence of examples with locative object markers for class 16 and 18 in (37) indicates that Citshwa does not have OMs for these classes. The answer to this parameter from Marten & Kula (2012) is therefore "no" for Citshwa.

3.2.4 (iv) Is object marking restricted to one object marker per verb?

In Citshwa, only one object can be realised as an object marker for each inflected verb. See the examples in (38) and (39):

(38) a. bava a-bhik-is-a zva-kuga nhanyana
1.father SM1.PRS-cook-CAUS-FV 8-food 1.girl
'the father made the girl cook the food'

b. bava a-**mu**-bhik-is-a zva-kuga nhanyana
1.father SM1.PRS-OM1-cook-CAUS-FV 8-food 1.girl
'the father made her (the girl) cook the food'

c. bava a-(*__zva__)-__mu__-bhik-is-a zva-kuga nhanyana
1.father SM1.PRS-(OM8)-OM1-cook-CAUS-FV 8-food 1.girl
Intd: 'the father made her (the girl) cook it (the food)'

d. bava a-(*__mu__)-__zva__-bhik-is-a zva-kuga nhanyana
1.father SM1.PRS-(OM1)-OM8-cook-CAUS-FV 8-food 1.girl
Intd: 'the father made her (the girl) cook it (the food)'

(39) a. mamani a-rim-el-a bava zvi-pfhaki.
1.mother SM1.PRS-cultivate-APPL-FV 1.father 8-maize
'the mother cultivates maize for the father'

b. mamani wa-(*__zvi__)-__mu__-rim-el-a bava zvi-pfhaki.
1.mother SM1.PRS-(OM8)-OM1-cultivate-APPL-FV 1.father 8-maize
'the mother cultivates it (the maize) for him (the father)'

c. mamani wa-(*__mu__)-__zvi__-rim-el-a bava zvi-pfhaki
1.mother SM1.PRS-(OM1)-OM8-cultivate-APPL-FV 1.father 8-maize
'the mother cultivates it (the maize) for him (the father)'

In (38) and (39), we have examples that illustrate that there is a space for only one object marker in the Citshwa verb structure.

3.2.5 (v) Can either benefactive or theme objects be expressed by an object marker in double object constructions?

In Citshwa, either the benefactive or theme object can be expressed by an object marker in double object constructions. This is illustrated by examples (40b) and (40c) below which illustrate that either the benefactive or the theme object can be object marked.

(40) a. bava a-nyik-a ti-manga mu-nghana.
1.father SM1.PRS-give-FV 10-peanuts 1-friend
'the father gave the friend peanuts'

b. bava wa-__ti__-nyik-el-a mu-nghana.
1.father SM1.PRS-OM10-give-APPL-FV 1-friend
'the father is giving them (the peanuts) on behalf of his friend'

c. bava wa-__mu__-nyik-el-a ti-manga.
1.father SM1.PRS-OM1-give-APPL-FV 10-peanuts
'the father is giving the peanuts for him (the friend)'

This means that Citshwa is a "symmetrical" language with respect to object marking in double constructions (cf. Bresnan & Moshi 1990).

3.2.6 (vi) Is an object marker required/optional/disallowed in object relatives?

The last parameter presented by Marten & Kula (2012) has do to with the availability of object markers in relative clauses. In Citshwa, object markers are obligatory only with the verb *-won-* 'to see'. Compare the examples (41) and (42).

(41) a-ma-din'wa a-nga-(**ma**)-xav-a mamani ma-andziha
AUG-6-orange SM1-PERF.REL-(OM6)-eat-FV 1.mother 6-sweets
'the oranges that mother bought (them) are sweet'

(42) a-madin'wa a-nga-*(**ma**)-won-a mamani ma-nandziha
AUG-6-orange SM1.PERF.REL-(OM6)-see-FV 1.mother 6-sweet
'the oranges that mother saw (them) are sweet'

The example presented in (41) illustrates that object markers are not obligatory in Citshwa. However, just like we saw when we were analyzing object marking in Cinyungwe, example (42), shows that it is obligatory to object mark the object argument in relative constructions in Citishwa. This can be related to what we described in section 3.2.1, in Citshwa the co-occurence of the object marker and the object argument is disallowed, making them optional. Table 2 summarizes the object marking properties in Citshwa.

Table 2: Parametric variation in object marking in Citshwa

(i)	Can the object marker and the object argument co-occur?	✓
(ii)	Is an object marker obligatory with particular object NPs?	✗
(iii)	Are there locative object markers?	✗
(iv)	Is object marking restricted to one object marker per verb?	✓
(v)	Can either benefactive or theme objects be expressed by an object marker in double object constructions?	✓
(vi)	Is an object marker required/optional/disallowed in object relatives?	✓
(vii)	Is an object marker obligatory with particular verb?	✗

3.3 Object marking in Ciwutee

In the present section we look at Ciwutee, spoken by 259,790 people in the central province of Manica.

3.3.1 (i) Can the object marker and the object argument co-occur in Ciwutee?

As we saw for Cinyungwe and Citshwa in §3.1 and §3.2 above, in Ciwutee the object marker and the corresponding object argument cannot co-occur out of the blue. It seems like there is both a specificity/givenness component in OM-doubling. See the examples (43b) and (43d).

(43) a. mhondolo y-a-rum-a mbudzi
9.lion SM9.PFV-bite-FV 9.goat
'the lion bit the goat'

b. mhondolo y-a-**yi**-rum-a (#mbudzi)
9.lion SM9.PFV-OM9-bite-FV 9.goat
'the lion bit it (the goat)'

c. mwaramu a-tem-a mu-ti
1.brother-in-low SM1.PFV-cut-FV 3-tree
'the brother-in-low cut the tree'

d. mwaramu a-**wu**-tem-a (#mu-ti)
1.brother-in-low SM1.PFV-OM3-cut-FV 3-tree
'the brother-in-low cut it (the tree)'

The examples in (43b) and (43d) illustrate that in Ciwutee the object marker and the object argument cannot co-occur in neutral context regardless of the animacy of the object argument. In addition, our consultant also suggested that it seems like all lexical objects behave similarly in that the object marker is prohibited to co-occur with the *in situ* object argument out of the blue. According to our speaker, the examples in (43b) and (43d) reflect this specificity and giveness reading of the object. This is the reason a Ciwutee speaker hearing this sentence out of context can ask "Which goat or tree are we talking about"? or "Why are you telling me that?". This restriction reminds us of what we described for Citshwa in §3.2.

Moreover, our consultant argued that there are contexts in which the examples in (43b) and (43d) presented above can be used by the speaker to illustrate that they have evidence, knows the person or the fact described, witnessed it (for more details about evidentiality in Bantu see Lippard et al. 2021). In such cases, the examples (44a) and (44b) repeated again from (43c) and (43d), can have the following translation in English.

(44) a. mhondolo y-a-**yi**-rum-a (#mbudzi)
9.lion SM9-PFV-OM9-bite-FV 9.goat
'the lion certainly bit it (the goat)'

b. mwaramu a-**wu**-tem-a (#muti)
1.brother-in-law SM1SG.PFV-OM3-cut-FV 3-tree
'the brother-in-law certainly cut it (the tree)'

As noted above, according to our informant, the Ciwutee speaker can use the OM-doubling structures to tell the hearer that they have evidence of what they

are talking about. The speaker is not expressing an opinion, they ar etelling the hearer what they know and so does not want to be challenged about the issue. If the other person insists, arguing about the same issue, this sentence can be used to say "hear want I am saying and let's end the conversation".

3.3.2 (ii) Is an object marker obligatory with particular object NPs?

As noted by Marten & Kula (2012) amongst others, Bantu languages differ with respect to the obligatoriness of co-occurrence of the OM with specific object argument. As has been described in a few other Mozambican Bantu languages such as Makhuwa (van der Wal 2015), Cuwabo (Guérois 2015) and Shimakonde (Ngunga et al. 2016), in some languages it is obligatory to object-mark class 1 and class 2 nouns (Makhuwa and Echuwabo) and animate objects (Shimakonde). This is not the case in Ciwutee where it is not obligatory to object mark particular objects. Example (45) illustrates that object marking is not obligatory with animate objects, while example (46) illustrates that object marking is not obligatory with inanimate objects.

(45) a. nd-a-won-a Zhambato
SM1SG-PFV-see-FV Zhambato
'I saw Zhambato'

b. nd-a-(**mu)**-won-a Zhambato.
SM1SG-PFV-(OM5)-see-FV Zhambato.
'I saw him (Zhambato)'

(46) a. nd-a-won-a bhuku.
SM1SG-PFV-see-FV bhuku
'I saw the book'

b. nd-a-(**ri**)-won-a bhuku.
SM1SG-PFV-(OM5)-see-FV book
'I saw it (the book)'

3.3.3 (iii) Are there locative object markers?

The presence or absence of locative markers in Ciwutee is the third parameter of variation examined by Marten & Kula (2012). Ciwutee has locative prefixes and they can be expressed by locative objects and similar to Cinyungwe, they can co-occur with their corresponding overt locative nouns. Consider the examples in (47):

(47) a. ku-munda ndi-no-*(ku)-ziy-a.
17-field SM1SG-PRS-(OM17)-know-FV
'in the field of cultivation (there), I know'

b. ku-munda ndi-no-(ku)-won-a.
17-field SM1SG-PRS-(OM17)-see-FV
'the field of cultivation I saw (it)'

c. mu-mvura ndi-no-(mu)-pind-a
18-marsh SM1SG-PRS-(OM18)-enter-FV
'in the marsh, I enter!'

d. pa-nyumba ndi-no-(pa)-gum-a
16-home 1SM1SG-PRS-(OM16)-arrive-FV
'at home, I arrive'

Ciwutee has the three locative object markers and they can be used to express locative objects. Examples (47a–d) show different verbs that illustrate that, in Ciwutee, locative objects can be expressed by locative prefixes. The examples also illustrate that it is obligatory to object mark locative objects when they occur in subject position. The verb *-won-* 'see' (47b) reminds us about what was described for Cinyungwe (example 28d) where we said that it was obligatory to object mark locative objects with the verb *-wona* 'to see'.

3.3.4 (iv) Is object marking restricted to one object marker per verb?

Ciwutee allows only one object marker per inflected verb. This is shown in the examples in (48b) and (48c) which demonstrate that in the Ciwutee's verb structure there is only one place for the OM.

(48) a. mbiya a-pas-a huku ma-gwere
1.grandma SM1.PFV-give-FV 9.chicken 6-maize
'the grandma gave the chicken maize'

b. mbiya *a-**yi**-**ma**-pas-a ma-gwere.
1.grandma SM1.PFV-OM9-OM6-give-FV 6-maize
'the grandma gave it (the chicken) maize'

c. mbiya *a-**ma**-**yi**-pas-a huku
1.grandma SM1.PFV-OM6-OM9-give-FV 9.chicken
'the grandma gave it (the maize) to the chicken'

It is important to note that the ungrammaticality of (48b) and (48c) is not related to the order of the objects, rather it is a strict restriction on the number of object markers possible in a verb form.

3.3.5 (v) Can either benefactive or theme objects be expressed by an object marker in double object constructions?

Ciwutee allows either benefactive or theme objects to be expressed by an object marker.

(49) a. Diminga w-aka-rim-ir-a mayi ci-mbamba.
Diminga SM1-PFV-cultivate-APL-FV 1.mother 7-beans
'Diminga cultivated beans for the mother'

b. Diminga w-aka-**mu**-rim-ir-a ci-mbamba.
Diminga SM1-PFV-OM1-cultivate-APPL-FV 7-beans
'Diminga (really) cultivated for her (the mother) beans'

c. Diminga w-aka-**ci**-rim-ir-a mayi.
Diminga SM1-PFV-OM7-cultivate-APPL-FV 1.mother
'Diminga (really) cultivated them (the beans) for the mother'

In (49a), the class 1 object marker (-*a*-) is co-referential with the object argument *mayi* 'mother' and in (49b), the class 7 prefix (-*ci*-) is co-referential with *cimbamba* 'beans'. Therefore, just like in Cinyungwe and Citshwa, in Ciwutee either benefactive or theme objects can be expressed by an object marker. This means that, Ciwutee is also a "symmetrical" language.

3.3.6 (vi) Is an object marker required/optional/disallowed in object relatives?

Different from Cinyungwe and Citshwa, in Ciwutee object markers are optional in object relatives, even with the verb -*won*- 'to see'.

(50) a. nyumba ya nd-a-(**yi**)-won-a nja Mazvarira
9.house REL SM1SG-PFV-(OM9)-see-FV COP Mazvarira
'the house that I saw (it) belongs to Mazvarira'

b. ma-khebe mayi a (a)-aka-(ma)-won-a akatapira
6-watermelon 1.mother REL SM1-PFV-(OM6)-FUT-see-FV sweet
'the watermelon that mother saw (it) was sweet'

The examples in (50a) and (50b) show that the verb -*won*- 'to see' does not need an object marker in the verb structure to render the sentence grammatical. Table 3 below summarizes what we have presented for object marking in the Ciwutee data so far.

Table 3: Parametric variation in object marking in Ciwutee

(i)	Can the object marker and the object argument co-occur?	✓
(ii)	Is an object marker obligatory with particular object NPs?	✗
(iii)	Are there locative object markers?	✓
(iv)	Is object marking restricted to one object marker per verb?	✓
(v)	Can either benefactive or theme objects be expressed by an object marker in double object constructions?	✓
(vi)	Is an object marker required/optional/disallowed in object relatives?	✓
(vii)	Is an object marker obligatory with particular verb?	✗

3.4 Object marking in Ciyaawo

After describing object marking in Cinyungwe, Ciwutee and Citshwa in the previous sections, in the present section we look at Ciyaawo data. Ciyaawo (P21 in Guthrie's (1967–1971) classification) is a Mozambican Bantu language spoken by 454,185 people in the Mozambican northern province of Niassa.

3.4.1 (i) Can the object marker and the object argument co-occur?

In Ciyaawo, there are no restrictions of co-occurrence of the object argument and the OM within the same sentence, as illustrated in (51).

(51) a. baaba a-dim-il-e yi-maanga.
dad SM1-cultivate-PFV-FV 8-maize
'dad has cultivated maize'

b. baaba a-**yi**-dim-il-e.
dad SM1-OM8-cultivate-PFV-FV
'dad has cultivated it (maize)'

c. baaba a-**yi**$_{\text{i}}$-dim-il-e yi-maanga$_{\text{i}}$
dad SM11-OM8-culitvate-PFV-FV 8-maize
'dad has cultivated the maize'

The examples in (51) are all grammatical and acceptable, although (51c) would probably be understood as emphatic to mean something like "dad has cultivated the maize very well". On the other hand, an OM such as *-yi-*, as in the verb structure in (51b), is usually included in the verb structure to respond to a question such as "What did dad do to the maize?".

3.4.2 (ii) Is an object marker obligatory with particular object NPs?

In Ciyaawo, there are no examples where the occurrence of OM is obligatory. That is, all transitive verbs can accommodate an object marker of any object NP regardless of their noun class. However, this is not obligatory under any circumstance.

(52) a. n'nyamaata ju-dim-il-e yi-maanga.
1.boy SM1-cultivate-PFV-FV 8-maize
'the boy has cultivated maize'

b. n'nyamaata ju-**yi**-dim-il-e.
1.boy SM1-OM8-cultivate-PFV-FV
'boy has cultivated it (maize)'

c. n'nyamaata ju-**yi**$_i$-dim-il-e yi-maanga$_i$
1.boy SM1-OM8-culitvate-PFV-FV 8-maize
'boy has cultivated the maize'

(53) a. ngweena ji-kamw-iil-e muu-ndu.
9.crocodile SM9-grab-PFV-FV 1-person
'the crocodile has grabbed a person'

b. ngweena ji-**n'**-kamw-iil-e
9.crocodile SM9-OM1-grab-PFV-FV
'the crocodile has grabbed a person'

c. ngweena ji-**n'**$_i$-kamw-iil-e muu-ndu$_i$.
9.crocodile SM1-OM1-culitvate-PFV-FV 1-person
'the crocodile has grabbed a person'

The examples (52a) and (53a) illustrate that in Ciyaawo, the OM is not obligatory with a specific object argument. That is to say that the presence of the OM in any transitive verb structures is not obligatory regardless of the class to which the noun belongs. When the OM occurs with the transitive verb, it may or may not co-occur with the lexical object seen in (52b, c) and (53b, c). The examples in (52b) and (53b) all correspond to questions like "What happened to the maize/

person?", while the examples in (52c) and (53c) respond to open questions like "What has happened?".

3.4.3 (iii) Are there locative object markers?

Just like in Cinyungwe and Ciwutee, in Ciyaawo, locative objects can be expressed by locative object markers and they can co-occur with the corresponding overt locative nouns. Consider the examples in (54):

(54) a. pa-cikoola$_i$, n-gu-**pa$_i$**-won-a.
16-school SM1SG-PRS-OM16-see-FV
Li.t 'at school, I see at'
Intd: 'I see the place of the school'

b. ku-musi$_i$, n-gu-**ku$_i$**-won-a
17-home SM1SG-PRS-OM17-see-FV
Lit. 'to home, I see to (it)'
Intd: 'I see there, the home'

c. mu-nyumba n-gu-**mu**-won-a
18-house SM1SG-PRS-OM18-see-FV
Lit. 'inside the house I see it'
Intd: 'I see the interior of the house'

(55) a. * pa-cikoola, n-gu-won-a.
16-school SM1SG1-PRS-see-FV
Lit. 'at school, I see'
Intd: 'I see the place of the school'

b. * ku-musi$_i$, n-gu-won-a
17-home SM1SG-PRS-see-FV
Lit. 'to home, I see to (it)'
Intd: 'I see there, the home'

c. * mu-nyumba n-gu-won-a
18-house SM1SG-PRS-see-FV
Lit. 'inside the house I see it'
Intd: 'I see the interior of the house'

In the examples in (54), the three locative prefixes are used as OMs. In (55), the omission of the locative object marker in the verb structure renders the sentence ungrammatical. This means that the verb *-wona* 'to see', and other verbs with the same lexical properties, require the object marker regardless of the respective

noun class to render the sentence grammatical. In this language, locative prefixes therefore behave in the same way as any other noun class prefixes.

3.4.4 (iv) Is object marking restricted to one object marker per verb?

Another parameter discussed in Marten & Kula (2012) concerns the number of object markers that can occur per inflected verb structure. In Ciyaawo, only one object marker is allowed per inflected verb. Consider the examples in (56):

(56) a. * mw-anace ju-ku-won-esy-a nguku yi-maanga.
1-child SM1-PRS-see-CAUS-FV 9.chicken 8-maize

b. mw-anace ju-ku-**ji**-(***yi)**-won-esy-a nguku yi-maanga.
1-child SM1-PRS-OM9-(OM8)-see-CAUS-FV 9.chicken 8-maize
Lit: the child is making it (the chicken) see the maize.
'the child is making the chicken see the maize'

c. mw-anace ju-ku-(***yi)**-(**ji)**-won-esy-a, nguku yi-maanga.
1-child SM1-PRS-(OM8)-(OM9)-see-CAUS-FV 9.chicken 8-maize
Lit: the child is making it (the chicken) see the maize.
'the child is making the chicken see the maize'

The example in (56a) shows once again that, when inflected, the verb *-won-* 'see' cannot occur without the obligatory presence of the OM in its structure. The data in (56a, b) illustrate that only one object marker is permitted in the verb structure. That is, in Ciyaawo, the co-occurrence of two OMs in the verb structure is forbidden.

3.4.5 (v) Can either benefactive or theme objects be expressed by an object marker in double object constructions?

In Ciyaawo, different from Cinyungwe, Citshwa and Ciwutee, theme objects cannot be expressed by an object marker in an applied construction. That is, in this language, the only object marker that is allowed to occur in the verb structure is the benefactive as shown below:

(57) a. Maama a-ku-**n**'-telec-el-a yi-maanga, mw-aanace.
1.mother 1SM-PRS-OM1-cook-APPL-FV 8-maize 1.child
'the mother is cooking maize for the child'

b. * Maama a-ku-**yi**-telec-el-a Siriza, yi-maanga.
1.mother 1SM-PRS-OM7-cook-APPL-FV Siriza 7-maize

The example (57b) illustrates that in Ciyaawo when the benefactive and the theme co-occur it is only the benefactive argument that can have a co-referent OM in the verb structure.

3.4.6 (vi) Is an object marker required/optional/disallowed in object relatives?

Ciyaawo functions as Ciwutee in relation to the use of object markers in object relatives which are generally optional, as illustrated in the following examples:

(58) a. ma-ticiti ga c-aa-ci-(**ga**)-dy-a maama ga
6-watermelon REL FUT-SM1-FUT-(OM6)-eat-FV 1.mother GEN
ku-dyoop-a
15-sweet-FV
'the watermelons that mum shall eat (them) are sweet'

b. * ma-ticiti ga c-aa-ci-won-a maama ga
6-watermelon REL FUT-SM1-FUT-OM6-see-FV 1.mother GEN
ku-dyoop-a
15-sweet-FV

c. ma-ticiti ga c-aa-ci-**ga**-won-a maama ga
6-watermelon REL FUT-SM1-FUT-OM6-see-FV 1.mother GEN
ku-dyoop-a
15-sweet-FV
'the watermelons that mum shall see (them) are sweet'

In Ciyaawo, the occurrence of the OM in the verb structure is optional (cf. 58a). It is important to note that the ungrammaticality of (58c) does not have to do with the absence of the OM in relative constructions as such, but with the fact that this verb is one which cannot occur without an OM.

Finally, we should add that, generally, optionality of the OM in the verb structure of most verbs is related to emphasis and what the speaker wants to express, as illustrated in (58). But this is different from (58c) which is marked as ungrammatical because of the specificity of the verb *-wona* whose structure requires the presence of an OM, be it in relative constructions or not.

3.4.7 (vii) Is an object marker obligatory with particular verbs?

In Ciyaawo, there are some verbs such as *-won-* 'see' and *-p-* 'give' which require an OM even if the lexical object occurs in the sentence. Object doubling

can occur with all transitive verbs. But its obligatoriness depends on the lexical properties of the verb. In (58) we have examples of verbs that must be categorized as [+OM] to which the lexical object is obligatorily added. This explains the ungrammaticality of (59a), (60a) and (61a).[1]

(59) Ciyaawo

a. * mw-anace ju-ween-i di-goombo.
1-child SM1.-see-FV 5-banana

b. mw-aanace ju-**di**$_i$-ween-i di-goombo$_i$.
1-child SM1.1-OM5-see-FV 5-banana
'the child has seen the banana'

c. mw-aanace ju-**di**-ween-i.
1-child SM1.1-OM5-see-FV
'the child has seen it'

(60) a. * uwe tu-p-eel-e mw-aanace mi-teela.
We SM1PL-give-PFV-FV 1-child 4-tree

b. uwe tu-**m**$_i$-p-eel-e mw-aanace mi-teela$_i$.
we SM1PL-OM1-give-PFV-FV 1-child 4-tree
'we have given the child the trees'

c. uwe tu-**m**-p-eel-e mi-teela.
we SM1PL-OM1-give-PFV-FV 4-tree
'we have given him (the child) the trees'

(61) a. * uwe tu-maany-i mw-aanace.
We SM1PL-know-(PFV) 1-child

b. uwe tu-**m**$_i$-maany-i mw-aanace$_i$.
we SM1PL-OM1-know-(PFV) 1-child
'we have known him the child the trees'

c. uwe tu-**m**$_i$-maany-i (mw-aanace$_i$)
we SM1PL-OM1-know-(PFV)- (1-child)
'we have known him (the child)'

In terms of the parameters under examination here, the answer for Ciyaawo is "yes" for all the parameters proposed by Marten & Kula (2012) except for (ii) and (v). Apart from that, we have shown that there are verbs like *-wona* 'see',

[1]Here we have imbrication, a phenomenon where, in certain verbs, the past tense marker (*il-*) is not suffixed to the verb root, it is imbricated within the verb root to yield the form *-ween.*

-pa 'give', *-manya* 'know' which are characterized by the obligatory presence of OM in their structure. We suggest that these verbs should be subcategorized as [+OM]. Table 4 summarizes object marking properties in Ciyaawo.

Table 4: Parametric variation in object marking in Ciyaawo

(i)	Can the object marker and the object argument co-occur?	✓
(ii)	Is an object marker obligatory with particular object NPs?	✗
(iii)	Are there locative object markers?	✓
(iv)	Is object marking restricted to one object marker per verb?	✓
(v)	Can either benefactive or theme objects be expressed by an object marker in double object constructions?	✗
(vi)	Is an object marker required/optional/disallowed in object relatives?	✓
(vii)	Is an object marker obligatory with particular verbs?	✓

3.5 Summary

This section has examined the morphosyntactic properties of object marking in four Mozambican languages taking into consideration Marten & Kula's (2012) parameters, as summarized in Table 5.

To summarise, only three from the seven parameters (ii), (iv) and (vi) have the same responses across all languages, parameter (ii) for which the value is NO across all languages of our sample and parameters (iv) and (vi) for which the value is YES across our sample. The four remaining parameters have one language whose response is different from the response of the other languages regardless of whether it is NO for Cinyungwe (i), Cithswa (iii), Ciyaawo (v) or YES for Ciyaawo (vii). Ciwutee is the only language which does not have any feature which is specific to it. This means that considerations like the language contact, multilingualism and language classification alone do not help to explain similarities or differences among the languages according to the different parameter values.

Considering the data from the four languages, we suggest that the obligatory requirement for an object marker [+OM] associated with some transitive verbs and structures should be added as a seventh parameter to the six parameters of variation in object marking in Bantu put forward by Marten & Kula (2012). Due to the existence of transitive verbs subcategorized as [+OM], we further encourage scholars to examine these parameters of variation in other Bantu languages in light of these features of variation.

Table 5: Object marking in the four analyzed Mozambican Bantu languages following Marten & Kula (2012)

Parameters of variation from Marten & Kula (2012)	Languages of the present study			
	Ciyaawo	Ciwutee	Cinyungwe	Tshwa
(i) Can the object marker and the object argument co-occur?	✓	✓	✗	✓
(ii) Is an object marker obligatory with particular object NPs?	✗	✗	✗	✗
(iii) Are there locative object markers?	✓	✓	✓	✗
(iv) Is object marking restricted to one object marker per verb?	✓	✓	✓	✓
(v) Can either benefactive or theme objects be expressed by an object marker in double object constructions?	✗	✓	✓	✓
(vi) Is an object marker required in object relatives?	✓	✓	✓	✓
(vii) Are there verbs whose inflection obligatorily require an OM in inflectional structure?	✓	✗	✗	✗

4 Conclusions

This paper has discussed object marking in four Mozambican Bantu languages, Cinyungwe, Citshwa, Ciwutee and Ciyaawo, based on Marten & Kula's (2012) parameters. In contrast to Ciyaawo, in Cinyungwe, Citshwa and Ciwutee the co-occurrence of the lexical object and OM in the same sentence is not allowed. Specifically, in Cinyungwe, the co-occurrence of the overt subject NP and the OM within the same sentence can happen only if the object is not *in situ.* In Ciwutee the co-occurrence of the object marker with the overt NP is allowed except in cases of emphasis or communicative strategies. In Citshwa OM-doubling the object marker and object argument results in a definiteness reading. The data illustrate that OM-doubling in Cinyungwe and Ciwutee is associated with an evidential reading, in a sense that the speaker is telling the hearer that s/he is sure of what s/he is talking about and so, s/he does not want to be contradicted (see Lippard et al. 2021 for more on this issue). On the basis of the data presented here, we also suggest that the feature [+OM] for some transitive verbs like *-manya* 'know', *-pa* 'give' and *-wona* 'see' should be added as the seventh parameter to the six parameters put forward by Marten & Kula (2012).

This research shows that of the four languages, only Ciyaawo has the value YES for the parameter (vii). Linking the Ciyaawo response for this parameter to what is happening in the relative constructions in the other three languages analyzed in this paper, we suggest that the verb *-wona* 'see' may have lost its [+OM] feature and remained only in the relative sentences. We need to undertake more research on this issue to check if we can find a trace of this feature in these languages using similar or other verbs because in Kilunguru (G30), for example, OM is obligatory with similar verbs *-ona* 'see'; *-ing'a* 'give' and with a different verb *-phika* 'find'.

Overall, this chapter has contributed to our understanding of the morphosyntax of four Bantu languages spoken in Mozambique, the broader properties of object marking in Bantu languages, as well as the use of a parametric approach (following Marten & Kula (2012)) to better understand variation within Bantu.

It has been noted that our aim was to discuss the Marten & Kula (2012) six parameters in four Mozambican languages. In the course of this, we have found a number of areas which require further investigation and attention. We leave for future work the discussion about the impact of the verb type on object marking, syntactic status of the OM, and (a)symmetry in double object constructions in the four languages analyzed in this paper.

Abbreviations

The following glosses are used in addition to the Leipzig Glossing Rules:

1, 5, 7, 9, 10 …	noun classes
FV	Final Vowel
Intd.	Intended meaning
OM	Object Marker
SM	Subject Marker
PFV	Perfective
PRS	Present
i	co-reference

Acknowledgements

We would like to thank Carlos Manuel for reading the first version of this paper for his comments and insights. We extend our gratitude to the reviewers for their valorous comments. We are also grateful to the comments of Hannah Gibson and Rozenn Gueróis that improved our discussion.

References

Aissen, Judith. 2002. Differential object marking: Iconicity vs economy. Ms. University of California, Santa Cruz.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Bresnan, Joan & Sam A. Mchombo. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63(4). 741–782.

Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2). 147–185.

Corbett, Greville. 2006. *Agreement*. Cambridge: Cambridge University Press.

Deen, Kamil Ud. 2006. Object agreement and specificity in early Swahili. *Journal of Child Language* 33. 223–246.

Guérois, Rozenn. 2015. *A grammar of Cuwabo (Mozambique, Bantu P34)*. Lyon: Université Lyon 2.

Guthrie, Malcolm. 1967–1971. *Comparative Bantu: An introduction to the comparative and prehistory of the Bantu languages*, vol. 1–4. London: Gregg International.

Henderson, Brent. 2006. Multiple agreement, concord and case checking in Bantu. In Olaoba F. Arasanyin & Michael A. Pemberton (eds.), *Multiple agreement, concord and case checking in Bantu*, 60–65. Somerville, MA: Cascadilla Proceedings Project.

Hualde, José I. 1989. Double object constructions in KiRimi. In Paul Newman & Robert D. Botne (eds.), *Current approaches to African linguistics*, vol. 5, 179–189. Dordrecht: Foris.

Iorio, David Edy. 2015. *Subject and object marking in Bembe*. Newcastle: University of Newcastle. (Doctoral dissertation).

Lippard, Hannah, Justine Sikuku, Crisófia Langa da Câmara, Madelyn Colantes, Jackson Kuzmik & Michael Diercks. 2021. Emphatic interpretations of object marking in Lubukusu and Cinyungwe. Paper presented at Miratives and evidentiality in Bantu online workshop.

Marten, Lutz & Nancy Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30(2). 237–253.

Marten, Lutz & Deograsia Ramadhani. 2001. An overview of object marking in Kiluguru. *SOAS Working Papers in Linguistics* 11. 259–275.

Marten, Lutz, Kristina Riedel, Silvester Ron Simango & Jochen Zeller. 2012. Preface. *Southern African Linguistics and Applied Language Studies* 30(2). III–VII. DOI: 10.2989/16073614.2012.737608.

Ngunga, Armindo. 2014. *Introdução à linguística Bantu*. Maputo: Imprensa Universitária.

Ngunga, Armindo, Fábio Bonfim Duarte & Quesler Fagundes Camargo. 2016. Differential object maring in Mozambican languages. In Doris Payne, Sara Pacchiaotti & Mokaya Bosire (eds.), *Diversity in African languages: Selected papers from the 46th Conference on African Linguistics*, 334–353. Berlin: Language Science Press. DOI: 10.17169/langsci.b121.489.

Ngunga, Armindo & Osvaldo G. Faquir. 2011. *Padronização da ortografia de Línguas Moçambicanas: Relatório do III seminário*. Maputo: Centro de Estudos Africanos (CEA).

Riedel, Kristina. 2009. *The syntax of object marking in Sambaa: A comparative Bantu perspective*. Leiden: LOT Publications.

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* 60(2–3). 277–296. http://dspace.bcu-iasi.ro/handle/123456789/15858.

van der Wal, Jenneke. 2016. Two curious gaps and one wrong analysis for Bantu object marking parameters. Paper presented at ACAL 47, UC Berkeley, 23–26 March 2016.

Zeller, Jochen. 2012. Object marking in isiZulu. *Southern African Linguistics and Applied Language Studies* 30(2). 219–235. DOI: 10.2989/16073614.2012.737600.

Zeller, Jochen. 2014. Three types of object marking in Bantu. *Linguistische Berichte* 239. 347–367.

Chapter 9

Verb extensions and morphosyntactic variation in Bantu: The case of Sumbwa

Kulikoyela Kahigi
Saint Augustine University of Tanzania, Mwanza

This study has two aims: (a) to describe verbal extensions in Sumbwa and their valence, and (b) to contribute to Bantu comparative data that may be used in research on Bantu morphosyntactic parametric variation as proposed in Guérois et al. (2017). The description covers all extensions that could be identified in the data at hand (e.g. Capus 1898, Kahigi 2008a,b), focusing on the forms and their various meanings, their valence possibilities, their productivity and co-occurrence constraints. The study reveals that most of the Proto-Bantu verb extensions reconstructed by Meeussen (1967) and Guthrie (1971) are still active in the language. Some extensions are found to be highly productive (applicative, passive, causative (also instrumental), associative, stative, and frequentative), others moderately or semi-productive (persistive, reversive, impositive and denominative) and quite a number may be regarded as being unproductive (associative/reciprocal *-an-*, reiterative, static, contactive and other minor extensions). The second aim of the study is to consider how Sumbwa compares to other Bantu languages by drawing on the parameters on verbal derivation identified in Guérois et al. (2017). Some of the findings agree with what is found in the majority of Eastern Bantu languages, e.g. the verb derivational strategies follow closely those mapped out by the PB reconstructions, except for a few innovations among the minor extensions (e.g. *-agan-*, *-agil-*). However, Sumbwa does not have *ba*-passives found in Bemba, nor does it have the CARP order as postulated by Hyman (2002) for Bantu. Some of the notable characteristics of Sumbwa verb extensions include the fact that (a) the causative and instrumental share extensions, (b) the associative markers include the post-verbal *-an-* and the pre-verbal *-i-*, which is homophonous with the reflexive; (c) the applicative conveys *benefactive, directive, location*, and *reason* meanings; (d) there is no systematic fixed order of extensions, except that in all co-occurrences, the passive comes last.

Kulikoyela Kahigi. 2024. Verb extensions and morphosyntactic variation in Bantu: The case of Sumbwa. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 229–275. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663777

1 Introduction

For the purposes of this study, we shall adopt Guthrie's (1962) use of the term "verb extension", which is interchangeable with "verb derivation" (cf. Katamba & Stonham 2006). As such, a verb extension is a suffix added to the verb root or base that changes the sense of the root/base.

Verb extensions are one of the most important features of Bantu languages and have been investigated and described since the inception of Bantu studies. A survey of descriptions of Bantu verb extensions across the years (e.g. Madan 1903, Ashton 1947, Johnson 1939, Guthrie 1962, Eastman 1967, Scotton 1967, Bokamba 1975, Khamis 1972, 1985, Rugemalira 1993, 2005, Schadeberg 2003) shows that verb extensions constitute one of the most important topics in Bantu linguistics. This is borne out by research on the topic, which has continued to produce valuable descriptive and theoretical contributions (cf. Baker 1985, Alsina 1999, Marten 2003, Hyman 2002, Katamba & Stonham 2006, Waweru 2005, Khumalo 2007, Chabata 2007, Kula & Marten 2010, Dom et al. 2018, etc.).

The focus of this study is on Sumbwa, a West Tanzania Bantu language, classified as F23[1] by Guthrie (1948, 1970: 11). Known as Sisuumbwa by its native speakers, this largely undescribed language is mainly spoken in Geita, Shinyanga, Tabora and Kagera regions. Other Sumbwa speakers, known as the Bayeke, are found in the DRC, in the current Yeke chiefdom, whose capital is Bunkeya, in Lualaba province (cf. Munongo & Grévisse 1967). The dialect addressed in this study is the Ushilombo/Lunzewe dialect spoken in Bukombe district, Geita region.

This study has two aims. The first aim is to describe verbal extensions in Sumbwa. I follow Guthrie's (1962) approach, which focuses on identifying the morphological shapes of the extensions, their meanings, and their syntactic effects (i.e. valence). In addition, a brief statement of the productivity of the extensions is provided. The concept of valence used here is the traditional one; it refers to the potential of the verb to take an argument (i.e. subject, direct object or indirect object) (cf. Humphreys 1999, Haspelmath & Müller-Bardey 2004). In most Bantu languages, some verb extensions trigger a change in valence by either adding or deducting an argument, while some extensions do not affect the basic valence of the verb. Thus, an extension may be referred to as valence-increasing, valence-decreasing or valence-maintaining (cf. Chabata 2007, Payne 1997, Hyman 2007). The concept of productivity is also used here in its traditional sense; a verb extension is viewed as being productive if it is used to coin new words (cf.

[1]This is according to the widely used Guthrie (1948) classification for identifying individual Bantu languages. In this classification, the Bantu area is divided into zones and the zones are divided into groups. Sumbwa belongs to Zone F, Group F20 (Sukuma-Nyamwezi Group).

Plag 2006: 121). Regarding Bantu extensions, productivity is considered to be a scalar concept, that is on a scale "from totally unproductive expansions occurring in just a few verbs to fully productive suffixes" (Schadeberg 2003: 73). Thus, in Sumbwa we have what we may call highly or fully productive extensions (e.g. the passive extension); moderately or semi-productive extensions like the frequentative *-agul-* which is mainly restricted to disyllabic roots; and non-productive extensions such as the static *-am-* that can hardly be used in coining new words in the language. Besides valence and productivity, this study also briefly examines the co-occurrence restrictions of the extensions.

The more general aim of the study is to contribute to Bantu comparative data that may be used in research on Bantu parametric morphosyntactic variation (cf. Marten et al. 2007, Guérois et al. 2017). Guérois et al. (2017) is a master list of 141 parameters in 12 morphosyntactic areas. These include nouns and pronouns, noun modifiers, nominal derivation, lexicon, verbal derivation, verbal inflection, relative clauses, clefts and questions, verbless clauses, simple clauses, constituent order, complex sentences, and expression of focus. Each parameter begins with a question, followed by possible answers, e.g.

> *Parameter* **36**. Canonical Passive: Is the canonical passive productively expressed through a verbal extension?
>
> Possible answers:
>
> null = unknown,
>
> no = another strategy is used to express passivisation, e.g. an impersonal construction…,
>
> yes = specify whether there is one or several possible forms.

The possible answers vary depending on the nature of the parameter.

One of the goals of the project is to collect Bantu morphosyntactic data with a view to identifying variation at the micro level. Of the 12 areas in the master list, the focus of the present study will be on area 5: verbal derivation (parameters number 36 to 48).

The second aim of this study, consequently, addresses these parameters, responding to the possible answers given for each parameter. Due to space limitations, the discussion is mostly restricted to Sumbwa data, although, occasionally, other Bantu languages will be referred to for comparison. The focus here will be to observe *whether* and *how* the proposed parameters involving verb extensions occur in Sumbwa. The data used in this study is mainly from Capus (1898)

and Kahigi (2008a,b); the first source is an earlier grammar while the other two include data collected directly from the field between 1976 and 2004.

The rest of the paper is organized as follows. §2 deals with the salient aspects of the verb extensions in the language: the extensions and, their meanings, syntactic effects (i.e. valence) and productivity. The co-occurrence constraints of the verbal extensions are presented in §3. §4 focuses on the parameters of morphosyntactic variation (Guérois et al. 2017), and, lastly, §5 presents the conclusion.

2 Sumbwa verb extensions

To facilitate comparison, we will use Meeussen's (1967: 92, 1969) and Guthrie's (1971: 144) verb extension reconstructions. Table 1 shows the modern Sumbwa reflexes of these reconstructions. Some less well-known extensions, i.e. *-agil-*, *-agan-*, *-al-*, and *-l-*, are not shown in Table 1 but will be covered in the discussion below.

What follows is a presentation and discussion of the verb extensions in terms of their morphological shapes, their meanings, syntactic effects (i.e. change in valence) and productivity.

2.1 The applicative: *-il-*

The applicative extension denotes an action performed on behalf of, on, at, towards an entity, etc. (cf. Madan 1903, Ashton 1947: 218–220, Rugemalira 2005: 46). The surface forms of the applicative extension are *-il-* and *-el-,* which are distributed in accordance with vowel harmony. Examples are given in Table 2.

The examples in Table 2 illustrate the various functions of the applicative in Sumbwa. Examples a and b represent the sense of 'do sth on behalf of' or 'for the benefit of'; example c represents the sense of 'location'; example d the sense of 'reason'; and example e the 'directional' sense. The applicative extension is a valence-increasing extension, i.e. it adds an extra argument to the verb, cf. example b which without the applicative extension would be *a-la-tem-a* (SM1-PST-cut-FV), 'he cut'. Verbs of all types (transitive, intransitive) take the applicative extension; consequently, it is one of the most productive extensions in the language.

2.2 The passive: *-u-, -iβu-*

The passive construction is found in many languages around the world. In Bantu, the Proto-Bantu passive extension **-u-*/**-ibu-* is still directly reflected in many ex-

Table 1: Bantu verb extension reconstructions and reflexes in Sumbwa. These PB reconstructions and their reflexes in Sumbwa do not include "imbricated" forms that characterize perfective constructions. These imbricated forms, which are found in many Bantu languages, are characterized by mutations which do not occur in simple verbal constructions and non-perfective contexts (cf. Berger 1937–1938, Givón 1970, Mould 1972, Bastin 1983, Kahigi 1989, Hyman 1995, Kula 2001).

Verb extension	Proto-Bantu form	Sumbwa form
Applicative[a]	*-id-	-il-[b]
Passive	*-u-, *-ibu-	-w-, iβw-
Causative	*-i, *ici-	-i-, -iisi- (also instrumental)
Impositive	*-ik-	-ik-
Stative	*-ik-/*-uk-/*-ad-	-ik-, -uk-, inkan-
Associative	*-an-	-an-, -i-
Reversive (active)	*-ud-, *-udud-	-ul-, -ulul
Reversive (stative)	*-uk-	-uk-
Persistive	*-idid-	-ilil-, -ilizi-
Frequentative (tr.)	*-agud-	-agul-
Frequentative (intr.)	*-aguk-	-aguk-
Denominative	*-ap-, *-p-, *-apad-	-h-, -ahal-
Reiterative	*-udud-	-uul-, -ul-
Static	*-am-	-am-
Contactive	*-at-	-at-

[a]The terms used here to refer to the extensions are by no means universal. Alternative terms for applicative include e.g. applied, directive, prepositional, dative; for persistive e.g. intensive, double prepositional; and for frequentative e.g. augementative (cf. Madan 1903, Johnson 1939, Ashton 1947, Guthrie 1962, Lodhi 2002, Schadeberg 2003).

[b]The extensions are represented in their basic form. The surface form is determined by vowel harmony, i.e. (i) the extension vowel /i/ is lowered to [e] in the environment of root vowels /ɛ/ or /ɔ/, e.g. /sɛk-il-a/ becomes [sɛkɛla] 'laugh for' and /βɔl-il-a/ becomes [βɔlɛla] 'rot for'; (ii) the extension vowel /u/ is lowered to [o] in the environment of root vowel /ɔ/, e.g. /dɔd-uulul-a/ becomes [dɔdɔɔlɔla] 'unsew'.

Table 2: Examples of the applicative extension

	Verb root	Gloss	Applicative stem	Examples
a.	*-nó-*	drink	*-nó-il-a [nwéélа]*	*a-la-mu-nó-il-a βusele* SM1-PST-OM1-drink-APPL-FV beer 'he drank beer for him'
b.	*-tem-*	cut	*-tem-il-a [temela]*	*a-la-mu-tem-il-a muti* SM1-PST-OM1-cut-APPL-FV tree 'he cut a tree for him'
c.	*-dod-*	sew	*-dod-il-a [dodela]*	*a-ø-dod-el-a kaaya* SM1-HAB-sew-APPL-FV home 'he sews wood at home'
d.	*-huul-*	whip	*-huul-il-a*	*a-ø-mu-huul-il-a βuzoβe* SM1-HAB-OM1-whip-APPL-FV laziness 'he whips her for laziness'
e.	*-iluk-*	run	*-iluk-il-a*	*a-li-iluk-il-a mu-numba* SM1-PST-run-APPL-FV 17-house 'he ran into the house'

isting languages, although some of them (e.g. the A70 group) use other extensions for the passive due to innovations that occurred in those languages (Bostoen & Nzang-Bie 2010).

Sumbwa has retained the reconstructed extensions: *-u-/-iβu-*, with the vowel /u/ becoming the glide /w/, a diachronic process common in Bantu; thus *-u-/-iβu-* becomes *-w-/-iβw-*. The short extension *-w-* occurs in the environment after consonant-final roots, while *-iβw-* occurs after vowel-final roots or extended stems. Examples are given in Table 3.

The passive construction licenses an optional prepositional phrase, which indicates the logical agent. Examples in (1) illustrate this point:

(1) a. *mu-kiima a-la-léét-a si-taβo*
1-woman SM1-PST-bring-FV 7-book
'A woman brought a book.'

Table 3: Examples of the passive extension

	Verb	Gloss	Passive	Gloss
a.	*-lim-*	cultivate; farm	*-lim-u-a [limwa]*	be cultivated
b.	*-tem-*	cut	*-tem-u-a [temwa]*	be cut
c.	*-dul-*	bore	*-dul-u-a [dulwa]*	be bored
d.	*-tah-*	draw (water)	*-tah-u-a [tahwa]*	be drawn
e.	*-li-*	eat	*-li-iβu-a [lííβwa]*	be eaten
f.	*-fil-*	take sb or sth somewhere	*-fil-u-a [filwa]*	be taken somewhere

b. *si-taβo si-la-léét-w-a ne mu-kiima*
7-book SM7-PST-bring-PASS-FV by 1-woman
'A book was brought by a woman.'

Apart from this standard construction, Sumbwa also has some other passive constructions that need to be noted. One of these is the 'passive form of infinitival nouns', usually placed in class 15 in the noun class system. Note the following examples in (2):

(2) a. *[kufwíílwa kwiinki kulalééta naku niinki]*
ku-fu-il-u-a ku-inki ku-la-léét-a naku n-inki
15-die-APPL-PASS-FV 15-much SM15-PST-bring-FV 10.misery 10-much
'Being bereaved many times brought a lot of misery.'

b. *[kulekaniisiβwa mukazi waamwe kulamusaayiisja]*
ku-lek-an-isi-iβu-a mu-kazi wa-amwe
15-leave-ASS-CAUS-PASS-FV 1-wife 1-his
ku-la-mu-saay-isi-a
SM15-PST-OM1-be.angry-CAUS-FV
'Separating from his wife made him angry.'

In the data we have, constructions of this type do not occur with the "by phrase" noted above. Another construction to note is the "passive with the locative noun". This is illustrated in (3):

(3) a. *[malaβo galapaambwamo munúúmba]*
ma-laβo ga-la-pamb-w-a-mo mu-numba
6-flower SM6-PST-decorate-PASS-FV-CL18 18-house
'Flowers were decorated in the house.'

b. *[munúúmba halapaambwamo malaβo]*
mu-numba ha-la-pamb-w-a-mo ma-laβo
18-house SM16-PST-decorate-PASS-FV-CL18 6-flower
Lit.: 'In the house there was decorated flowers'

These two examples illustrate what is known as locative inversion. In the first example, the locative noun *munúúmba* occurs after the passivized verbal construction *galapaambwamo*; in the second example, the locative noun is in subject position. The final *-mo* is a locative enclitic. What needs to be noted is that in both sentences the passive construction does not need to, but can have a "by phrase", as illustrated in (4) below:

(4) a. *[malaβo galapaambwamo munúúmba ne mukiima]*
ma-laβo ga-la-pamb-w-a-mo mu-numba ne mu-kiima
6-flower SM6-PST-decorate-PASS-FV-CL18 18-house by 1-woman
'Flowers were decorated in the house by a woman'

b. *[munúúmba halapaambwamo malaβo ne mu-kiima]*
mu-numba ha-la-pamb-w-a-mo ma-laβo ne mu-kiima
18-house SM16-PST-decorate-PASS-FV-CL18 6-flower by 1-woman
Lit.: 'In the house there was decorated flowers by a woman'

A final construction to note is the "passive with the past participle". This construction is illustrated in example (5) below:

(5) *mu-gunda gu-li βu-lim-w-e*
3-farm SM3-be PP-cultivate-PASS-FV
'The farm is cultivated.'

As can be noticed here, the passive past participle in the language is characterized by the verb *-li* 'be' followed by main verb with the structure: *βu-*VRT-PASS*-e*. In the data we have, this type of construction is also not followed by the "by phrase". It is used when one intends to imply a "state" of an entity.

The passive extension is a valence-decreasing extension, i.e. it deducts an argument from the verb as the examples in (1a) and (1b) show. In addition, all transitive verbs take the passive extension; it is thus a productive extension in the language.

2.3 The causative: *-i-/-iisi-*

The extension generally denotes 'causing someone to perform some action'. Historically, the extension *-i-* has caused what is known as spirantization in Bantu

(cf. Schadeberg 1995, Bostoen 2008). This has resulted in morphophonemic alternations in root-final position in all roots or bases ending in /p, b, l, t, d, k, g/. These change to [f, v, z, s, z, s, z] respectively. These changes are illustrated in Table 4 below. Root-final consonants which change are shown as well as consonants which do not change. Notice that all roots with final vowels take the *-iisi-* extension, while some roots (e.g. *-βamb-* 'peg out') take both extensions. All instances of *i* and *u* occurring at morpheme boundaries change to the corresponding glides [j] and [w], respectively.

Table 4: Examples of the causative extension

	Verb root	Causative *-i-*		Causative *-iisi-*	Causative gloss
		Change of root-final C	No change of root-final C		
a.	*-βamb-*	*-βamv-i-a*		*-βamb-iisi-a*	cause to peg out
b.	*-puup-*	*-puuf-i-a*			cause to be light
c.	*-tem-*		*-tem-i-a*	*-tem-iisi-a [teméésja]*	cause to cut
d.	*-lil-*	*-liz-i-a*			cause to cry
e.	*-sees-*			*-sees-iisi-a [seeseesja]*	cause to spill
f.	*-kan-*		*-kan-i-a*	*-kan-iisi-a*	cause to groan or creak
g.	*-swiiz-*			*-swiiz-iisi-a*	cause to filter or strain
h.	*-dod-*	*-doz-i-a*		*-dod-iisi-a [dodéésja]*	cause to sew
i.	*-hit-*	*-his-i-a*			cause to pass
j.	*-ak-*	*-as-i-a*			cause to burn
k.	*-og-*	*-oz-i-a*		*-og-iisi-a [ogéésja]*	cause to bathe
l.	*-saay-*			*-saay-iisi-a*	cause to be angry
m.	*-oβah-*		*-oβah-i-a*		cause to fear
n.	*-no-*			*-no-iisi-a [nwéésja]*	cause to drink
o.	*-li-*			*-li-iisi-a*	cause to eat
p.	*-gu-*			*-gu-iisi-a*	cause to fall

The only root-final consonants that do not change are *m, s, z, n,* y, and *h*. As can be seen in examples (Table 4a–p), an extra argument has been added to the basic valence of the respective verbs, e.g. *a-la-gu-a* (SM1-PST-fall-FV) 'she fell' with the causative extension becomes: *a-la-mu-gu-iisi-a* (SM1-PST-OM1-fall-CAUS-FV) 'he caused her to fall'. As to productivity, the causative extension is highly productive; it applies to intransitive and transitive verbs.

2.4 The instrumental: *-i-, -iisi-*

In Sumbwa, the causative extensions *-i-/-iisi-* are also instrumental, as the following examples show:

(6) a. *[alamulíísja mwaana]*
a-la-mu-li-iisi-a mu-ana
SM1-PST-OM1-eat-CAUS-FV 1-child
'She caused the child to eat.' (i.e. she fed the child)

b. *[alalíísja siliko]*
a-la-li-iisi-a si-liko
SM1-PST-eat-CAUS-FV 7-spoon
'She ate with a spoon.'

As can be seen here, there is a crucial difference between the first and the second sentence, despite the fact that both use the same extension *-iisi-*. In (6a), *-iisi-* is used in its causative sense, while in (6b) it is used in its instrumental sense. This is clearly a case of "causative-instrumental syncretism". This case is also found in other Bantu languages. For example, Jerro (2017) discusses a similar case in Kinyarwanda, and Wald (1998) argues for a split between Bantu languages that use the applicative and those that use the causative extension to mark instrumentals; he hypothesizes that the latter is an innovation.

In Sumbwa, the instrumental is a valence-increasing extension, as shown in Table 4b above. It is also quite productive.

2.5 The impositive: *-ik-*

In Bantu, *-ik-* is also the extension for the stative/neuter, which is described below. The impositive differs from the stative in that it has a 'causative' meaning. The action of the verb results in "causing something or somebody to be in some position or state". In Table 5 are some examples.

Table 5: Examples of the impositive extension

	Verb root	Gloss	Impositive	Gloss
a.	*-tumb-*	increase (INTR)	*-tumb-ik-a*	gather together in a heap
b.	*-loβ-*	get wet	*-loβ-ik-a*	soak
c.	*-oβah-*	be afraid	*-oβah-ik-a*	frighten
d.	*-om-*	dry (INTR)	*-om-ik-a*	dry sth over fire
e.	*-hengam-*	tilt (INTR)	*-hengam-ik-a*	cause to tilt sideways

Since it is causative in meaning, the impositive extension is valence-increasing. However, it is not as pervasive as the normal causative *-i-/-isi-*; it may be characterized as slightly productive.

2.6 The stative: *-ik-*

The stative is also referred to as "neuter" (Schadeberg 2003). This extension denotes a state already completed or still going on. It also denotes "potentiality", depending on the context. Most verbs use *-ik-*, but a few use *-inkan-*. It is the *-ik-* extension that is still productive.

Table 6: Examples of the stative extension

	Verb root	Gloss	Stative	Gloss
a.	*-lim-*	cultivate; farm	*-lim-ik-a*	be cultivated
b.	*-tem-*	cut	*-tem-ek-a*	be cut
c.	*-nó-*	drink	*-nó-ik-a*	be drinkable
d.	*-dul-*	bore	*-dul-ik-a*	be bored
e.	*-bhon-*	see	*-bhon-inkan-a*	be seen
f.	*-mani-*	know	*-mani-inkan-a*	be known
g.	*-suh-*	forget	*-suh-inkan-a*	be forgotten

Note that the Sumbwa *-ik-/-inkan-* extension is also found in other Bantu languages such as Swahili (e.g. *-pik-* 'cook'; *-pik-ik-a* [pikika] 'be cooked'; *-on-* 'see'; *-on-ikan-a* [onekana] 'be visible'). The stative is valence-decreasing, e.g. *a-la-dul-a lubaβo* (SM1-PST-bore-FV a plank) 'he made holes into a plank' becomes *lubaβo lu-la-dul-ik-a* 'a/the plank was bored'. It is a characteristic of the Bantu stative that the agent is never expressed.

2.7 The associative: *-i-/-an-*

We use the term "associative" following Ashton (1947) and Maganga & Schadeberg (1992: 164). Ashton states:

> The term "Associative" is used instead of the more generally accepted term "Reciprocal" as found in the *Standard Swahili-English Dictionary*, for in addition to reciprocity -NA expresses other aspects of association such as concerted action, interaction and interdependence (and in some cases disassociation) (1947: 240).

In Sumbwa, there are two affixes that denote action performed mutually or associatively, *-i-* and *-an-*. The first, *-i-*, is a pre-verb root affix that is also used as a reflexive marker. Thus, it is polysemous and highly productive. Examples for the associative *-i-* are in Table 7.

Table 7: Examples of the associative *-i-*

	Verb	Gloss	Associative *-i-*	Gloss
a.	*ku-huul-a*	to hit	*ku-i-huul-a*	to hit each other
b.	*ku-taahi-a*	say farewell	*ku-i-taahi-a*	to say farewell to each other
c.	*ku-gú-a*	to fall	*ku-i-gú-il-a*	to fall on each other
d.	*ku-li-a*	to eat	*ku-i-li-a*	to eat each other

Due to the ambiguity of the *-i-* affix, the above associative constructions could also be glossed as 'to hit oneself', 'to bid farewell to oneself', 'to fall on oneself', and 'to eat oneself', respectively. It is important to note that the meaning of the construction with the *-i-* affix will depend on the linguistic context. A reflexive meaning will always imply that the subject NP and the object NP are identical, as illustrated in (7) below:

(7) a. *mu-ana$_i$ a-la-huul-a mu-ana$_i$* (Subject NP = Object NP)
1-child$_i$ SM-PST-hit-FV 1-child$_i$

b. *mu-ana$_i$ a-la-i$_i$-huul-a*
1-child$_i$ SM-PST-REF$_i$-hit-FV
'the child hit him/herself'

As can be seen here, while the reflexive in (7b) allows a singular subject (i.e. an agent which is at the same time a patient), the reciprocal in (8) below can only allow a plural subject, which is consistent with reciprocity or associativeness (i.e. two or more individuals doing the same thing to each other).

(8) *βa-ana$_i$ βa-la-i-huul-a*
2-child$_i$ SM-PST-REF$_i$-hit-FV
'the children hit each other'

A further point to note here is that Sisumbwa is not the only language in zone F to express associativeness using the pre-verbal *-i-*. Expression of associativeness using *-i-* has also been found in Rimi (Olson 1964: 172–174), Rangi (Stegen 2002: 144) and Sukuma (Batibo 1985: 172–173). The extent to which this phenomenon is widespread in Bantu awaits investigation.

The second affix, *-an-*, competes with *-i-* as an associative marker. Our analysis shows that it is unproductive; reciprocity/associativeness in Sumbwa is more frequently expressed by the pre-verbal *-i-*. Some examples illustrating *-an-* are given in Table 8.

Table 8: Examples of the associative *-an-*

	Verb root	Gloss	Associative *-an-*	Gloss
a.	*-taag-*	throw away	*-taag-an-a*	leave or abandon each other
b.	*-las-*	shoot using a bow	*-las-an-a*	shoot at each other; fight using bows and arrows
c.	*-som-*	stab	*-som-an-a*	stab each other
d.	*-bhut-*	give birth	*-bhut-an-a*	give birth in great numbers
e.	*-tol-*	backbite	*-tol-an-a*	backbite each other

An inspection of the "*-an-* associative" in actual speech and in Kahigi (2008a) shows that the verb roots targeted are either -CV(V)C- (many) or -CV(V)CVC- (fewer). There are, in addition, a few verbs that form their associative forms with *-aan-* instead of *-an-*. Examples of these are given in Table 9.

There are also a few examples which appear to have an associative meaning but are not related to any corresponding verb roots. These are shown in Table 10.

Table 9: Examples of the associative *-aan-*

	Verb root	Gloss	Associative *-aan-*	Gloss
a.	*-lek-*	leave	*-lekaana*	leave each other
b.	*-sang-*	find	*-sangaana*	meet
c.	*-gabh-*	divide	*-gabhaana*	share

Table 10: Examples of the associative *-aan-*/*-an-*

	Verb root	Gloss	Associative	Gloss
a.	*-fu-*	die	*-fw-aan-a*	quarrel
b.	*-nó-*	drink	*nw-aan-a*	become friends or well-mixed
c.	*-lag-*	say goodbye to king	*lag-an-a*	promise
d.	*-sas-*	make bed	*sas-an-a*	take from each other by force
e.	*-tong-*	claim for payment of debt	*tong-an-a*	quarrel

Although the forms *-fu-*, *-nó-*, *-lag-*, *-sas-*, and *-tong-* by themselves are found in the language, their current meanings have nothing to do with the associative forms.

As examples in (9) below show, the associative *-i-*/*-an-* is valence-decreasing; the affixes can only license an external argument.

(9) a. *βa-la-i-las-a*
SM2-PAST-REC-shoot.with.arrows-FV
'They shot each other with arrows.'

b. *βa-la-las-an-a*
SM2-PAST-REC-shoot.with.arrows-FV
'They shot each other with arrows.'

2.8 The reversive: *-ul-*/*uk-*~*-uul-*/*-uuk-*~*-ulul-*/*-uluk-*

Although there is variation in the form of the reversive extension, the overall meaning is the same: it denotes the opposite of the meaning of the verb root.

The reversive involves both the active and stative/neuter forms.[2] Examples are given in Table 11 and Table 12.

Table 11: Examples of the reversive active *-ulul-*

	Verb root	Gloss	Reversive active	Gloss
a.	*-anz-*	spread e.g. bedclothes	*-anz-ulul-a*	take off sth spread
b.	*-βamb-*	peg out	*-βamb-ulul-a*	unpeg
c.	*-tung-*	thread	*-tung-ulul-a*	unthread
d.	*-fung-*	close	*-fung-ulul-a*	open
e.	*-gongom-*	bend; stoop	*-gongom-ul-a*	raise
f.	*-semb-*	tie with rope/ bandage	*-semb-ul-a*	unwrap; untie

Table 12: Examples of the reversive stative *-uluk-*

	Verb root	Gloss	Reversive stative	Gloss
a.	*-anz-*	spread e.g. bedclothes	*-anz-uluk-a*	(of bed) have bedclothes taken off
b.	*-βamb-*	peg out	*-βamb-uluk-a*	become unpegged
c.	*-tung-*	thread	*-tung-uluk-a*	be become unthreaded
d.	*-fung-*	close	*-fung-uluk-a*	be opened
e.	*-gongom-*	bend; stoop	*-gongom-uk-a*	be raised
f.	*-semb-*	tie with rope/ bandage	*-semb-uk-a*	become unwrapped or untied

All the above cases show that the reversive extension is suffixed directly to the verb root. There are some cases, however, where the basic verb root does not exist in its simple form. Here we have what we may call "complex verb roots"; but

[2]The *-ulul-* extension may also be considered to be a "double reversive" *-ul-ul-* (an intensive reversive), while *-uluk-* may be regarded as a combination of the reversive extension *-ul-* followed by the reversive stative *-uk-*.

the reversive suffix is never attached to them directly. The examples in Table 13 illustrate this point.

Table 13: Examples of the reversive *-ul-*

	Verb root	Gloss	Reversive	Gloss
a.	*-anikil-*	spread out to dry	*-an-ul-a*	take in sth spread out to dry
b.	*-hanik-*	hang up	*-han-uul-a*	take down sth suspended
c.	*-higik-*	arrange cooking stones	*-hig-ul-a*	move (e.g. cooking stones)
d.	*-siβik-*	tether	*-siβ-ul-a*	untether

In Table 13, comparison of the reversive forms with the complex verb roots shows that the "complex verb roots" have been reanalyzed; they have been shortened to *-an-, -han-, -hig-,* and *-siβ-,* respectively, before the attaching of the reversive extension. The "shortened" roots in the reversive are probably older forms which are no longer used. Swahili has a similar pattern, e.g. *-anik-* 'spread out to dry', *-an-u-a* 'take in sth spread out to dry'; *-angik-* 'hang up; suspend', *-ang-u-a* 'pick; knock down'.

Regarding valence, the active reversive *-ul-* is valence-maintaining, while the stative *-uk-* is valance-decreasing. Overall, the extension is slightly productive.

2.9 The persistive *-ilil-, -ilizi-*

Guthrie (1971: 144) uses the term "persistive" for **-idid-*, which is reflected in Sumbwa as *-ilil-*. The form is literally the doubling of the applicative *-il-* which is why Johnson (1939) called it "double-prepositional". The extension, however, does not have any applicative meaning; rather, it denotes action performed persistently or continuously but intensively. This is probably why Ashton (1947: 214, 243–245) uses the term "augmentative".

The second extension, *-ilizi-*, is assumed to be a direct outcome of the spirantization that occurred from a combination of **-idid- + -i-* (causative).

Both of these extensions occur in many Eastern Bantu languages, although only slightly productively. For instance, in Swahili, where the **-idid-* extension is no longer very productive, we have: *-end-* 'go' > *-end-ele-a* 'progress', *-shik-* 'hold'> *-shik-ili-a* 'hold on tightly or insist', *-pend-* 'like' > *-pend-ele-a* 'favour';

and for *-*idid*- + -*i*- we have: -*end*- 'go' > -*end-elez-a* 'cause to progress', etc. (Ashton 1947: 214, 243–245).

In Sumbwa, -*ilil*- occurs mostly with the meaning of 'persistence, continuity or intensity', while -*ilizi*- occurs with two meanings: 1) doing sth persistently, continuously or intensively, and 2) doing sth persistently, continuously or intensively for payment. Given these meanings, the extension does not trigger any change in the basic valence of the verb. We exemplify each of these below.

2.9.1 The persistive -*ilil*-

This extension is illustrated in the examples in Table 14.

Table 14: Examples of the persistive -*ilil*-

	Verb root	Gloss	-*ilil*-	Gloss
a.	-*ling*-	look	-*ling-ilil-a*	look at for a long time
b.	-*lind*-	await	-*lind-ilil-a*	await for a long time
c.	-*lond*-	follow	-*lond-elel-a*	follow for a long time
d.	-*sek*-	laugh	-*sek-elel-a*	laugh for a long time
e.	-*kwáát*-	hold, seize	-*kwáát-ilil-a*	hold firmly for a long time

We also note probable instances of lexicalization in the two examples in Table 15.

Table 15: Examples of lexicalized persistive -*ilil*-

	Verb root	Gloss	persistive -*ilil*-	Gloss
a.	-*mani*-	know	-*mani-ilil-a*	get accustomed to
b.	-*fuk*-	pour (water)	-*fuk-ilil-a*	irrigate

It should be noted that the meanings of the extended verb bases -*mani-ilil*- and -*fuk-ilil*-, though relatable to the meaning 'doing sth persistently, continuously or intensively', may be argued to be qualitatively different from the meanings of the verb roots -*mani*- and -*fuk*-.

2.9.2 The persistive: -*ilizi*-[1]

The extension -*ilizi*-[1] retains much of the meaning 'doing sth persistently, continuously or intensively', as the examples in Table 16 show.

Table 16: Examples of persistive *-iliz-*$_1$

	Verb root	Gloss	*-ilizi-*^{1}	Gloss
a.	*-an-*	groan	*-an-ilizi-a*	groan for a long time; yell
b.	*-anguh-*	hasten	*-anguh-ilizi-a*	hasten overmuch
c.	*-gum-*	harden	*-gum-ilizi-a*	persevere
d.	*-gelek-*	put sth on top of another	*-gelek-elezi-a*	pile up to the top
e.	*-kooβ-*	look for	*- kooβ-elezi-a*	search for a long time

2.9.3 The persistive: *-ilizi-*^{2}

The extension *-ilizi-*^{2}, besides the meaning 'doing sth persistently, continuously or intensively', has the meaning 'doing some work continuously for payment'. Examples are given below:

Table 17: Examples of persistive *-iliz-*$_2$

	Verb root	Gloss	*-ilizi-*^{2}	Gloss
a.	*-diim-*	herd	*-diim-ilizi-a*	herd for payment
b.	*-hakul-*	harvest honey from beehive	*-hakul-ilizi-a*	harvest honey for payment
c.	*-tumam-*	work	*-tumam-ilizi-a*	work for payment
d.	*-fufúúl-*	clear farm ready for planting	*-fufúúl-ilizi-a*	clear farm for payment
e.	*-lim-*	cultivate	*-lim-ilizi-a*	cultivate for payment

The use of this extension for the meaning 'doing some work continuously for payment' is still moderately productive, i.e. it can be used with any verb describing work that one does for payment.

2.10 Frequentative *-agul, -aguk-*

These are moderately productive extensions, not only in Sumbwa but also in related languages, e.g. Sukuma (Richardson & Mann 1966) and Nyamwezi (Mag-

anga & Schadeberg 1992). I follow Guthrie (1971: 144) in using the term "frequentative" for these extensions. Other terms have also been used: for example, augmentative (Lodhi 2002) and iterative-separative (Maganga & Schadeberg 1992: 167).

I take **-agud-* to be present in Proto-Bantu as reconstructed by Guthrie (1971: 144). I also assume that the stative form **-aguk-* was also present.[3] In Sumbwa, these extensions are reflected as *-agul-* and *-aguk-*.

The usual meaning for these extensions is 'doing something quickly or hurriedly, excessively or clumsily, and repeatedly'; *-agul-* has an active meaning, and does not add any argument to the basic valence of a verb, while *-aguk-* has a stative meaning, and deducts an argument from the basic valence of a verb. The various contexts in which these extensions occur are spelt out below.

2.10.1 Examples involving both extensions

These examples include verbs which are transitive when used with *-agul-* but become intransitive when used with *-aguk-*. Consider Table 18:

Table 18: Examples involving frequentative *-agul-*, *-aguk-*

	Verb root	Gloss	Frequentative active	Gloss	Frequentative stative	Gloss
a.	*-simb-*	dig	*-simb-agul-a*	dig intensively	*-simb-aguk-a*	become dug up intensively
b.	*-bel-*	break	*-bel-agul-a*	break into small pieces	*-bel-aguk-a*	be broken into small pieces
c.	*-kat-*	cut	*-kat-agul-a*	cut into small pieces	*-kat-aguk-a*	be cut into small pieces
d.	*-kuul-*	extract; uproot	*-kuul-agul-a*	extract/ uproot repeatedly	*-kuul-aguk-a*	become extracted/ uprooted intensively
e.	*-dul-*	bore	*-dul-agul-a*	pierce with many holes	*-dul-aguk-a*	be riddled with many holes

[3] Another possibility would be to assume **-agul-* and **-aguk-* to have evolved as a combination of **-ag-* (frequentative) and **-ul-* (intensive, active) and **-uk-* (intensive, stative). The extension **-ag-/-ang-* is glossed as "repetitive" and noted to behave "tonally as an extension" but functions also as an inflectional suffix with the meaning "durative/habitual" (Schadeberg 2003: 72).

2.10.2 Examples involving *-agul-* only

There are many verbs which take the *-agul-* extension, but not *-aguk-* . In Table 19 are a few examples. As can be noted here, the verbs involved are all transitive verbs which do not allow the *-aguk-* extension.

Table 19: Examples involving *-agul-* only

	Verb root	Gloss	Frequentative active	Gloss
a.	*-βoh-*	tie	*-βoh-agul-a*	tie clumsily and quickly
b.	*-βeez-*	carve	*-βeez-agul-a*	carve clumsily
c.	*-tuk-*	insult	*-tuk-agul-a*	insult excessively
d.	*-tah-*	draw (water, etc.)	*-tah-agul-a*	draw (water, etc.) excessively or quickly
e.	*-moog-*	shave	*-moog-agul-a*	shave clumsily

2.10.3 Examples involving *-aguk-* only

Notice that all the verbs involved here are intransitive, and the suffixing of *-aguk-* to the verb root results in some sort of 'state' or 'condition'.

Table 20: Examples involving *-aguk-* only

	Verb root	Gloss	Frequentative stative	Gloss
a.	*-lul-*	be bitter	*-lul-aguk-a*	become excessively bitter
b.	*-gin-*	be fat	*-gin-aguk-a*	become excessively fat
c.	*-duuh-*	be blunt	*-duuh-aguk-a*	become blunt quickly
d.	*-nunk-*	smell	*-nunk-aguk-a*	stink

2.10.4 Cases involving -CVCVC- verb roots

All the above examples involve -CVC- verb roots, which represent the overwhelming majority of verbs targeted by these extensions. Occurrence of these

Table 21: Cases involving -CVCVC- verb roots

	Verb root	Gloss	Frequentative active	Gloss
a.	*-tafun-*	chew	*-tafun-agul-a*	chew intensively or clumsily
b.	*-heken-*	chew	*-heken-agul-a*	chew excessively

extensions with -CVCVC- verb roots is not commonly observed, although the examples in Table 21 have been attested.

It should be noted that the majority of -CVCVC- verb roots may take any of the extensions after the elision of the final VC. In Tables 22 and 23 are some examples.

Table 22: Examples of frequentative *-agul-* in -CVCVC- verb roots

	Verb root	Gloss	Frequentative active	Gloss
a.	*-hepul-*	cause to be hungry	*-hep-agul-a*	cause to be excessively hungry
b.	*-hogol-*	break sth off	*-hog-agul-a*	break sth off quickly
c.	*-gangul-*	crack sth	*-gang-agul-a*	crack sth quickly
d.	*-tandul-*	tear	*-tand-agul-a*	tear quickly or excessively

As can be noted here, the verb roots *-hepul-, -hogol-, -gangul-, -tandul-, -sambul-*, and *-konyol-* are not used with the extensions;[4] instead, the truncated forms, i.e. *-hep-, hog-,* etc. are used.

2.10.5 Cases of lexicalization of *-agul-, -aguk-*

Cases of lexicalization include words whose verb roots do not have any meaning connection with the extended form. These words include the examples in Table 24. The asterisk (*) in the verb-root column indicates that these are not attested verb roots but "reconstructed" ones.

[4]It is possible that the *-ul-* in *-hep-ul-, -hog-ol-, -tand-ul-,* etc. was once used as an extension in prehistory.

Table 23: Examples of frequentative *-aguk-* in -CVCVC- verb roots

	Verb root	Gloss	Frequentative stative	Gloss
a.	*-hepul-*	cause to be hungry	*-hep-aguk-a*	become excessively hungry
b.	*-hogol-*	break sth off	*-hog-aguk-a*	become broken off quickly
c.	*-gangul-*	crack sth	*-gang-aguk-a*	become cracked quickly or excessively
d.	*-tandul-*	tear	*-tand-aguk-a*	become torn quickly or excessively

Table 24: Cases of lexicalization of *-agul-* and *-aguk-*

	Verb root	Frequentative active	Gloss	Frequentative stative	*Gloss*
a.	**-ken-*	*-ken-agul-a*	destroy; spoil	*-ken-aguk-a*	become destroyed/ spoiled
b.	**-pos-*	*-pos-agul-a*	break into many pieces	*-pos-aguk-a*	become broken into many pieces
c.	**-tamp-*	*-tamp-agul-a*	pierce repeatedly with pointed weapon	*-tamp-aguk-a*	become riddled with piercings
d.	**-hunz-*	*-hunz-agul-a*	exhaust	*-hunz-aguk-a*	be exhausted

The asterisk (*) in the verb-root column indicates that these are not attested verb roots but "reconstructed" ones.

2.11 The denominative *-h-*, *-ahal-*

These extensions are used in the derivation of verbs from nouns or adjectives. In Table 25 are examples.

There are many more examples. The above derived examples may in turn accept other extensions, such as causative, applicative and passive.

Table 25: Examples of the denominative *-h-* and *-ahal-*

	Adjective or nominal base	Gloss	Derived verb	Gloss
a.	*-angu*	quick	*angu-h-a*	hurry
b.	*-gazi*	wide	*gazi-h-a*	become wide
c.	*-ganzi*	favourite	*ganzi-h-a*	become favourite
d.	*-bhanz*	brave	*bhanzi-h-a*	become brave
e.	*-ingi*	many	*ingi-h-a*	become many
f.	*-daasa*	sterile (of animals)	*daas-ahal-a*	become sterile
g.	*-guzu*	strength	*-guzu-ahal-a [guzuhala]*	become strong

2.12 The reiterative: *-ul-, uul-*

This term is Guthrie's (1971: 144), and the implication is one of "added quantity or quality or intense effort". The actual meaning will depend on the meaning of the root. In Sumbwa, there appears to be two forms of this extension: *-uul-* and *-ul-*. Examples for the first one are in Table 26.

Table 26: Examples of the reiterative *-uul-*

	Verb root	Gloss	Reiterative *-uul-*	Gloss
a.	*-kam-*	squeeze (as when milking cow)	*-kam-uul-a*	squeeze tightly (as when preparing juice with hands)
b.	*-kemb-*	trim; pare	*-kemb-uul-a*	trim evenly
c.	*-tah-*	draw (e.g. water)	*-tah-uul-a*	scoop or draw in large quantities
d.	*-han-*	admonish	*-han-uul-a*	advise strongly
e.	*-sem-*	bevel	*-sem-uul-a*	cut evenly a large portion of

As can be seen, the meanings include ‘squeeze tightly’, ‘trim evenly’, ‘draw in large quantities’, ‘advise strongly’, and ‘cut a large portion of’. There are also a few items where the *-ul-* instead of *-uul-* is used, as in Table 27.

Table 27: Examples of the reiterative *-ul-*

	Verb root	Gloss	Reiterative *-ul-*	Gloss
a.	*-simb-*	dig	*-simb-ul-a*	uproot
b.	*-hel-*	grind coarsely	*-hel-ul-a*	grind excessively coarsely
c.	*-seng-*	cut	*-seng-ul-a*	cut trees for building

The reiterative *-ul-* and *-uul-* extension is only slightly productive and does not affect the basic valence of the verb.

2.13 The static *-am-*

The general meaning for this extension is ‘assume or be in a position or state’. Consider the examples in Table 28.

Table 28: Examples of the static *-am-*

	Verb/nominal root	Gloss	Static *-am-*	Gloss
a.	*-fuk-*	kneel	*-fuk-am-a*	kneel obediently; menstruate
b.	*i-papa*	wing	*-pap-am-a*	beat (of bird’s wing); palpitate
c.	*-gazi*	wide	*-gaz-am-a*	widen
d.	*-gond-*	bend	*-gond-am-a*	bend
e.	*-hanga*	alive	*-hang-am-a*	live for a long time

In Table 28, examples a and d involve verb roots; example b involves a noun, and c and e involve adjectival roots. The static meaning of the derived verb is quite clear.

There are also some examples involving verb roots whose meanings have been lost but may be recoverable through connection with related static and impositive derived forms. In Table 29 are examples.

All the static forms in Table 28 and Table 29 may take the applicative and causative extensions. As can be noted, the extension is only slightly productive.

Table 29: Examples of the static *-am-* involving verb roots whose meanings have been lost

	Verb root	Gloss	Static *-am-*	Gloss	Impositive forms
a.	*-gol-*	*	*-gol-am-a*	INTR bend	*-go-lek-a* (TR bend)
b.	*-heng-*	*	*-heng-am-a*	INTR tilt	*-heng-ek-a* (TR tilt)
c.	*-send-*	*	*-send-am-a*	be leaning	*-send-ek-a* (TR lean sth against sth else)
d.	*-in-*	*	*-in-am-a*	bend, stoop	*-in-ik-a* (TR lay over on one side)

2.14 The contactive: *-at-*

This extension is not productive. The original meaning of the extension implies some contact by an agent on a beneficiary or patient. In Table 30 are the few examples available in our data.

Table 30: Examples of contactive *-at-*

	Verb root	Gloss	Contactive *-at-*	Gloss
a.	*-kwa-*	---	*-kwa-at-a*	hold
b.	*-kumb-*	cover	*-kumb-at-a*	embrace
c.	*-fumb-*	close	*-fumb-at-a*	embrace
d.	*-lam-*	---	*-lam-at-a*	stick firmly; adhere

In these examples, only *-fumb-* and *-kumb-* have meanings that may be grasped by a native speaker. The other verb roots have no meaning that may be related to the meaning of the extended bases. The extension does not change the basic valence of the verb.

2.15 Other less-known verb extensions

In the data we have, there are four non-productive extensions, *-agil-*, *-agan-*, *-al-* and *-l-*.

2.15.1 *-agil-*

This is probably a combination of **-ag-* (a repetitive extension, cf. Schadeberg 2003: 96) and the applicative **-id-*. In Sumbwa, it appears to have a 'persistive meaning', as the examples in Table 31 show.

Table 31: Examples involving *-agil-*

a.	*-sunt-*	limp	*sunt-agil-a*	limp along
b.	*-yomb-*	speak	*yomb-agil-a*	talk too much
c.	*-kand-*	step on	*kand-agil-a*	walk fast in hot sun
d.	*-kump-*	stumble	*kump-agil-a*	stumble along
e.	*-don- ?*	?	*don-agil-a*	start to walk (infant)
f.	*-met- ?*	shine	*met-agil-a*	strut about
g.	*-kum- ?*	gather	*kum-agil-a*	move
h.	*-zwi-*	ideophone	*zwi-agil-a [zwiigila]*	squack (like a baby)

Notice that the only examples where there is a close relation between the basic root and the extended base are the first four. The last example *-zwi-* is ideophonic: it imitates the cry of an infant. In the data we have, there are only about ten words with the *-agil-* extension.

2.15.2 *-agan-*

This also appears to be a persistive extension. There are not many examples (Table 32).

Table 32: Examples involving *-agan-*

a.	*βumb-*	mould (with clay)	*βumb-agan-a*	stick together (soil)
b.	*-om- (intr)*	harden; dry	*om-agan-a*	harden from drying; solidify
c.	*-vimb-*	swell	*vimb-agan-a*	swell very much
d.	*-tab-*	confuse	*tab-agan-a*	get quite confused
e.	*-tonto(lok)-*	weaken	*tonto-gan-a*	weaken further

As can be observed here, the examples show a clear meaning relationship between the root and the extended base. There is, however, one example where the *-agan-* extension functions as a denominative extension:

(10) *ma-paβa* mischief *-paβ-agan-a* be mischievous

In this example, *-paβa* 'mischief' is a nominal stem; when suffixed with *-agan-* it changes into a verb.

2.15.3 *-al-*

The **-ad-* > *-al-* extension was recognized by Meeussen (1967: 90), but he notes that it 'appears partly as an expansion, partly as a suffix with ill-defined meaning'. Examples given include: **-dúad-* 'be ill', **-démad-* 'be crippled', **-ikal-* 'sit' whose Sumbwa reflexes are*: -lúal-, -lémal-, -ikal-,* respectively. Schadeberg (2003: 72) calls it "extensive", by which he means 'to be in a spread-out position'. This meaning is probably borne out by such Sumbwa words as *-samb-aal-a* 'spread', but there are not many. However, inspection of the various examples in Sumbwa shows that there is no single meaning that may be attributed to this extension. For some of the examples, the meaning of this extension is 'change into a state', as illustrated in Table 33 below:

Table 33: Examples involving *-al-* 'change into a state'

	Verb/nominal root/ stem	Gloss	*-al-*	Gloss
a.	*syaha* (n)	anger	*syah-al-a [syaahala]*	be angry
b.	*-humb-*	be stupid	*-humba-al-a [humbaala]*	lose one's mind
c.	*-lema* (n)	lame	*-lem-al-a [lemala]*	become lame

For the remaining few examples, there are different senses attached to the extension. For instance, the extension has an intensive meaning in the examples in Table 34.

Table 34: Examples involving *-al-* 'intensive'

a.	*-sees-*	pour out, e.g. water	*seesek-al-a [seesekala]*	pour out completely
b.	*-siis-*	spoil (TR)	*siisik-al-a [siisikala]*	spoil completely

In our last example, *-al-* acts as a denominative suffix, creating a verb which indicates action:

(11) *i-suβa* urine container *-suβa-al-a* [suβaala] urinate

In this example, *-suβa* 'urine container' is a nominal stem; if suffixed with *-al-* it changes into the verb *-suβa-al-a*.

2.15.4 *-l-*

This extension is also noted by Meeussen (1967: 91), and he gives examples such as **-ganud-* 'narrate' < **-ganú* 'tale', *-púmúd-* 'breathe', **-púmu* 'breath, rest', **-pokud-* 'make blind' < **-poku* 'blind', etc. In Sumbwa, examples showing the *-l-* extension are presented in Table 35.

Table 35: Further examples involving *-l-* (with various meanings)

a.	*-hofu*	'blind'	*-hofu-l-a* ⟶ *[hofula]*	'be blind'
b.	*-panti*	'deaf'	*-panti-l-a* ⟶ *[pantila]*	'become deaf'
c.	*-sefu*	'nausea'	*-sefu-l-a* ⟶ *[sefula]*	'nauseate'

As can be seen in Table 35, the meaning of the examples a and b is 'change into a state', but the meaning of example c *-sefu-l-a* 'to nauseate' is causative.

3 Co-occurrence constraints

Co-occurrence constraints, otherwise referred to as "suffix ordering" constraints in the literature, have been the subject of intense discussion in Bantu linguistics studies for quite some time (cf. e.g. Baker 1985, Alsina 1999, Hyman 2002). The main debate is whether there are Pan-Bantu constraints that govern multiple affixation. Three main approaches may be identified: a semantic or compositional approach (whereby affix order is based on 'relevance' – the most relevant is closer to the verb root, and the least farthest from the verb root – cf. Bybee (1985)); a syntactic approach (whereby affix ordering reflects syntactic derivation, cf. Baker's (1985) "mirror principle") and the strictly morphological approach (whereby suffix ordering is strictly governed by morphological criteria in the majority of cases, while exceptions are handled in reference to semantic or syntactic criteria cf. Hyman 2002). For our purposes, we consider Hyman's (2002) morphocentric approach to be germane as a point of departure. We summarize

the main ideas, and consider whether the proposed orderings are supported by the Sumbwa data.

Hyman (2002), using the Optimality Theory framework, assumes that Bantu suffix ordering is determined by the ranking of two licensors:

1. CARP (CAUS – APPL – RECP – PASS) Template – which licenses suffix ordering in most Bantu languages; quite general and highly ranked;

2. Non-templatic constraints (i.e. semantic compositionality or MIRROR constraints which deal with all cases which do not follow the CARP template).

To formulate these postulates, Hyman used data from Chichewa, Kinande, Chibemba, Chimwiini, Luganda, Ciyao, Emakua, Nyakyusa, Tonga and other Bantu languages. Table 36 shows some examples of suffix orders licensed by the two proposed licensors: the CARP template and the non-templatic constraints. The suffix order examples are from Chichewa: *-mang-* = 'tie', *-its-* = CAUS, *-ir-* APPL, *-an-* = RECP.

Table 36: Examples of suffix ordering in Bantu

Suffix orders governed by the CARP template		Suffix orders governed by non-templatic constraints	
CAR	*mang-its-ir-an* 'cause to tie for each other'	CRA	*mang-its-an-ir-an-* 'cause to tie for each other'
CA	*mang-its-ir* 'cause to tie for'		
AR	*mang-ir-an* 'tie for each other'		

Now, what is the situation like in Sumbwa? One important difference which sets Sumbwa (and other similar languages) apart from languages like Chichewa is that the former does not have a productive reciprocal/associative extension *-an-*; instead, reciprocity is expressed by the pre-verb root *-i-*, which is also a reflexive marker (cf. §2.7). Due to its pre-verb root position, the reciprocal *-i-* cannot participate in suffix ordering. We have to keep this in mind as we present the suffix ordering facts as they pertain to Sumbwa.

The first attempt to state Sumbwa suffix ordering constraints was in Kahigi (2008b: 71). Below I present a modified statement of these constraints:

1. An affix type cannot be repeated in the same verb stem (as Table 36 shows, there is repetition of *-an-* in Chichewa *-mang-its-an-ir-an-*; this does not occur in Sumbwa). The only exception observed has to do with one verb, *-zi-* 'go', and the applicative extension, as shown below:

(12) a. *a-la-zi-a*
SM1-PST-GO-FV
'he went'

b. *a-la-zi-il-a* *si-ntu*
SM1-PST-GO-APPL-FV 7-thing
'he went for sth'

c. *a-la-mu-zi-il-il-a* *si-ntu*
SM1-PST-OM1-GO-APPL-APPL-FV 7-thing
'he went for sth for him'

2. The maximum number of affixes that can co-occur in a verb stem is four. Example:

(13) REV + FRE + PERS + PASS
dod-ool-agul-iliz-ibhw-a > [dodoólagulizibwa]
SEW-REV-FRE-PERS-PASS-FV
'be caused to sew quickly for pay'

3. The Passive may follow the Bare Verb Root, applicative, instrumental, persistive, frequentative and causative.

(14) a. *-kat-u-a ⟶ [katwa]* 'be cut' VR+PASS
b. *-kat-il-u-a ⟶ [katilwa]* 'be cut for' APPL+PASS
c. *-kat-iisi-iβu-a ⟶ [katiisiβwa]* 'be cut with' INST+PASS
d. *-vig-ilizi-iβu-a ⟶ [vigiliziβwa]* 'be squeezed tightly' PERS+PASS
e. *-kat-agul-u-a ⟶ [katagulwa]* 'be cut repeatedly' FRE+PASS
f. *-kat-iisi-iβu-a ⟶ [katiisiβwa]* 'be caused to cut' CAUS+PASS

4. The associative/reciprocal *-i-* may occur with the following: Bare Base, APPL, FRE, PERS. It never occurs with PASS and ST/NEU. As noted earlier (cf. §2.7), this associative marker is also the reflexive marker; hence, all such constructions are ambiguous.

(15) a. *-i-kat-a → [ikata]* 'cut each other'/'cut oneself'
b. *-i-kat-il-a → [ikatila]* 'cut for each other'/'cut for oneself'
c. *-i-kat-agul-a → [ikatagula]* 'cut each other repeatedly'/'cut oneself...'
d. *-i-kwat-ilil-a → [ikwaatilila]* 'hold each other tightly'/'hold oneself ...'

The associative *-an-* is predominantly restricted to -CVC- verb roots. In the data we have, there are only a few examples that show co-occurrence with other extensions:

(16) a. *-lek-an-iisi-iβu-a* → [lekaniisiβwa]
leave-RECP-CAUS-PASS-FV
'be separated from each other'
b. *-βi-h-il-an-a* → [βiihilana]
bad-DEC-APPL-RECP-FV
'be bad for each other' (i.e. 'be angry with each other')
c. *-li-iisi-an-a* → [liisjana]
eat-CAUS-RECP-FV
'feed each other'

The meaning in (16b) suggests that the example *-βiihilana* ('be angry with each other') appears to be lexicalized.

5. In all co-occurrence cases, the Passive occurs last before the final element, FV.

The statements in 1 – 5 may be summarized as follows:

1. APPL + APPL
2. REV + FRE + PERS + PASS
3. a) BB + PASS
   b) APPL + PASS
   c) INST + PASS
   d) FRE + PASS
   e) CAUS + PASS
4. Constructions with *-i-* (RECP) may allow APPL, FRE, PERS, but not PASS, ST/NEU. On the other hand, constructions with *-an-* may allow the following orderings:

a) RECP + CAUS + PASS

b) DEN + APPL + RECP

c) CAUS + RECP

Considerations in this section lead to the following conclusions:

1. Sumbwa does not have a single example illustrating a complete CARP ordering.
2. The only examples that could be taken to partially follow the CARP template are: *-βi-h-il-an-a* (APPL + RECP), *-li-iisi-an-a* (CAUS + RECP) and PASS, which occurs last in the ordering.
3. The remaining examples do not fit in the CARP template.

4 Sumbwa verb extensions and parameters of Bantu morphosyntactic variation

Having presented the verb extensions in Sumbwa in §2 and the co-occurrence constraints in §3, we are now in a position to deal with the parameters of Bantu morphosyntactic variation as presented in Guérois et al. (2017). As pointed out in the introduction (§1), the relevant parameters are in §5 of the master list, i.e. parameters 36-48 which deal with verbal derivation. The parameters have to do with the canonical passive, the 'impersonal' passive, agent noun phrase, bare agent, reciprocal, other functions of *-an-*, causative, instrumental causative, applicative, applicative functions, multiple applicative extensions, neuter/stative, and the order of suffixes. The objective of the exercise is to provide data that may be used in identifying micro-variation among Bantu languages with respect to the proposed parameters. Some of the questions have already been answered in §2. In this section, we summarize the relevant points and provide further discussion of any points not covered in previous sections.

4.1 Canonical passive (Parameter 36)

A canonical passive is taken to be a normal passive which is a "construction by which the subject of an active clause is demoted to an oblique or remains unexpressed, while the object is promoted to subject status" (cf. Kula & Marten 2010: 2). It is thus a result of classical passivization, which involves a transitive verb, and which can be expressed in a rule format as NP_1 + V + NP_2 $\longrightarrow$ NP_2 +

V-*w*-*a* (*na* + NP_1), describing the canonical Swahili passive, where -*w*- represents the passive extension (with its allomorphs), -*a* the FV (with its allomorphs), and *na* the preposition that is the head of the optional oblique NP.

In Sumbwa, as in most Bantu languages, the canonical passive is expressed through a verbal extension. As shown in §2, passives in Sumbwa are marked by -*u*- (occurring after consonantal-final verb roots) and -*iβu*- (occurring after vowel-final verb-roots). Examples shown in §2.2 summarize the facts on Sumbwa passivization.

4.2 "Impersonal" passives (Parameter 37)

The so-called impersonal passives are non-canonical. A case in point is the *ba*-passive construction in Bemba, a language of Zambia, where "… the active clause subject, as in typical passives, is demoted to an oblique position introduced by a preposition or remains unexpressed. The preferred preposition to introduce agents is *ku-*/*kuli-* 'by', while *na* 'by/with' is more frequent with instruments" (Kula & Marten 2010: 118). An example of the *ba*-passive is given in (17) below, where (17a) is active while (17b) is passive:

(17) Kula & Marten (2010: 119)

a. *umw-áàna bá-alí-mu-ít-a ku mu-mbúlu*
1-child SM2-PAST-OM1-call-FV by 3-wild.dog
'The child was called by the wild dog.'

b. *bá-alí-it-a umw-áàna ku mu-mbúlu*
SM2-PAST-call-FV 1-child by 3-wild.dog
'The child was called by the wild dog.'

A characteristic of the passive in (17b) is that "the theme argument is not clearly promoted to subject position: It remains in situ in post-verbal position" (Kula & Marten 2010: 119). This construction does not occur in Sumbwa.

4.3 Agent noun phrase (Parameter 38)

The agent noun phrase in Sumbwa is introduced by the preposition *ne*.

(18) a. *mu-ana a-la-tem-a mu-ti*
1-child SM1-PST-cut-FV 3-tree
'the child cut a tree'

b. *mu-ti gu-la-tem-w-a ne mu-ana*
3-tree SM1-PST-cut-PASS-FV by 1-child
'a tree was cut by the child'

It is important to note that the agent noun phrase may be dropped if the focus is on the patient that is the new subject:

(19) *mu-ti* GU-LA-TEM-W-A
3-tree SM1-PST-cut-PASS-FV
'a tree was cut'

There are also other constructions where the agent noun phrase is not needed, as noted in §2.2.

4.4 Bare agent (Parameter 39)

Can the preposition *ne* be omitted and the passive construction remain grammatical? In Sumbwa, such omission will always result in ungrammatical sentences, and is not allowed, as the following examples shows:

(20) a. *mu-ti gu-la-tem-w-a ne mu-ana*
3-tree SM1-PST-cut-PASS-FV by 1-child
'a tree was cut by the child'

b. * *mu-ti gu-la-tem-w-a mu-ana*
3-tree SM1-PST-cut-PASS-FV 1-child
'a tree was cut by the child'

As can be noted here, (20a) with *ne* is grammatical, while (20b) without is not.

4.5 Reciprocal (Parameter 40)

As shown in §2.7, there are two reciprocal/associative markers in Sumbwa, *-i-* and *-an-*, the former occurring in pre-verbal position and the latter in post-verbal position. It should be noted that the marker *-i-* is the more frequent one. There are a few examples which use the extension *-aan-* instead of *-an-*. As already noted, *-i-* is also a reflexive marker, making it ambiguous. Examples are:

(21) a. *βa-la-li-il-an-a*
SM2-PST-eat-APPL-REC-FV
'they ate at each other's home'

b. * *βa-la-i-li-a* *[βalíílja]*
SM2-PST-PAST-REC-eat-FV
'they ate each other'

As can be noted here, which of the two reciprocal/associative marker is used is not a free choice. The choice depends on several factors, some of which are:

1. The meaning of the verb involved; if the meaning is incongruous as in (21b) above the polysemous *-i-* is avoided
2. is restricted to shorter verb roots, usually -CVC-, and occasionally -CVCVC-

4.6 Other functions of the associative (Parameter 41)

In addition to the reciprocal function of the associative *-i-* and *-an-*, there are some examples which indicate the comitative function, as follows:

(22) a. *βa-la-gaβ-aan-a*
SM2-PST-divide-ASS-FV
'they shared'
b. *βa-la-lek-aan-a*
SM2-PST-leave-ASS-FV
'they separated'

There is also one example which does not indicate either reciprocal or comitative function:

(23) *a-la-zí-an-a* *i-kóóti*
SM2-PST-go-ASS-FV 5-coat
's/he took a coat with her/him' (literally: 's/he went with a coat')

4.7 Causative (Parameter 42)

The causative extensions are *-i-* and *-iisi-* (cf. §2.3). Briefly, these forms are distributed as follows: *-i-* occurs in verb roots with final consonants. It is accompanied by spirantization of /p, b, t, d, l, k, g/ into [f, v, s, z, z, s, z], respectively. *-iisi-* occurs in verb roots with final vowels or consonants. In the examples in Table 37, *a-* is the class 1 subject marker, and *-la-* is the past tense marker.

As can be seen in Table 37, the causative *-i-* in a–c involves two rules: spirantization (/b/ ⟶ [v], /d, g/ ⟶ [z] and gliding /i/⟶[j].

Table 37: Examples of the causative *-i-* and *-iisi-*

	Verb root	Causative *-i-*	Causative *-iisi-*
a.	*-ββáámb-* 'peg out'	*a-la-ββáámb-i-a* → *[álaββáámvja]* 'he caused (sth) to be pegged out'	*a-la-ββáámb-iisi-a]* →*[alaββáámbíisja]* 'he caused (sth) to be pegged out'
b.	*-dod-* 'sew'	*a-la-dod-i-a* → *[aladozja]* 'he caused (sth) to be sewn'	*a-la-dod-iisi-a* → *[aladodeesja]* 'he caused (sth) to be sewn'
c.	*-og-* 'take a bath'	*a-la-og-i-a* → *[aloozja]* 'he bathed (sb)'	*a-la-og-iisi-a* → *[aloogeesja]* 'he bathed (sb)'

4.8 Instrumental causative (Parameter 43)

As already noted in §2.4, the extensions *-i-* and *-iisi-* are used for the causative and instrumental. Consider the following example, which was given in §2.4 and is repeated here as (24).

(24) a. *[alamulíísja mwaana]*
a-la-mu-li-iisi-a *mu-ana*
SM1-PST-OM1-eatCAUS-FV 1-child
'She caused the child to eat.' (i.e. she fed the child)

b. *[alalíísja siliko]*
a-la-li-iisi-a *si-liko*
SM1-PST-eat-CAUS-FV 7-spoon
'She ate with a spoon.'

The first sentence is a causative construction while the second is an instrumental.

4.9 Applicative (Parameter 44)

Applicative constructions are formed by using the extension *-il-*, as was noted in §2.1.

4.10 Applicative functions (Parameter 45)

Of the five functions noted by Ashton (1947: 218–221) for the Swahili applicative, at least four may be recognized for Sumbwa, as exemplified in the examples given in §2.1, some of which are repeated in Table 38 for convenience.

Table 38: Examples of functions of the applicative

	Verb root	Applicative	Example
a.	*-tem-a* 'cut'	*-tem-il-a [temela]*	*a-la-mu-tem-il-a muti* SM1-PST-OM1-cut-APPL-FV tree 'he cut a tree for her' (Benefactive)
b.	*-iluk-a* 'run'	*-iluk-il-a*	*a-la-iluk-il-a mu-numba* SM1-PST-run-APPL-FV 17-house 'he ran into the house' (Directional)
c.	*-húúl-a* 'whip'	*-húúl-il-a*	*a-ø-mu-húúl-il-a* SM1-HAB-OM1-whip-for-FV *βuzoβe* laziness 'he whips her for laziness' (Reason)
d.	*-dod-a* 'sew'	*-dod-il-a [dodela]*	*a-Ø–dod-il-a kaaya* SM1-HAB-sew-APPL-FV home 'he sews at home' (Location)

The 'reason' meaning expressed in c in Table 38 is also found in found in 'why' questions such as *a-Ø-mu-húúl-il-a si?* 'Why does he whip her?'. The location meaning in d agrees with the corresponding question *a-Ø-dod-il-a hi?* 'Where does he sew?'.

This multiplicity of functions of the applicative extension, recognized quite early by Bantuists (cf. Madan 1903: xii, Ashton 1947: 218–221), is true of many Eastern Bantu languages.

4.11 Multiple applicative extensions (Parameter 46)

In Sumbwa, as in other Bantu languages, there is what appears to be a "multiple applicative" extension due to the fact that it is a reduplication of the usual applicative extension, *-il-*. In this study, following Guthrie (1971: 144), we have called it the persistive *-ilil-* (cf. §2.9.1). In the Swahili-English Dictionary of 1939, Johnson called it "double prepositional".

Its function is to express intensity, repetition or completeness. It does not allow addition of an argument other than the one licensed by the verb root. This is shown in the following examples:

(25) a. *a-la-kwáát-a si-ntu*
SM1-PST-hold-FV 7-thing
'he held a thing'

b. *a-la-mu-kwáát-il-a si-ntu*
SM1-PST-OM1-hold-APPL-FV 7-thing
'he held a thing for him'

c. *a-la-kwáát-ilil-a si-ntu*
SM1-PST-hold-PERS-FV 7-thing
'he held the thing tightly'

In (25a), the verb root *-kwáát-* allows an argument *sintu* (thing). In (25b), the applicative extension *-il-* allows an extra argument, marked as *-mu-* i.e. the object marker, while in (25c) the persistive *-ilil-* does not allow any extra argument other than the one allowed by the verb root *-kwáát-*. So, in general, the persistive does not allow addition of an argument.

In Sumbwa, however, we find one exception, which was given in (12) and is repeated as (26) below:

(26) a. *a-la-zi-a*
SM1-PST-GO-FV
'he went'

b. *a-la-zi-il-a si-ntu*
SM1-PST-GO-APPL-FV 7-thing
'he went for sth'

c. *a-la-mu-zi-il-il-a si-ntu*
SM1-PST-OM1-GO-APPL-APPL-FV 7-thing
'he went for sth for him'

In (26a), the verb root *-zi-* 'go' does not allow an extra argument because it is intransitive. But the example in (26b), which is applicative, allows addition of one argument. In (26c), there is an addition of the applicative extension which is accompanied by the addition of an argument, *-mu-*, as beneficiary.

4.12 Neuter/stative (Parameter 47)

Bantuists have attributed two functions to the stative extension: to express state without implying agency, and to express "potentiality" (cf. Ashton 1947: 227–228).

In this study, we have covered the following extensions which express stative meanings in Sumbwa:

1. The usual stative marker *-ik-* (cf. §2.6)
2. The reversive stative *-uk-/-uuk-/-uuluk-* (cf. §2.8)
3. The frequentative stative *-aguk-* (cf. §2.10).

Since these extensions have been dealt with at length in the foregoing, reference should be made to the respective sections.

4.13 Order of suffixes (Parameter 48)

Co-occurrence constraints have been stated in §3. Here we shall be brief. Is there a specific order for Sumbwa productive verbal extensions? Does Sumbwa have the causative-applicative-reciprocal-passive (CARP) order postulated by Hyman (2002) for Bantu?

It is important to note, first, that, of the four extensions involved, i.e. causative, applicative, reciprocal, and passive, only three (causative, applicative and passive) enjoy very high productivity. The fourth, the associative/reciprocal, is expressed by two separate forms, *-i-* and *-an-*, of which the former is a pre-verb-root affix (and highly productive) and the latter is an unproductive extension. Given the fact that *-an-* is restricted mostly to -CVC- verb roots, and is currently unproductive, it becomes evident that *-an-* cannot have the freedom to combine freely with the other extensions.

Now, what are the orders that are allowed? These orders were spelt out in §3 above, but for convenience we present a few grammatical and ungrammatical examples here to show the orders allowed and not allowed:

(27) a. *-li-il-an-a*
eat-APPL-REC-FV
'eat at each other's home' APPL-REC

b. *-li-iisi-an-a*
eat-CAUS-REC-FV
'feed each other' CAUS-REC

c. * *-li-iisi-il-an-iβu-a*
eat-CAUS-APPL-REC-PASS-FV
'be caused to feed each other' CAUS-APPL-REC-PASS (CARP)

d. *-lek-an-iisi-iβu-a*
leave-REC-CAUS-PASS-FV
'be caused to leave each other' REC-CAUS-PASS

e. *-dod-ool-agul-iliz-ibhu-a*
sew-REV-FRE-APPL-CAUS-PASS-FV
'be caused to sew clumsily and quickly' REV-FRE-PERS-PASS

f. * *-dod-iisi-il-an-ibhu-a*
sew-CAUS-APPL-REC-PASS-FV
'be caused to sew each other' CAUS-APPL-REC-PASS (CARP)

The orders that are allowed are those in (27a, 27b, 27d, 27e). The orders in (27c) and (27f), based on the CARP hypothesis, are ungrammatical. Whether the orders above reflect non-templatic constraints as spelt out in Hyman (2002) is an issue for further study.

5 Conclusion

This study has revealed the following important facts about verb extensions in Sumbwa: their productivity, co-occurrence constraints, their valence possibilities, and their behavior in relation to the parameters of morphosyntactic variation proposed by Guérois et al. (2017).

The study reveals that most of the Proto-Bantu verb extensions reconstructed by Guthrie and Meeussen are still active in the language. The extensions may *roughly* be categorized into three groups: highly productive, moderately productive and least productive. The highly productive extensions are the applicative, passive, causative (also instrumental), stative, and frequentative. The pre-verbal associative *-i-* is also highly productive, while the status of the associative *-an-* needs further investigation. The moderately productive extensions are the persistive, reversive, impositive and denominative. The least productive are reiterative, static, contactive and other minor extensions. Overall, the productivity ranking

is similar to that in other Bantu languages (cf. Maganga & Schadeberg 1992, Rugemalira 1993, 2005, Schadeberg 2003, Stegen 2002, Waweru 2005, Chabata 2007).

The study also shows that Sumbwa verb extensions may be categorized as either valence-increasing, valence-decreasing or valence-maintaining, as in Table 39.

Table 39: Extensions categorized in terms of valence

Valence-increasing	Valence-decreasing	Valence-maintaining
Applicative	Passive	Reversive active
Causative	Stative/Neuter	Persistive
Instrumental	Associative	Reiterative
Impositive	Frequentative Stative	Frequentative active
	Reversive Stative	

The categorization in Table 39 generally reflects the valence possibilities for most Bantu languages.

Furthermore, in the answer to the questions in the Master List of the Parameters of Morphosyntactic Variation, the study has revealed some interesting facts that may be useful in Bantu comparative morphosyntax. These are summarized in Table 40.

Most of the characteristics listed in Table 40 are found in the majority of Bantu languages. There are, however, characteristics that are peculiar to Sumbwa (and other languages similar to it). These include:

1. The reflexive-reciprocal syncretism marked by the pre-verb root affix *-i-*. Reciprocity is expressed in a productive way by *-i-*, while the Proto-Bantu reciprocal extension *-an-* occurs only in restricted contexts. This characteristic is not restricted to Sumbwa; it has been reported in Rimi (F.32; Olson 1964) and Rangi (F.33; Stegen 2002). Other zone F languages should be investigated in connection with this feature to find out whether it is a characteristic for the zone.

2. The causative-instrument syncretism marked by the causative extensions *-i-* and *-iisi-*. This syncretism has been discussed in the literature (cf. Wald 1998). In this case, Sumbwa belongs to Bantu languages which no longer uses *-il-* to mark the instrumental role.

Table 40: Characteristics of the Parameters manifested in Sumbwa

No.	Topic	Important characteristics
36	Canonical Passive	1. It is expressed through a verbal extension *-u-/-iβu-* 2. There is no other strategy to express passivization.
37	"Impersonal" passive	There are no *ba*-passives
38	Agent Noun Phrase	The Agent NP in a passive construction is introduced by the preposition *ne*
39	Bare agent	The preposition which introduces the agent cannot be omitted.
40	Reciprocal	Through the use of the pre-verbal *-i-* (which is also the reflexive marker) and the suffix *-an-*
41	Other functions for *-an-*	Yes, it has the comitative function
42	Causative	It is expressed through suffixes *-i-* and *-iisi-*
43	Instrumental Causative	Yes, the causative extension is also used to introduce prototypical instruments
44	Applicative	Applicative constructions are formed through the use of the suffix *-il-*
45	Applicative functions	In addition to benefactive meaning, applicative constructions convey the following meanings: directional, location, reason.
46	Multiple applicative extensions	1. What appears to be a case of multiple applicative extension (i.e. *-ilil-*), is in fact a persistive extension. 2. The only possible case of multiple applicative extension is *-zi-il-il-a* 'go for (sth) for (sb)
47	Neuter/Stative	In addition to the stative *-ik-*, the language has the reversive stative and the frequentative stative.
48	Order of suffixes	1. CARP is not possible in the language 2. There is no systematic fixed order 3. The Passive always occurs last.

3. Sumbwa data do not support the CARP template as formulated in Hyman (2002). This is probably because the productive affix for reciprocity/associativeness is no longer *-an-* but *-i-* which occurs in pre-verb-root position and is not a suffix.

As a final remark, we need to stress the limited nature of this study and that all the above issues (and others listed in the above table) are of interest to Bantu comparative linguistics and require further in-depth investigation.

Acknowledgements

I am grateful to the editors and anonymous reviewers for comments which have improved parts of this chapter.

Abbreviations

ACT	active	INTR	intransitive
AM	agreement marker	OM	object marker
APPL	applicative	PASS	passive
ASS	associative	PERS	persistive
BB	bare base	PST	past
CAUS	causative	RECP	reciprocal
CLI	clitic	REIT	reiterative
DEN	denominative	REV	reversive
FRE	frequentative	SM	subject marker
FRE-ACT	frequentative-active	ST	stative
FRE-ST	frequentative stative	TAM	tense-aspect-modal marker
FV	final vowel	TR	transitive
HAB	habitual	TRV	transitive verb
IMP	impositive	VR	verb root
INST	instrumental		

References

Alsina, Alex. 1999. Where is the mirror principle? *The Linguistic Review* 16. 1–42.
Ashton, Ethel O. 1947. *Swahili grammar*. 2nd edn. London: Longman.
Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16. 373–415.

Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en bantou* (Serie in-80, Sciences Humaines 114). Tervuren: Annales du Musée Royal de l'Afrique Centrale.

Batibo, Herman. 1985. *Le Kesukuma (langue Bantu de Tanzanie: Phonologie, morphologie*. Paris: Éditions Recherche sur le civilisations.

Berger, Paul. 1937–1938. Die mit B. -île gebildeten Perfektstämme in der Bantusprachen. *Zeitschrift für Eingeborenen-Sprachen* XXVII. 81–122.

Bokamba, Eyamba. 1975. On the syntax and semantics of derivational verb suffixes in Bantu languages. In Frances Ingemann (ed.), *Proceedings of the 1975 Mid-America Linguistics Conference*, 38–50. Lawrence: Linguistics Dept., University of Kansas.

Bostoen, Koen. 2008. Bantu spirantization: Morphologization, lexicalization and historical classification. *Diachronica* 25(3). 299–356.

Bostoen, Koen & Yolande Nzang-Bie. 2010. On how 'middle' plus 'associative/reciprocal' became 'passive' in the Bantu A70 languages. *Linguistics* 48(6). 1255–1307.

Bybee, Joan. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Capus, Auguste. 1898. Grammaire de Shisumbwa. *Zeitschrift für Afrikanische und Oceanische Sprachen* IV. 1–123.

Chabata, Emmanuel. 2007. *The Nambya verb with special emphasis on the causative*. Oslo: University of Oslo. (Doctoral dissertation).

Dom, Sebastian, Leonid Kulikov & Koen Bostoen. 2018. Valency-decreasing derivations and quasi-middles in Bantu: A typological perspective. *Southern African Linguistics and Applied Language Studies* 36(3). 165–173.

Eastman, Carol. 1967. *An investigation of verbal extensions in Kenya coastal dialects of Swahili with special emphasis on KiMvita*. Madison, WI: University of Wisconsin. (Doctoral dissertation).

Givón, Talmy. 1970. On ordered rules and the modified base of ChiBemba verbs. *African Studies* 29. 47–54.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list*. Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.

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

Guthrie, Malcolm. 1962. The status of radical extensions in Bantu languages. *Journal of African Languages* 1(3). Printed in: *M. Guthrie 1970 Collected papers in Bantu linguistics*, 92–110. London: Gregg International Publishers, 202–220.

Guthrie, Malcolm. 1970. *Collected papers in Bantu linguistics*. London: Gregg International Publishers.

Guthrie, Malcolm. 1971. *Comparative Bantu*, vol. 2. London: Gregg International Publishers.

Haspelmath, Martin & Thomas Müller-Bardey. 2004. Valence change. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphology: An international handbook on inflection and word formation*, vol. 2, 1130–1145. Berlin: Walter de Gruyter.

Humphreys, R. L. 1999. Valency changing alternations. In Keith Brown & Jim Miller (eds.), *Concise encyclopedia of grammatical categories*, 391–400. Oxford: Elsevier.

Hyman, Larry M. 1995. Minimality and prosodic morphology in Cibemba imbrication. *Journal of African Languages and Linguistics* 16. 3–39.

Hyman, Larry M. 2002. Suffix ordering in Bantu: A morphocentric approach. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2002*, 245–281. Dordrecht: Kluwer.

Hyman, Larry M. 2007. Niger-Congo verb extensions: Overview and discussion. In Doris L. Payne & Jaime Peña (eds.), *Selected Proceedings of the 37th Annual Conference on African Linguistics*, 149–163. Somerville MA: Cascadilla Proceedinsg Project.

Jerro, Kyle. 2017. The causative-instrumental syncretism. *Journal of Linguistics* 53(04). 1–38.

Johnson, Frederick. 1939. *A Standard Swahili-English dictionary*. Oxford: Oxford University Press.

Kahigi, Kulikoyela. 1989. Developments affecting the Bantu perfective in *–ile*: Another view. *Kiswahili* 56. 92–107.

Kahigi, Kulikoyela. 2008a. Derivation in Sisumbwa. *Languages of Tanzania (LOT) Occasional Papers in Linguistics* 3. 53–81.

Kahigi, Kulikoyela. 2008b. *Sisumbwa-Kiswahili-English/English-Sisumbwa-Kiswahili dictionary* (Lexicon Series 11). Dar es Salaam: Languages of Tanzania (LOT) Project, University of Dar es Salaam.

Katamba, Francis X. & John Stonham. 2006. *Morphology*. 2nd edn. London: Macmillan Press.

Khamis, Abdu. 1972. *A consideration of grammatical roles in Swahili*. York: University of York. (MA thesis).

Khamis, Abdu. 1985. *Swahili verb derivation*. Honolulu, HI: University of Hawaii. (Doctoral dissertation).

Khumalo, Langa. 2007. *An analysis of the Ndebele passive construction*. Oslo: University of Oslo. (Doctoral dissertation).

Kula, Nancy. 2001. Imbrication in Bemba. In Elisabeth Hume, Norval Smith & Jeroen van de Weijer (eds.), *Surface syllable structure and segment sequencing* (HIL Occassional Papers 4), 102–116. Leiden: Holland Institute of Generative Linguistics (HIL).

Kula, Nancy & Lutz Marten. 2010. Argument structure and agency in Bemba passives. In Karsten Legere & Christina Thornell (eds.), *Bantu languages: Analyses, description and theory* (East African Languages and Dialects 20), 115–130. Cologne: Rüdiger Köppe.

Lodhi, Abdulaziz Y. 2002. Verbal extensions in Bantu: The case of Swahili and Nyamwezi. *Africa and Asia* 2. 4–26.

Madan, Arthur Cornwallis. 1903. *Swahili-English dictionary*. Oxford: Clarendon Press.

Maganga, Clement & Thilo C. Schadeberg. 1992. *Kinyamwezi: Grammar, texts, vocabulary*. Köln: Rüdiger Köppe.

Marten, Lutz. 2003. The dynamics of Bantu applied verbs: An analysis at the syntax-pragmatics interface. In Kézié Lébikaza (ed.), *Actes du 3e Congrès Mondial de Linguistique Africaine, Lome, 2000*, 207–221. Köln: Köppe.

Marten, Lutz, Nancy Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105. 253–338.

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

Meeussen, Achille E. 1969. *Bantu lexical reconstructions*. Tervuren: Royal Museum for Central Africa.

Mould, Martin. 1972. On reconstructing the modified base of Bantu verbs. *Studies in African Linguistics* 3. 107–125.

Munongo, Antoine Mwenda & Fernand Grévisse. 1967. *Pages d'histoire Yeke (by Munongo), followed by Les Yeke (by Grevisse)*. Elizabethville: CEPSI (Centre d'Etude des Problèmes Sociaux Indigènes).

Olson, Howard. 1964. *The phonology and morphology of Rimi*. Hartford: The Hartford Seminary Foundation. (Doctoral dissertation).

Payne, Thomas. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

Plag, Ingo. 2006. Productivity. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., 121–128. Oxford: Elsevier.

Richardson, Irvine & Michael Mann. 1966. A vocabulary of Sukuma. *African Language Studies* 7. 1–79.

Rugemalira, Josephat M. 1993. *Runyambo verb extensions and constraints on argument structure*. Berkeley, CA: University of California. (Doctoral dissertation).

Rugemalira, Josephat M. 2005. A grammatical sketch of Runyambo. In David P. B. Massamba (ed.), *Occasional papers in linguistics*, vol. 1, 38–74. Dar es Salaam: Languages of Tanzania (LOT) Project, University of Dar es Salaam.

Schadeberg, Thilo C. 1995. Spirantization and the 7-to-5 vowel merger in Bantu. *Belgian Journal of Linguistics* 9. 71–84.

Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 71–89. London & New York: Routledge.

Scotton, Carol. 1967. *Aspects of the Swahili extended verb system with special reference to some deep structure syntactic and semantic restrictions*. Madison, WI: University of Wisconsin. (Doctoral dissertation).

Stegen, Oliver. 2002. Derivational processes in Rangi. *Studies in African Linguistics* 31(1/2). 129–154.

Wald, Benji. 1998. Issues in the North/South syntactic split of East Bantu. In Ian Maddieson & Thomas J. Hinnebusch (eds.), *Language history and linguistic description in Africa*, 95–106. Trenton, NJ: Africa World Press.

Waweru, Martin. 2005. *Gikuyu verbal extensions: A minimalist analysis*. Nairobi: Kenyatta University. (Doctoral dissertation).

Chapter 10

Tense and aspect marking in Bantu languages of the Morogoro region, Tanzania

Malin Petzell[a] & Peter Edelsten[b]
[a]University of Gothenburg [b]SOAS University of London

A review of the verbal morphology of several Bantu languages of the Morogoro region, Tanzania, reveals surprising diversity in both their distribution and meaning. Bantu languages are renowned for their rich verbal morphology, including remoteness distinctions in the tenses. However, some Bantu languages of the region have essentially only two tenses (past and non-past), limited aspectual distinctions, and some have no negative tense/aspect markers. This chapter summarises our current knowledge of the tense/aspect systems of five Bantu languages of the Morogoro region: Kagulu (G12), Luguru (G35), Kami (G36), Ndamba (G52) and Pogoro (G51). In particular, the chapter reviews the distribution and meaning of these morphological distinctions, the abundance versus scarcity of specific tense/aspect markers, and the methods of expressing negation.

1 Introduction

1.1 Background

This chapter provides an analysis of the tense/aspect systems of five selected Bantu languages of the Morogoro region, Tanzania. The Morogoro region spreads from the area north of Morogoro town to the southern part of the Kilombero valley. Tanzania has about 100 Bantu languages (Maho & Sands 2002), most of them being poorly documented. Of the 100 Bantu languages spoken in Tanzania, 10 are spoken mainly in the Morogoro region:

Malin Petzell & Peter Edelsten. 2024. Tense and aspect marking in Bantu languages of the Morogoro region, Tanzania. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 277–321. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663779

- Kagulu (G12)
- Ngulu (G34)
- Luguru (G35)
- Kami (G36)
- Kutu (G37)
- Vidunda (G38)
- Pogoro (G51)
- Ndamba (G52)
- Sagala (G61)
- Mbunga (P15)

In addition, other languages which are spoken primarily in neighbouring regions but which have a significant presence in Morogoro region include:

- Zigula (G31)
- Kwere (G32)
- Zaramo (G33)
- Hehe (G62)
- Bena (G63)
- Ngoni (N12)
- Ngindo (P14)

Additionally, Swahili (G42) is spoken all over the country, and all consultants in this study are bilingual in Swahili.

Bantu languages are known for their rich verbal morphology, including elaborate sets of tense/aspect markers. The five chosen languages, although fairly closely related, show variation not only in the number of markers but also in their function. We will describe and analyse the tense/aspect marking in these languages based on models of Bantu verbal morphology, including tense/aspect,

by Meeussen (1967) and Guthrie (1967–1971). More recently, Nurse (2008) and Nurse & Devos (2019) present a review of tense/aspect data for 100 Bantu languages from across the entire Bantu language area, providing an analysis of the main patterns found and some proposals for their diachronic evolution, which we will also refer to.

The context and rationale for the present study was that both authors had been studying Bantu languages in different parts of the Morogoro region, and they decided to attempt to synthesise their results in one significant area of Bantu grammar: tense/aspect. The expectation was that, given their close proximity, the tense/aspect systems of the selected languages would show some similarities. In fact, they show a surprising amount of diversity.

1.2 The languages selected for the study

The five languages analysed in this study were selected on the basis that they are distributed across the region and that they might therefore be expected to expose variation in structures found throughout the region. The approximate locations of the five selected languages, Kagulu (G12), Luguru (G35), Kami (G36), Ndamba (G52) and Pogoro (G51), are shown in Figure 1.

In this chapter, all data in the examples are derived from the authors' fieldwork unless otherwise stated.

Kagulu (G12) is a Bantu language spoken in and around the Kagulu or *Itumba* mountains in the north-west of the region. The language is estimated to have between 240,000 (Petzell 2008) and 336,000 speakers (LOT 2009). Some speakers use the autonym *Chimegi* to refer to their language, while others prefer *Chikagulu*, since *Megi* is a derogatory term used by Maasai speakers, meaning 'non-Maasai' (Mol 1996: 251). The most prestigious Kagulu dialect stems from the mountains and is referred to as (*Chi*)*Tumba*. Data are sourced from Petzell (2008), supplemented from the authors' more recent fieldwork (2009–2020).

Luguru (G35) is a Bantu language spoken in the Luguru mountains south of Morogoro town. It is reported to have 400,000 speakers (LOT 2009) and it is a dominant language in the region. Data are sourced from Mkude (1974) and Petzell (2020), supplemented from the authors' fieldwork.

Mkude (1974) identifies two dialects of Luguru (highland and lowland), which are not well documented. An MA thesis (Moses 2018) questions this division and reaches the conclusion that there are indeed different dialects of Luguru, but that the division between highland and lowland is not clear (Moses 2018: 66). The dialects, which are mutually intelligible, are somewhat different in phonology and lexicon (Moses 2018).

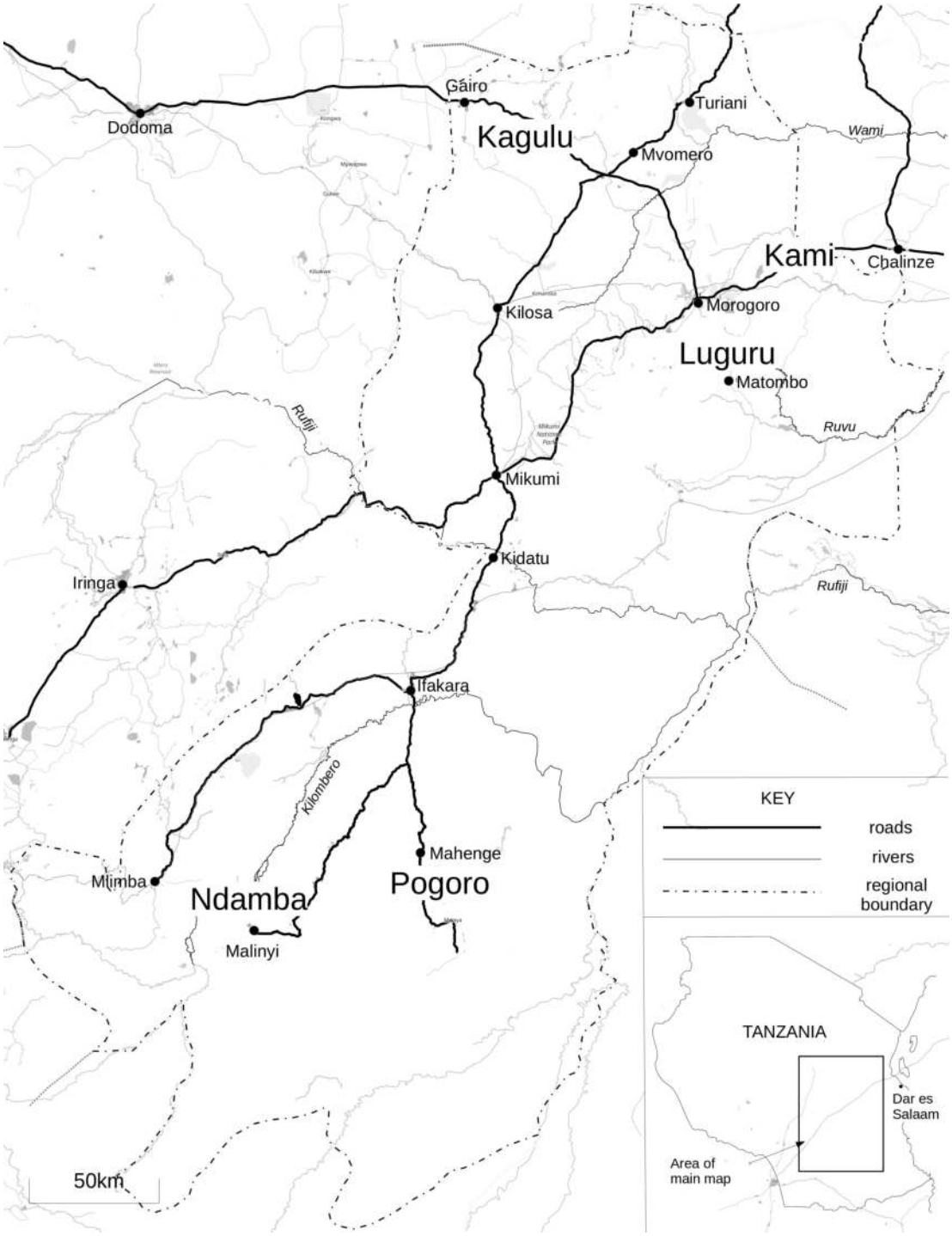


Figure 1: Languages of the study. Data sourced from © open-streetmap.org

Kami (G36) is a highly endangered Bantu language spoken around Mikese, east of Morogoro town. It is reported to have only 5000 speakers (LOT 2009). This figure refers to the number of persons who consider themselves to be Kami speakers, but it does not say anything about the competence of those speakers. There are significantly fewer than 5000 fluent speakers left, which was corroborated during field trips in the area. The youngest consultant we found was in his thirties, and he could understand but not speak Kami, which indicates that the language is not being transmitted to the next generation. Data are sourced from Petzell & Aunio (2019), supplemented by the authors' fieldwork.

Pogoro (G51) is a Bantu language spoken in the Pogoro Mountains in the south-east of the region. It is estimated to have 200,000 speakers (LOT 2009). Data are sourced from Hendle (1907), supplemented from the authors' fieldwork. Given the age of Hendle's work, his main conclusions about morphosyntax seem to correlate remarkably well with data collected recently, over 100 years later. Less clear is the current validity of the translation of many of the words in the word list, but this may be as much to do with the evolution of their semantics in their German translations as in the original Pogoro.

Ndamba (G52) is a Bantu language spoken in the Kilombero Valley in the south-west of the region. It is estimated to have between 55,000 (Lewis 2009) and 196,000 speakers (LOT 2009). Data are sourced from Edelsten & Lijongwa (2010), supplemented from the authors' fieldwork.

The variant of Ndamba documented by Novotná (2005) shows some differences from that documented by Edelsten & Lijongwa (2010). In particular, Novotna describes phonological features such as verb final *-i* and the loss of inflectional future tenses. These differences may show an influence from Pogoro, which may have contributed to Ndamba and Pogoro being grouped together by Guthrie (1948) as the G50 group of languages, and to the comment by Nurse (2008: Appendix 1, p. 177) that "G51 and G52 are quite similar".

Edelsten & Lijongwa's (2010) data, however, point towards Ndamba being somewhat more distinct from Pogoro, with complex verbal tense/aspect morphology, as discussed in §3.5, more reminiscent of neighbouring G60 languages such as Bena (Morrison 2011) and Hehe (Nurse 2008: Appendix 1, pp. 178–180).

1.3 Structure of the chapter

Following this introductory section, §2 discusses the verbal template used for the analysis, and how tense/aspect and related morphemes fit into the template in the selected languages. The objective is to provide a consistent basis for comparing the morphological structure of verbs in the languages of the study, while at

the same time reviewing whether the generally accepted template proposed by Meeussen (1967), as amended by Nurse (2008), is consistent with our template.

This is followed in §3 with a discussion of tense/aspect in Bantu languages in general, followed by subsections for each of the five selected languages.

§4 discusses related verbal categories in the selected languages, including imperative, subjunctive, conditional, habitual and negative, followed by a final section which draws conclusions from the analysis and provides suggestions for further research.

2 The verbal template

Bantu languages are often analysed as using morphological verbal templates, into which various affixes fit (Meeussen 1967, Nurse 2008). One of the reasons for using a template is to show how the affixes concatenate, since the order of affixes is typically strict. The ordering of syntactic elements, on the other hand, is typically much less restricted.

The exact specification of the template slots varies across Bantu languages, but the five selected languages show some uniformity. To compare the verbal morphology of the five selected languages, the template shown in Table 1, based on Meeussen (1967: 108–111), is used in this chapter.

Table 1: The verbal template

Template slot	Abbreviated to
Pre-subject marker	PRE.SM
Subject marker	SM
Post-subject marker	POST.SM
First tense/aspect marker	TA1
Object marker	OM
Verb root	ROOT
Extensions	EXT
Second tense/aspect marker	TA2
Passive suffix	PASS
Final vowel	FV
Post-final marker	POST.FM

Meeussen distinguishes two tense/aspect slots, “formative” and “limitive”, occurring before the object marker slot, but in Table 1, we have combined them in slot 4, as does Nurse (2008: 40).

Nurse (2008: 40) also combines TA2 with FV, such that FV then assumes a complex multi-morpheme role. For clarity, we have kept these slots separate. This is discussed further in §2.10.

(1) shows the use of the verbal template slots with data from Ndamba.[1]

(1) Ndembo a-ka-mu-somol-el-ile ngwena lwimbo. (Ndamba)
'The elephant sang the crocodile a song.'

from which *a-ka-mu-somol-el-ile* may be analysed as:

(2)

morphemes:	a-	ka-	mu-	somol-	el-	ile
slots:	SM-	TA1-	OM-	ROOT-	EXT-	TA2
gloss:	SM1-	PST-	OM1-	SING-	APPL-	PFV

's/he sang him/her a song'

Table 2 compares how the template slots are used in the five selected languages. Details and examples of how each of the slots is used are described in subsequent sections. The table shows that the way these slots are used is more varied than expected, given the proximity of the selected languages; this is discussed further in §3 and §4.

The following sections discuss how the template slots are used in the five selected languages.

2.1 Slot 1: Pre-subject marker (PRE.SM)

This slot is used for various pre-verbal affixes, the most common being the conditional/temporal marker (all five languages), the negative affix (four out of five languages) and the relative object marker (three out of five languages). Use of this slot for the two latter affixes is posited by Meeussen (1967: 108) for Proto-Bantu. Furthermore, Nurse (2008: 32) points out that negative and relative object markers are the affixes most commonly marked in this slot.

The negative markers found in this slot are discussed in §4.6.

Kagulu, Luguru and Pogoro all use a relative morpheme in this slot. Kagulu uses a relative morpheme based on *-o-*, which agrees with the noun class of the relativised object, as shown in examples (3), (4) and (5).

[1]Most Bantu languages are tonal (Marlo & Odden 2019). However, none of the languages selected for the study employs grammatical or lexical tone, and none of the examples in this chapter are therefore marked for tone.

Table 2: Use of template slots in the five selected languages

	Slot	Kagulu	Kami	Luguru	Ndamba	Pogoro
1	Pre-subject marker	relative object marker *-o-* temporal *fo-* negative marker 1 tense/aspect marker	negative marker temporal *fi-*	temporal *-(h)a-* relative object marker negative marker	conditional/ temporal *pa-*	tense/aspect marker negative marker *na-* relative object marker temporal *pa-*
2	Subject marker	subject marker	subject marker	subject marker	subject marker	subject marker
3	Post-subject marker	negative marker 2	temporal *-(h)a-*[a]			
4	First tense/aspect marker (TA1)	tense/aspect marker conditional *-ngh'a-*	tense/aspect marker	tense/aspect marker	tense/aspect marker	tense/aspect marker
5	Object marker	object marker reflexive/ reciprocal *-ki-*	object marker reflexive/ reciprocal *-i-*	object marker reflexive *-i-*	object marker reflexive *-i-*	object marker reflexive *-i-*

[a]It is impossible to be specific as to whether this morpheme is in slot 3 or slot 4 since it cannot co-occur with any TA marker.

Table 2: Use of template slots in the five selected languages (continued)

	Slot	Kagulu	Kami	Luguru	Ndamba	Pogoro
6	Verb root	verb root	verb root	verb root	verb root	verb root
7	Extensions	extensions	extensions	extensions	extensions	extensions
8	Second tense/ aspect marker (TA2)	habitual/ progressive *-ag- -ile*	habitual/ progressive *-ag- -ile*	habitual/ progressive *-ag- -ile*	habitual/ progressive *-agh-* perfective *-ile*	perfective *-iti*
9	Passive suffix	passive *-igw-*	passive *-igw-*	passive *-igw-*	*passive -w-*	
10	Final vowel (FV)	final vowel *-a* subjunctive *-e*	final vowel *-a* subjunctive *-e*	final vowel *-a* subjunctive *-e*	final vowel *-a* subjunctive *-e*	final vowel *-a* subjunctive *-e* future tense marker *-i*
11	Post-final marker	plural *-i*	plural *-ni*	plural *-ni* habitual/ progressive *-ag*	relative	

(3) Kagulu
yo-cha-mw-end-ile
REL.OM1-SM1PL.PST-OM1-love-PFV[2]
's/he who we loved'

(4) Luguru (Mkude 1974: 179)
mw-alimu, i-chi-tabu **chi**-a-mu-ing'-ile i-mw-ana
1-teacher AUP-7-book REL.**OM7**-SM1-OM1-give-PFV AUP-1-child
'the book which the teacher gave to the child'

(5) Pogoro
chi-gota chi-gu-kop-iti
7-chair REL.OM7-SBJ.2SG-buy-PRF
'the chair you bought'

In Kami, there is no specific marking of object relatives, as shown in examples (6) and (7).

(6) Kami
chi-nu chi-no wa-chi-sol-a wa-uz-a
7-thing 7-DEM.PROX SM2-SM7-take-FV SM2-sell-FV
'This thing (which) they took, they sold.'

(7) u-mw-ele a-kom-ile nguku Rahma.
AUP-3-knife SM1-kill-PFV 9. chicken NAME
'The knife with which Rahma killed the chicken.'

Ndamba also does not use slot 1 for a relative marker. Instead, it uses a post-verbal relative morpheme, as shown in example (8). This does not, however, seem to be derived from the pre-verbal relative morphemes used in other neighbouring languages like Bena and Ngoni (Morrison 2011, Ngonyani 2003).

(8) Ndamba
li-piki tu-ka-li-dumul-ile-**lyo** li-ka-pand-il-w-e na
5-tree SM1PL-PST-OM5-cut-PFV-REL.5 SM5-PST-plant-PFV-PASS-FV by
tati.
1a.father
'The tree which we cut down was planted by father.'

[2]Note that *-ile* no longer functions as a perfective in Kagulu.

In Kami, slot 1 may be used for the *fi-* conditional/temporal marker, as shown in example (9). However, this is less frequent than the *(h)a-* conditional/temporal marker in slot 3/4, as shown further below in example (14). The two markers *fi-* and *(h)a-* are mutually exclusive.

(9) Kami
fi-wa-tow-ile ngoma …
COND-SM2-play-PFV 9.drum
'when they played the drum …'

In Luguru, slot 1 can contain either a conditional/temporal marker or a negation. The two cannot co-occur. Either the conditional/temporal marker is used as shown in example (10), or an adverbial is used together with the negation marker, as shown in example (11). The conditional/temporal is further discussed in §4.4.

(10) Luguru
ha-ni-gend-ile ha-tali …
COND-SM1SG-go-PFV 16-distance
'if/when I had walked a long distance …'

(11) Luguru
kama si-gend-ile …
COND NEG.SM1SG -go-PFV
'if/when I did not go …'

In Ndamba, slot 1 is used for the *pa-* conditional/temporal marker, as shown in example (12).

(12) Ndamba
pa-tu-yend-ile pa-tali …
COND-SM1PL-go-PFV 16-far
'if/when we have walked far …'

In Kagulu, the conditional/temporal marker, or the relative object marker when present, appears before negation, as shown in (13), although a concatenation of markers such as this is rare, and periphrastic constructions are preferred.

(13) Kagulu
fo-si-cha-lut-e, wa-na wa-onel-a
COND-NEG-SM1PL.PST-go-FV 2child SM2-rejoice-FV
'When we could not go, the children were happy.'

In Pogoro, several different morphemes may appear in this slot, including future tense markers, negative *na-*, the relative object marker, and conditional/temporal *pa-*, but it is unclear from the source data in what order they may co-occur.

2.2 Slot 2: Subject marker (SM)

All five selected languages require a subject marker in this slot, except for imperatives, as is normally the case in Bantu languages (Meeussen 1967: 108). Examples of imperatives with no subject marker are given in §4.1.

2.3 Slot 3: Post-subject marker (POST.SM)

In Kagulu, this slot is used for negative markers, which may appear before or after the subject marker (see §4.6). In Kami, the slot is also used for the conditional/temporal marker *(h)a-* (which can occur in slot 1 as well), as shown in example (14) and discussed further in §4.4.

(14) Kami
wa-**ha**-to-a ngoma …
SM2-COND-play-FV 9.drum
'when they play the drum …'

2.4 Slot 4: First tense/aspect marker (TA1)

This is the principal slot for inflectional tense/aspect markers in all the selected languages.

2.5 Slot 5: Object marker (OM)

All five selected languages use this slot for an optional object marker which, in most cases, agrees with the noun class of the object. An exception is seen for animate objects in Ndamba, which take class 1/2 agreement. Limited data would suggest that Pogoro also follows a system of animate agreement similar to that of Ndamba.

The slot is also used for the reflexive marker *-i-* in Ndamba, *-ki-* in Kagulu and *-li-* in Pogoro, as shown in examples (15) to (17). In Kagulu, Kami and Luguru the reflexive marker also acts as the reciprocal marker, as shown in example (17). When the reflexive marker is used with plural subjects, there is ambiguity between the reciprocal and reflexive meaning. The forms can be disambiguated by stress, or of course by a reciprocal independent pronoun (i.e. 'each other').

(15) Ndamba, reflexive
ka-**i**-gom-ile → ke-gom-ile
SM1.PST-REFL-hit-PFV
'S/he hit him/herself'

(16) Pogoro, reflexive
ka-**li**-kom-iti
SM1-REFL-hit-PFV
'S/he hit him/herself'

(17) Kagulu, reciprocal and reflexive
wa-nhu wa-**ki**-end-a
2-person SM2-RECP-love-FV
'People like each other' [The stress is on the morpheme *-ki-*] or 'People like themselves' [The stress is on the verb *-enda*]

2.6 Slot 6: Verb root

The verb root appears in this slot.

2.7 Slot 7: Extensions (EXT)

In all five selected languages, this slot is used for one or more derivational morphemes. In some cases, the extensions have become unproductive and appear only in specific lexicalised verb stems. The main productive extensions which appear in the selected languages are the applicative (*-il-* or *-el-*), causative (*-iz-*, *-is-*, *-ez-* or *-es-*), stative (*-ik-* or *-ek-*) and associative (*-an-*) extensions. A more complete description of verbal extensions is outside the scope of this chapter. A useful summary of Bantu verbal extensions may be found in Schadeberg & Bostoen (2019).

(18) to (21) illustrate the main productive extensions in the selected languages.

(18) Kagulu, applicative
ya-ku-chi-golos-**el**-a
SM1-PRS-OM1PL-do-APPL-FV
's/he is working for us'

(19) Kami, causative
Ni-mw-ang'-**iz**-a ma-zi m-bwanga.
SM1SG-OM1-drink-CAUS-FV 6-water 1-boy
'I made the boy drink water'

(20) Luguru, stative
I-chi-dole che mu-gheni chi-ben-**ek**-a.
AUP-7-finger of 1-stranger SM7-break-STAT-FV
'The stranger's finger is broken'

(21) Ndamba, associative
va-ku-tov-**an**-a
SM2-PRS-hit-RECP-FV
'they are fighting (each other)'

2.8 Slot 8: Second tense/aspect marker (TA2)

This slot is also used in all the languages for the *-ag(h)-* habitual/progressive marker. Meeussen (1967: 110) states that habitual/progressive marking is the primary use of TA2 in Proto-Bantu.

This slot is also used for the suffix *-ile* (*-iti* in Pogoro), which is perfective in Ndamba and Pogoro but only used in dependant clauses in Kagulu, Kami and Luguru. Meeussen (1967: 111) places this in the FV slot. This implies that the *-ag(h)-* suffix in TA2 could co-exist with *-ile* in FV. Nevertheless, this is not the case in any of the selected languages; the *-ag(h)-* and *-ile* morphemes are mutually exclusive in all of them. We have therefore placed both *-ag(h)-* and *-ile* in TA2.

2.9 Slot 9: Passive suffix (PASS)

This slot is used for the passive derivational suffix (*-igw or -w-)* in all the selected languages except Pogoro, as illustrated in examples (22) and (23).

(22) Kagulu
cho-kol-**igw**-a
SM7.FUT-catch-PASS-FV
'it will be trapped'

(23) Ndamba
u-bagha u-ku-telek-**w**-a
14-food SM14-PRS-cook-PASS-FV
'the food is being cooked'

There is no passive marker in Pogoro. Instead, a periphrastic construction with an impersonal third person plural subject marker is used, as shown in example (24).

(24) Pogoro
wa-m-fir-a nene
SM2-OM1SG-love-FV PRO.1SG
'I am loved' lit. 'they love me'

Stappers (1967) proposes that the passive suffix was *-*u*- in Proto-Bantu. Meeussen (1967: 92) states that *-*u*- has the last position, following the pre-final (our slot 8), but does not assign it a specific slot. Similarly, Nurse (2008: 37) states that the passive marker is usually the last "extension" following the pre-final, but again does not assign it a separate slot. For our analysis, however, we assume that the passive marker appears in a separate slot, thus creating a second derivational slot. This is further corroborated by the fact that the passive can co-occur with other extensions (although semantic restrictions apply).

Nurse (2008: 37) and Meeussen (1967: 92) both point out that a tense/aspect morpheme in TA2 may merge with a following passive marker, leaving the final vowel of the morpheme in the FV slot. Examples (25) and (26) illustrate this using data from Ndamba.

(25) Ndamba
/-ile- + -w-/ ⟶ -il-w-e
lw-imbo lu-ka-somol-il-w-e
11-song SM11-PST-sing-PFV-PASS-FV
'the song was sung'

(26) Ndamba
/-agha- + -w-/ ⟶ -egh-w-e
lw-imbo lu-ka-somol-egh-w-e
11-song SM11-PST-sing-PROG-PASS-FV
'the song was being sung'

An alternative view of this process is that the passive marker is underlyingly an extension appearing as the last extension in slot 7 and that the merging process is as illustrated in example (27).

(27) Ndamba
/-w- + -ile-/ ⟶ -il-w-e
lw-imbo lu-ka-somol-il-w-e
11-song SM11-PST-sing-PFV-PASS-FV
'the song was sung'

This process is an example of a phonological process termed "imbrication" (Bastin 1983, Kula 2001, Chebanne 1993), in which, under certain conditions, a verb-final inflectional morpheme moves to a position prior to the last consonant of the extended base, as shown in (28) for Tswana and (29) for Bemba.

(28) Tswana (Chebanne 1993: 4)
/-rek-w-ile/ → -re-il-w-e
buy-PASS-PFV
'be bought'

(29) Bemba (Kula 2002: 153)
/βúng-il-ile/ → βúlung-i:l-e
mould-APPL-PFV
'has moulded for'

2.10 Slot 10: Final vowel (FV)

The final vowel is normally *-a* in all five languages, as illustrated in (30) for Kagulu. This is the unmarked default in most Bantu languages (Nurse 2008: 261). However, FV appears as *-e* in the subjunctive in all five languages, as illustrated in (31) for Ndamba. In Pogoro, FV also appears as *-i* as a future tense marker, as illustrated in (32).

(30) Kagulu, present indicative
Di-bwa di-ku-diy-**a**
5-dog SM5-PRS-eat-FV
'the dog eats'

(31) Ndamba, subjunctive
tu-telek-**e**
SM1PL-cook-SBJV
'let us cook'

(32) Pogoro, future indicative
ha-ga-fir-**i**
FUT-SM1-love-FV
's/he will love'

2.11 Slot 11: Post-final vowel suffix (POST.FM)

Three of the languages (Kagulu, Luguru and Kami) use this slot for a *-ni* plural suffix in imperatives (see e.g. (33)), a feature claimed by Meeussen (1967: 111) to

be Proto-Bantu. Ndamba and Pogoro do not use this slot for plural imperatives, instead relying on a plural subject marker (see §4.1). Ndamba uses the slot for a relative marker, as illustrated in example (34), and Pogoro has nothing in this slot.

(33) Kagulu, plural imperative marker
Ni-ingh'h-e-**ni**
OM1SG-give-SBJV-PL
'you (PL) give me …'

(34) Ndamba, relative marker
va-yis-ile-**vo** nalelo
SM1PL-arrive-PFV-REL2 today
'they who have arrived today'

2.12 Conclusions about the template

A verbal template was established for comparing the verbal morphology of the five languages in the study. This template closely follows the template proposed by Meeussen (1967) and amended by Nurse (2008), the main differences being

- Meeussen's "formative" and "limitive" slots are combined to form a "first tense/aspect marker" TA1
- A separate derivational slot is included for the passive suffix.

3 Tense/aspect

This section discusses how tense and aspect are represented in the languages of this study. The section starts with a general introduction to tense and aspect in Bantu, followed by a sub-section for each of the five selected languages. These are followed by further sub-sections dedicated to two specific topics: the suffix *-ile* and periphrastic constructions, followed by a preliminary summary of the data from the five languages. Periphrastic constructions are very common in Bantu languages and are typically used in languages where the inflectional tense/aspect system is inadequate, as discussed in §3.6.

Negative tenses are subsequently discussed in §4.6.

3.1 Models of Bantu tense/aspect

Many Bantu languages have multiple past and future tenses. Nurse (2008: 103) estimates that 80% of Bantu languages have more than one past tense and nearly 50% have multiple future tenses. Botne & Kershner (2008: 147) describe how research comparing the tense/aspect markers of Bantu languages has mostly attempted to fit them into a standard model, based primarily on absolute and relative time-scales, but that this approach has tended to obscure more nuanced semantic details of these systems.

One approach to analysing the Bantu tenses is to distinguish "tense" and "aspect" (Dahl 1985, Nurse 2008). In this model, there are two dimensions: "tense" encodes the absolute time-scale of an event or action and "aspect" describes details of how that event or action takes place within a specific time-scale. Botne & Kershner (2008) makes use of this tense/aspect model to form a system of dimensions in which absolute timescales are represented as one dimension (the P-domain) and other contrasts are represented as multiple D-domain dimensions which operate at different points along the P-domain.

In many Bantu languages, tense and aspect are marked in the two distinct slots of the verbal template: TA1 and TA2 respectively. The sections below describe how these slots are used to express tense/aspect in the five languages of the study.

3.2 Tense/aspect morphology in Kagulu

Kagulu has three specific tense markers appearing in the TA1 slot: *-ku-* non-past (i.e. present or future), *-ka-* future and *-o-* future. The *-o-* marker merges with the preceding SM to produce modified subject markers such as *cho-* (class 7 *chi+o*). The three future forms appear to be in free variation and there is no apparent distinction in meaning (for a discussion of this, see Petzell 2008: 108–109). In addition to these forms, the past imperfective has *ha-* in PRE.SM, while the past perfective carries no overt marker. A summary of Kagulu tense/aspect markers is shown in Table 3.

3.3 Tense/aspect morphology in Kami

Kami marks non-past (present or future) with *-o-*, which merges with the preceding SM to produce a modified SM such as *to-* (*tu+o*). Past tense (perfective and imperfective) has a null marker in the TA1 slot. A summary of Kami tense markers is shown in Table 4.

Table 3: Kagulu inflectional tense markers

Tense	PRE.SM	TA1	Example
Non-past (present or future)		*ku*	*chi-**ku**-lut-a* SM1PL-PRS-GO-FV 'we go/will go'
Future1		*o* ([o] merges with SM)	*ch**o**-lut-a* SM1PL.FUT-GO-FV 'we will go'
Future2		*ka*	*chi-**ka**-lut-a* SM1PL-FUT-GO-FV 'we will go'
Past perfective		*ø*	*chi-**ø**-lut-a* SM1PL-PST.PFV-GO-FV 'we have gone/we went'
Past imperfective	*ha*	*ø*	***ha**-chi-**ø**-lut-a* IPFV-SM1PL-PST-GO-FV 'we were going/we went'

Table 4: Kami inflectional tense markers

Tense	TA1	Example
Non-past (present or future)	*o* ([o] merges with SM)	*t**o**-gend-a* SM1PL.NON_PST-GO-FV 'we are going' / 'we will go'
Past tense (perfective and imperfective)	*ø*	*tu-**ø**-himb-a simo* SM1PL-PST-dig-FV 9.hole 'We (have) dug a hole.'

3.4 Tense/aspect morphology in Luguru

In Luguru, the present tense is marked with *-o-* (which merges with the preceding SM), the future tense is marked with *-tso*[3] -, and the past tense (perfective and imperfective) has a null marker in the TA1 slot.

Apart from these inflectional markers, there is another verbal formative, the temporal/aspectual status of which is not clear. This formative *tsa-* (also realised as *dza-*) encodes some type of shared knowledge or shared reference, and conveys meanings such as 'at a specific time', 'at a place', 'as we know', or even 'for that reason' (Petzell 2020). It is used primarily in past-time contexts and refers to something like a 'definite span' of time or space, or to more abstract notions, e.g. reasons and expectations. For example, compare (35) with (36).

(35) Luguru
ni-gend-a
SM1SG-go-FV
'I went'

(36) Luguru
tsa-ni-gend-a
at.that.time/because-SM1SG-go-FV
'at that time/because I went.'

A summary of Luguru tense markers is shown in Table 5.

The future tense marker *-tso-* sometimes surfaces as *-tsa-* when followed by the morpheme *ku-*.

Two other markers, *-za-* and *-ya-*, are mentioned by Mkude (1974: 77, 101), but these appear to have become grammaticalised as future markers in current Luguru. Mkude refers to them as "verb like operators" and states that they represent motion towards and away from the speaker, i.e. 'come' and 'go' respectively. We assume that *-za-* combines with non-past *-o-* to form *-zo-*, realised as future *-tso-*, as is shown in the example in Table 5. It can either mean 'we will speak' or rarely, depending on the context, 'lest we speak'. The other morpheme, *-ya-*, does not exist in our data and is rejected by our consultants.

[3]We believe the future marker has been grammaticalized from the present *-o-* combined with a remnant of the verb *-za* 'come'. The marker has several allomorphs that vary in spelling: *-dzo-* and *-nz'o-* being the most common (see also (96)–(98)).

Table 5: Luguru inflectional tense markers

Tense	TA1	Example
Present	*o* ([o] merges with SM)	*tw**o**-ghend-a* SM1PL.PRS-go-FV 'we are going'
Future	*tso (tsa)*	*tu-**tso**-long-a* SM1PL-FUT-speak-FV 'we will speak'
Past tense (perfective and imperfective)	*ø*	*tu-**ø**-himb-a simo* SM1PL-PST-dig-FV 9.hole 'We (have) dug a hole.'

3.5 Tense/aspect morphology in Ndamba

Ndamba has inflectional markers for seven distinct tenses: three past tenses, one present tense and three future tenses. All these tense markers are assembled from combinations of TA1 morphemes and the *-ile* suffix in TA2. Table 6 shows a summary of Ndamba inflectional tense markers.

Three of the tenses (future indefinite and future and past emphatic) use a tense/aspect marker in TA1 that is used to express a level of certainty. It is possible that these are related to or derived from degrees of remoteness, but we do not have any data to be conclusive about this.

A way of analysing these tense/aspect markers might be to treat them as evidentiality markers as part of the TAME framework (Dahl 2013). In this framework, evidentiality is added as an additional category to the usual verbal categories of tense, aspect and mood. Evidentiality marking indicates how certain the speaker is about the source of information (the evidence) used to make a statement. Dahl states, based on data from WALS (de Haan 2013), that evidentiality markers are "almost entirely absent in Africa".

Another approach might be to treat these tense/aspect markers as having a modal meaning, as does Fleisch (2000) for the "definite future" tense of the Angolan language Luchazi (K13), as illustrated in (37).

Table 6: Ndamba inflectional tense markers

Tense	TA1	TA2	Example
Present	*ku*		*tu-**ku**-telek-a* SM1PL-PRS-cook-FV 'we cook'
Near future	*ta*		*tu-**ta**-telek-a* SM1PL-FUT.NEAR-cook-FV 'we will cook (in the near future)'
Future indefinite	*ala*		*tw-**ala**-telek-a* SM1PL-FUT.IND-cook-FV 'we will cook (at some undefined time in the future)'
Future emphatic	*aa*		*tw-**aa**-telek-a* SM1PL-FUT.EMPHATIC-cook-FV 'we will definitely cook'
Perfect	*ø*		*tu-**ø**-telek-a* SM1PL-PRF-cook-FV 'we have cooked'
Past	*ka*	*ile*	*tu-**ka**-telek-**ile*** SM1SG-PST-go-PFV 'we cooked'
Past emphatic	*aa*	*ile*	*tw-**aa**-telek-**ile*** SM1PL-PST.EMPHATIC-cook-PFV 'we definitely cooked'

(37) Luchazi (Fleisch 2000: 150)
nji-kụ́kù-y-a ku-Venduka
SM1SG-DEF_FUT-go-FV 17-Windhoek
'I will definitely go to Windhoek / I will have to go to Windhoek'

Another interesting aspect of the Ndamba tense/aspect markers is that they may be grouped into symmetrical pairs of past and future. For example, the two emphatic tenses, marked by *-aa-* and *-aa-* + *-ile,* show a symmetry in which the same tense marker is used for both tenses, the contrast being achieved by adding *-ile* for the past tense.

This symmetrical contrast is analogous to that found in Nugunu (A62), which has eight tenses, including three future and three past tenses (Botne & Kershner 2008: 161, based on data from Gerhardt 1989). The future and past tenses form three pairs of near, mid and far past/future tenses respectively, in which each past/future tense marker pair uses the same basic tense morpheme, modified with a nasal prefix to convert the future version into the past tense. For example, the mid-future tense marker, high-toned *á,* becomes past tense by prefixing a nasal, as shown in examples (38) and (39). The non-hyphenated orthography is taken from the source.

(38) Nugunu (Gerhardt 1989: 321)
a **á** bolá
SM1 **PST2** arrive
's/he arrived'

(39) Nugunu (Gerhardt 1989: 326)
a **ná** bola
SM1 FUT**2** arrive
's/he will arrive'

Another symmetrical contrast may also be seen with the Ndamba *-ka-* + *-ile* past tense, in which dropping the final *-ile* generates an imperfective meaning of an event that started in the past and continues into the future, as shown by comparing (40) with (41).

(40) Ndamba past imperfective
tu-**ka**-telek-a
SM1SG-PST-go-FV
'we are still cooking'

(41) Ndamba past perfective
tu-**ka**-telek-**ile**
SM1SG-PST-GO-PFV
'we cooked'

3.6 Tense/aspect morphology in Pogoro

The Pogoro tense markers appear in three separate slots: PRE.SM, TA1 and TA2, as shown in Table 7. Present tense carries no marking in any of the three slots. Past is marked with *-iti* in TA2. There are two future tenses: near future is marked with *za-* in PRE.SM, while far future has *naga-* or *ha-* in PRE.SM and *-i* as FV. In addition, there are two secondary TA1 morphemes: inceptive *-mku-* and counter-expectational *-na*.

3.7 Loss of the suffix *-ile* in Kagulu, Kami and Luguru

The distribution of the "perfective" suffix *-ile* is restricted in Kagulu, Kami and Luguru, and it has lost its primary function of marking perfectivity. In Kagulu, Kami and Luguru, *-ile* is used only in conditional/temporal constructions, negative and relative clauses (Petzell 2008: 126, Petzell & Aunio 2019: 581–582, 588). That a morpheme is retained in subordinate clauses only is not unusual since subordinate clauses are considered more conservative (cf. Bybee 2002, among others). The usage of *-ile* in subordinate clauses is exemplified with Luguru in (42), where the first verb takes conditional/temporal marking plus *-ile* and the second verb is an (imperfective) negative. This contrasts with the use of *-ile* in the G50 group, where it is used as a productive perfective marker in Ndamba, and (as *-iti*) for past tense in Pogoro.

(42) Luguru
Ha-fvik-ile si-lim-**ile** bae.
TEMP.SM1-arrive-PFV NEG.1SG-cultivate-PFV NEG
'When s/he arrived, I was not cultivating.'

Other G30 languages such as present day Zaramo (G33) have also lost the principal use of *-ile* as marking perfective (Petzell, field data; Brad Harvey, pers. comm.). This behaviour was also attested in Nurse's data from the 1970s (2008: Appendix 1, pp.169–170). Guthrie (1948: 49) also remarks on the unusual behaviour of *-ile* in some of the G30 languages, noting that the marker does not occur in "regular" affirmative sentences. Furthermore, in Mkude's (1974) grammatical sketch of Luguru there is only one occurrence of *-ile* in an affirmative clause,

Table 7: Pogoro inflectional tense markers

Tense	PRE.SM	TA1	TA2	Example
Present		*ø*		*ga-**ø**-fir-a* SM1SG-**PRS**-love-FV 's/he loves'
Near future	*za*	*ø*		***za**-gu-ø-gend-a* FUT-SM2SG-FUT-GO-FV 'you (sg) will go'
Far future	*naga or ha*	*ø*		***naga**-ga-ø-fir-i* FUT-SM1-FUT-love-FV 's/he will love'
Past perfect		*ø*	*iti*	*ka-ø-gend-**iti*** SM1-PFV-GO-PFV 's/he has gone'
Inceptive		*mku*		*na-**mku**-fir-a* SM1SG-**begin**-love-FV 'I am beginning to love'
Counter-expectational		*na*		*na-**na**-m-on-i* SM1SG-not_yet-OM1-see-FV 'I cannot yet see him/her'

shown in (43). This, however, is translated as an applicative by our consultants; see (44).

(43) Luguru, (Mkude 1974: 81)
a-lim-**ile**
SM1-cultivate-PFV
's/he dug'

(44) Luguru
a-lim-**il**-e
SM1-cultivate-APPL-FV
's/he dug (for somebody or at a place)'

What is more, another Luguru consultant explains the *-ile* marker in example (44) as having a conditional meaning: 'where/when s/he dug'. What is clear is that *-ile* is rejected as a perfective marker in affirmative clauses in today's Luguru.

3.8 Periphrastic constructions

Comparison of periphrastic tenses (referred to by Nurse 2008: 46 as "compound constructions") may be hampered by uneven levels of detail in the descriptions of the languages. Nevertheless, it is interesting to examine the range of periphrastic constructions used in the five languages under study to find patterns of similarity or difference.

In Kagulu, Kami and Luguru, several periphrastic tense/aspect constructions are used. One of the most common verbs used in periphrastic constructions is *kuwa* 'to be', as shown in example (45), which can be used for the habitual, among other functions.

(45) Kagulu, imperfective
Ya-ku-uw-a ya-sok-a ku-lang-a filamu.
SM1-NON_PST-be-FV SM1-(be)come_tired-FV 15-watch-FV 9.film
'S/he gets tired whenever she watches a film.'

Other verbs are used as well, such as modal[4] *-daha* 'be able' (in Kagulu), *-kala* 'remain' for past constructions in Kami and Luguru, and modal *-weza* 'can' (in Kami), as shown in (46) to (50).

(46) Kagulu, modal (Petzell 2008: 187)
Wa-gamb-a si-chi-ku-dah-a ku-seng-a.
SM2-speak-FV NEG-SM1PL-PRES-be_able-FV 15-cut-FV
'They said we cannot cut/cultivate. '

(47) Kami, past (Petzell & Aunio 2019: 583)
To-kal-a tu-lim-a.
SM1PL.NON_PST-remain-FV SM1PL-cultivate-FV
'We (had) cultivated.'

(48) Luguru, past
Tu-kal-a tu-bigh-a.
SM1PL-remain-FV SM1PL-dance-FV
'We had danced.'

[4] The term "model verb" is used here in the conventional sense as being a non-affirmative verb expressing mood, often used as an auxiliary (Crystal 2003: 295).

(49) Kami, modal (Petzell & Aunio 2019: 584)
No-wez-a ku-fik-a?
SM1SG.NON_PST-can-FV INF-arrive-FV
'Can/may I get (there)?'

(50) Luguru, modal
Two-dah-a ku-himb-a pondo.
SM1PL.PRS-can-FV INF-dig-FV 5.hole
'We can dig a hole.'

Other periphrastic constructions are made up of a defective verb, *ng'(h)ali* 'be still', as shown in (51). In Kagulu, and occasionally in Kami, it also conveys the meaning of 'not yet', as shown in (52). In agreement with Nurse, we assume that *ng'ali* contains a negation, *ng'(h)a,* and the copula *li* 'be' (Nurse 2008: 173).

(51) Kami, persistive
Di-tunda di-ng'ali dyo-d-igw-a.
5-fruit SM5-be_still 5.NON_PST-eat-PASS-FV
'The fruit is still edible.'

(52) Kagulu, persistive
Ni-ng'hali ku-lim-a.
SM1SG-be_still INF-cultivate-FV
'I have not yet cultivated.'

In Ndamba, a periphrastic future tense may be constructed from *-daya* 'want', as shown in (53).

(53) Ndamba
Va-henja va-ku-day-a va-yis-e chilawu
2-guest SM2-PRS-like-FV SM2-come-SBJV tomorrow
'The guests will arrive tomorrow.'

va-ku-day-a may be contracted to a cliticised prefix *da-,* as shown in example (54), showing a process of grammaticalisation. Some speakers defined this as their preferred or only method of constructing the future tense, suggesting that the use of the system of inflectional future tenses described above in §3.5 may be in the process of disappearing.

(54) Ndamba
Va-henja **da**-va-yis-e chilawu
2-guest FUT-SM2-come-SBJV tomorrow
'The guests would like to / will arrive tomorrow.'

Additional tense/aspect constructions may be formed in Pogoro using adverbial or conjunctional particles, as shown in (55) and (56).

(55) Pogoro, temporal conditional
hangu gu-on-i wa-ndu
when SM2SG-see-FV 2-person
'when you see the people …'

(56) Pogoro, far past
ka-lewer-a kala
SM1-forbid-FV long_ago
's/he forbade it'

Adverbial *kala* in Pogoro, as seen in example (56), may derive from Proto-Bantu **yikala* 'be, live, stay' (Nurse & Philippson 2006: 166). A similar construction is available in Ndamba, as shown in example (57).

(57) Ndamba
tu-ka-telek-ile kala
SM1SG-PST-go-PFV already/long_ago
'we have already cooked / we cooked long ago'

These examples contrast with the use of *kala* 'remain' as an auxiliary in Kami and Luguru, as shown in examples (47) and (48).

In conclusion, comparing the five languages, there seem to be some similarities in periphrastic constructions between the three northern languages, Kagulu, Kami and Luguru, but the two southern languages, Ndamba and Pogoro, are different.

3.9 Summary of tense/aspect morphology

The tense/aspect morphology of the five selected languages described above show that there are three groups of languages.

The first group, consisting of the two G30 languages, Kami and Luguru, exhibit notably little tense/aspect morphology. They essentially have just one tense marker, based on *-o-,* which is used for non-past, apart from Luguru that also has a future marker (*-tso-*). In this group of languages, there is only one past tense, which in turn doubles as a perfective and which carries no overt marking (Bar-el & Petzell 2021). In addition, the use of the "perfective" marker *-ile* has disappeared in these languages, except in certain specific contexts such as

dependant clauses. Our data also appear to show that these languages make use of periphrastic constructions to express tense/aspect, enhancing their reduced systems of inflectional markers.

The question is why these languages have such reduced verbal tense morphology compared with most other Bantu languages? Nurse (2008: 103) proposes that this is a result of a two-stage historical process. Proto-Bantu initially had a very rudimentary inventory of tenses, possibly only one past and one future tense. In the first stage of transformation, innovations increased this inventory, resulting in the complex tense systems seen in many Bantu languages today. Some languages, however, went through a second stage of transformation in which multiple tenses reduced back to a minimal set. Nurse's evidence for this is that there is little uniformity across the Bantu languages with reduced tense systems. He goes on to hypothesise that the unusual null marked past tense in Kami and Luguru (and occasionally Kagulu) derives from so called "vowel copy forms" (Nurse 2008: 84–85).

The second group consists of two G50 languages, Ndamba and Pogoro, which lie in the southern part of the region and have richer sets of tenses, typical of Bantu languages. Nonetheless, the Ndamba data show that these tense distinctions are based less on temporal remoteness and more on degrees of certainty.

The final group consists of the Kagulu language (G12). This language lies somewhere between the two other groups in terms of the complexity of its system of tenses, while there is no morphological encoding of degrees of certainty.

4 Other related markers

This section addresses aspects of verbal morphology in the five languages not covered in §3. The reason for including a discussion of other markers at this stage is that they often interact with the tense/aspect system, such that it becomes difficult to delineate structures which are specific to tense/aspect. For example, §4.5 describes the use of conditional affix *-ng'a-*, which typically takes the place of a tense marker. In his cross-linguistic review of tense/aspect systems, Dahl (1985) concludes that Bantu languages have the most complex tense/aspect systems of the languages included in his review. In particular, prefix positions assigned for tense/aspect markers are often also used for other categories which, in other languages, are typically expressed by adverbs (Dahl 1985: 176).

As with the preceding section, the objective is to review similarities and differences in structures used by the five languages in the study. This review is presented in sections covering the verbal categories of imperative, subjunctive, conditional, temporal, habitual/progressive/intensive and negative.

4.1 Imperative

The constructions of imperatives are similar across the five languages. In all languages there is a contrast between an emphatic imperative with no SM and FV *-a*, and a "polite" imperative formed from the subjunctive *-e* (Nurse 2008: 28, Devos & Van Olmen 2013). (58) shows the emphatic imperative and (59) shows the polite imperative.

(58) Kagulu, imperative
Leuk-**a**!
go_away-FV
'go away!'

(59) Kagulu, polite imperative
Ni-tamil-**e**!
OM1SG-tell-SBJV
'Tell me!'

All five languages require a subject or object marker to precede the verb stem when the polite imperative is used in the singular, as shown in examples (60) to (64).

(60) Luguru
Mu-himb-e i-vi-adzi
SM2SG-dig-SBJV AUP-7-potato
'Dig up (pl.) the potatoes'

(61) Kagulu
ni-lim-e
SM**1**SG-cultivate-SBJV
'I should cultivate'

(62) Kami
M-kem-e!
OM1-call-SBJV
'Please call (him/her)!'

(63) Ndamba
wu-gholok-e
SM2SG-get-up-SBJV
'get up!'

(64) Pogoro
gu-fir-e!
SM2SG-love-SBJV
'love!'

For plural imperatives, three languages (Kagulu, Kami and Luguru) use verb-final *-ni,* as illustrated in (65), whereas Ndamba and Pogoro use a plural SM, as illustrated in (66). Kagulu and Luguru may also make use of a plural SM as an alternative to the *-ni* suffix, as illustrated in (67).

(65) Kami, plural polite imperative
Himb-e-**ni** vi-bogwa!
dig-SBJV-PL 8-potato
'Dig up (pl) the potatoes!'

(66) Ndamba, plural polite imperative
Mu-telek-e!
SM2PL-cook-SBJV
'You (PL) cook!'

(67) Kagulu, plural polite imperative
Mu-kumul-e!
SM2PL-open-SBJV
'You (PL) cook!'

Nurse (2008: 39) states that the *-ni* suffix is the most common form of plural negative in Bantu languages.

4.2 Subjunctive

All five selected languages have a verb final *-e* for subjunctive, as illustrated in (68) to (72). (68) and (69) illustrate the use of subjunctive forms in non-affirmative subordinate clauses. (70) to (72) illustrate the use of the subjunctive for hortatives. These two uses of the subjunctive are also found in other Bantu languages (cf. Nurse & Devos 2019)

(68) Kami
no-lond-a ni-lim-**e** m-gunda w-angu
SM1SG.NON_PST-want-FV SM1SG-cultivate-SBJV 3-farm 3-POSS.1SG
'I want to cultivate my farm.'

(69) Luguru
no-bama-a ni-lim-**e** m-gunda gw-angu
SM1SG.PRS-want-FV SM1SG-cultivate-SBJV 3-farm 3-POSS.1SG
'I want to cultivate my farm.'

(70) Kagulu
ni-lim-**e**
SM1SG-cultivate-SBJV
'I should cultivate'

(71) Ndamba
tu-telek-**e**
SM1PL-cook-SBJV
'let us cook'

(72) Pogoro
ni-fir-**e**
SM1SG-love-SBJV
'I may love'

4.3 Conditional

The conditional is often marked morphologically in Bantu languages, usually in the TA1 slot (Nurse 2008: 34). Variations of the conditional affix *-ng'a-*, which is reconstructed for Proto-Bantu (Meeussen 1967: 113), are seen in all the languages in this study except Pogoro. In Kagulu, Kami and Luguru, *-ng'ha-* is used for 'if …' conditional clauses, as shown in (73). In Kami, *-ng'-* together with an *-ile* suffix is used in past conditional clauses, as shown in (74).

(73) Kagulu
u-**ng'ha**-ij-a
SM2SG-COND-come-FV
'if you come …'

(74) Kami
kama u-**ng'**-ez-ile
if SM2SG-COND-come-FV
'if you came …'

In Ndamba, *-nga-* is used in both the antecedent and consequent of hypothetical conditional 'if … then … would …' statements, as shown in (75).

(75) Ndamba
ma-huka gha-**nga**-dumuk-ile ndi-**nga**-gha-gol-ile
6-hoe SM6-COND-break-PFV SM1SG-COND-OM6-mend-PFV
'if the hoes were broken, I would mend them'

Instead of *-ng'a-*, Pogoro, uses the affix *-ya-* for conditional 'if …', as shown in example (76).

(76) Pogoro
na-**ya**-m-fir-a m-dalla a-yu
SM1SG-COND-OM1-love-FV 1-woman DEM-1.1
'if I loved that woman …'

Apart from these conditional markers, there are also non-hypothetical conditional/temporal markers meaning 'if/when …', as described in the §4.4.

4.4 Conditional/temporal 'when …'

Bantu languages often have a marker which may be used both for conditional 'if' and temporal 'when' (Doke 1935: 75). This is the case for the languages in this study, all of which use a morpheme in the PRE.SM slot for conditional/temporal 'if/when …', as shown in examples (77) to (81).

(77) Kagulu
fo-chi-ku-mal-a …
COND-SM1PL-PRS-finish-FV
'if/when we finish …'

(78) Kami
fi-wa-tow-ile ngoma …
COND-SM2-play-FV 9.drum
'if/when they played the drum …'

(79) Luguru
ha-ni-gend-ile ha-tali …
COND-SM1SG-go-PFV 16-distance
'if/when I had walked a long distance …'

(80) Ndamba
pa-tu-yend-ile pa-tali …
COND-SM1PL-go-PFV 16-far
'if/when we have walked far …'

(81) Pogoro (Hendle 1907: 52)
pa-ga-fik-iti
COND-SM1-arrive-PFV
'if/when s/he arrived …'

While Luguru, Ndamba and Pogoro make use of what looks like the noun class prefix of class 16, as shown in (79), (80) and (81) respectively, Kagulu (77) and occasionally Kami (78) use morphemes that can be traced to noun class 8. The origin of the Kagulu *fo-* marker shown in (77) is the most unclear, since the *fo-* also appears to contain the reference marker *-o-* plus noun class 8 *fi-*. The anaphoric marker *-o-* is often used in Bantu languages to refer to something previously mentioned in the discourse (Güldemann 2002: 275).

When Kami uses the less frequent class 8 *fi-*, it appears in slot 1, as shown in (78), while the more commonly used class 16 *ha-* (often realised only as *a-*) appears in slot 3, as shown in (82).

(82) Kami
wa-**(h)a**-to-a ngoma …
SM2-16-play-FV 9.drum
'when they play the drum …'

Furthermore, the same Kami speaker may use the *fi-* prefix and *-ha-* morphemes interchangeably, as shown in examples (78) and (82) (both examples given during the same elicitation session). This type of variation is not unusual for Kami – being a small and endangered language, it has borrowed many forms from neighbouring and dominating languages such as Luguru and Swahili (Petzell & Aunio 2019).

4.5 Habitual/progressive/intensive

All the languages except Pogoro have an *-ag(h)-* affix which may be used for habitual, progressive, imperfective, continuous or intensive. This affix appears in the TA2 post-extension slot in all four languages, as shown in examples (83) and (84).

(83) Kagulu
Ha-ka-ij-**ag**-a.
PST-SM1-come-HAB-FV
's/he came (regularly)'

(84) Luguru
Tu-gend-**ag**-a chila mara Dar_es_Salaam.
SM1PL-go-HAB-FV every time place_name
'We go to Dar es Salaam frequently.'

In Pogoro there is a progressive affix *-aŋku-* which appears in TA1, as shown in (85).

(85) Pogoro (Nurse 2008: Appendix 1, p.176)
tw-**aŋku**-hemer-a
SM1PL-PROG-buy-FV
'we are buying'

The derivation of Pogoro *-aŋku-* is unclear, and may not be a variant of *-ag(h)-*, given that *-ag(h)-* variants usually appear in TA2 (Meeussen 1967: 110).

Some Bantu languages use *-ang-* rather than *-ag(h)-* for progressive/intensive (Nurse 2008: 263). Ndamba and Pogoro, however, use both variants, showing a distinction between habitual/progressive/imperfective *-ag(h)-* and augmentative/intensive *-ang-*, as illustrated in (86) and (87) for Ndamba.

(86) Ndamba
a-ku-va-tov-**agh**-a
SM1-PRS-OM2-hit-HAB-FV
's/he usually beats them' or 's/he is beating them'

(87) Ndamba
a-ku-va-tov-**ang**-a
SM1-PRS-OM2-beat-AUG-FV
's/he is beating them intensively'

Another distinction between *-ang-* and-*agh-* in Ndamba is that *-ang-* behaves more like a derivational extension than *-ag(h)-,* which behaves as expected for an inflectional affix. For example, *-ang-* is affected by reduplication processes, as shown in (88).

(88) Ndamba
a-ku-va-tov-**ang**-a-tov-**ang**-a
SM1-PRS-OM2-hit-AUG-FV-hit-AUG-FV
's/he is continuously and intensively beating them'

However, *-ag(h)-* is not affected by reduplication processes, as shown in (89).

(89) Ndamba
a-ku-yend-a-yend-**agh**-a
SM1-PRS-GO-FV-GO-HAB-FV
's/he usually walks'

Furthermore, the two morphemes *-ang-* and *-ag(h)-* can be used together, as seen in (90).

(90) Ndamba
a-ku-va-tov-**ang**-**agh**-a
SM1-PRS-OM2-beat-AUG-HAB-FV
's/he usually beats them intensively'

This co-occurrence of *-ang-* and *-ag(h)-* is also observed in Bena (G63), a neighbouring language to Ndamba, as shown in example (91).

(91) Bena (Nurse 2008: 37)
ndi-laa-gul-**ang**-**ag**-a
SM1SG-FUT-buy-AUG-HAB-FV
'I'll be buying in quantities'

4.6 Negatives

This section discusses how negatives are formed in the selected languages, and how these interact with the tense/aspect system. A summary of how negative strategies interact with Bantu tense/aspect systems is provided by Nurse (2008: 180–184), who identifies six strategies. The two most common strategies are to use a negative morpheme in the pre-subject marker (PRE.SM) or post-subject marker (POST.SM) slots. This follows a pattern common in Bantu languages (Guérois et al. forthcoming). Three of the selected languages (Kagulu, Kami and Luguru) have inflectional negatives using these strategies, while the other two (Pogoro and Ndamba) do not, relying instead on periphrastic forms.

None of the languages uses a strategy of having specific negative tense/aspect morphemes that alternate with their non-negative counterparts, a strategy identified by Nurse (2008: 34) with an example from Nen (A44), which is spoken in Cameroon. A further example of this strategy is the Swahili (G42) past perfect *-me* / *-ja* alternation, illustrated in (92) and (93).

(92) Swahili
wa-**me**-kul-a
2.SM-PRF-eat-FV
'they have eaten'

(93) Swahili
ha-wa-**ja**-kul-a
NEG-2.SM-NEG.PRF-eat-FV
'they have not eaten'

The three languages with inflectional negatives (Kagulu, Kami and Luguru) use the PRE.SM slot for their negative morphemes, which include *si-, hu-, ha-*, and *ng'h-*. One of these (Kagulu) also allows the negative markers to occur in the POST.SM position; see Table 8 below for a discussion of this variation. According to Nurse (2008: 180), use of the POST.SM slot for the negative marker is the most common pattern across the Bantu languages. Table 8 summarises how negatives are formed in the five selected languages, with examples for each language.

Kagulu generally uses the negative marker *si-*, but for second person singular and class 1, *ng'h-* is used. This systematic alternation between two negative morphemes is not unusual and can be traced back to Proto-Bantu. When the former marker is used, its position in the verb phrase is in free variation. The *si-* morpheme appears either in the PRE.SM slot, as shown in (94), or in the POST.SM slot, as shown in (95), depending on the speaker's dialect or even idiolect (Petzell 2010). Moreover, the same speaker can switch slots in the middle of an utterance without any apparent change in meaning. This type of variation is highly unusual, not only for this region, but for Bantu languages in general.

(94) Kagulu
si-chi-ka-lim-a
NEG-SM1PL-FUT-cultivate-FV
'we will not cultivate'

(95) Kagulu
chi-**si**-ka-lim-a
SM1PL-NEG-FUT-cultivate-FV
'we will not cultivate'

Kami and Luguru use the negation markers *si-, hu-* and *ha-* for first, second, and third person animates (i.e. in class 1) respectively. These negation markers merge with the subject marker, as shown in (96) to (98).

(96) Luguru
si-nz'o[5]-lim-a u-m-gunda
NEG.SM1SG-FUT-cultivate-FV AUP-3-field
'I shall not cultivate the field'

Table 8: Negatives

Tense	**Kagulu**	**Kami**	**Luguru**	**Ndamba**	**Pogoro**
Imperative	*ng'ha* in PRE.SM	periphrastic	periphrastic	periphrastic	*na* (in PRE.SM)
Present / non-past	*si* or *ng'ha* in PRE.SM + *ku* in TA1 or *si* in POST.SM + *ku* in TA1	*ha*, *hu*, *si* in PRE.SM	*ha*, *hu*, *si* in PRE.SM		periphrastic
Future	*si* in PRE.SM + *ka*/*ku* in TA1 or *si* in POST.SM + *ka* in TA1		*ha, hu, si* in PRE.SM + *tso* in TA1		
Past	*si* or *ng'ha* in PRE.SM + a *in* TA1 + *ile* or *si* in PRE.SM or POST.SM + *ile* or *s* in POST.SM + *a* in TA1 (+ *ile*)	*ha*, *hu*, *si* in PRE.SM + *ile*	*ha*, *hu*, *si* in PRE.SM + *ile*		

(97) Luguru
hu-nz'o-lim-a u-m-gunda
NEG.SM2SG-FUT-cultivate-FV AUP-3-field
'you (sg) shall not cultivate the field'

(98) Luguru
ha-nz'o-lim-a u-m-gunda
NEG.SM**1**-FUT-cultivate-FV AUP-3-field
's/he shall not cultivate the field'

For all other persons and noun classes, *ha-* is used in the PRE.SM slot. Mkude (1974: 100) states that the Luguru negative marker is *ng'(a)- instead of ha-*, which is also occasionally found in our data (see (99) and (100)). Our hypothesis is that *ha-* is a phonological (or possibly dialectal) variant of the same morpheme, conceivably due to the influence of Swahili.

(99) Luguru
ng'a-wa-mw-on-ile
NEG-SM1PL-OM1-see-FV
'they did not see him/her'

(100) Luguru
ha-wa-mw-on-ile
NEG-SM1PL-OM1-see-FV
'they did not see him/her'

Ndamba and Pogoro do not have inflectional negatives, and instead use periphrastic negatives, as illustrated in (101) and (102).

(101) Ndamba
N-gu-yend-a **duhu**
SM1SG-PRS-go-FV not
'I am not going'

(102) Pogoro
gwa-fir-a **ndiri**
SM2SG-love-FV not
'you (sg) do not love'

Pogoro does, however, have an inflectional negative imperative morpheme *na-* which appears in the pre-subject marker slot, as shown in example (103).

[5]This is an allomorph of the future marker *-tso-*, see footnote 4.

(103) **na**-gu-fir-a!
NEG-SM2SG-love-FV
'do not love!'

4.7 Summary of other related markers

Much of the data in this section shows consistency between the five languages in the study. Imperatives use the same structure across the five languages, with the exception of methods of expressing plurals. All the languages use verb-final *-e* for subjunctives. All five languages use a conditional marker in place of the primary tense/aspect marker in TA1, and a conditional/temporal marker in the pre-subject marker slot. All the languages expect Pogoro use variants of the habitual/progressive/intensive marker *-ah(h)-* as the secondary tense/aspect marker in TA2. The main differences found in this section relate to how negatives are formed, with several different strategies being used.

5 Conclusions

The tense/aspect systems of the five selected languages from the Morogoro region show surprising diversity. One of the languages (Ndamba) has a typical Bantu inflectional system of multiple past and future tenses, while the G30 group of languages (Luguru and especially Kami) have a greatly reduced tense/aspect system, relying heavily on periphrastic forms. The other two languages (Kagulu and Pogoro) are intermediate in terms of tense/aspect system complexity, but they are still fairly reduced compared to most other Bantu languages. A common theme across all five languages is that none has a tense/aspect system showing sharp time distinctions, as documented for many Bantu languages (Nurse 2008: 88–94). These reduced systems, especially the ones with the neutralised past/perfective, are not recognised in the literature. We are still looking into how this affects the aspectual categories and vice versa, and how much temporal (and aspectual) information is conveyed through other constituents such as adverbials.

A particular aim of this study was to look at how negative tenses are handled. This revealed two patterns:

Firstly, the three northern languages (Kagulu, Luguru and Kami) have systems based on pre-verbal markers, while the two southern languages (Ndamba and Pogoro) have no inflectional negatives, relying on periphrastic forms.

Secondly, an interesting feature is that in the three languages with inflectional negatives, *-ile* surfaces only in non-affirmative contexts, supporting the view that it has lost its primary role of marking past perfective in these languages.

Aspects of the study merit further investigation. As these are under-described languages, the amount of morphological description and analysis available for the languages is limited, although increasing. In particular, the descriptions of their tense/aspect markers lack contextual information. It would be interesting to collect more data on the contexts in which the tense/aspect markers are used. This may help us go further into temporal interpretations for a deeper understanding beyond the standard paradigms.

The available data would suggest that tense/aspect marking is evolving in all the languages of the study due to increased contact with other languages, particularly, but not exclusively, Swahili. Swahili is the dominant language in Tanzania, spoken by nearly everyone, including all of our consultants, and given that it is a related Bantu language, it is unsurprising that other local languages are evolving to incorporate aspects of Swahili. That said, the intense contact does not necessarily imply accommodation to the dominant language, Swahili; it may also be non-accommodation, as described by Petzell & Kühl (2017). They analyse the overuse of a nominal marker in Luguru as *stability despite contact* due to *covert prestige*.

It may be interesting to document the evolution of markers more thoroughly by comparing current data with older data in a more systematic manner. Furthermore, for the languages that currently display little overt tense/aspect marking, it may be interesting to see if other strategies are emerging and if periphrastic constructions are becoming more common.

Finally, a specific topic worthy of further investigation would be an exploration of the semantics of the Ndamba tense/aspect markers within the context of the evidentialty component of the TAME framework (Dahl 2013).

Acknowledgements

We are immensely grateful to the speakers of the five languages who generously gave of their time and insights, and the anonymous reviewers for valuable comments. We wish to acknowledge, as well, Riksbankens jubileumsfond for financial support.

Abbreviations

In addition to the abbreviations listed below, numbers in abbreviations refer to noun classes.

APPL	applicative	PASS	passive
AUG	augmentative	PFV	perfective
AUP	augment prefix	PL	plural
CAUS	causative	PRF	perfect
COND	conditional	PRO	pronoun
DEM	demonstrative	PROG	progressive
EXT	extension	PROX	proximate
FUT	future	PST	past
FUT.IND	future indefinite	RECP	reciprocal
FUT.NEAR	near future	REFL	reflexive
FV	final vowel	REL	relative
HAB	habitual	REL.OM	relative object marker
IPFV	imperfect(ive)	SBJV	subjunctive
INF	infinitive	SG	singular
neg	negative	SM	subject marker
NON_PST	non-past	STAT	stative
OM	object marker		

References

Bar-el, Leora & Malin Petzell. 2021. (Im)perfectivity and actionality in East Ruvu Bantu. In Thera Marie Crane, Johanna Nichols & Bastian Persohn (eds.), *The role of the lexicon in actionality, Special issue of language typology and universals*, vol. 74. DOI: 10.1515/stuf-2021-1044.

Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en bantou* (Serie in-80, Sciences Humaines 114). Tervuren: Annales du Musée Royal de l'Afrique Centrale.

Botne, Robert & Tiffany L. Kershner. 2008. Tense and cognitive space: On the organization of tense/aspect systems in Bantu languages and beyond. *Cognitive Linguistics* 19(2). 145–218.

Bybee, Joan. 2002. Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. In Joan Bybee & Michael Noonan (eds.), *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*, 1–17. Amsterdam: John Benjamins.

Chebanne, Andy M. 1993. *The imbrication of suffixes in Setswana*. Paper presented at the 24th Annual Conference on African Linguistics. Columbus, Ohio.

Crystal, David. 2003. *A dictionary of linguistics and phonetics*. 5th edn. Oxford: Blackwell.

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

Dahl, Östen. 2013. Tense-aspect-mood-evidentiality (TAME) and the organization of human memory. In Karina Veronica Molsing & Ana Maria Tramunt Ibaños (eds.), *Time and TAME in language*, 22–52. Newcastle: Cambridge Scholars Publishing.

de Haan, Ferdinand. 2013. Semantic distinctions of evidentiality. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/77 (21 October, 2019).

Devos, Maud & Daniël Van Olmen. 2013. Describing and explaining the variation of Bantu imperatives and prohibitives. *Studies in Language* 37(1). 1–57.

Doke, Clement Martyn. 1935. *Bantu linguistic terminology*. London: Longmans, Green & Co.

Edelsten, Peter & Chiku Lijongwa. 2010. *A grammatical sketch of Chindamba*. Köln: Köppe.

Fleisch, Axel. 2000. *Lucazi grammar: A morphosemantic analysis*. Köln: Köppe.

Gerhardt, Phyllis. 1989. Les temps en Nugunu. In Daniel Barreteau & Robert Hedinger (eds.), *Descriptions de langues camerounaises*, 315–331. Paris: Agence de Coopération Culturelle et Technique & ORSTOM.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. Forthcoming. A parametric approach to negation in Bantu languages. In Eva-Marie Bloom Ström, Hannah Gibson, Rozenn Guérois & Lutz Marten (eds.), *Current approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.

Güldemann, Tom. 2002. When "say" is not say: The functional versatility of the Bantu quotative marker with special reference to Shona. In Tom Güldemann & Manfred von Roncador (eds.), *Reported discourse: A meeting ground for different linguistic domains*, 253–288. Amsterdam: John Benjamins.

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

Guthrie, Malcolm. 1967–1971. *Comparative Bantu: An introduction to the comparative and prehistory of the Bantu languages*, vol. 1–4. London: Gregg International.

Hendle, P. J. 1907. *Die sprache der Wapogoro (Deutsch-Ostafrika) nebst einem Deutsch-Chipogoro und Chipogoro-Deutschen wörterbuche*. Berlin: Reimer.

Kula, Nancy. 2001. Imbrication in Bemba. In Elisabeth Hume, Norval Smith & Jeroen van de Weijer (eds.), *Surface syllable structure and segment sequencing* (HIL Occassional Papers 4), 102–116. Leiden: Holland Institute of Generative Linguistics (HIL).

Kula, Nancy. 2002. *The phonology of verbal derivation in Bemba*. Leiden: University of Leiden. (Doctoral dissertation).

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

Lewis, Paul M. (ed.). 2009. *Ethnologue: Languages of the world*. 16th edn. Dallas, TX: SIL International. http://www.ethnologue.com/) (7 January, 2011).

Maho, Jouni Filip & Bonny Sands. 2002. *The languages of Tanzania: A bibliography*. Gothenburg: Acta Universitatis Gothoburgensis.

Marlo, Michael R. & David Odden. 2019. Tone. In Derek Nurse Mark van de Velde Koen Bostoen & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 150–171. London & New York: Routledge.

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

Mkude, Daniel J. 1974. *A study of Kiluguru syntax with special reference to the transformational history of sentences with permuted subject and object*. London: University of London. (Doctoral dissertation).

Mol, Frans. 1996. *Maasai language and culture dictionary*. Lemek: Maasai Centre Lemek.

Morrison, Michelle Elizabeth. 2011. *A reference grammar of Bena*. Houston, TX: Rice University. (Doctoral dissertation).

Moses, Godian. 2018. *Are highland Luguru and lowland Luguru distinct dialects?* Dar es Salaam: University of Dar es Salaam. (MA thesis).

Ngonyani, Deo. 2003. *A grammar of Chingoni*. München: Lincom.

Novotná, Jana. 2005. *A grammar of Ndamba*. St. John's. (Doctoral dissertation).

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

Nurse, Derek & Maud Devos. 2019. Aspect, tense and mood. In Derek Nurse Mark van de Velde Koen Bostoen & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 173–203. London & New York: Routledge.

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

Petzell, Malin. 2008. *The Kagulu language of Tanzania: Grammar, texts and vocabulary*. Köln: Rüdiger Köppe.

Petzell, Malin. 2010. Further analysis of negation in Kagulu. In Karsten Legère & Christina Thornell (eds.), *Bantu languages: Analyses, description and theory*, 209–216. Köln: Köppe.

Petzell, Malin. 2020. An analysis of the verbal formative *tsa* in Luguru. In Malin Petzell, Leora Bar-el & Lotta Aunio (eds.), *The semantics of verbal morphology in under-described languages, Special issue of studia orientalia electronica*, vol. 8, 119–133.

Petzell, Malin & Lotta Aunio. 2019. Kami G36. In Derek Nurse Mark van de Velde Koen Bostoen & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 563–590. London & New York: Routledge.

Petzell, Malin & Karoline Kühl. 2017. The influence of non-linguistic factors on the usage of the pre-prefix in Luguru. *Linguistic Discovery* 15(1). 35–48.

Schadeberg, Thilo C. & Koen Bostoen. 2019. Word formation. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 172–203. London: Routledge.

Stappers, Leo. 1967. Het passief suffix *-u-* in de Bantoe-talen. *Africana Linguistica* 3. 137–145.

Part III

Comparative and historical approaches

Chapter 11

A comparative sketch of TA markers in Kilimanjaro Bantu: In search of the directionality of semantic shift and micro-parametric correlation

Daisuke Shinagawa
ILCAA, Tokyo University of Foreign Studies

This chapter presents a comparative overview of the tense and aspect (TA) systems in Kilimanjaro Bantu languages (KB), including those from which comprehensive information about the TA system has not been available in the literature. Fundamental description about the TA system of the eight varieties of KB, namely Rwa, Siha, Mashami, Kibosho, Uru, Vunjo, Rombo-Mkuu, and Gweno, reveals a general picture of geographical distribution and formal correspondences of shared TA markers. Based on the systematic correspondences, which can be described as grammaticalisation chains, we further discuss historical processes of semantic development of shared TA markers, as well as possible typological generalizations lying behind the observed variation of the TA systems in KB.

1 Introduction

This chapter provides comparative lists of tense and aspect markers (TAM) from eight Kilimanjaro Bantu languages (KB) covering all of the three major subgroups, namely Western Kilimanjaro (WK), Central Kilimanjaro (CK), and Rombo, with the goal of presenting an overall picture of the distribution and semantic variation of common TA markers in KB[1]. Based on the data, I will discuss the

[1]For a linguistic overview, see Philippson & Montlahuc (2003) and Shinagawa (forthcoming[b]).

Daisuke Shinagawa. 2024. A comparative sketch of TA markers in Kilimanjaro Bantu: In search of the directionality of semantic shift and micro-parametric correlation. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 325–349. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663781

diachronic processes and typological implications of TA systems, namely i) the semantic development and grammaticalisation processes of TA markers across boundaries between languages and between subgroups, and ii) highlight micro-typological correlations between morphosyntactic parameters including those related to TA systems.

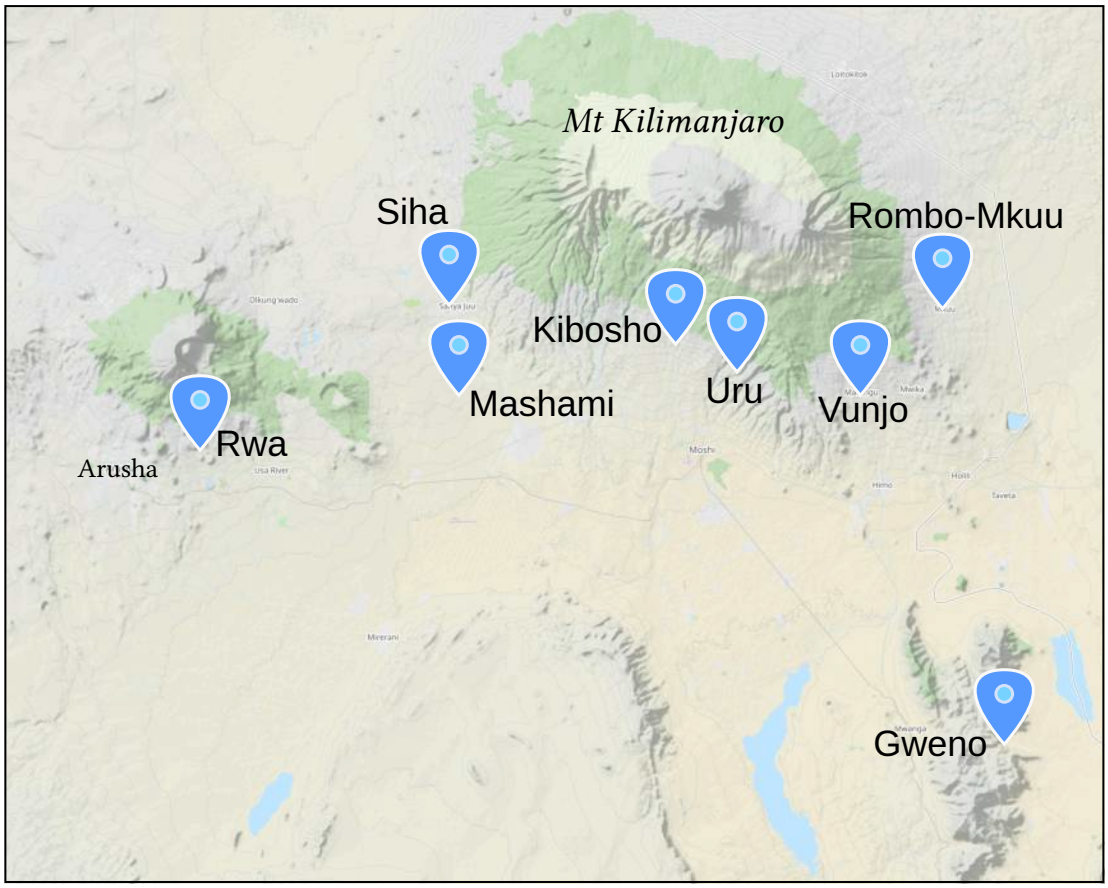
The languages presented in this chapter constitute two groups. The first group includes the languages described by the author with a full list of core TA markers, which includes Rwa (E621A[2], WK), Siha (E621C, WK), Uru (E622D, CK), and the Mkuu variety of Rombo (E623C). The second group consists of languages with reliable descriptions of TA forms in the existing literature; this group includes Mashami (E621B, WK; Rugemalira & Phanuel 2012), Kibosho (Kiw'oso, E621D, WK; Kagaya 1989), Vunjo (E622C, CK; Nurse 2003b, Moshi 1994), and Gweno (E65; Philippson & Nurse 2000).

The study of TA systems in KB was pioneered by Derek Nurse. His "Tense and aspect in Chaga" (Nurse 2003b) provides a comprehensive overview of TA concepts and forms found in KB with a special focus on Vunjo. Following in this vein, this study intends to provide descriptive data for the languages which are only referred to in a limited way in his study (especially WK languages), as well as to present a more in-depth account of both historical and typological aspects of TA systems in KB.

It should be noted here that since the primary focus of this chapter is on the form-meaning correspondences of common TA markers in KB, the following topics are not included in the scope of this study: i) compound tenses, ii) forms with so-called limitatives (cf. Meeussen 1967), iii) TA in negative constructions, and iv) modality markers.[3] Although these points are of importance for a comprehensive

[2]The five-digit codes shown in parentheses after language names are from the updated list of Guthrie codes by Maho (2009).

[3]Though these topics require further investigation in future research, the following points can be briefly mentioned as somewhat common features in KB: i) As for compound tenses, most of KB utilises the construction for aspectual forms with future tense, where the lexical stem meaning 'find' or 'get' (e.g. *koóya* or related forms in WK) is used as an auxiliary conveying future tense. ii) The limitative *ka-* is widely attested in KB with the meaning of consecutive; however, in Uru, it also expresses "ironical negation", the use of which may have developed from "T inceptive" 'already, not yet' in Meeussen (1967: 109). iii) Negation is also an essential part of the TA system and there seems to be crosslinguistic variation in terms of strategies of negation marking in relation to tense and aspect categories. For example, in Siha, NEG_2 *ta-*, which is generally used as a non-main clause negation marker in other KB languages, has even spread to main clause verb forms; [Sih.] *tikaváa* 'We will hit' vs. *ti**ta**kaváa pfo* 'We will not hit'. iv) Most languages can take elements grammaticalized from 'come' and 'go' verbs in the TA slot, which essentially denote the modal concept that can be labelled as "certainty". I will only briefly mention the forms when they are interpretable as denoting TA notions.



Maps © www.thunderforest.com, Data © www.osm.org/copyright

Figure 1: Geographical distribution of the KB languages examined.

understanding of TA expressions, it should be noted that the core system of TA forms can be structurally described independently from these elements.[4]

This chapter is organised as follows. First, comparative lists of common TA markers from the above-mentioned languages are presented in §2, followed by a comparative analysis of semantic variation and its interrelation with each TA marker in §3. Based on these observations, the grammaticalisation processes and micro-typological correlations found in the TA systems of KB will be further discussed in §4. Conclusions are presented in §5.

[4] As Nurse (2003b: 73) mentions, generally in KB, negation in independent clauses is morphologically marked by a negative particle which has little influence on the morphological structure of the verb, i.e., it is relatively independent from broader TA-marking strategies.

2 Comparative lists of TAM

This section presents comparative lists of TAM of selected sample KB languages. Where available, sample sentences are tone-marked based on surface realizations, while the location of underlying lexical high tones is shown by an underscore. The lists are presented in the form of tables in which the morphological structure of the verb is indicated in rows and the core TA categories are shown in columns.

The morphological template of the verb in KB basically follows the typical Bantu structure, i.e., {Preinitial=SM-NEG_2-TAM-OM≠Stem-Final=Postfinal} (cf. Rose et al. 2002), out of which the structurally essential parts are SM, TAM, Stem, and Final. In addition to these elements, some of the lists include Preinitial[5] and Postfinal slots, if they are relevant to the expressions of the core TA categories in the language in question.

Another point to be noted regarding the structure relates to the multiplicity of TAM slots. As the literature shows, a string of TAMs in a single verb structure is quite typical in KB. Reflecting this morphological feature, the list has three slots for TAMs to capture the gradual nature of the TAMs in a simplified way, i.e., TAM_0 and TAM_2 are positively defined on the scale and TAM_1 is defined as "in-between". TAM_0 includes forms which are phonologically fused with the preceding SM (resulting in a monosyllabic cluster) and structurally self-standing, i.e., they can be realised without the co-occurrence of other TAMs, while TAM_2 forms tend to be realised in combination with preceding TAM(s), especially when denoting a past reference, and their lexical sources are relatively clear, i.e., they can be regarded as recently grammaticalised.

As for the core TA concepts, eight categories have been chosen to ensure sufficient semantic coverage and a clear formal distinction of the TAMs, namely two tense categories (past and future), four aspectual categories (progressive, anterior, completive, and habitual), and two combined categories (present and past statives). The so-called general present tense is not included in the list because, as Nurse (2008: 115–117) explains, the present time reference is normally expressed with an aspectual focus, with a typical situation involving the progressive aspect with a dynamic verb like "She is walking" or a stative aspect with an inchoative verb like "He sleeps". Thus, a pure exponent of the perfective (i.e., aspectually unmarked) present is not always clearly identifiable, or in other words the cate-

[5]The Mashami table includes the Preinitial slot, simply because all the examples suitable for examining TA expressions are presented as forms with the element in the original source, i.e., it does not (necessarily) mean that the Preinitial (generally understood as a Focus marker) directly affects the TA notion of the verb.

gory may be described as "empty" in TA matrices in many Bantu languages (see also Nurse 2003b: 77).

As shown in the following sections, all of the languages examined in this study have more than two distinct pasts and the non-WK languages have multiple futures. The four aspectual categories can be distinguished by TAMs throughout KB (although anterior and completive are not necessarily distinguished in some languages). Statives are adopted here as it may help identify the existence of a TA category that can be (tentatively) labelled as past imperfective, which may not be a common TA (combined[6]) concept in Bantu languages in general, but is widely attested in KB at least as a category frequently morphologised by a TA marker (cf. Nurse 2003b: 80). Examining these categories will thus help us identify how a TA marker diachronically expands or shifts its semantic coverage across KB languages.

2.1 Rwa

Rwa, the geographically western-most WK language, has three morphologically distinct pasts and a single future which is marked by Final *-áa*, apparently originating from a Prefinal-Final cluster **-ag-a*. Its tonally modified form *-aá* is used as a habitual marker. The formal distinction between PST_1 and PST_3 is only tonal, i.e., PST_3 has a grammatically assigned high tone[7] on the final vowel of the stem, e.g., [PST_1] *va-a-ʃí≠kab-ís-a m-biíri* vs. [PST_3] *va-a-ʃí≠kab-ís-á m-biíri* {SM2-$PST_{1/3}$-OM7≠hit-CAUS#CPx9-stick}[8] 'they hit (something in cl.7) by a stick'. However, if the verb stem is monosyllabic, the tonal difference is neutralised and the distinction is made by replacing the Final *-á* with *-é*[9]: [Rwa] *t-ā≠l-**ā***{SM1sg-PST_1≠eat-F} "we ate (PST_1)" vs. *t-ā≠l-**ē***{SM1sg-PST_3≠eat-F} "we ate (PST_3)". Though the historical background of PST_2 *nde-*[10] is uncertain, one may relate it to *le-* as a common past marker or the segmentally identical *ndé-*

[6] As a TA concept, this may not necessarily be a "combined" category. Rather, the form in question may be regarded as a special form of past marker that appears in various contexts of imperfective aspects. For further discussion, see §3.1.5.

[7] This high tone may cause high tone plateauing, i.e., tone-less TBUs between this and an immediately preceding high tone may be realised as high flat tones, e.g., *avatobírílíá* 'S/he made (something) for them'.

[8] Both SM2 and OM7 have an underlying high tone and the former is realised on the following syllable, while the latter is deleted by the so-called anti-Meeussen's Rule (i.e., HH > LH).

[9] From a historical point of view, it may be suggested that this element might be regarded as an irregular manifestation of **-ịle*, whose regular realisation is *-íe*.

[10] Note that *nde-* of PST2 is structurally different from *nde-* as a modality marker, relatively recently grammaticalised from the verb 'go', which can only appear with other tense markers.

in Gweno, which is a "verb-focusing" near past (cf. Philippson and Nurse [2000: 254]; see also §2.8).

The common markers *keé-* (from **kad* 'sit'), *ṃ-*, and *maa-* (both from**mad* 'finish') are attested as denoting progressive, anterior, and completive, respectively. Stative is marked by *-ié* which originates from **-ile* and its past tense is expressed by TAM *i-* with a lengthened final vowel (shown as *=V* in Table 1). This morphological template of past marking occurs repeatedly not only in various imperfective aspects of regular verbs including progressive, anterior/completive, and habitual, but also in non-verbal predicates such as existentials: [Rwa] *ti-∅≠ifó* {SM1pl-PRS≠EXT} 'we are (in a specific place)' vs. *tí-**i**≠ifo=**ó*** 'we were (in a specific place)'. Thus, it seems possible to posit a TA category morphologised by this template, which will be tentatively referred to as past imperfective. As shown in the following sections, this TA-combined category is often grammaticalized in other KBs as well.[11,12]

Table 1: List of the core TA markers in Rwa

TA category	SM	TAM_0	TAM_1	TAM_2	stem	F	PoF	translation
PST$_1$	*t-*	*a-*			*kab*	*-á*		'We hit'
PST$_2$	*ti-*		*nde-*		*kab*	*-á*		'We hit'
PST$_3$	*t-*	*á-*			*káb*	*-á*		'We hit'
FUT	*ti-*				*kab*	*-áa*		'We will hit'
PRS.STAT	*ti-*				*lol*	*-ié*		'We see/are watching'
PST.STAT	*tí-*		*í-*		*lol*	*-ié*	*=V*	'We saw/were watching'
PROG	*ti-*			*keé-*	*kab*	*-á*		'We are hitting'
ANT	*t-*	*a-*	*ṃ-*		*kab*	*-á*		'We have hit'
COMP	*t-*	*a-*	*ṃ-*	*maa-*	*kab*	*-á*		'We have finished hitting'
HAB	*ti-*				*kab*	*-āā*		'We hit (regularly)'

[11]According to Nurse (2003b: 77), the category marked by *we-* in Vunjo seems to correspond to this combined category. On the other hand, the recurrent morphologisation of this rather uncommon category can support Nurse's (ibid.: 85) claim from a cross-KB perspective that "TA categories are more stable than the morphemes which carry them."

[12]Examples in the lists presented in Tables 1–4 are provided in morphophonemic description and those in the body of the text are in phonemic description. Others are quoted without modification from the source. For detailed information of the phonemic inventories of Rwa, Uru, and Rombo, see Shinagawa (forthcoming[b]). The phonemic inventories of Mashami, Kibosho, and Gweno are presented in their original sources. A Cross-KB comparative list of phonemes is provided in Philippson & Montlahuc (2003).

In addition to these TAMs, Rwa has two markers grammaticalised from 'come' and 'go' (referred to here as COM and GOM), which are *ʃe-* and *nde-*, respectively (N.B. this is a distinct form from the *nde-* of PST_2). Although the former, in particular, is often used for future marking in some languages of CK and Rombo, these two markers clearly denote "certainty" as a modal concept in this language[13]: [Rwa] *fua y-a-**ndé**≠nis-á* {SM9-PST_1-GOM≠rain-F} 'It (certainly) rained/It rained (as expected)' vs. *fua y-a-**ʃé**≠nis-á* {SM9-PST_1-COM≠rain-F} 'It rained (unexpectedly)', where *nde-* expresses a past event ("raining") as more certain (i.e., it is recognised by the speaker that the event certainly or expectedly happened), while *ʃe-* indicates less certainty and/or expectedness. Note also that they are used in a past tense (marked by *a-*), suggesting that both COM and GOM are highly grammaticalised as modality markers.

2.2 Siha

As in Rwa, Siha also has a tripartite past and a single future. Since this distinction is also attested in Mashami (see §2.5), it may be regarded as a typical tense division in WK. Note, however, that some exponents are different from those in Rwa (see Table 1). PST_1 is marked by *le-*, a common past marker throughout KB, while PST_2 uses the same form with an extra lengthened vowel in the Final, which can be regarded as semantically parallel to and morphologically homogenous with *=V* of the past stative in Rwa. Stativeness can be expressed by *-ile* or *-i*, a possible shortened allomorph of **-ile*, and its past is also marked by an extra lengthened vowel.

As for aspectual forms, progressive is marked by *li-*, which is also a common marker regarded as being grammaticalised from a copula **li* 'be'. Otherwise, the aspectual exponents are similar to Rwa, i.e., completive is marked by *me-* (<**mad*) and *-aa* (<**-ag-a*) is used as habitual, though its tonal behaviour differs.

2.3 Uru

The tense system of Uru, a CK language, shows configurational differences from that of WK, since both past and future[14] are divided into two sub-categories with

[13] Interestingly, this semantic interpretation of GOM and COM is apparently the opposite to those found in Vunjo. For example, where COM marks a "more definite" intention while GOM marks a "less definite" intention (cf. Nurse 2003b: 87). For further discussion, see Shinagawa (forthcoming[b]).

[14] The number of tense categories of future is sometimes unclear since futurity can also be expressed by present progressive (cf. "present-used-as future" in Nurse 2003b) and this can be applied to other CK languages and Rombo. However, it is relatively clear that future tense is exclusively indicated by a single form (mostly with *-aa*) in WK.

Table 2: List of the core TA markers in Siha

TA category	SM	TAM_0	TAM_1	TAM_2	stem	F	PoF	translation
PST_1	*ti-*		*(l)e-*		*kā́v*	*-á*		'We hit'
PST_2	*ti-*		*(l)e-*		*kāv*	*-á*	*=V*	'We hit'
PST_3	*t-*	*á-*			*kā́v*	*-á*		'We hit'
FUT	*ti-*				*lóli*	*-áa*		'We will see'
PRS.STAT	*to-*				*ón*	*-i*		'We see/are watching'
PST.STAT	*tó-*				*on*	*-í*	*=V*	'We saw/were watching'
PROG	*ti-*		*li-*		*kā́v*	*-á*		'We are hitting'
ANT	*ti-*		*(l)e-*		*kā́v*	*-á*		'We have hit'
COMP	*ti-*		*(l)e-*	*me-*	*kā́v*	*-á*		'We have finished hitting'
HAB	*ti-*				*loli*	*-aa*		'We see (regularly)'

an extra future tense limitedly realised in main clause verbs. However, their exponents are interrelated with those in WK. The past markers *e-*(PST_2) and *le-* (PST_1) are parallel to PST_3 and PST_2 of Mashami, while *i-* of FUT_1 is regarded as a cognate of progressive *li-* in Siha, which is justified in terms of historical sound change (loss of intervocalic /l/, cf. Nurse 2003b: 79), as well as the fact that the semantic development process from progressive to future is a universally attested grammaticalisation pattern (cf. Bybee et al. 1994). FUT_2 *ʧi-* is a common future marker in CK which is regarded as having been grammaticalised from **ci* 'know' (cf. Nurse 2003b: 76). Stativeness is marked by *-ie* or *-i* as in Siha, while its past is marked by *e-* of PST_2. Progressive is marked by *ke-* (<**kad*) as in Rwa, contrary to its habitual usage in another CK language, Vunjo. The anterior marker *a-* (realised as *o-* in Table 3 as a result of vowel coalescence) is consistently fused with SM, as is widely attested in other CK languages and Rombo. Habitual has no segmental exponent in the TAM slots and is marked only by a high tone assigned to Final.[15]

It should be mentioned here that apart from the forms listed in Table 3, there is another morpheme that can be slotted in the TAM position. The morpheme *we-* is mentioned in Nurse (2003b) as "anomalous" in that in Vunjo it seems to indicate aspectual concepts (or more precisely the pastness of imperfective aspects), while

[15]The structural interpretation of this form can be rather controversial. It is reasonable to regard this as a "zero form," which "refers to timeless action, an activity which does or can occur over a vast present" (Nurse 2003b: 81). However, it is also possible to regard this as a descendant form of *∅≠(Stem)-*ag-a*, which will be discussed further in §3.3.

Table 3: List of the core TA markers in Uru

TA category	SM	TAM_0	TAM_1	TAM_2	stem	F	translation
PST_1	*lú-*		*le-*		*ólok*	*-a*	'We fell down'
PST_2	*lw-*	*é-*			*olók*	*-a*	'We fell down'
FUT_1	*lu-*		*í-*		*káp*	*-a*	'We will hit'
FUT_2	*lú-*		*tʃi-*		*káp*	*-a*	'We will hit'
FUT/COND	*a´*		*e-*		*wón*	*-a*	'(If...) s/he would see'
PRS.STAT	*lú-*				*won*	*-í*	'We see'
PST.STAT	*lw-*	*é-*			*won*	*-íé*	'We saw'
PROG	*lú-*			*ke-*	*káp*	*-a*	'We are hitting'
ANT	*l-*	*ó-* ({*a-*})			*wón*	*-a*	'We have seen'
COMP	*l-*	*o-* ({*a-*})		*ɱ-*	*wón*	*-a*	'We have finished seeing'
HAB	*lú-*				*ɾeí*	*-á*	'We write (regularly)'

it appears in the leftmost position of the TA string, where tense markers are usually slotted (ibid.: 77). However, in Uru *we-* appears between the TAM_1 and TAM_2 positions and seems to denote a kind of predicate focus function as in; *ɲálĕ:tʃa* 'S/he came' vs *ɲálewê:tʃa* 'S/he also came' (subject additive focus) or 'S/he came again' (event recurrence). This will be further discussed in §3.2.

2.4 Rombo-Mkuu

In the current classification, e.g., Maho (2009) based on the classifications of Nurse (1981) and Philippson & Montlahuc (2003), Rombo as a subgroup is further classified into (at least) four varieties, namely Useri, Mashati, Mkuu, and Keni (from North to South). However, its dialectal variation seems more diverse and complicated than this division suggests.

The tense system is basically comparable with that of Uru, i.e., both past and future are bipartite, though there may be more exponents denoting futurity. FUT_1*ʃe-* is regarded as having grammaticalised from *ʃa* 'come', denoting a near future reference (possibly with epistemic modal connotations). The inherited form of **-ile* is used as part of past tense marking, while stativeness is marked by its shortened form. Its past tense is indicated by a possibly relevant form of *we-* in Vunjo. The aspectual patterns are similar to Siha, except that (present) habitual is marked by *e-*, which is segmentally identical to FUT_2 *e-*. However, it should be noted that the two markers occupy different slots, as shown in Table 4.

Table 4: List of the core TA markers in Rombo-Mkuu

TA category	SM	TAM_0	TAM_1	TAM_2	stem	F	translation
PST$_1$	*dú-*		*le-*		*lolj*	*-a*	'We saw'
PST$_2$	*dú-*				*lol*	*-íé*	'We saw'
FUT$_1$	*dú-*			*ʃe-*	*ɾ̲und*	*-a*	'We will work'
FUT$_2$	*du-*		*é-*		*ɾ̲und*	*-a*	'We will work'
PRS.STAT	*dú-*				*k̲und*	*-i*	'We want'
PST.STAT	*dú-*		*ve-*		*k̲und*	*-i*	'We wanted'
PROG	*du-*		*í-*		*el̲eke*	*-a*	'We are heading for'
ANT	*dw-*	*á-*			*lolj*	*-a*	'We have seen'
COMP	*dw-*	*á-*		*me-*	*lolj*	*-a*	'We have finished seeing'
HAB	*dw-*	*é-*			*k̲ab*	*-a*	'We hit (regularly)'

2.5 Mashami (Rugemalira & Phanuel 2012)

As examined in previous sections, it is clear from the data presented in Table 5 that the tense system of Mashami, as expected, clearly follows the WK type. Moreover, the TA system as a whole seems quite similar to that of Rwa.

Table 5: List of the core TA markers in Mashami

TA category	PreI	SM	TAM_0	TAM_1	TAM_2	stem	F	translation
PST$_1$	*n=*	*lw-*	*á-*			*many*	*-a*	'We knew'
PST$_2$	*n=*	*lú-*		*le-*		*mány*	*-a*	'We knew'
PST$_3$	*n=*	*lw-*	*é-*			*mány*	*-a*	'We knew'
FUT	*n=*	*ʃí-*				*kór*	*-aa*	'I will cook'
PRS.STAT	*n=*	*lu-*				*salal*	*-ye*	'We are standing'
PST.STAT	*n=*	*lu-*		*é-*	*ké-*	*many*	*-a*	'We were understanding'
PROG	*n=*	*lú-*			*ké-*	*many*	*-a*	'We are knowing' [sic]
ANT^A/COMP	*n=*	*lw-*	*á*		*m-*	*many*	*-a*	'We have already known'
HAB	*n=*	*ʃí-*				*kor*	*-aa*	'I cook'

[a]According to the source, anterior can be marked without *a-*, as in *ku-n≠shani-shi-a≠kya* {SM2sg-ANT-come#FOC-SM1sg-PST1≠be cured} 'Since you have come, I am safe (Sw: Kwa kuwa umekuja, nimepona)'.

However, there are two points to be noted which are not explicitly shown in Table 5. First, TAM_1*e-* as a past marker of stative verbs also marks the past-

ness of imperfective forms such as habitual: [Mas.] *n=lu-**é**≠many-aa* {FOC=SM1pl-PST.IMPF≠know-HAB} "We used to know". Second, final *-aa* (<*-*ag-a*) can be used not only as a future marker but also to indicate progressive meaning: [Mas.] *ni-a-koy≠aa # i≠kor-**aa*** {FOC-SM3sg≠find-FUT # INF≠cook-PROG} "He will be cooking".

2.6 Kibosho (Kagaya 1989)

Though available data are rather limited, it can be said that Kibosho,[16] a WK language, shows a somewhat unique pattern.[17] First, the tense distinction seems to be a bipartite past and a single future, which is slightly different from the 3:1 pattern in the other WK languages examined in this study. Second, the descendant of *-*ag* does not denote futurity as in other WK, but indicates progressive aspect, which is, however, observed in limited contexts in Mashami. While Kagaya (1989: 829) describes future as marked by the vowel lengthening of SM (shown as *V-* in Table 6), this element may well be identified as *i-* in other languages and the same element is attested in progressive forms as well, according to my own data.[18]

Table 6: List of the core TA markers in Kibosho

TA category	SM	TAM_0	TAM_1	TAM_2	stem	F	translation
PST_1	*l-*	*o-* ({a-})			*ch*	*-a*	'We arrived' (Hodienal)
PST_2	*lu-*		*le-*		*ch*	*-a*	'We arrived' (Remote past)
FUT	*lú-*		*V-* ({i-})		*som*	*-a*	'We will arrive'
PRS.STAT	*lu-*				*ke*	*-i*	'We are (at a place)'
PST.STAT	*lw-*	*e-*			*ke*	*-i*	'We were (at a place)'
PROG	*lú-*		*V-* ({i-})		*som*	*-áa*	'We read/are reading'
HAB	*lú-*				*som*	*-aá*	'We read (regularly)'

It is also to be noted that this language, too, seems to morphologise past imperfective, which is marked by *e-*: [Kib.] *n̩=lw-**e**≠som-aa* 'We were reading', where *e-*encodes the past tense of the progressive aspect.

[16]Though Kibosho is normally classified as WK in the linguistic literature, it is regarded as a variety of Vunjo in local narratives, according to Kagaya (2006).

[17]Note however, that there seem to be dialectal patterns which are more complicated than that which is presented here, e.g. ∅≠(Stem)-*ie* is also attested as a remote past form or as a variant of hodiernal past, while future may be denoted by a null-marked form or by *ende-* as GOM. I acknowledge Gérard Philippson for this information.

[18]The examples of FUT, PROG, and HAB in Table 6 were confirmed through elicitation with a native speaker in his 30s in my field research carried out in August 2018.

2.7 Vunjo (Nurse 2003b)

Vunjo exhibits a typical CK-Rombo pattern of a tense system with a bipartite distinction both in past and future. The exponent of FUT$_1$ is *ci-*, which is shared with Uru. Though it is not shown in Table 7, there is a form which is realised in the past tense of some imperfective forms, which is *we-* as mentioned in §2.3: [Vun.] *lu-we-i≠kap-a* {SM1pl-P.I.-PROG≠hit-F} 'We were hitting'. As observed above, *we-* in this example may well be regarded as a form functionally equivalent to *ve-* in Rombo and *i-* in Rwa, which denotes a certain range of past imperfective. As for aspectual categories, it should be noted that progressive *i-* is shared with Rombo, while habitual is marked by *ke-*, which is identical to the progressive marker in Uru.

Table 7: List of the core TA markers in Vunjo

TA CATEGORY	SM	TAM$_0$	TAM$_1$	TAM$_2$	stem	F	translation
PST$_1$	*l-*	*o-*			*kap*	*-a*	'We hit'
PST$_2$	*lu-*		*le-*		*kap*	*-a*	'We hit'
FUT$_1$	*lw-*	*(e-)*		*ci-*	*kap*	*-a*	'We will hit'
FUT$_2$	*lw-*	*e-*			*kap*	*-a*	'We will hit'
PROG	*lw-*		*i-*		*kap*	*-a*	'We are hitting'
ANT	*lu-*				*kap*	*-ie*	'We have hit'
COMP?	*l-*	*o-*		*m-*	*kap*	*-a*	'We have/had already hit'
HAB	*lu-*			*ke-*	*kap*	*-a*	'We hit regularly'

2.8 Gweno (Philippson & Nurse 2000)

According to Philippson & Nurse (2000), the community of Gweno speakers has long lived in Northern Pare, remote from Kilimanjaro, which may have caused the language to retain some archaic features and to develop differently from other languages of the three main subgroups.

The configuration of the Gweno tense system (Table 8) seems identical to the CK-Rombo type, where both pasts are morphologised by the **-ile* form. The TA forms which are more or less relatable to common forms in other KB languages include *le-* and less reliably *nde-* and *ʧe-*. Progressive is not marked by a specific morpheme but is expressed as a "focused form of general present". Though habitual *ʧi-* appears to be identical to FUT$_1$ in Vunjo, it is pointed out in Philippson and Nurse (ibid.: 255) that it is a shortened form of the verb *ʧiβia* 'be accustomed to'.

Table 8: List of the core TA markers in Gweno

TA category	Pre-Ini	SM	TAM_0	TAM_1	TAM_2	stem	F	translation
PST_1/ANT		*fu-*				*βúk*	*-íe*	'We left'
PST_2		*ni-*		*lé-*		*ɣend*	*-ie*	'I went'
FUT_1		*fw-*	*a-*	*ɣe-*		*ʃiɣ*	*-a*	'We will look for'
FUT_2		*fw-*	*a-*	*tʃe-*		*ɣu*	*-a*	'We will buy'
PROG[A]	*ḿ=*	*fw-*	*â*			*ɣu*	*-a*	'We are buying'
ANT/COMP		*í*		*nde-*	*(mi-)*	*pfw*	*-á*	'It has died'
HAB		*ni*			*tʃi*	*ɣend*	*-a*	'I go'

[a]Progressive aspect seems to be expressed in several different ways in Gweno. According to Philippson & Nurse (2000: 253), in addition to the form cited in Table 8, there is another progressive marker, *ky-a-*, whose lexical origin is not explicitly stated in the source.

3 Semantic correspondence of each TAM

Based on the above observations, the distribution of all the common TA markers in the eight sample languages examined in this study are summarised in Table 9.

The following sections provide brief notes on the variation of TA categories that each TAM covers in the different languages.

3.1 Inherent markers: TAM_0 and TAM_1

3.1.1 *a-*

This prefix is attested in all the languages examined. While its conceptual coverage in each language varies from past through anterior and even to future, the geographical distribution of each type appears to overlap with the boundaries of the subgroups, i.e., it is used as past in WK (N.B. two tonally distinctive markers in Rwa), anterior in CK and Rombo, and "general present-future" (including progressive when the verb is in Focus) in Gweno.[19]

3.1.2 *e-*

The data suggest that two distinct TA categories can be encoded by the prefix *e-*. One is past imperfective, as a frequently morphologised category across KB, as

[19]For this apparently uncommon grammaticalisation process, see Nurse (2003b: 74–75).

Table 9: Comparative list of TAMs in eight sample languages of KB

	WK				CK		Rombo	Gwe.
	Rwa	Sih.	Mas.	Kib.	Vun.	Uru	Mkuu	
a_1- *a_2-*	PST.N PST.R	PST.R	PST.N	PST.N	PST.N	ANT	ANT	PRS /FUT /CONT
e_1-		P.I.	PST.R/ P.I.	P.I.		PST.R/ P.I.		
e_2-					FUT.R	FUT	FUT/ HAB	
V-				FUT				
le-	PST.M	PST.N/ M	PST.M	PST.R	PST.R	PST.N	PST.N	PST.R
li-	P.I.	CONT			CONT	FUT.N	CONT	
we-, *ve-*					P.I.?	FOC	P.I	PST.M?
ci-					FUT.N	FUT.R		
ke-	CONT		CONT		HAB	CONT		CONT[a]
(ker- *i-)*					CONT	CONT		
ʃe-	CERT↓				CERT↑		FUT	FUT.R
nde-	CERT↑				CERT↓			
ɱ-, *mi-*	ANT	ANT	ANT/ COMP		COMP	COMP	COMP	ANT
maa-	COMP	COMP[b]	COMP					
-ile	STAT	STAT	STAT	STAT	ANT/ STAT	PST.ST	PST.R	PST
-i		STAT		STAT		STAT	STAT	
-ag	FUT/ HAB	FUT/ HAB	FUT/ HAB	CONT/ HAB		HAB		
=V	P.I.	PST.R?						

[a]Used as an auxiliary verb.
[b]Used as an auxiliary verb.

attested in Uru (as past tense of stative, progressive, and anterior) and in Kibosho (as past tense of progressive): [Uru] *lw-é-m̩≠ɾeí-a* {SM1pl-P.I.-ANT≠write-F} "We had written", [Kib.] *n̩=lw-e-som-aa* {FOC=SM1pl-P.I.≠read-PROG} "We were reading". The other is a series of markers denoting the future-habitual references as in Vunjo (remote future), Uru (future [with conditional]), and Rombo (remote future): [Uru] *kaʃiká ʃimbó, n=á-**e**≠wón-a m̩meéku* {FOC=SM3sg-FUT.COND≠see-F} "If s/he arrives in Shimbwe (a name of a village), s/he will meet an old man".

What is significant here is that in Uru and Rombo-Mkuu, the segmentally identical exponent appears in different TAM slots for marking different TA concepts, e.g., [Mkuu] *dw-é≠kab-a* "We hit (regularly)" vs. *du-é≠ɾund-a* "We will work". This may suggest that while the two TA categories are regarded as close to each other, they are still systematically differentiated, and that the multiple TAM slots are used in order to morphologise the semantic difference of the TA categories.

3.1.3 *le-*

This marker is attested in all the languages examined except Rwa and consistently denotes past tense throughout KB. It should be noted that in Siha, *le-* is also realised as *e-* especially when preceded by a SM with the syllable structure /Ci/ or /Cu/: [Sih.] *ti-le≠káv-á ~ te-e≠káv-á* "We hit (near past)".

3.1.4 *li-*

li-, or its weakened form *i-*, is also a common TAM attested across KB and its typical function is progressive, as seen in Siha, Vunjo (cf. labelled as "present-used-as-future" in Nurse 2003b), and Rombo-Mkuu. Uru also uses this prefix to encode a future reference more clearly. However, as mentioned in §2.1, *i-* in Rwa covers the past reference of imperfective aspects including progressive, anterior, habitual, stative, and even copula constructions.

3.1.5 *we-*

The forms originating from **wa* 'be, become' are attested in CK and Rombo, and they are realised in past tense forms of various imperfective aspects such as anterior and progressive in Vunjo (*we-*), and the past of stative verbs in Rombo (*ve-*). In contrast, *we-* in Uru clearly shows a different direction of its grammaticalisation process, i.e., it appears to have developed as a marker of predicate focus. The interpretation that *we-* in Uru marks predicate focus can be justified by the fact that it is not allowed to appear in the "out-of-focus" positions such as before a clause-final negation particle (as mentioned in §2.3) or in relative clauses, where

the preverbal clitic *ni=*, which denotes "main clause-ness" in the language, is also avoided, as in [Uru] *a-le-tʃ-a=se* 'The one who came again' vs. **a-le-we-tʃ-a* (for further discussion of *ɲi=*, see Shinagawa forthcoming(a)).[20]

3.2 Recently grammaticalised markers: TAM_2

3.2.1 **ci*

This marker is observed only in CK and scarcely found elsewhere. Its function is consistently future marking and it can be regarded as having grammaticalised from the verb **ci* 'know' (Nurse 2003b: 76), suggesting that it can be viewed as a rather recent innovation in the CK area. An apparent homophonic form (*tʃi-*) is also attested in Uru and Gweno, denoting habitual in both languages. However, as Philippson & Nurse (2000: 255) clarify, *tʃi-* should be regarded as a shortened form of a verb stem *tʃiβia* 'be accustomed to', since in Uru, too, speakers replace habitual *tʃi-* with a periphrastic structure involving *tʃiβia*.

3.2.2 **kad*

TAMs grammaticalised from **kad* 'sit, stay' are broadly attested in WK (Rwa, Mashami), in CK (Vunjo, Uru) and in Rombo. Though the semantic category it denotes can be basically recognised as progressive aspect, it is also used as a habitual marker, as attested in Vunjo. If we follow the general tendency of the direction of grammaticalisation paths involving both concepts, it can be presumed that habitual is further grammaticalised from progressive (cf. Heine & Kuteva 2002: 93, Haspelmath 1998: 48). This grammaticalisation path from progressive to habitual (then to future) will be discussed again in §4.1.2.

3.2.3 **mad*

TAMs originating from **mad* 'finish' stably denote anterior meanings (then to completive in some WK languages) and distributed throughout KB.[21] Two distinctive forms *maa-* vs. *m̩-* have developed in some WK languages, where the former is obviously a later innovation than the latter and seems to be in complementary distribution with anterior (to past) use of **-ile*.

[20] See Hyman & Watters (1984) for various examples of structural co-occurrence restrictions on (inherent) focused forms in Bantu. For a cross-Bantu discussion of the developmental process of progressive into focus, see Güldemann (2003) and Gibson (2019).

[21] *me-* in Rombo can be regarded as a relic of an anterior form of **mad* (**mad-i̩de* > *me*), which is also attested in Old Moshi (Gérard Philippson, p.c.).

3.3 Suffixes

3.3.1 *-*ag*

As previously pointed out by Philippson & Montlahuc (2003), the reflexes of *-*ag* are distributed only in WK. In the data, it is attested (as *-aa* with various tonal realization) in Kibosho, Siha, Mashami, and Rwa, i.e., exactly covering the WK area. As Shinagawa (2015) suggests, the absence of *-*ag* in non-WK languages may at least partly be explained by the lack (or weakness) of vowel length contrast in those languages. In this sense, it may be possible to regard the present habitual form of Uru, which has no segmentally overt TAM and is only marked by a high tone on the final vowel,[22] as a shortened form of a relic of *-*ag-a*. Though its semantic coverage is essentially habitual and future as described in the literature, it is worth mentioning that *-aa* in Kibosho denotes progressive aspect. The historical process of the semantic split of this marker will be further discussed in §4.1.2.

3.3.2 *-*ile*

Reflexes of *-*ile*, mostly realised as *-ie*, as well as its historical allomorphic form *-i*, are widespread throughout KB. While *-ie* denotes anteriority or past tense in CK,[23] Rombo, and Gweno, it encodes stativity in WK. For example, in Rwa, this element derives stative verb stems that are paradigmatically differentiated from default (*-a* ending) verb stems in that they follow different past-marking paradigms. In the following examples, the past tense is marked by the combination of TAM *i-* and a lengthened final vowel just as in existential predicates (see §2.1): [Rwa] *ti-∅≠tisiɾ-ié* {SM1pl-PRS≠write-STAT} "We have written (resultant state)" vs. *tí-í≠tisiɾ-ié=e* "We had written (resultant state)". Parallel morphosyntactic behaviour is observed in stative *-i* in Rombo-Mkuu, where the past imperfective *ve-* denotes the past tense of stative verbs: [Mkuu] *du-∅≠kund-i* "We love" vs. *dú-ve≠kund-i* "We loved".

[22] Of course, the form can be simply seen as a so-called "null form" which is unmarked for tense and aspect, thus denoting the general present or a generic situation (cf. Nurse 2008: 117–118). More investigation on this issue is needed.

[23] in Vunjo seems to exhibit an intermediate situation, i.e., Nurse (2003b) describes its function as anterior (presupposing that inchoative verb stems such as *won* 'see,' *lal* 'lie, sleep' etc. express a stative meaning when attached to the anterior marker), while Moshi (1994) rather consistently insists that, though it is greatly influenced by the lexical meanings of the verb, the core semantics of *-ie* forms can be regarded as stativity.

3.3.3 Vowel copy suffix

In some western WKs, there is a Postfinal clitic which is a copy of the Final (realised as a lengthened vowel), denoting the past tense of specific aspect forms. As stated in §2.1 and §3.3.2 the past tense of stative predicates is expressed by a combination of past imperfective *i-* and this vowel copy (VC) suffix in Rwa. Moreover, Siha apparently expands its use into default verb forms to encode remoteness of past tense (cf. the formal distinction between PST_1 vs. PST_2 in Table 2).

4 Grammaticalisation chains and a microparametric approach to regionally shared features

4.1 Grammaticalisation chains

As presented in the previous sections, there are several TAMs which are shared by different languages in the same subgroup or even across sub-group boundaries. TA categories encoded by such TAMs can be either shared throughout languages (e.g., *maa-* in WK is consistently used as a completive marker) or gradually shift from language to language. In the latter case, a gradual and systematic shift of the semantic coverage of common TAMs may clarify the grammaticalisation process of each TAM and shed light on the historical development of TA systems in KB. Figure 2 is a simplified illustration of this type of correspondence, found between Vunjo and Rwa.

The corresponding relation shown in Figure 2 clearly demonstrates that habitual and progressive are interconnected by a shared TAM grammaticalised from **kad* 'sit, stay' and, in parallel with the connection, the progressive *i-* of Vunjo in turn corresponds to the past imperfective in Rwa. Similarly, the common TAM *ɱ-* denotes completive in Vunjo and anterior in Rwa, while anterior *-ie* in Vunjo encodes stative in Rwa, and so on. The following sections will focus on two clusters of TA categories, in which such corresponding connections across languages are clearly observed. Note, however, that the following discussion will focus only on the interrelation between the three main subgroups, due to the limited materials on Gweno itself and its historical linguistic relations with other KB languages.

4.1.1 PST-ANT-STAT continuum

As shown in Table 10, a semantic area ranging over past, anterior, and stative forms a cluster of TA categories interconnected by common TAMs such as *a-*, *le-*, *e-*, *-ie*, and various markers originating from **mad* 'finish'. If we regard the

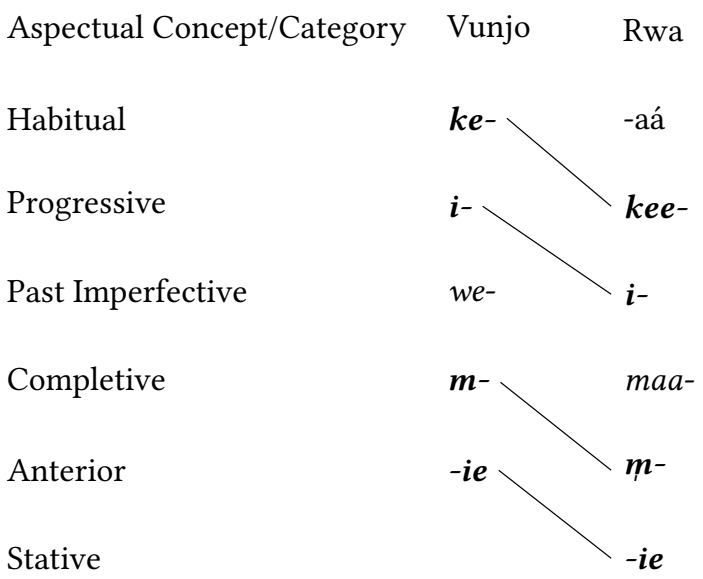


Figure 2: Schematised interrelation of TAMs between Vunjo and Rwa

prototypical semantic area of *-*ie* as anterior, the following process can be suggested.

In WK, probably motivated by the connotation of resultant state that likely emerged from the concept of anterior (cf. Nurse 2003a: 96), the semantic area of *-ie* shifted to stative; and successively *m̩-*, most probably grammaticalised as completive, was pulled to anterior. The apparently redundant innovation of *maa-* may be explained by the empty gap made after the movement of *m̩-*. In contrast, *-ie* in Rombo-Mkuu moved in the other direction to past, which seems to be a more usual semantic shift of **-ile*, while *a-*occupies the semantic area of anterior.

The TAM *a-* in Vunjo and WK, on the other hand, plays the role of past-marking as in many other Bantu languages (cf. Nurse 2008: 82). With the exception of Siha, where *a-* (with a high tone) denotes the remotest past, the relative order of temporal distance denoted by these TAMs seems stable throughout different TA systems, i.e., *a-* occupies the temporal space closest to the present (or time of utterance), followed by *le-*, and *e-*occupies the furthest.[24]

4.1.2 PROG-FUT-HAB continuum

The other TA cluster connected by common TAMs ranges over the progressive, future, and habitual areas. To clarify this interrelation, it would be necessary

[24]The order of temporal distance between *a-* and *le-* has already been proposed by Nurse (2003b: 74).

Table 10: Formal correspondence of TAMs denoting PST-ANT-STAT continuum

	WK				CK		Rombo
	Rwa	Mas	Sih.	Kib.	Vun.	Uru	Mkuu
PST.R	*a-*	*e-*	*a-*	*le-*	*le-*	*e-*	*-ie*
PST.M	*nde-*	*le-*	*le-* (=*V*)				
PST.N	*a-*	*a-*	*le-*	*a-*	*a-*	*le-*	*le-*
STAT	*-ie*	*-ie*	*-ie*	*-ie*	*-ie*	*-ie*	*-i*
ANT	*ṃ-*	*m-*	*m(e)-*			*a-*	*a-*
COMP	*ṃ-maa-*	*ṃ-maa*	*ṃ-maa*		*m-*	*a-ṃ-*	*a-me-*

to identify the original semantic area of **-ag-a* in the TA system at the Proto-KB stage. Though it is widely assumed that its prototypical meaning throughout Bantu may be a broad range of imperfective, with habituality or iterativity as core concepts (cf. Meeussen 1967: 110, Nurse 2008: 262–263), the data examined here suggest that while *-aa* denotes habitual and future in Rwa, Siha, and Mashami, it is also used as a progressive marker in Kibosho. Moreover, its progressive use is also, though limitedly, attested in Mashami, where habitual and future usages are also attested, i.e., the Mashami system can be regarded as an intermediate stage of the semantic shift of *-ag*. If this is the case, then it is reasonable to posit that the direction of semantic change may have started from progressive to habitual or future and not vice versa, as suggested by a widely attested tendency of the grammaticalisation path developed from present progressive to habitual or future (cf. Bybee et al. 1994: 158). Along the same lines, Haspelmath (1998) also explains how an old present form (**-ag* in this case) changes into a new future as a side effect of the emergence of a new present (*ke(e)-* as an apparently recent innovation of grammaticalisation) in a wide variety of languages. If we follow these assumptions, the distribution summarised in Table 11 would suggest the following process.

In Rwa and Mashami, **-ag-a* shifted and split into future and habitual due to the innovation of a novel progressive marker originating from **kad*.[25] In Siha, the same process was initiated by the progressive *li-*, which plays the same role in Vunjo, where *ke-* shifted to habitual aspect. As suggested in §2.3 and §3.3.1, *-á* of habitual in Uru might be regarded as a relic of **-ag-a*, which has completely

[25]For more detail on the semantic change of **-ag* in Rwa, see Shinagawa (2009).

vanished in non-WK areas probably due to the lack or weakness of vowel length contrast. Though the process is still uncertain, it is worth mentioning that in Rwa a reflex of **li*, which is used as a progressive marker in most languages, denotes past imperfective, which is encoded by *we-* in Vunjo and Rombo-Mkuu and by *e-* in other languages.

Table 11: Formal correspondence of TAMs denoting PROG-FUT-HAB continuum

	WK				CK		Rombo
	Rw.	Ms.	Sh.	Kb.	Vn.	Ur.	Mkuu
FUT2	*-áa*	*-áa*	*-áa*	(SM-)*V-*	*e-*	*ci-*	e_1-
FUT1					*ci-*	*i-*	*ʃe-*
HAB	*-aá*	*-aa*	*-aa*	*-áa*	*ke-*	*-á*	e_0-
PROG	*kée-*	*ke-*	*li-*	*-aá*	*i-*	*ke-*	*i-*
P.I.	*i-*	*e-*	*e-*	*e-*	*we-*	e_0-	*ve-*

4.2 Microparametric approach to regional features

This section briefly presents a provisional sketch of micro-typological correlations between morphosyntactic parameters related to TA systems. As observed in §2, tense-marking systems in KB can be divided into two types, i.e., one is the WK type with a tripartite past and a single future, and the other is the non-WK type with a bipartite system for both past and future. If we focus on the future marking, the former can be classified as the mono-future type, while the latter as the pluri-future type. This distinction largely overlaps with the distribution of **-ag-a*, i.e., the languages classified as the mono-future type are those with **-ag*, while the others lack this element.

The first point of note is that this structural difference itself may suggest a typological correlation, i.e., if a language is the pluri-future type, then the language is likely to denote future time reference with pre-stem TAMs, while if a language is the mono-future type, then future tense is likely to be marked in the Final slot. This provisional correlation may be justified by the difference of "openness" of the morphological slots, i.e., pre-stem slots are relatively open for newly grammaticalised elements, while the Final slot is a rather closed slot for a limited number of inflectional elements.

The second point involves the relationship with the degree of grammaticalisation of the markers COM and GOM. As presented in §2.1, these markers are essen-

tially used to denote modality in Rwa and the situation seems similar in Vunjo (Nurse 2003b).[26] However, in pluri-future type languages, COM is more like a future tense marker (e.g., Rombo-Mkuu is a typical case, see §2.4), while GOM is asymmetrically less grammaticalised, e.g., as shown in §2.1. On the other hand, GOM is used even with past tense forms in Rwa, while its cooccurrence with a past marker is not grammatically acceptable in Rombo-Mkuu. While this may suggest a correlation between types of future tense categorisation and grammaticalisation types of 'come' and 'go' verbs, it should be noted that languages without a fully developed GOM also tend to have a fully developed focus-marking system with *ní-* (cf. Shinagawa 2015, forthcoming(a)).

5 Concluding remarks

This study has presented an overview of the distribution of common TAMs in KB and their semantic correspondences across the languages of this group, some of which have scarcely been examined in previous studies. Based on the systematic correspondences revealed by the data, it was shown that there are two clusters of TA categories, namely past-anterior-stative and progressive-future-habitual. These categories are interconnected across the boundaries of languages or subgroups through grammaticalisation chains of common TAMs. The findings have also suggested that some regionally shared features may also be regarded as reflecting micro-typological correlations, i.e., a typological distinction between a mono-future (WK) type and a pluri-future (non-WK) type may be correlated with different components of the TA-marking system such as grammaticalization types of COM and GOM and/or the existence of *ni-* as a focus-marking strategy in a broad sense.

However, there are various issues to be investigated further. First, comprehensive descriptions of the TA systems of under-described languages are needed to fill the gaps with reliable data. Second, the scope of description should also be expanded to cover the entire range of TA expressions including compound tenses, TA in dependent clauses, negative clauses. Finally, it would be valuable to explore the whole range of morphosyntactic microvariation of each language in order to investigate the possible micro-parametric correlations between TA systems and other logically independent properties of grammar, which may shed light on shared principles of KB grammar underlying the group's internal linguistic diversity.

[26] However, their meaning seems discrepant between WK and Vunjo, i.e., GOM denotes strong certainty in WK, while it denotes relatively weak certainty in Vunjo compared to COM (cf. Shinagawa forthcoming(b)).

Acknowledgements

This work is made possible by financial support from JSPS's Grants-in-Aid for Scientific research (C) #16K02630 and #19K00568, and is part of the academic results of ILCAA's joint research project "Typological Study of Microvariation in Bantu (Phase 2)." My sincere thanks go to anonymous reviewers for their helpful comments and especially Gérard Philippson, whose detailed comments and suggestions helped me to avoid errors in the analysis and interpretation of the data. However, of course, remaining errors or misunderstandings are my own.

Abbreviations

1, 2, 3 …	Class numbers
1sg/pl …	Person + Singular/Plural
ANT …	Anterior
(↑/↓) CERT …	(high/low) Certainty
COM	grammaticalised marker from 'come'
COMP	Completive
PROG	Progressive
EXT	Existential
FOC	Focus marker
$\text{FUT}_{(1\sim2)}$	Future (near ~ remote)
G	Glide
GOM	grammaticalised marker from 'go'
H	High tone
HAB	Habitual
NEG_2	Secondary Negative
P.I	Past Imperfective (Past tense for Imperfective aspect forms)
PRS	Present
$\text{PST}_{(1\sim3)}$	Past (near ~ remote)
SM	Subject Marker
STAT	Stative
TAM	Tense and Aspect marker
V	Vowel (including a copied vowel of a preceding element)
VC	Vowel Copy (clitic)
VLC	Vowel Length Contrast
-	Affix boundary
=	Clitic boundary

≠	Verb stem boundary
#	Word boundary

Tonal annotation (broad phonetic):

[á]	high
[ˈá]	upstepped high (descriptive expression of [ő]: super high)
[ˌá]	downstepped high
[ā]	middle
[â]	falling
[ǎ]	rising

References

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar*. Chicago, IL: The University of Chicago Press.

Gibson, Hannah. 2019. The grammaticalisation of verb-auxiliary order in East African Bantu: From information structure to tense-aspect. *Studies in Language* 43(4). 757–799.

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.

Haspelmath, Martin. 1998. The semantic development of old presents: New futures and subjunctives without grammaticalization. *Diachronica* XV(1). 29–62.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Hyman, Larry M. & John R. Watters. 1984. Auxiliary focus. *Studies in African Linguistics* 15(3). 233–273.

Kagaya, Ryohei. 1989. Chaga-go (the Chaga language). In Takashi Kamei, Rokurō Kôno & Eiichi Chino (eds.), *Gengogaku daijiten*, vol. 2 (The Sansiedo Encyclopedia of Linguistics), 826–831. Tokyo: Sanseidō.

Kagaya, Ryohei. 2006. *A Kiw'oso vocabulary*. Tokyo: Research Institute of Languages, Cultures of Asia & Africa.

Maho, Jouni Filip. 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 (20 March, 2013).

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

Moshi, Lioba. 1994. Time reference markers in KiVunjo-Chaga. *Journal of African Languages and Linguistics* 15. 127–159.

Nurse, Derek. 1981. Chaga-Taita. In Thomas J. Hinnebusch, Derek Nurse & Martin Mould (eds.), *Studies in the classification of Eastern Bantu languages*, 127–180. Hamburg: Helmut Buske Verlag.

Nurse, Derek. 2003a. Aspects and tense in Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 91–102. London: Routledge.

Nurse, Derek. 2003b. Tense and aspect in Chaga. *Annual Publications in African Linguistics* 1. 69–90.

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

Philippson, Gérard & Marie-Laure Montlahuc. 2003. Kilimanjaro Bantu (E60 and E74). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 475–500. London: Routledge.

Philippson, Gérard & Derek Nurse. 2000. Gweno: A little known language of Northern Tanzania. In Kulikoyela Kahigi, Yared Magori Kihore & Maarten Mous (eds.), *Lugha za Tanzania/Languages of Tanzania*, 231–284. Leiden: CNWS Publications.

Rose, Sarah, Christa Beaudoin-Lietz & Derek Nurse (eds.). 2002. *A glossary of terms for Bantu verbal categories, with special emphasis on tense and aspect*. Munich: Lincom Europa.

Rugemalira, Josephat M. & Benedictor Phanuel. 2012. *A grammatical sketch of Kimashami* (LOT publications 2). Dar es Salaam: Languages of Tanzania (LOT) Project, University of Dar es Salaam.

Shinagawa, Daisuke. 2009. A rare story of the emergence of the future? A hypothesis on the historical development of Proto-Bantu **-ag* in Rwa (Bantu, E61). *HERSETEC: Journal of Hermeneutic Study and Education of Textual Configuration* 2. 137–150.

Shinagawa, Daisuke. 2015. Vowel length and TMA micro-variation in Kilimanjaro Bantu. *Asian and African Languages and Linguistics* 9. 5–21.

Shinagawa, Daisuke. Forthcoming(a). A micro-parametric approach to focus marking *ní* in Kilimanjaro Bantu languages. In Lutz Marten, Rozenn Guérois, Hannah Gibson & Eva-Marie Bloom-Ström (eds.), *Morphosyntactic variation in Bantu*. Oxford: Oxford University Press.

Shinagawa, Daisuke. Forthcoming(b). Chaga languages. In Lutz Marten, Nancy Kula, Jochen Zeller & Ellen Hurst (eds.), *Oxford guidebook of Bantu languages*. Oxford: Oxford University Press.

Chapter 12

The historical development of the reflexive-reciprocal polysemy in Hehe

Lengson Ngwasi[a] & Abel Mreta[a]
[a]University of Dar es Salaam

This chapter describes the encoding of reflexive and reciprocal events in Hehe, a Bantu language spoken in Tanzania. It is argued that the reflexive prefix has historically developed into a reciprocal marker, thus, replacing the reflex of the Proto-Bantu reciprocal suffix *-*an*-. As such, the reflexive prefix encodes both reflexive and reciprocal meanings. The data presented and analyzed in this chapter show that there are some remnants of the Proto-Bantu reciprocal suffix *-*an*- in a very few list of verbs encoding inherent reciprocal events, suggesting that this suffix was productive at some point in the history of the Hehe language. The analysis of the development from reflexive to reciprocal marker follows the three stages of grammaticalization theory proposed by Heine (1993) and applied in the analysis of German reflexive and reciprocal constructions by Heine & Narrog (2009). Following Ngwasi (2021), it is shown in this chapter that, unlike German, Hehe attests a fourth stage in the grammaticalization from reflexive to reciprocal marker. The fourth stage is evidenced by the recruitment of the reflexive prefix encoding events, such as chaining and associative, that are closer to the reciprocal prototype.

1 Introduction

Many Bantu languages distinguish two morphemes for encoding reflexive and reciprocal events in terms of their forms and their morphological distribution (see Meeussen 1967, Schadeberg 2003, Schadeberg & Bostoen 2019). The reflexive events are most often encoded by a reflexive prefix that occurs in the OM slot, located immediately before the verb root in the morphological structure of the verb. The reflexive prefix's shape can be a single vowel, such as -*i*- in Hehe

Lengson Ngwasi & Abel Mreta. 2024. The historical development of the reflexive-reciprocal polysemy in Hehe. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 351–371. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663783

(G62) (Msamba 2013, Ngwasi 2016, 2021), or CV, such as -*ki*- in Kagulu (G12) (Petzell 2008) and some languages of the Kikongo Language Cluster (KLC) (Dom & Kulikov 2019). In turn, reciprocal events are predominantly encoded by a suffix which has the form -*an*-, or a compound form involving -*an*-, e.g., *angan*- or -*asan*- , as in the KLC (Bostoen et al. 2015); or in Runyambo (JE21) (Rugemalira 1993). Interestingly, Polak (1983) (see also Marlo 2015) notes that in some Bantu languages, the reflexive and reciprocal events are encoded by the same verbal morpheme. On the one hand, there are languages where the reflexive prefix has been recruited to encode reciprocal events (e.g., Bolia (C35b), Chokwe (K11), Ganda (JE15), Lunda (L52) etc.), whereas, in other languages, the reciprocal suffix encodes reflexive events (e.g., Ewondo (A72) and Tsogo (B31)) (Marlo 2015). The first case is more widespread while the second case is extremely rare in Bantu languages. This chapter focuses on the first case by describing how the reflexive prefix has developed to encode reciprocal events, thus being polysemous in Hehe, particularly in Dzungwa[1] dialect. The analysis for the development from reflexive to reciprocal follows the three stages of grammaticalization theory proposed by Heine (1993) and applied by Heine & Narrog (2009) in the grammaticalization of the reflexive marker to reciprocal marker in German. I will add the fourth stage that has not been applied by Heine & Narrog (2009) in the analysis of the grammaticalization from reflexive to reciprocal marker in German, but this stage has been applied by Ngwasi (2021). It will be argued that the reflexive-reciprocal marker in Hehe is a result of the grammaticalization process leading from reflexive to reciprocal marker, taking over the role of the reflex of the Proto-Bantu reciprocal suffix *-*an*- which is no longer productive in this language. The remnants of the Proto-Bantu reciprocal suffix *-*an*- are found in a very few verbs encoding inherent reciprocal events.

Before introducing the language under study, we first define the terms used in this chapter, which are: (i) prototypical reflexive event or situation, (ii) prototypical reciprocal event or situation, and (iii) inherent reciprocal event or situation. By a prototypical reflexive event, we refer to a two-participant event type where the agent and the patient/theme refer to the same participant (see Faltz 1985, Haspelmath forthcoming, Kemmer 1993). In other words, as Moyse-Faurie (2008: 107) points out, prototypical reflexive events express actions or events that one usually performs on other entities being performed on oneself, as exemplified in (1).

[1]The data presented in this chapter were collected at Bomalang'ombe village, one of the villages where the Dzungwa dialect is spoken. Data from the other dialect come from Msamba (2013). In the rest of this chapter, we will use the term Hehe or Dzungwa when referring to the Dzungwa dialect, and where data are cited from the other dialect, we will use the term "Standard" Hehe.

(1) John hit **himself** with a hammer.

A prototypical reciprocal event encodes a similar or symmetric relation between two participants A and B, where A acts on B and B acts on A (see Haspelmath 2007, Kemmer 1993, König & Kokutani 2006). It should be noted that the prototypical reciprocal events are neither necessarily nor very frequently semantically reciprocal (see Dom et al. forthcoming, Haiman 1983). And they include events such as 'punching each other', 'seeing each other', 'hitting each other', 'cutting each other', 'killing each other', etc. The example in (2) illustrates the construction encoding a prototypical reciprocal event for English, where John and Bill are mutually involved in the punching action.

(2) John and Bill punched **each other**.

An inherent or natural reciprocal event is an event type that necessarily or very frequently expresses reciprocal situations (see Kemmer 1993, König & Kokutani 2006, Nedjalkov 2007). Kemmer (1993: 104) lists verbs which encode inherent reciprocal events cross-linguistically, which are: verbs of antagonistic actions ('fight', 'quarrel', 'wrestle'), verbs of affectionate actions ('kiss', 'embrace', 'make love'), verbs of encountering and associations ('meet', 'greet', 'shake hands'), verbs of actions denoting unintentional physical contact ('bump into', 'collide'), verbs of physical convergence or proximity ('touch', 'join', 'unite', 'be close together'), verbs of exchanging ('trade', 'share', 'divide', 'split'), verbs of agreement/disagreement ('converse', 'argue', 'gossip', 'correspond'), and verbs of similarity/dissimilarity ('resemble'). The examples in (3) illustrate a construction encoding inherent reciprocal event for English.

(3) John and Bill **met**.

With this brief introduction of reflexive and reciprocal events, we turn to the introduction of Hehe language. Hehe is spoken mainly in the Iringa region of Tanzania. It is classified as G62 by Guthrie (1948, 1967–1971) and Maho (2009), and it is closely related to other G60 languages, such as Sangu (G61), Bena (G63), Pangwa (G64), Kinga (G65), Wanji (G66), and Kisi (G67). Hehe was reported to have approximately 598,839 native speakers by LOT (2009), but recently, Ethnologue Languages of the World reports the number of native speakers to be approximately 1,210,000, as of 2016 (Eberhard et al. 2020). In terms of dialects, there is no agreement among scholars on the number of dialects of Hehe. For instance, Madumula (1995) identifies five dialects, Mpalanzi (2010) identifies three dialects, while Haonga (2013) identifies two dialects called "Standard" Hehe and

Dzungwa (also called Tsungwa by its native speakers). We follow Haonga's (2013) analysis of the dialectal variation of Hehe since it is the only source that is solely based on linguistic evidence, i.e., phonological, morphological, syntactic, and semantic evidence. As already noted above, this chapter focuses on the Dzungwa dialect with sporadic reference to the other dialect, the so-called "Standard" dialect, where the data are accessible.

Like many other Bantu languages, Hehe is "verby" in the sense that the verb root can be attached with several morphemes for various inflectional and derivational functions (see Nurse 2008). The structure of Hehe verbs can be demonstrated by examples (4)–(7), elaborating the templatic structure shown in Table 1 below, as extracted from Ngwasi (2016).

Table 1: The structure of Hehe verbs (Ngwasi 2016: 50)

1	2	3	4	5	6	7	8	9	10	11
REL	NEG1	SM	NEG2	TAM	OM/REFL-REC	VR	EXT	PFV	FV	CLIT

(4) *yesiakutsági*
ye-si-a-ku-ts-ág-i
REL-NEG1-SM1-TAM-come-HAB-FV
'S/he who does not normally come.'

(5) *witóve*
u-i-tóv-e
SM1-REFL-beat-IMP/SBJ
'Beat yourself'

(6) *alakulimítsa*
a-la-ku-lim-íts-a
SM1-NEG2.SBJ-TAM-cultivate-CAUS-FV
'S/he should not make you cultivate.'

(7) *vaitseengíte*[2]
va-i-tseeng-íte
SM2-OM9-build-PFV
'They have built it.'

[2]It should be noted that the class 9 object prefix *-i-*, unlike the reflexive prefix *-i-*, does not trigger the deletion of the vowel *a* of the subject marker *va-* (see Ngwasi (2016) on vowel deletion and glide formation triggered by the reflexive prefix in Hehe).

As can be seen from Table 1, the productive reciprocal marker occupies slot 6, the slot for OM and reflexive markers in many Bantu languages. As will be argued later in this chapter, the reciprocal marker occupies this slot as a result of the historical development (grammaticalization) whereby the reflexive prefix has undergone grammaticalization and has taken over the role of the reflex of the Proto-Bantu reciprocal suffix *-*an*-. As such, both reflexive and reciprocal meanings are productively expressed by the same morpheme, the reflexive prefix -*i*-, occupying the OM slot, as will be discussed further in §2.

The remainder of this chapter is divided into four sections. §2 provides an overview of the construction types where the Hehe reflexive prefix has various functions, particularly those encoding reflexive and reciprocal events. §3 introduces grammaticalization theory and discusses the rise of the reflexive-reciprocal polysemy in Hehe, as explained from a grammaticalization perspective. §4 briefly highlights the loss of the reflex of the Proto-Bantu reciprocal suffix *-*an*- in Hehe and the emergence of the reflexive prefix -*i*- as a new means of encoding reciprocal events. §5 concludes the discussion.

2 Reflexive-reciprocal polysemy: An overview of construction types

This section describes various constructions where the reflexive prefix -*i*- encodes exclusive reflexive events, ambiguous reflexive-reciprocal events, and exclusive reciprocal events. The construction types which we focus on in this section are infinitive constructions (§2.1), constructions with singular subjects (§2.2), and constructions with plural subjects with ambiguous reflexive-reciprocal interpretation and those with exclusive reciprocal interpretation (§2.3).

2.1 Infinitive constructions

The reflexive prefix -*i*- encodes ambiguous reflexive-reciprocal meanings or exclusively reciprocal meaning in infinitive constructions, as can be exemplified by the examples in (8) and (9). As shown in (8), the reflexive prefix -*i*- has an ambiguous reflexive-reciprocal interpretation. This is because the verb in this construction is neither necessarily nor frequently semantically reciprocal, while in (9), the reflexive prefix -*i*- has only a reciprocal interpretation because the verb is semantically reciprocal (see Nedjalkov 2007 for an overview of cross-linguistic encoding of inherent reciprocal events).

(8) *kw**i**bumíla*
kú-**i**-bumíl-a
INF-REFL-REC-hit-FV
'to hit oneself' or 'to hit each other'

(9) *kw**i**huungíla*
kú-**i**-huungíl-a
INF-REC-greet-FVFV
'to greet each other'

2.2 Constructions with singular subjects

The reflexive prefix *-i-* renders only reflexive meaning with constructions having singular subjects in Hehe, as exemplified in (10). In fact, the constructions with singular subjects have only reflexive meaning because the plurality of the participants, which is a key defining property of constructions encoding reciprocal events, is not available. As already noted above in §1, the definition of reciprocal events or situations requires plural participants (see Frajzyngier 2000, Heine & Miyashita 2008, Lichtenberk 2000).

(10) *Juma ak**i**bumyé*[3]
Juma a-ka-**i**-bumíl-íle
Juma SM1-PST-REFL-hit-PFV
'Juma hit himself.'

It should be noted that constructions with singular subjects can optionally occur with emphatic reflexive pronouns for emphasis in Hehe, as illustrated in (11). The emphatic reflexive pronouns, just like in English, can also follow the subject NP it emphasises.

(11) *Juma ak**i**bumyé **yimwene***
Juma a-ka-**i**-bumíl-íle **yimwene**
Juma SM1-PST-REFL-hit-PFV EMPH
'Juma hit himself.'

[3]It should be noted that the perfective suffix *-íle* triggers imbrication with some verb roots or stems in Hehe, as can be seen in example (10). See Bastin (1983) and Hyman (1995) for a detailed discussion on imbrication in Bantu.

2.3 Constructions with plural subjects

The reflexive prefix *-i-* encodes ambiguous reflexive-reciprocal meaning in constructions with plural subjects and verbs which do not trigger inherent reciprocal interpretation. This is unlike the constructions with singular subjects discussed in §2.2 which have reflexive interpretation only. This is illustrated by the example in (12), where the reflexive prefix *-i-* has an ambiguous reflexive-reciprocal interpretation.

(12) *Kiliani na Naftali vak**i**bumyé*
Kiliani na Naftali va-ka-**i**-bumíl-íle
Kiliani COM Naftali SM2-PST-REFL-REC-hit-PFV
'Kiliani and Naftali hit each other.' or 'Kiliani and Naftali hit themselves.'

Hehe speakers use emphatic pronouns to remove this ambiguity and rule out a reciprocal interpretation in favour of the reflexive interpretation in constructions with plural subjects, as (13) exemplifies. If the intended meaning is the reciprocal interpretation, the speakers can employ discontinuous reciprocal constructions to rule out the reflexive interpretation, as can be seen in (14). It should be noted that in discontinuous reciprocal constructions, one of the two participants follows a verb and is introduced by a comitative preposition (see Dimitriadis 2004, 2008, Haspelmath 2007), i.e., the comitative *na* in the case of Hehe in (14). The example (13) shows that the emphatic reflexive pronoun functions as the disambiguator for the reflexive interpretation just as in other languages cross-linguistically, such as French *eux-mêmes* and German *sich* (see Cable (2014) for examples), while the discontinuous reciprocal construction in (14) functions as a disambiguation strategy for the reciprocal interpretation (see Dimitriadis 2004, Seidl & Dimitriadis 2003 for this disambiguation strategy in Swahili (G42) and German).

(13) *Kiliani na Naftali vak**i**bumyé* ***vavene***
Kiliani na Naftali va-ka-**i**-bumíl-íle **vavene**
Kiliani COM Naftali SM2-PST-REFL-hit-PFV EMPH.REFL
'Kiliani and Naftali hit themselves.'

(14) *Kiliani ak**i**bumyé* ***na Naftali***
Kiliani a-ka-**i**-bumíl-íle **na Naftali**
Kiliani SM1-PST-REC-hit-PFV COM Naftali
'Kiliani and Naftali hit each other.'

It is important to note that, unlike some Bantu languages where both singular and plural subject markers are acceptable in the case of discontinuous reciprocal constructions (cf. Mwera (P22) and Cilubà (L31a), see Bostoen et al. 2015: 763–764, Schadeberg & Bostoen 2019: 183), in Hehe, the subject marker continues to show singular agreement with the remaining lexical NP in the subject position.

The reflexive prefix *-i-* is also used as a productive means of encoding inherent reciprocal events in constructions with plural subjects and verbs that are semantically frequently or necessarily reciprocal. The example (15) illustrates the reflexive prefix *-i-* encoding inherent reciprocal events with plural subject NPs.

(15) *Juma na Ali vak**i**húunje*
Juma na Ali va-ka-**i**-huungíl-íle
Juma COM Ali SM2-PST-REC-greet-PFV
'Juma and Ali greeted each other.'

There are some verbs which trigger inherently reciprocal interpretation that have retained the reflex of the Proto-Bantu reciprocal suffix **-an-*. Such verbs are listed in (16) below. It should be noted that the verbs *-leka* 'leave/abandon' and *-hwaana* 'resemble' also take the reflexive prefix *-i-* in the synchronic state of the language. In addition, the verb *-hwaana* does not occur without the reciprocal suffix **-an-*. As such, the reciprocal suffix *-an-* is fossilized (has become part of the verb stem) in this verb. The same fossilized reciprocal suffix is observed on the verb *-taang'ána* 'meet', which also requires the reflexive prefix to be present in order to express reciprocity. Thus it is the reflexive prefix that encodes reciprocal meaning in such cases.

(16) Verbs with the reflex of the Proto-Bantu reciprocal suffix **-an-*

*kúgav**á**na*	*kú-gav-án-a*	'to share'	< *kúgava*	'to distribute'
*kúlek**á**na*	*kú-lek-án-a*	'to divorce'	< *kúleka*	'to leave/ abandon'
*kúhwa**á**na*	*kú-hwaán-a*	'to resemble'		
*kúloong**á**na*	*kú-loong-án-a*	'to chat'	< *kúloonga*	'to talk'
*kw**í**taang'**an**a*	*kú-i-taang'án-a*	'to meet'		

Besides encoding inherent reciprocal events, the reflexive prefix *-i-* also encodes other events such as chaining reciprocal and associativity. By chaining reciprocal events, following Kemmer (1993), we refer to events that involve an ordered sequence or series of participants who are in a certain relation, while associative events refer to events or actions that are carried out jointly. The following examples in (17) and (18) illustrate constructions encoding chaining and associative events, respectively.

(17) *Avanyashule vak**i**fwaatíte*
a-va-nyashule va-ka-**i**-fwaat-íte
AUG-2-student SM2-PST2-REC-follow-PFV
'The students followed each other.'

(18) *Avanu vakitaanzíle*
a-va-nu va-ka-**i**-taang-íle
AUG-2-person SM2-PST-REC-help/do-PFV
'The people did together.' (lit. 'The people worked together')

In general, we can conclude that the reflexive prefix *-i-* is a productive means of encoding reflexive, reciprocal, chaining, and associative events in Hehe. Within the domain of reciprocal events, it is used to encode both prototypical and inherent reciprocal events, as well as other events related to the reciprocal prototype, such as chaining and associative events. Having described the way, the reflexive prefix *-i-* encodes these events, we turn to the discussion of the historical development of the reflexive-reciprocal polysemy in Hehe in §3.

3 The historical development of the reflexive-reciprocal polysemy in Hehe: A grammaticalization perspective

This section discusses the historical development of the reflexive-reciprocal polysemy in Hehe by applying Heine's (1993) Overlap Model on the grammaticalization from reflexive to reciprocal markers. This Model has been applied in other languages, particularly German by Heine & Miyashita (2008), Heine & Narrog (2009) who examine the grammaticalization from the reflexive pronoun *sich* to reciprocal marker. The Model presupposes three synchronic stages, reflecting a historical development leading from reflexive to reciprocal marker. In addition to the three stages of the Model, following Heine (2002) (see also Heine & Kuteva 2007), we add the fourth stage called *conventionalization*. Before discussing the stages of grammaticalization from reflexive to reciprocal marker, we briefly define grammaticalization and the mechanisms of change by providing examples from other languages on the grammaticalization from reflexive to reciprocal marker.

3.1 Grammaticalization and its parameters

The term "grammaticalization" has been used in linguistics in two ways. First, it is used to refer to a process of language change. Second, it is used to refer to

the theoretical framework that is used to account for the processes of language change (see Campbell & Janda 2001, Heine 2003, Heine & Narrog 2009). According to Croft (2006), to understand what grammaticalization means, we need to understand first the processes that create the grammar of a particular language. In general, as Heine & Narrog (2009) define it, grammaticalization is a process in which lexical items become grammatical items, or grammatical items become more grammatical. From this definition, there are two types of grammaticalization. First, there is primary grammaticalization which involves a change from lexical to grammatical items. Second, secondary grammaticalization which involves a change from already grammatical(ized) items to more grammatical ones. This chapter is based on secondary grammaticalization because there is no lexical source reconstructable for Proto-Bantu as a source of the reflexive marker in Bantu.

As for grammaticalization as a theoretical framework, it is meant to explain what causes grammaticalization, and how grammatical or more grammatical categories are developed and structured in languages (Heine 2003, Heine & Kuteva 2007). Thus, it is an explanatory tool for the grammaticalization phenomenon.

As a process of change, grammaticalization involves four parameters, namely: "extension", "desemanticization", "decategorialization", and "erosion". Although these parameters are mainly associated with primary grammaticalization, they are worth exploring because they have been used to explain the grammaticalization from reflexive to reciprocal where lexical sources are attested. They can also equally be used with secondary grammaticalization in many respects. Each of these parameters is explained in the following paragraphs as applied in the grammaticalization from reflexive to reciprocal marker in other non-Bantu languages, in particular, German.

The first parameter, *extension*, involves the rise of new grammatical meanings for a particular form, especially in a new context (semantic component). This is to say, the linguistic item with its meaning receives a new meaning in another context (context-induced reinterpretation) (see Heine 2002, Heine & Dunham 2010, Heine & Kuteva 2007). It also involves the extension of the use of a linguistic item in its usual or primary context to a new set of context(s) (Heine & Dunham 2010), such that it is no longer limited to a particular defined context (text-pragmatic component). As Heine & Kuteva (2007) argue, all these come out due to some sociolinguistic component whereby speakers, usually a group, start employing a new usage or meaning of the existing linguistic item, and later on adopted by the entire speech community. The German reflexive pronoun *sich*, for example, was extended to encode reciprocal events in constructions with plural subjects or antecedents (see Heine & Miyashita 2008, Heine & Narrog 2009). As can be seen

in (19), the reflexive pronoun *sich* has a reflexive meaning only, but in (20), it is reinterpreted as encoding also reciprocal meaning since the context – the plurality of the participants – leads to its reinterpretation as a reciprocal marker while maintaining its source meaning, i.e., reflexive. Hence, the construction becomes ambiguous between the source meaning and the new meaning.

(19) *Er wusch* ***sich***
He wash.PST REFL
'He washed (himself).'

(20) *Sie wuschen* ***sich***
They wash.PST.PL REFL-REC
'They washed themselves.'
'They washed each other.' (Heine & Narrog 2009: 410)

The second parameter, *desemanticization* or *semantic bleaching*, refers to the process whereby a linguistic item loses its old or source meaning or use due to the reinterpretation in the new context of use (Heine & Dunham 2010, Heine & Narrog 2009). This parameter follows from extension because the extended linguistic item may lose part of its primary meaning in specific contexts. With respect to the reflexive pronoun *sich* in German, it loses the reflexive meaning when used with verbs that trigger inherent reciprocal interpretation, as exemplified in (21).

(21) *Sie küssten* ***sich***
They kiss.PST.PL REFL
'They kissed (each other).' (Heine & Narrog 2009: 410)

The third parameter, *decategorialization*, involves the loss of the morphosyntactic characteristics of the linguistic item after being desemanticized. This means that the morphosyntactic properties which the linguistic item had before its extension and desemanticization are no longer available in the new usage context. This may include, among others: (i) Loss of ability to be inflected; (ii) Loss of ability to take on the derivational morphology; (iii) Loss of ability to take modifiers; (iv) Loss of independence as an autonomous linguistic item, leading to an increased dependence on some other linguistic item; (v) Loss of syntactic freedom of a linguistic item, such as, the ability to be moved in a sentence; (vi) Loss of ability to be referred to anaphorically; and (vii) Loss of membership to a grammatical paradigm (see Heine 2003, Heine & Dunham 2010). In addition to these, Heine & Miyashita (2008) attribute decategorialization to a limited set of contexts, both

syntactic and pragmatic, where the grammaticalized item can occur. According to Heine & Miyashita (2008: 196–197), the most widespread decategorialization involving reflexive markers that become reciprocal markers is the constraint on the category of number. In other words, the reflexive-reciprocal marker, when used to encode reciprocal meaning, becomes restricted to "a smaller set of syntactic and pragmatic contexts" compared to when it is used as a reflexive marker. As such, the reciprocal interpretation is restricted to constructions with plural subjects or antecedents only. This is to say, for example, the pronoun *sich* in German can only be interpreted as encoding reciprocal meaning with plural subjects. In contrast, with singular subjects, it continues to encode reflexive meaning.

The fourth parameter, *erosion* or *phonetic reduction*, refers to the loss in phonetic substance of the linguistic item undergoing a change in grammaticalization. This may involve the loss of an entire syllable, phonetic simplification, loss of phonetic autonomy as well as the adaptation to adjacent phonetic units, or loss of segmental properties such as stress, tone, or intonation (see Heine & Dunham 2010, Heine & Miyashita 2008, Heine & Narrog 2009). For the German reflexive pronoun *sich*, Heine & Narrog (2009) argue that it loses the stress that it bears when encoding reflexive events when used to encode reciprocal events.

It is argued that grammaticalization is a continuous process or a "chain-like" development in the sense that it usually follows the parameters from extension to phonetic reduction (see Heine 2000, Heine & Kuteva 2007, Heine & Narrog 2009). However, it should be noted that the grammaticalization process can stop at any point of development, and it does not necessarily replace older linguistic forms expressing the same grammatical meaning (see Heine 2000, Heine & Kuteva 2007, Hopper 1991). With this note on the mechanisms or parameters of grammaticalization, we turn to the grammaticalization of the reflexive prefix *-i-* in Hehe in §3.2.

3.2 The stages of grammaticalization of the reflexive prefix *-i-* in Hehe

The four stages of grammaticalization involving the reflexive markers mentioned at the beginning of §3 are explained in this subsection with reference to the data presented in §2. The data presented in §2 where the reflexive prefix *-i-* has other functions are analyzed from the grammaticalization perspective.

The first stage (stage I) in this Model is called the "Initial stage". In this stage, as Heine (2002) argues, the linguistic item has its original meaning, and it is not restricted in terms of contexts where it can occur. This stage in Hehe, in the synchronic state of the language, is represented by the constructions with reflexive interpretation only (those with singular subjects), but it can be hypothesized that

before its grammaticalization to reciprocal marker, it was not restricted to constructions with singular subjects. This stage is illustrated by the constructions with singular subjects and the verbs that do not trigger inherent reciprocal interpretation, as in the example (10) above.

The second stage (stage II) is called "bridging context". The linguistic item gets reinterpreted with reference to the source meaning and the new target meaning. Thus, the linguistic item becomes ambiguous. For Hehe, this stage is represented by constructions with an ambiguous reflexive-reciprocal interpretation, i.e., the constructions with plural subjects and the constructions with infinitive prefixes with verbs of prototypical two-participant event verbs, as in the examples (8) and (12) above. In fact, the constructions in this stage differ from the constructions in stage I in that the subjects in these constructions are plural, and the constructions with an infinitive prefix. This stage is an intermediate stage for grammaticalization from reflexive to reciprocal marker. As examples (8) and (12) show, the constructions are simultaneously interpreted with reference to the source or original meaning (reflexivity) and the target or new meaning (reciprocity).

The third stage (stage III) is called "switch context". In this stage, the linguistic item is interpreted with the new or target meaning only (Heine 2002: 85). In other words, the source meaning is no longer accessible. This stage in Hehe is represented by constructions with plural subjects (and infinitive constructions), just like the ones in stage II, but the difference is based on the type of verbs used at this stage. Unlike the verbs used at stage II, the verbs used at stage III constructions trigger an inherently reciprocal interpretation with the reflexive prefix *-i-*. In switch contexts, the target function or meaning, encoding reciprocal in this case, is the only available interpretation. In other words, there is no source function at this stage (the reflexive function of the prefix *-i-* is excluded at this stage). So, the reflexive interpretation of the reflexive prefix *-i-* is infelicitous in stage III. It is inappropriate for the examples (9) and (15) above to mean 'to greet oneself', or 'Juma and Ali greeted themselves'. The only appropriate interpretation of this construction is reciprocal, i.e., 'to greet each other' or 'Juma and Ali greeted each other'.

The fourth stage (stage IV) is called the "conventionalization stage". In this stage, as Heine (2002: 86) argues, the linguistic item may be used in other new contexts because it is no longer restricted to its source function. In Hehe, the reflexive prefix *-i-* is also recruited to encode chaining and collective or associative events, apart from encoding prototypical and inherent reciprocal events, as we have already seen in examples (17) and (18) above. This is because the language speakers have conventionalized it to be their new means of encoding reciprocal events. Thus, it is also extended to encode other less core reciprocal functions of

the reflex of the Proto-Bantu reciprocal suffix *-*an*-, in particular the constructions with verbs that trigger chaining and associative reciprocal interpretation.

The four stages of grammaticalization of the reflexive prefix -*i*- in Hehe are summarized in Table 2 below, following Heine (2002).

Table 2: The stages of grammaticalization from reflexive to reciprocal of the reflexive prefix -*i*-

Stage	Context	Resulting interpretation
I. Initial	Not restricted	Reflexive
II. Bridging context	Plural subjects/Infinitive prefix, prototypical two-participant event verbs	Reflexive-reciprocal
III. Switch context	Plural subjects/Infinitive prefix, verbs resulting to inherent reciprocal interpretation	Reciprocal
IV. Conventionalization	Plural subjects/Infinitive prefix, verbs resulting to chaining reciprocal interpretation, and associative interpretation	Chaining reciprocal, associative interpretation

4 The loss of the reflex of the Proto-Bantu reciprocal suffix *-*an*- and the emergence of the reflexive-reciprocal polysemy

A number of facts indicate that in an earlier stage, Hehe conformed to the common Bantu situation, in that it had the reflexive prefix for encoding reflexive events and the reciprocal suffix for encoding reciprocal events. First, the fact that there are some verbs with the reflex of the Proto-Bantu reciprocal suffix *-*an*-, as shown in (16) above, is a piece of evidence that the reciprocal suffix was a productive reciprocal marker in Hehe. Second, Msamba (2013) argues that while most speakers of "Standard" Hehe prefer to use the reflexive prefix -*i*- instead of the reflex of the Proto-Bantu reciprocal suffix *-*an*- (the suffix -*an*-) to express reciprocity, a few speakers, especially elders, still use the reciprocal suffix with some verbs, as shown in Table 3 below. This indicates that even in this dialect,

the reflexive prefix is becoming conventionalized as a productive means of encoding reciprocal events, replacing the reflex of the Proto-Bantu reciprocal suffix *-*an*-.

Table 3: The coexistence of the reflexive prefix and reciprocal suffix in encoding reciprocal events in "Standard" Hehe (Msamba 2013: 59)

Verb stem	Gloss	Reflexive	Gloss	Reciprocal	Gloss
-*tova*	'beat'	-***i***tova	'beat oneself'	-tov***an***a	'beat each other'
-*heka*	'laugh'	-***i***heka	'laugh at oneself'	-hek***an***a	'laugh at each other'
-*kwega*	'pull'	-***i***kwega	'pull oneself'	-kweg***an***a	'pull each other'
-*homba*	'pay'	-***i***homba	'pay oneself'	-homb***an***a	'pay each other'

It is important to note that the grammaticalization from reflexive to reciprocal described in this chapter and summarized in Table 2 for Hehe should be regarded as a means of creating a new grammatical item for encoding reciprocal events, taking over the role of the reflex of the Proto-Bantu reciprocal suffix *-*an*-. It has been hypothesized by Schladt (1998) that the Proto-Bantu reciprocal suffix *-*an*- developed from the comitative marker *na*. He argues that the development from the comitative marker to reciprocal suffix resulted from a serial construction following the grammaticalization chain: V-a *na* > V-a-*na* > V-*an*-a (note: V-a stands for the verb root + the default final vowel). This hypothesis has been adopted in other work on Bantu languages, i.e., Schadeberg & Bostoen (2019) and Bostoen et al. (2015).[4]

The fact that there is evidence for the existence of the reflex of the Proto-Bantu reciprocal suffix *-*an*- in Hehe, means that it can be concluded that this suffix went through the grammaticalization chain hypothesized by Schladt (1998) before it fell out of favour by Hehe speakers.

According to Hopper (1991: 22–23), when a linguistic item is taking over the functional role of another linguistic item, it is expected that the new item and the

[4]Another suffix (verb extension) that has been hypothesized to have developed from a lexical source is the extensive suffix *-*al*-. According to Schadeberg (2003), it is from the lexical item *-*jal*- 'to spread'. For other verb extensions, there is no suggested lexical sources (see Schadeberg 2003 and Schadeberg & Bostoen 2019: 174).

old item may coexist for a certain period. This means that the new linguistic item does not immediately replace an already existing item. In the case of the grammaticalization of the reflexive prefix *-i-* in Hehe, the verbs where the reciprocal suffix *-an-* is still existent synchronically (cf. example (16)) offer evidence of the coexistence stage in the history of the language. In addition to the examples in (16), the data in Table 3 from Msamba (2013: 59) showing the coexistence of the reflexive prefix *-i-* and the reciprocal suffix *-an-* especially in the speech of elders in "Standard" Hehe illustrate this phenomenon. Similar coexistence has been reported by Morrison (2011: 249) for Bena (G63) (a language which is spoken in close geographic proximity to Hehe), as can be exemplified in (22).

(22) Bena (G63) (Morrison 2011: 249)

a. *Tuhwíwona*
tu-hu-**i**-won-a
SM2-E-REFL-REC-see-FV
'We see each other/We see ourselves.'

b. *Twíwonana*
tu-i-won-**an**-a
SM2-PRS-see-REC-FV
'We see each other.'

In general, the coexistence of the reflexive prefix *-i-* and the reciprocal suffix *-an-* in encoding reciprocal events provides evidence that the reciprocal suffix *-an-* had been productively used as a reciprocal marker, and the reflexive prefix *-i-* is now taking over the role of the reciprocal suffix in Bena. The reflexive prefix might ultimately be the only productive means of encoding reciprocal events as has happened in Hehe.

5 Conclusion

Based on the Hehe data presented and analyzed in this chapter, it is evident that the reflexive prefix *-i-* has developed from being a dedicated reflexive marker into a polysemous marker encoding both reflexive and reciprocal events. We have argued that the various present-day uses of the reflexive prefix can be interpreted as distinct stages illustrating the diachronic grammaticalization process leading from a prototypical reflexive marker to a reciprocal marker. The reflex of the Proto-Bantu reciprocal suffix **-an-*, which occurs throughout Bantu languages as a productive reciprocal marker, is still found with some verbs encoding inherent reciprocal events. However, the grammaticalization of the reflexive prefix *-i-*

is becoming dominant to such an extent that it is also used and preferred with some of these archaic lexicalized reciprocal verbs (e.g., *kwítaang'ána* 'to meet'). Finally, we have shown that the reflexive prefix *-i-* after grammaticalizing and becoming a new productive means of encoding reciprocal events has been extended to encode chaining and associative events, the events which are closer to the reciprocal prototype. These two events are also encoded by the reflexes of the Proto-Bantu reciprocal suffix **-an-* in the languages where the reciprocal suffix is still productive.

Acknowledgements

We thank our language consultants, who dedicated their time to us for data elicitation. We also appreciate the comments and insights from the reviewers that helped us reach this final version. This work could not have been possible without their thoughts and insights. Lastly, we appreciate the work by the editors of this volume from the beginning of this project until its completion.

Abbreviations and symbols

AUG	Augment	OM	Object marker
CL	Noun class	PFV	Perfective
CLIT	Clitic	PL	Plural
COM	Comitative	REC	Reciprocal
E	Epenthetic morpheme	REFL	Reflexive
EMPH.REFL	Emphatic pronoun	REL	Relative
EXT	Verb extension	SG	Singular
FV	Final vowel	SM	Subject Marker
INF	Infinitive	TAM	Tense, Aspect, Mood
NEG	Negative marker	>	to

References

Bastin, Yvonne. 1983. *La finale verbale -ide et l'imbrication en bantou* (Serie in-80, Sciences Humaines 114). Tervuren: Annales du Musée Royal de l'Afrique Centrale.

Bostoen, Koen, Sebastian Dom & Guillame Segerer. 2015. The antipassive in Bantu. *Linguistics* 53(4). 731–772. DOI: 10.1515/ling-2015-0016.

Cable, Seth. 2014. Reflexives, reciprocals and contrast. *Journal of Semantics* 31. 1–41.

Campbell, Lyle & Richard Janda. 2001. Introduction: Conceptions of grammaticalization and their problems. *Language sciences* 23. 93–112.

Croft, William. 2006. Evolutionary models and functional-typological theories of language change. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 68–91. Malden: Blackwell.

Dimitriadis, Alexis. 2004. *Discontinuous reciprocals*. Ms., Utrecht Institute of Linguistics, Utrecht. https://staticweb.hum.uu.nl/medewerkers/alexis.dimitriadis/papers/discon-long-ms04.pdf.

Dimitriadis, Alexis. 2008. Irreducible symmetry in reciprocal constructions. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 375–410. Berlin: Mouton de Gruyter.

Dom, Sebastian, Heidi Goes & Koen Bostoen. Forthcoming. Multiple reciprocity marking in the Kikongo language cluster: Functional distribution and origins. In Eva-Marie Bloom Ström, Hannah Gibson, Rozenn Guérois & Lutz Marten (eds.), *Current approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press.

Dom, Sebastian & Leonid Kulikov. 2019. *A comparative case-study of the Kikongo language cluster*. Paper presented at the 52 Annual Meeting of the Societas Linguistica Europaea, Leipzig University (German). Bantu reflexives.

Eberhard, David M., Simons F. Gary & Charles D. Fenning (eds.). 2020. *Ethnologue: Languages of the world*. 23rd edn. Dallas, TX: SIL International. http://www.ethnologue.com.

Faltz, Leonard. 1985. *Reflexivization: A study in universal syntax*. New York: Garland.

Frajzyngier, Zygmunt. 2000. Coding of the reciprocal function: Two solutions. In Zygmunt Frajzyngier & S. Curl Traci (eds.), *Reciprocals: Forms and functions*, vol. 41, 179–194. Amsterdam: John Benjamins.

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

Guthrie, Malcolm. 1967–1971. *Comparative Bantu: An introduction to the comparative and prehistory of the Bantu languages*, vol. 1–4. London: Gregg International.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819.

Haonga, Ernest. 2013. *Lahaja za Kihehe: Mtazamo wa kiisimu*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Haspelmath, Martin. 2007. Further remarks on reciprocal constructions. In P. Vladimir Nedjalkov (ed.), *Reciprocal constructions*, 2087–2115. Amsterdam: John Benjamins.

Haspelmath, Martin. Forthcoming. Comparing reflexive constructions in the world's languages. In Katarzyna Janic, Nicoletta Puddu & Martin Haspelmath (eds.), *Reflexive constructions in the world's languages*. Berlin: Language Science Press.

Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. Oxford: Oxford University Press.

Heine, Bernd. 2000. Polysemy involving reflexive and reciprocal markers in African languages. In Zygmunt Frajzyngier & S. Curl Traci (eds.), *Reciprocals: Forms and functions*, vol. 41, 1–30. Amsterdam: John Benjamins.

Heine, Bernd. 2002. On the role of context in grammaticalization. *Typological studies in language* 49. 83–102.

Heine, Bernd. 2003. Grammaticalization. In Joseph D. Brain & Janda D. Richard (eds.), *The handbook of historical linguistics*, 575–601. Oxford: Blackwell.

Heine, Bernd & Margaret Dunham. 2010. Grammaticalization in Bantu languages with special reference to Swahili. In Karsten Legère & Christina Thornell (eds.), *Bantu languages*, 31–45. Cologne: Rüdiger Köppe.

Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar: A reconstruction*, vol. 9. Oxford: Oxford University Press.

Heine, Bernd & Hiroyuki Miyashita. 2008. The intersection between reflexives and reciprocals: A grammaticalization perspective. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, vol. 192, 169–224. Berlin: Mouton de Gruyter.

Heine, Bernd & Heiko Narrog. 2009. Grammaticalization and linguistic analysis. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 401–423. Oxford: Oxford University Press.

Hopper, Paul J. 1991. On some principles of grammaticization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. 1, 17–35. Amsterdam: John Banjamins.

Hyman, Larry M. 1995. Minimality and prosodic morphology in Cibemba imbrication. *Journal of African Languages and Linguistics* 16. 3–39.

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

König, Ekkehard & Shigehiro Kokutani. 2006. Towards a typology of reciprocal constructions: Focus on German and Japanese. *Linguistics* 44(2). 271–302.

Lichtenberk, Frantisek. 2000. Reciprocals without reflexives. In Zygmunt Frajzyngier & S. Traci Curl (eds.), *Reciprocals: Forms and functions*, vol. 2, 31–62. Amsterdam: John Benjamins.

LOT. 2009. *Atlasi ya lugha za Tanzania*. Dar es Salaam: University of Dar es Salaam.

Madumula, Joseph. 1995. *Proverbs and sayings: Theory and practice with examples from the Wahehe of Southern Highlands of Tanzania*. Dar es Salaam: TUKI.

Maho, Jouni Filip. 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 (20 March, 2013).

Marlo, Michael R. 2015. Exceptional properties of the reflexive in Bantu languages. *Nordic Journal of African Studies* 24(1). 1–22.

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

Morrison, Michelle Elizabeth. 2011. *A reference grammar of Bena*. Houston, TX: Rice University. (Doctoral dissertation).

Moyse-Faurie, Claire. 2008. Constructions expressing middle, reflexive and reciprocal situations in some Oceanic languages. In Ekkehard König & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 105–168. Berlin: Mouton de Gruyter.

Mpalanzi, Lameck. 2010. *An analysis of tense and aspect system in Kihehe: A case of affirmative and negative construction*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Msamba, Enitha. 2013. *Object marking in Kihehe*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Nedjalkov, Vladimir P. 2007. Overview of the research: Definition of terms, framework, and related issues. In Vladimir P Nedjalkov (ed.), *Reciprocal constructions*, 3–113. Amsterdam: John Benjamins.

Ngwasi, Lengson. 2016. *Reflexive marking in Kihehe*. Dar es Salaam: University of Dar es Salaam. (MA thesis).

Ngwasi, Lengson. 2021. *The multiple functions of the reflexive prefix in Hehe, Sukuma, Nilamba, and Nyaturu*. Gothenburg: University of Gothenburg. (Doctoral dissertation).

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

Petzell, Malin. 2008. *The Kagulu language of Tanzania: Grammar, texts and vocabulary*. Köln: Rüdiger Köppe.

Polak, Louise. 1983. Le réfléchi en bantou. *Africana Linguistica* 9(1). 271–304.

Rugemalira, Josephat M. 1993. *Runyambo verb extensions and constraints on argument structure*. Berkeley, CA: University of California. (Doctoral dissertation).

Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 71–89. London & New York: Routledge.

Schadeberg, Thilo C. & Koen Bostoen. 2019. Word formation. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 2nd edn., 172–203. London: Routledge.

Schladt, Mathias. 1998. Reciprocals in Bantu languages: A case of grammaticalization. *Afrikanistische Arbeidspapieren* 53. 5–25.

Seidl, Amanda & Alexis Dimitriadis. 2003. Statives and reciprocal morphology in Swahili. *Typologie des langues d'Afrique et universaux de la grammaire* 1. 239–284.

Chapter 13

Morphosyntactic variation in Old Swahili

Lutz Marten[a], Hannah Gibson[b], Rozenn Guérois[c,d] & Kyle Jerro[b]

[a]SOAS University of London [b]University of Essex [c]LLACAN CNRS [d]University of KwaZulu-Natal

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

1 Introduction

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

Lutz Marten, Hannah Gibson, Rozenn Guérois & Kyle Jerro. 2024. Morphosyntactic variation in Old Swahili. In Hannah Gibson, Rozenn Guérois, Gastor Mapunda & Lutz Marten (eds.), *Morphosyntactic variation in East African Bantu languages: Descriptive and comparative approaches*, 373–420. Berlin: Language Science Press. DOI: 10.5281/zenodo.10663785

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

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

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

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

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

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

2 Methodological background

2.1 Methodological approach

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

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

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

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

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

2.2 Old Swahili

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

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

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

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

2.3 Standard Swahili

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

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

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

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

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

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

3 Parameters of variation

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

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

Table 2: Parameters of variation for Old and Standard Swahili

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

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

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

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

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

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

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

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

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

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

These pronominal stems mark agreement in noun class with the possessee, as shown in (3).[2]

(3) Standard Swahili possessive pronouns

a. nyumba y-ake
9.house 9-POSS3SG
'her/his house'

b. wa-toto w-ao
2-child 2-POSS3PL
'their children'

c. m-pango w-angu
3-plan 3-POSS1SG
'my plan'

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

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

(4) Standard Swahili possessive suffixes

a. dada-ke
9.sister-POSS3SG
'her/his sister'

b. mw-enz-angu
1-friend-POSS1SG
'my friend'

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

[2]Unless otherwise indicated, Standard Swahili examples are our own. We are grateful to Ida Hadjivayanis for discussing relevant Swahili examples with us.

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

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

Analytic forms are very similar in form and function to Standard Swahili. Miehe (1979: 166) calls these forms "disjunct" forms:[3]

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

In contrast to these pronominal, disjunct forms, there are suffixed forms which Miehe (1979: 159) calls "conjunct" forms:[4]

(6) Old Swahili possessive suffixes

a. wa-na-w-e
2-child-2-POSS1
'her sons' (Miehe 1979: 162)

b. rafiki-o (< rafiki-(y)-o)
9.friend-9.POSS2SG 9.friend-9-POSS2SG
'your friend' (Miehe 1979: 161)

c. mu-lango-w-o
3-door-3-POSS2SG
'your door' (Miehe 1979: 161)

d. mahali-p-e
16.place-16-POSS1
'his position' (Miehe 1979: 162)

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

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

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

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

3.2 Demonstrative morphology (Parameter 20)

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

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

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

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

(8) Standard Swahili demonstratives:
   a. proximal: *h* + V(owel) + C(oncor)d
   b. distal: Cd + *le*
   c. referential: *h* + V + Cd + *o*

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

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

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

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

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

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

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

a. m-tu hu-y-o
1-person DEM-CD1-DEM
'this (aforementioned) person'

b. ma-ji ha-y-o
6-water DEM-CD6-DEM
'this (aforementioned) water'

c. vi-ti hi-vy-o
8-chair DEM-CD8-DEM
'these (aforementioned) chairs'

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

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

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

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

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

(13) Old Swahili demonstratives: *s/h* + V + Cd

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

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

(14) Old Swahili proximal demonstratives (Type 1): *s/h* + V + Cd

a. hu-yu binti
DEM-CD1 1.daughter
'this daughter' (Miehe 1979: 143)
b. ngamia su-yu
9.camel DEM-CD1
'this camel' (Miehe 1979: 142)

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

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

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

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

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

a. za-le zi-tunda
CD8-DEM 8-fruit
'those fruits' (Miehe 1979: 138)

b. u-lee isilamia
CD1-DEM 1.muslim
'that Muslim' (Miehe 1979: 138)

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

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

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

b. s-u-yo yatima
DEM-CD1-DEM 9.orphan
'this orphan' (Miehe 1979: 144)

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

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

(17) Old Swahili proximal demonstratives (Type 2): (*-s/h*) + (V) + Cd + *no*
Hu-yu-no si malaika
DEM-CD1-DEM NEG.COP 1.angel
'whether this is (not) an angel at all' (lit. 'this one is not an angel') (Miehe 1979: 146)

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

Table 3: Demonstrative forms in Standard Swahili and Old Swahili

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

(22) Standard Swahili agentive borrowings

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

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

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

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

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

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

(24) Standard Swahili agentive nominalisation in *-i*

a. msomi 'scholar' (< soma 'read')
b. mcheshi 'joker, jester' (< cheka 'laugh')
c. fundi 'technician, artisan, expert' (historically from °funda 'learn')

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

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

(25) Putative Standard Swahili agentive nominalisation in *-i*

a. *mchori (cf. mchoraji 'painter/artist')
b. *mwendeshi (cf. mwendeshaji 'driver', and also dereva 'driver')

Furthermore, there are agent nominals in *-aji* which denote professions, so the interpretation of 'someone doing X habitually' appears to arise mainly in contrast with another form, in *-i* or a loanword.[7]

(26) Standard Swahili agentive nominalisation in *-aji* denoting professions

a. mchekeshaji 'comedian' (cheka 'laugh')
b. mtungaji 'designer' (tunga 'compose, design')
c. mchezaji mpira 'footballer' (cheza mpira 'play football')
d. mshonaji 'tailor' (shona 'sew')

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

(27) Standard Swahili agentive nominalisation in *-aji* in verb + object nominalisations

a. mwuzaji kompyuta 'computer salesperson'
b. mwuzaji bima 'insurance salesperson'

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

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

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

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

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

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

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

(29) a. Kesi hi-yo i-me-fungul-i-w-a **na** m-ke w-ake
9.case DEM-CD9 SM9-PERF-open-APPL-PASS-FV COM 1-wife 1-POSS3SG
'The case was opened by his wife.'

b. Wa-me-shik-w-a **na** njaa
SM2-PERF-hold-PASS-FV COM 9.hunger
'They were grabbed by hunger.'

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

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

(30) a. mw-ema, m-za-w-a **na** w-ema
1-good 1-bear-PASS-FV COM 14-good
'the good one, born in goodness' (Miehe 1979: 196)

b. mahari a-l-o-pa-w-a **ni** Jabiri
9.bride_price SM1-PST-REL-give-PASS-FV COP Jabiri
'the bride-price set by him by Jabir' (Miehe 1979: 196)

c. me-zing-iw-a **ni** mal'una
SM1.PERF-surround-PASS-FV COP 10.cursed
'He was surrounded by the cursed.' (Miehe 1979: 196)

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

3.5 Negative imperatives (Parameter 58)

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

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

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

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

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

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

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

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

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

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

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

(35) a. si-pend-a
NEG-like-FV
'Don't like.' (Miehe 1979: 251)

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

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

3.6 The formation of the perfect (Parameter 68)

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

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

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

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

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

(37) Old Swahili perfects in *-ile*

a. ni-kom-**ile** ku-kutubu
SM1SG-finish-PERF 15-write
'I have finished writing.' (Miehe 1979: 179)

b. u-tu-p-**ile** kuwwa
SM1-OM1PL-give-PERF 9.power
'He has given us power.' (Miehe 1979: 180)

c. na ratabu u-ni-p**ee**
CONJ dates SM1-OM1SG-give.PERF
'and dates he gave me' (Miehe 1979: 178)

d. Athumani ondosh-**ile** …
Athumani SM1.leave-PERF
'Othman went …' (Chuo cha Herekali, Knappert 1967: 159)

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

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

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

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

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

(38) Standard Swahili perfect in *me-*
Wa-tu wa-**me**-fik-a
2-person SM2-PERF-arrive-FV
'People have arrived.'

(39) Standard/Colloquial Swahili perfects in *sha-*

a. A-**me-kwisha** (ku-)imb-a
SM1-PERF-finish (INF-)sing-FV
'S/he has finished singing ~ has already sung.'

b. A-**me-sha**-imb-a
SM1-PERF-COMPL-sing-FV
'S/he has already sung.'

c. Ni-**sha**-fahamu
SM1SG-PERF-understand
'I have (already) understood.'

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

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

3.7 Preverbal TAM slots (Parameter 73)

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

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

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

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

(41) Old Swahili combination of *(a)li-* and *me-*

a. **ali**-**me**-iṯiṉḏ-a nḏia
PAST-PERF-block-FV 9.road
'(S/he) blocked the way.' (Miehe 1979: 219)

b. a-**li** **me**-keti nyumba-ni
SM1-PAST PERF-sit house-LOC
'He was sitting in his house.' (Miehe 1979: 219)

(42) Old Swahili combination of *(a)li-* and *ki-*
N-**ali**-**ki**-kw-evuz-a mno
SM1SG-PAST-SIT-OM2SG-search-FV very
'I was looking for you a lot.' (Miehe 1979: 220)

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

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

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

Table 4: Standard Swahili TAM markers

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

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

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

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

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

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

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

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

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

(44) Standard Swahili object marking (cf. Marten et al. 2007: 263/4)

a. ni-li-**m**-p-a
SM1SG-PAST-OM1-give-FV
'I gave him/her.'

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

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

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

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

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

(45) Old Swahili emphatic object marking

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

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

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

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

b. ni-**mu**-dhamini-**ye** jaza
SM1SG-OM1-guarantee-OM1 reward
'I will guarantee him reward.' (Miehe 1979: 100)

c. a-ka-**zi**-angush-a-**zo**
SM1-CONS-OM10-throw.down-FV-OM10
'and he threw them down' (Steere 1884: 108, in Miehe 1979: 101)

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

3.9 Summary of comparative results

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

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

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

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

Table 5: Parameters of variation for Old and Standard Swahili

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

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

4.1 Qualitative differences and the standardisation of Swahili

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

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

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

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

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

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

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

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

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

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

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

4.2 Quantitative differences and comparative Bantu contexts

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

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

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

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

Table 6: Languages of the quantitative comparison

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

5 Conclusions

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

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

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

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

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

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

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

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

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

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

Acknowledgements

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

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

Abbreviations

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

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

References

Ashton, Ethel O. 1947. *Swahili grammar*. 2nd edn. London: Longman.

Baldi, Sergio. 2012. Arabic loans in East African languages through Swahili: A survey. *Folia Orientalia* 49. 37–52.

Balestrieri, Fiona Agnes. 2017. *An analysis of Swahili lexical borrowings into Shambala, Nyamwezi, and Haya: 1880s to present*. SOAS University of London. (MA thesis).

Beaudoin-Lietz, Christa, Derek Nurse & Sarah Rose. 2004. Pronominal object marking in Bantu. In Akinbiyi Akinlabi & Oluseye Adesola (eds.), *Proceedings of the 4th World Congress of African Linguistics, New Brunswick 2003*, 175–188. Cologne: Köppe.

Bertoncini Zúbková, Elena, Mikhail Gromov, Said Khamis & Kyallo Wamitila. 2009. *Outline of Swahili literature*. 2nd edn. Leiden: Brill.

Blommaert, Jan. 2014. *State ideology and language in Tanzania*. 2nd edn. Edinburgh: Edinburgh University Press.

Dom, Sebastian & Koen Bostoen. 2015. Examining variation in the expression of tense/aspect to classify the Kikongo language cluster. *Africana Linguistica* 21. 163–211.

Drolc, Ursula. 2000. Zur Typology des Perfekts (am Beispiel des Swahili). In Werner Breu (ed.), *Probleme der Interaktion von Lexik und Aspekt (ILA)*, 91–112. Tübingen: Niemeyer.

Edelsten, Peter & Chiku Lijongwa. 2010. *A grammatical sketch of Chindamba: A Bantu language (G52) of Tanzania*. Cologne: Köppe.

Fleisch, Axel. 2005. Agent phrases in Bantu passives. In Erhard Voeltz (ed.), *Studies in African linguistic typology*, 93–111. Amsterdam: John Benjamins.

Gibson, Hannah, Gastor Mapunda, Lutz Marten, Sheena Shah & Julius Taji. 2019. Variation in double object marking in Swahili. In Daisuke Shinagawa & Nico Nassenstein (eds.), *Swahili Forum 26, Special Issue "Variation in Swahili"*, 142–165.

Grollemund, Rebecca, Simon Branford, Koen Bostoen, Andrew Meade, Chris Venditti & Mark Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences* 112(43). 13296–13301.

Guérois, Rozenn. Forthcoming. Passive constructions. In Lutz Marten, Nancy Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages*. Oxford: Oxford University Press.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation: Draft master list*. Report produced as part of Leverhulme project "Morphosyntactic Variation in Bantu: Typology, contact and change". DOI: 10.5281/zenodo.3958997.

Guthrie, Malcolm. 1967–1971. *Comparative Bantu: An introduction to the comparative and prehistory of the Bantu languages*, vol. 1–4. London: Gregg International.

Harries, Lyndon. 1962. *Swahili poetry*. Oxford: Clarendon Press.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

Jerro, Kyle. 2018. Linguistic complexity: A case study from Swahili. In Jason Kandybowicz, Travis Major & Harold Torrence (eds.), *The Selected Proceedings of the 45th Annual Conference on African Linguistics: African Linguistics on the Prairie*, 3–20. Berlin: Language Science Press.

Knappert, Jan. 1967. *Traditional Swahili poetry: An investigation into the concepts of East African Islam as reflected in the Utenzi literature*. Leiden: Brill.

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

Mwaliwa, Hanah Chaga. 2018. Modern Swahili: The integration of Arabic culture into Swahili literature. *Tydskrif vir Letterkunde* 55(2). 120–133.

Nicolle, Steve. 2012. Semantic-pragmatic change in Bantu *-no* demonstrative forms. *Africana Linguistica* 18. 193–234.

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

Nurse, Derek & Thomas J. Hinnebusch. 1993. *Swahili and Sabaki: A linguistic history*. Berkeley: University of California Press.

Rugemalira, Josephat M. 2007. The structure of the Bantu noun phrase. *SOAS Working Papers in Linguistics* 15. 135–148.

Schadeberg, Thilo C. 1989. The three relative constructions in Swahili (Kisanifu). In Marie-Françoise Rombi (ed.), *Le swahili et ses limites: Ambiguité des notions reçues*, 33–40. Paris: Editions Recherche sur les Civilisations.

Schadeberg, Thilo C. 1990. Schon – noch – nicht – mehr: Das Unerwartete als grammatische Kategorie im KiSwahili. *Frankfurter Afrikanistische Blätter* 2. 1–15.

Schadeberg, Thilo C. 1992. *A sketch of Swahili morphology*. 3rd edn. Cologne: Köppe.

Shinagawa, Daisuke & Yuko Abe (eds.). 2019. *Descriptive materials of morphosyntactic microvariation in Bantu*. Tokyo: Research Institute for Languages, Cultures of Asia & Africa, Tokyo University of Foreign Studies.

Steere, Edward. 1870. *A handbook of the Swahili language as spoken at Zanzibar*. London: Bell & Daldy.

Steere, Edward. 1884. *A handbook of the Swahili language as spoken at Zanzibar*. 3rd edn. London: Society for the Promotion of Christian Knowledge.

Swadesh, Morris. 1952. Lexico-statistical dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society* 96. 452–463.

Taylor, William Ernest. 1891. *African aphorisms or, saws from Swahili-land*. Collected, translated and annotated by the Rev. W. E. Taylor. London: Society for Promoting Christian Knowledge.

Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an eng variable*, vol. i, 98–109. Oxford: Oxford University Press.

Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics with special focus on Bantu and Mande*, 425–441. Tervuren: Royal Museum for Central Africa.

van der Spuy, Andrew. Forthcoming. A history of research into Bantu linguistics. In Lutz Marten, Nancy Kula, Jochen Zeller & Ellen Hurst (eds.), *The Oxford guide to the Bantu languages*. Oxford: Oxford University Press.

Vierke, Clarissa. 2011. *On the poetics of the Utendi: A critical edition of the nineteenth-century Swahili poem "utendi wa haudaji" together with a stylistic analysis*. Zurich: LIT.

Voeltz, F. K. Erhard. 1980. The etymology of the Bantu perfect. In Luc Bouquiaux (ed.), *L'expansion bantue*, 487–492. Paris: Centre National de la Research Scientifique.

Whiteley, Wilfred Howell. 1955. Kimvita: An inquiry into dialectal status and characteristics. *Journal of the East African Swahili Committee* 25. 10–39.

Whiteley, Wilfred Howell. 1969. *Swahili: The rise of a national language*. London: Methuen.

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

Zawawi, Sharifa. 1979. *Loan words and their effect on the classification of Swahili nominals*. Leiden: Brill.

Name index

Abe, Yuko, 2, 376
Aboh, Enoch O., 98
Ahn, Dorothy, 53, 54, 64
Aissen, Judith, 194, 195
Alexiadou, Artemis, 98
Alsina, Alex, 184, 230, 256
Ashton, Ethel O., 53, 54, 156, 230, 232, 233, 240, 244, 245, 265, 267, 376, 378, 380, 381, 385, 394, 406, 414
Asiimwe, Allen, 18, 20, 28, 34, 45, 49–55, 57–59, 64–68, 72–74, 76, 77, 91, 92, 94, 98, 99
Aunio, Lotta, 281, 300, 302, 303, 310

Baker, Mark, 192, 230, 256
Baldi, Sergio, 376
Balestrieri, Fiona Agnes, 374
Bar-el, Leora, 304
Barlew, Jefferson, 129
Bastin, Yvonne, 233, 292, 356
Batibo, Herman, 241
Bax, Anna, 2
Beaudoin-Lietz, Christa, 1, 400
Beermann, Dorothee, 51, 58
Benton, Richard A., 27
Berger, Paul, 233
Bernsten, Jan, 44
Bertoncini Zúbková, Elena, 377
Bleek, Wilhelm H. I., 1, 131
Blommaert, Jan, 377, 406
Bloom Ström, Eva-Marie, 58
Bokamba, Eyamba, 230
Bokamba, Georges D., 91, 92
Bostoen, Koen, 2, 7, 234, 237, 289, 351, 352, 358, 365, 374
Botne, Robert, 294, 299
Bresnan, Joan, 1, 58, 110, 116, 118, 120, 145, 155, 157, 178, 186, 187, 192, 210
Buell, Leston, 129, 140, 141, 184
Bybee, Joan, 256, 300, 332, 344

Cable, Seth, 357
Campbell, Lyle, 360
Capus, Auguste, 229, 231
Carlson, Robert, 27, 33
Carstens, Vicki, 118, 119, 121
Chabata, Emmanuel, 230, 269
Chaula, Edigio Hangaika Yamligile, 130
Chebanne, Andy M., 292
Cocchi, Gloria, 58
Contini-Morava, Ellen, 184
Corbett, Greville, 192
Creissels, Denis, 116
Croft, William, 360
Crystal, David, 302

Dahl, Östen, 294, 297, 305, 317
de Blois, Kornelis Frans, 64, 66, 85, 88, 89, 91, 95, 96, 98, 105
de Haan, Ferdinand, 297
Deen, Kamil Ud, 192

Demuth, Katherine, 1, 110, 145
Denny, J. Peter, 27
Derbyshire, Desmond C., 33
Devos, Maud, 109, 123, 279, 306, 307
Dewees, John William, 66
Dicks, Ian D., 20
Diercks, Michael, 2
Diessel, Holger, 24, 25, 27, 46, 52, 69, 72–74
Dimitriadis, Alexis, 357
Dixon, Robert Malcolm Ward, 24, 25, 60, 61, 74
Doke, Clement Martyn, 309
Dollar, Shawn, 20
Dom, Sebastian, 230, 352, 353, 374
Drolc, Ursula, 395
Du Plessis, Jacobus Albertus, 20, 45, 47, 57
Dunham, Margaret, 360–362
Duranti, Alessandro, 157, 184

Eastman, Carol, 230
Eberhard, David M., 353
Edelsten, Peter, 3, 281, 410

Faltz, Leonard, 352
Faquir, Osvaldo G., 201, 206
Felberg, Knut, 86, 88, 91
Fennig, Charles D., 44
Fillmore, Charles J., 72, 73
Fleisch, Axel, 116, 117, 297, 299, 392
Floor, Sebastian, 34
Frajzyngier, Zygmunt, 356

Gerhardt, Phyllis, 299
Gibson, Hannah, 113, 115, 340, 400, 402
Givón, Talmy, 233
Goodness, Devet, 91, 92
Greenberg, Joseph H., 33
Grégoire, Claire, 110, 116, 117, 119
Grévisse, Fernand, 230
Grollemund, Rebecca, 3, 407
Guérin, Valerie, 61
Guérois, Rozenn, 2, 3, 5, 6, 8, 85, 86, 109, 110, 116–118, 122, 125, 129, 130, 132, 134, 141, 156, 213, 229, 231, 232, 260, 268, 312, 373, 375, 376, 379, 392, 407, 413, 414
Guillemin, Diana, 74
Güldemann, Tom, 115, 117, 310, 340
Gunnink, Hilde, 45, 63, 78, 125
Guthrie, Malcolm, 130, 131, 216, 229, 230, 232, 233, 244, 247, 251, 266, 279, 281, 300, 353, 407

Haiman, John, 353
Hamlaoui, Fatima, 140
Hammarström, Harald, 2
Haonga, Ernest, 353, 354
Harford, Carolyn, 135
Harries, Lyndon, 129, 377
Harris, Martin, 33
Harvey, Andrew, 2
Haspelmath, Martin, 230, 340, 344, 352, 353, 357
Heath, Jeffrey, 27
Heine, Bernd, 3, 7, 340, 351, 352, 356, 359–364, 395, 396
Henderson, Brent, 198
Hendle, P. J., 281, 310
Himmelmann, Nikolaus P., 46, 74, 77
Hinnebusch, Thomas J., 2, 3, 398, 399
Hopper, Paul J., 362, 365
Hualde, José I., 194
Humphreys, R. L., 230

Hyman, Larry M., 91, 92, 157, 229, 230, 233, 256, 257, 267, 268, 271, 340, 356

Iorio, David Edy, 193
Ismail, Joseph Hokororo, 130

Janda, Richard, 360
Jerro, Kyle, 238, 413
Johnson, Frederick, 230, 233, 244

Kagaya, Ryohei, 111, 326, 335
Kahigi, Kulikoyela, 135, 229, 232, 233, 241, 257
Kaji, Shigeki, 98, 99
Kanerva, Jonni M., 58, 110, 116, 118, 120, 145
Karwemera, Festo, 52, 54, 67, 77
Katamba, Francis X., 91, 92, 113, 230
Kemmer, Suzanne, 352, 353, 358
Kershner, Tiffany L., 294, 299
Khamis, Abdu, 230
Khumalo, Langa, 1, 230
Kießling, Roland, 2
Kimambo, Gerald E., 94, 99, 103
Kioko, Angelina Nduku, 116
Kirwan, Brian Edmond Renshaw, 47, 49, 55, 56
Knappert, Jan, 396
Kokutani, Shigehiro, 353
König, Ekkehard, 353
Krapf, Johann Ludwig, 395
Kröger, Oliver, 130
Krumm, Bernard, 376, 390
Kühl, Karoline, 91, 92, 99, 317
Kula, Nancy, 6, 86, 103, 155, 191, 192, 194–200, 204, 205, 209, 210, 213, 219, 221–224, 230, 233, 260, 261, 292, 400
Kulikov, Leonid, 352
Kusters, Wouter, 404, 413
Kuteva, Tania, 340, 359, 360, 362, 396
Kutsukake, Sayaka, 410

Ladefoged, Peter, 44
Lambrecht, Knud, 33
Languages of Tanzania Project (LOT), 111, 279, 281
Legère, Karsten, 89, 91
Lehmann, Christian, 33
Levinson, Stephen C., 33, 72–74
Lewis, Paul M., 44, 281
Lichtenberk, Frantisek, 356
Lijongwa, Chiku, 281, 410
Lippard, Hannah, 212, 224
Lodhi, Abdulaziz Y., 233, 247, 376, 390
LOT, 353
Lusekelo, Amani, 85, 86, 88, 94, 99, 100, 102–105
Lyons, Christopher, 31, 34, 35, 46, 72, 74, 86, 94, 97, 98, 101, 102, 106

Machobane, 'Malillo 'Matsepo, 132, 136
Madan, Arthur Cornwallis, 230, 232, 233, 265
Madumula, Joseph, 353
Maganga, Clement, 240, 246, 247, 269
Maho, Jouni Filip, 20, 44, 111, 117, 121, 156, 277, 326, 333, 353
Makanjila, Dominick, 19, 32
Malinga, R. M. M., 78
Mallya, Aurelia, 111, 115, 122, 131
Mann, Michael, 246
Mapunda, Gastor, 130

Marconnes, Francisque A., 27
Marlo, Michael R., 155, 182, 184, 187, 283, 352, 400
Marten, Lutz, 2, 3, 6, 58, 86, 103, 104, 109, 110, 115–122, 129, 132, 135, 136, 140, 142, 155, 157, 158, 191, 192, 194–200, 204, 205, 209, 210, 213, 219, 221–224, 230, 231, 260, 261, 375, 395, 396, 400, 405, 407, 413
Mazrui, Alamin, 406
Mbaabu, Ireri, 376, 377
Mbope, Ngusubila, 85, 88, 105
Mcha, Yohana, 131
Mchombo, Sam A., 1, 192
McWhorter, John, 404, 413
Meeussen, Achille E., 95, 113, 131, 229, 232, 255, 256, 279, 282, 283, 288, 290–293, 308, 311, 326, 344, 351, 387, 388, 393
Meinhof, Carl, 1, 115
Miehe, Gudrun, 376, 377, 382, 383, 385–389, 393–396, 398, 401, 402, 405, 414
Miyashita, Hiroyuki, 356, 359–362
Mkude, Daniel J., 132, 136, 279, 286, 296, 300, 301, 315
Mlacha, S. A. K, 406
Mmusi, Sheila, 1, 110, 145
Mol, Frans, 279
Moltmann, Friederike, 27
Montlahuc, Marie-Laure, 325, 330, 333, 341
Morimoto, Yukiko, 1
Morris, Henry Francis, 47, 49, 55, 56
Morrison, Michelle Elizabeth, 281, 286, 366
Moses, Godian, 279
Moshi, Lioba, 1, 110, 122, 132, 155, 157, 178, 186, 210, 326, 341
Mould, Martin, 233
Moyse-Faurie, Claire, 352
Mpalanzi, Lameck, 353
Mpiranya, Fidèle, 132
Mpofu-Hamadziripi, Nomalanga, 135
Mreta, Abel, 32
Msamba, Enitha, 352, 364–366
Mtenje, Atikonda, 2
Mubangizi, K. Benedicto, 55, 57, 69, 74, 75
Mugane, John M., 116, 376–378
Mugumya, Levis, 65, 68, 75, 77
Müller-Bardey, Thomas, 230
Mulokozi, Mugyabuso M., 377
Munongo, Antoine Mwenda, 230
Muriungi, Peter, 184
Mushi, C., 111
Muzale, Henry R. T., 2, 405
Mwaliwa, Hanah Chaga, 376
Mwamzandi, Mohamed Y., 100, 102

Narrog, Heiko, 351, 352, 359–362
Ndomba, Rodrick G., 94
Nedjalkov, Vladimir P., 353, 355
Ngoboka, Jean Paul, 2, 119, 142
Ngonyani, Deo, 130, 286
Ngunga, Armindo, 20, 135, 191, 192, 201, 206, 213
Ngwasi, Lengson, 351, 352, 354
Nicolle, Steve, 3, 18, 45, 53, 54, 56, 74, 77, 387
Novotná, Jana, 281
Nurse, Derek, 2, 3, 20, 279, 281–283, 291–294, 300, 302–308, 311–313, 316, 326–333, 336, 337,

339–341, 343, 344, 346, 354, 398, 399, 405
Nzang-Bie, Yolande, 234

Odden, David, 283
Olomi, Rogati, 111
Olson, Howard, 241, 269

Payne, Doris L., 27, 33
Payne, Thomas, 230
Peña, Jaime, 27, 33
Persohn, Bastian, 85, 86, 88, 89, 91, 93, 96, 98, 105, 109, 123
Petzell, Malin, 2, 91, 92, 99, 109, 110, 116, 120, 279, 281, 294, 296, 300, 302–304, 310, 313, 317, 352
Phanuel, Benedictor, 326, 334
Philippson, Gérard, 2, 20, 304, 325, 326, 330, 333, 336, 337, 340, 341
Plag, Ingo, 231
Polak, Louise, 184, 352

Ramadhani, Deograsia, 104, 192, 194
Ranero, Rodrigo, 155, 160, 164, 171, 172, 177, 186
Reh, Mechthild, 395
Rehg, Kenneth L., 27
Repp, Sophie, 86, 98
Richardson, Irvine, 246
Riedel, Kristina, 1, 86, 103, 104, 109, 110, 120, 122, 191–193, 196, 198
Rijkhoff, Jan, 98
Rose, Sarah, 328
Rubongoya, L. T., 57, 64
Rugemalira, Josephat M., 1, 99, 109, 117, 129, 131, 136, 184, 230, 232, 269, 326, 334, 352, 386

Salzmann, Martin David, 141
Samson, Ridder, 115, 116
Sands, Bonny, 277
Schadeberg, Thilo C., 115, 116, 230, 231, 233, 237, 239, 240, 246, 247, 254, 255, 269, 289, 351, 358, 365, 378, 381, 384, 389, 399, 401, 402, 405, 406, 414
Schladt, Mathias, 365
Schuh, Russell G., 27
Scotton, Carol, 230
Seidl, Amanda, 357
Sengo, Tigiti S. Y., 377
Shinagawa, Daisuke, 2, 325, 330, 331, 340, 341, 344, 346, 376
Sikuku, Justine, 192
Simons, Gary F., 44
Sohl, Damian G., 27
Ssekiryango, Jackson, 155, 156, 160, 164, 165, 172, 177, 182, 186, 187
Stappers, Leo, 291
Steere, Edward, 378, 402
Stegen, Oliver, 241, 269
Stonham, John, 230
Stucky, Susan U., 135, 136
Swadesh, Morris, 408

Taji, Julius, 32, 35, 36, 47, 63, 64, 78, 104
Taylor, Charles, 47, 49–51, 54–58, 64, 66
Taylor, William Ernest, 375, 377
Trudgill, Peter, 404, 413

Valinande, Nzama K., 45
Van de Velde, Mark, 24–26, 28, 29, 31, 35, 45, 53, 63, 64, 99, 109, 113, 386

van der Spuy, Andrew, 110, 116, 413
van der Wal, Jenneke, 2, 19, 32, 35, 37, 53, 54, 63, 64, 69, 115, 129, 140, 142, 155, 178, 193, 194, 213
Van Olmen, Daniël, 306
Vierke, Clarissa, 375, 377
Visser, Marianna, 20, 45, 47, 86, 91, 94, 99, 103, 104
Voeltz, F. K. Erhard, 395

Wald, Benji, 20, 45, 47, 64, 238, 269
Watters, John R., 340
Waweru, Martin, 230, 269
Whiteley, Wilfred Howell, 20, 376–378, 395, 406

Yoneda, Nobuko, 410

Zawawi, Sharifa, 390
Zeller, Jochen, 2, 58, 63, 64, 119, 129, 155, 191–194, 198

Morphosyntactic variation in East African Bantu languages

The approximately 500 Bantu languages spoken across vast areas of Central, Eastern and Southern Africa are united by the presence of a number of broad typological similarities, including, for example, complex noun class system and agglutinative verbal morphology. However, the languages also exhibit a high degree of micro-variation. Recent work has demonstrated fine-grained morphosyntactic variation across many Bantu languages focusing on grammatical topics such as double object constructions, inversion constructions, or object marking, adopting formal, comparative and typological perspectives.

Continuing in this vein, this volume builds on the momentum of the dynamic field of morphosyntactic variation in Bantu and contributes to the growing body of work which examines morphosyntactic variation, with a regional focus on the Bantu languages of East Africa. The East African region is characterized by high linguistic complexity in terms of the number of languages spoken, in terms of the four different linguistic phyla present, and in terms of the inherent sociolinguistic dynamics.

The current volume explores this complexity further by bringing together studies which investigate features of morphosyntax of an individual language as well as those which develop an in-depth examination of a single morphosyntactic phenomena in a small sample of languages.

The book seeks also to add to the descriptive status of the languages under examination, as well as raising questions relating to language, language contact, language change, and micro-variation in related languages spoken in close geographic proximity.