COMPARATIVE TENSE AND ASPECT IN THE MARA BANTU LANGUAGES: TOWARDS A LINGUISTIC HISTORY

by

JOHN B. WALKER

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

LINGUISTICS

We accept this thesis as conforming to the required standard

…….

Jamin Pelkey, PhD; Thesis Supervisor

…….

Sean Allison, PhD; Second Reader

…….

Oliver Stegen, PhD; Third Reader

TRINITY WESTERN UNIVERSITY

October 24, 2013

Abstract

The Mara region of Tanzania is a densely populated area that contains at least 22 Bantu speech varieties in addition to the Nilotic languages Luo and Taturu (Hill et al. 2007; Mitterhofer & Robinson 2012). In Maho's updated version of Guthrie's geographically influenced classification system, the Bantu speech varieties of the region are divided into the JE25 and JE40 groups (Maho 2009). To a large degree this classification corresponds with the proposed genetic linguistic sub-groups of "Mara" (most of JE40) and "Suguti" (JE25), which are purportedly related to each other under the East Nyanza branch of Great Lakes (GL) Bantu based on lexical similarity in core vocabulary, shared lexical innovation/borrowing, and insight from other scientific fields like archaeology and palynology (Schoenbrun 1990). Additionally, Schoenbrun's (1990) study on GL Bantu further subdivides "Mara" into North Mara and South Mara groups. More recent studies interested in comparative linguistics amongst the Mara languages have focused on vowel systems (Higgins 2011: 267-275) and discourse (Rundell 2012).

The primary goal of this current research is to compare the TA systems amongst the Mara languages with the aim of finding any shared "individual-identifying" innovations (Nichols 1996); these can then serve as a basis for subgrouping linguistic varieties and gaining a better understanding of historical relationships between them. A secondary goal is to provide a preliminary linguistic description of the TA systems of the Mara languages, which is an area of study that has previously been lacking. The five core speech varieties under examination are Ikizu (JE402, [ikz]), Ikoma (JE45, [ntk]), Kabwa (JE405, [cwa]), Simbiti (JE431, [ssc]), and Zanaki (JE44, [zak]). Additional data is drawn from previous studies (Nurse 2007: 127-129 based on Whiteley 1960; Cammenga 2004; Mwita 2008) on two other Mara languages, Gusii (JE42, [guz]) and Kuria (JE43, [kuj]), to broaden the scope of the comparative analysis.

The primary data for this study was gathered in each of the five core languages through the elicitation of 91 Swahili sentences and an oral text based on events in a short film called "Pear Story" (University of California at Berkeley 1975; Chafe 1980). Based on other TA studies in Bantu (Nurse 2008; Botne 2013; Botne & Kershner 2008), this data is presented and compared in three sections. The first section covers various degrees of past and non-past tenses, as well as referential tenses referring to sequential events in a narrative. Following this, a number of aspectual properties of Bantu (including Anterior/Perfect, Progressive/Continuous, Habitual, Persistive and Inceptive) are observed in the Mara languages. The third section looks at several modal categories (Subjunctive, Conditional, Hypothetical and Potential) and other related verbal categories (Negation and Relatives).

This mass of data is then analyzed in light of other comparative studies in GL Bantu (Schoenbrun 1990; Nurse & Muzale 1999; Nurse 2008) to determine whether there is evidence from TA comparisons to support the hypothesized Mara genetic grouping and also to ascertain the relationship between the North Mara and South Mara languages. The final conclusion of the study is that, although the languages all show evidence of a TA system related to Proto-GL, there is not a sufficient base of shared individual-identifying innovations that unite the TA systems of all the Mara languages. Instead, there is stronger evidence for a split between the North Mara (Kuria, Simbiti, and Kabwa) and the South Mara (Ikizu, Zanaki, and Ikoma) sub-groups, with some overlap between the systems of Simbiti (North Mara) and Ikoma (South Mara). This lends favor to an interpretation of a short-lived Proto-Mara language that soon became a dialect chain with Proto-North and Proto-South varieties still in contact with each other. After a while these became distinct languages with their own internal dialect variation, but the varieties closest to the contact zone (pre-Simbiti and pre-Ikoma) shared some common features with each other.

Acknowledgements

The process of writing a thesis has been full of frustrations and joys, but the journey has also produced great personal growth. This study is by far the biggest single task I have ever set out to accomplish and its completion is due to the input, advice and encouragement of many people.

First of all, I am very grateful to my thesis advisor, Dr. Jamin Pelkey, under whose tutelage I have had the privilege of learning many important lessons both in linguistics and in life. His teaching has provided me with a solid philosophy of language that undergirds a more holistic approach to my work in linguistics. Additionally, his advice on specific matters of detail on this thesis has improved the final product.

I am also appreciative of the input I have received from my other readers, Dr. Sean Allison and Dr. Oliver Stegen. They have brought their distinctive areas of expertise to bear on this thesis, which has stretched me beyond where I would have otherwise gone. I thank Sean for his comments on descriptive linguistic standards and ideas for how to incorporate that into this research. I thank Oliver for his suggestions about academic writing as well as his wealth of knowledge on Bantu languages that far exceeds my own. I also am grateful to both men for their recommendations on further background research that has proven invaluable in increasing the academic quality of the thesis.

I am incredibly thankful for the five Tanzanian language consultants I worked with from the Ikoma, Ikizu, Kabwa, Simbiti, and Zanaki language groups. Unfortunately, due to confidentiality agreements required by research ethics, I cannot disclose their names. Without their help there would be no thesis, so I look forward to working with them more in the future and giving them full credit for further input on their languages.

I also received input from linguistic experts on the Mara languages and I thank them heartily. I thank Holly Robinson (*née* Higgins) for her general encouragement and also for her insight into the Ikoma language, which has proven difficult to analyze. I thank Hazel Gray for sharing her knowledge on the Ikizu and Zanaki languages, even in the midst of working on her own thesis. I also thank Tim Roth for spurring me on in the task and for giving input on how to integrate and apply theoretical insight to Bantu historical and comparative linguistics.

Additionally, I want to thank those who were instrumental in facilitating different aspects of my Thesis Defense so that a teleconference could be held from three different locations on two continents. These people include Rod Casali, Dave Jeffery, Diane Beaton and Andrew Sweet.

On the non-linguistic side of things, I have received invaluable emotional support from my family. My parents, Doug and Marilyn Walker, and my brother and sister-in-law, Kevin and Janelle, have continually shown an interest in my work and prayed for me. Most of all, I thank my dear wife, Ellen, for her unceasing support of me in the seemingly endless task. Though I know she was often tired of my busy schedule and ready for me to be finished, she still helped me to set aside time to work on the thesis. I love you dear and appreciate you more than you can imagine!

Most importantly, abundant thanksgiving is due to the One, creator and king, who enabled me to think and write in the first place and to whom all true theoretical insight points. With His enabling, I hope that this is the foundation for a fruitful academic career in linguistics.

List of Figures

List of Tables

Abbreviations

< >	Orthographic Transcription
/ /	Phonemic Transcription
[ ]	Phonetic Transcription
Ø	Null morpheme
AGR	Agreement (any number following this refers to a specific noun class)
ANT	Anterior (related to the Perfect)
APPL	Applicative
ASC	Associative
AUG	Augment vowel (any number following this refers to a specific noun class)
CAUS	Causative
CCONJ	Coordinating Conjunction
CK	Central Kenya
CNS	Consecutive
COND	Conditional
COP	Copula
DDEM	Distal Demonstrative
EFT	English Free Translation
EMPH	Emphatic
EWW	English Word-for-Word translation
F_1	Immediate Future Tense
F_2	Hodiernal Future Tense
F_3	Crastinal Future Tense
F_4	Remote Future Tense
FUT	Future Tense (general)
GL	Great Lakes
HAB	Habitual
HYP	Hypothetical
IMP	Imperative
INCE	Inceptive
INF	Infinitive
IPA	International Phonetic Alphabet
IPFV	Imperfective
LOC	Locative (any number following this refers to a specific noun class)
MC	Main Clause
MCA	Main Clause Affirmative verbs
MMG	Morpheme-by-Morpheme Gloss
NAR	Narrative Tense
NC	Noun Class prefix (usually followed by a specific number)
NEG	Negative
NEG_2	Secondary Negative slot (of the verb)
NPST	Non-Past
NUGL	*New Updated Guthrie List* (Maho 2009)
OBJ	Object agreement (any number following this refers to a specific noun class)
OM	Object Marker slot (on the verb)

P_1	Immediate Past Tense
P_2	Hodiernal Past Tense
P_3	Hesternal Past Tense
P_4	Remote Past Tense
PASS	Passive
PB	Proto-Bantu
PDEM	Proximal Demonstrative
PER	Persistive Aspect
PL	Plural
POSS	Possessive
POT	Potential
PRF	Perfect
PRS	(Vast) Present
Pre-SM	Pre-Subject Marker slot (of the verb)
PROG	Progressive
PRON	Pronoun (general)
PST	Past tense (general)
R	Root (of verb)
RC	Relative Clause
RCA	Relative Clause Affirmative verbs
RDEM	Referential Demonstrative
RDPL	Reduplication
RECP	Reciprocal
REFL	Reflexive
REL	Relative
SBJ	Subject agreement (any number following this refers to a specific noun class)
SBJV	Subjunctive
SFT	Swahili Free Translation
SG	Singular
SM	Subject Marker slot (of the verb)
STAT	Stative
SWW	Swahili Word-for-Word translation
TA(M)	Tense/Aspect/(Mood)
UT	Utterance Time
V	Vowel (unspecified)

Table of Contents

1. INTRODUCTION

On the eastern shore of Lake Victoria, the Mara region of Tanzania covers just over 8,400 square miles of land south of the Kenyan border, stretching from the lake in the west to the Serengeti National Park in the east. The region is popularly referred to as "the Region of 24 languages" (Hill et al. 2007: 5) and maps based on recent surveys show at least 22 speech varieties (Hill et al. 2007: 21; Mitterhofer & Robinson 2012: 4; see also map in Section 1.5.1). Due to this densely situated language diversity, the area has great potential for historical-comparative linguistic studies.

Most of the speech varieties in the Mara region are of Bantu origin and fall under the classification in the *New Updated Guthrie List* (Maho 2009) as either JE25 or JE40[1] (see further discussion in Sections 1.3 and 1.5 below). Although this classification is understood to be heavily based on geography and is therefore useful mainly as a reference tool (Nurse 1999: 1), previous historical work has also posited that the JE25 and JE40 languages (with the notable exclusion of the JE41[2] varieties) form two distinct genetic sub-groupings linguistically. The JE25 speech varieties are called "Suguti" and the JE40 varieties are referred to as the "Mara" languages (Nurse & Muzale 1999: 519; Schoenbrun 1990). Since this thesis is concerned with genetic relationships between languages, these sub-groups will also be referred to herein as Mara and Suguti. A detailed list of the languages that compose these groups is found in Section 1.5.

Regarding genetic affiliation at a higher level, the Mara and Suguti sub-groups have been proposed as the only two branches of a larger genetic grouping called East Nyanza (Schoenbrun

[1] In addition to the Bantu languages there are two Nilotic speech varieites included in a recent language map of the region: Luo and Taturu (Hill et. al 2007: 21).
[2] These varieties include Logooli (JE41), Idaxo (JE411), Isuxa (JE412), and Tiriki (JE413). Nurse includes these varieties with the JE30 languages as the Luhya sub-group (1999: 8-9). See Section 1.5.2 for further discussion.

1990: 135-136), and East Nyanza is in turn one of the groups under Great Lakes[3] (GL), which corresponds to a large degree with zone J in the *NUGL* (Nurse 1999: 7).

Since the Mara languages are the focus of this thesis, it is important to understand the internal composition of this sub-group of GL. Schoenbrun (1990) has given evidence from lexical similarity in core vocabulary, shared lexical innovation/borrowing, and phonological changes for two branches: North Mara and South Mara. Additionally, his research consults extra-linguistic sources from other fields of science like archaeology and palynology to provide evidence for a historical settlement scenario of Bantu people in the area. Section 1.5.2 presents a more detailed description of this research and its implications for genetic classification.

The research in this thesis corroborates the existence of both North and South Mara sub-groups through study of the Tense/Aspect (TA) systems in several of the Mara languages; however, the evidence for grouping the North Mara and South Mara groups together at the next highest Mara node is more scant. This casts some doubt on a Proto-Mara phase of development following both the split of Great Lakes (GL) Bantu and the subsequent East Nyanza phase. At the very least, the TA evidence appears to limit the Proto-Mara language to a short time frame that was insufficient to create many innovations in its TA system. It is my view that dialect diversification of North and South daughter varieties occurred early and eventually these became distinct languages which then developed their own internal dialect variation. The North and South dialect clusters continued to stay in contact with each other; and this situation has persisted until the present. Due to this development, the varieties closest to the contact zone share some common TA features with each other.

[3] The term Lacustrine is also used in the literature (i.e. Mould 1976), but Schoenbrun treats Lacustrine as a subset of his Great Lakes and proposes this latter title as a more comprehensive term (1990: 132-149).

The remainder of the thesis will present the data and reasoning that has led me to these conclusions.

1.1 Section Overview

The remaining sections in the Introduction focus on setting the background necessary for understanding a historical-comparative approach to TA in the Mara languages. Section 1.2 puts the whole study in perspective by giving a preview of the research and its results. Following that, further background information, including a review of relevant literature, is organized into three distinct sections: Bantu language classification and historical-comparative linguistics (1.3), the study of TA in Bantu (1.4), and the language situation in the Mara region (1.5). The Introduction closes with an overview of the body of the thesis.

1.2 Preview of Research and Results

I present this work as a step towards more solidly sub-classifying the Mara languages based on a broader base of linguistic criteria and gaining a better understanding of the linguistic history of the Mara region of Tanzania. As a secondary goal, this research aims to provide a preliminary description of the Tense/Aspect (TA) systems in five of the lesser studied Mara speech varieties, with additional reference to Mood/Modality and other verbal categories. The five core[4] languages of this study are Ikizu (JE402, [ikz]), Ikoma (JE45, [ntk]), Kabwa (JE405, [cwa]), Simbiti (JE431, [ssc]), and Zanaki (JE44, [zak]).[5] In addition to these five, linguistic data from other sources is presented for a comparison of several additional Mara speech varieties like Kuria (JE43, [kuj]) and Gusii (JE42, [guz]). A map of the relative locations of all these languages is featured as Figure 4 in Section 1.5 where the classificatory status and relevant background of the Mara languages is considered in greater depth.

[4] These five languages will be referred to frequently as the "core languages" in this thesis.
[5] The notations following each language in parentheses include the *NUGL* reference number first, followed by the ISO 639-3 language code in brackets.

A couple caveats about terminology are in order when talking about TA in Bantu. First, other verbal categories like Mood, Focus and Negation cannot be completely ignored in a discussion of TA since they are inflectional categories that are indicated at similar slots on the verbal stem (see Section 1.4 for more on this topic). For this reason, it is sometimes necessary to refer to all these verbal categories using the cover designation Tense/Aspect/Mood (or simply TAM, for short). Even though Mood and other categories will be covered to some extent, and indeed they are the focus of Section 5, my research is generally concerned with the predominant TA categories and the way they are marked in the morphosyntax. Mood and other verbal categories will be treated as secondary, providing additional support for comparisons between languages, but also pointing to the need for further research in those areas. Second, there is potential ambiguity between the terminology for an individual morpheme that indicates a specific TAM marking and the entire construction of which it is part. I will use the term "form" to refer to the latter (i.e. a specific set of morphemes that forms a construction and has a specific range of semantic/pragmatic usage). It is the forms that work together systematically, each covering a slightly different area of the cognitive space for TA (see further explanation of this in Section 1.4.1). Additionally, each TA form is made up of "formatives," which are individual morphemes used in at least one form and frequently in more than one form. In brief, forms are made up of formatives; more will be mentioned about this in regard to slots on the verbal word in Figure 2 of Section 1.4.2.

In order to better understand the relationship between the morphophonological realization, standard function, and extended usage of a given TAM form, I collected both sentence elicitations and textual data in each of the five core languages. Sections 2.2 and 2.3 delve into the data gathering techniques I used and their implications. Some of the elicited

sentences are used as examples throughout the body of the paper, but the main data gathered from them is summarized in the TA charts for each core language (Table 12 through Table 16) in Appendix A. The textual data was gathered through the recording of a narrative based on a short video. This recording was then transcribed in the orthography of each language and given morpheme-by-morpheme glosses. The transcriptions, glosses, and translations of the five narrative texts comprise Appendices B through F.

Based on this raw data and further research with the language consultants from each of the core languages, the morphophonological and syntactical realizations of a number of important TAM categories are compared across the five languages. This comparison takes into account morphophonological changes that have occurred to the surface form in one language that obscures its similarity to a related form in another language. To the extent that extended usage of a particular form is witnessed in the data, this is also considered in order to establish which formative morphemes are shared by the Mara speech varieties and which are innovations in a certain subset of the Mara languages. The focus of the comparisons between languages is historical-comparative in nature and not typological-comparative, which is an important distinction that will be explained further in Section 1.3.2.

These findings then contribute to a broader comparison of the TAM categories in Mara with realizations of similar TAM categories in other branches of Great Lakes Bantu. This comparison aids in assessing which features can be considered inherited from Proto-GL and which are innovations that provide evidence for how a theoretical Proto-Mara TAM system would have differed from that of Proto-GL. However, through the study, the viability of a Proto-Mara TAM system that is unique from the Proto-GL system is cast into doubt. The evidence presented in this thesis points to a different conclusion; namely, that the Mara language

TAM systems divide into North Mara varieties (including Kabwa, Simbiti and Kuria) and South Mara varieties (including Ikoma, Ikizu and Zanaki). There is evidence that both the North Mara and South Mara groups exhibit characteristics of broader GL languages, and so their higher level association is not in doubt. However, the details of their relationship to each other at a post-Proto-GL (i.e. Proto-East Nyanza or Proto-Mara) phase of language evolution are not readily transparent; so further analysis of the North and South Mara languages in relation to the Suguti[6] languages will be necessary to better understand the relationship of all the languages included under the East Nyanza label.

Within the Mara languages, there is overlap between specific features of individual TAM systems across the North and South sub-branches, particularly between Simbiti (from North Mara) and Ikoma (from South Mara). From this, I deduce that there has been long-term contact between the two sub-branches, and this lends favor to an interpretation of a short-lived Proto-Mara language that existed after the split from Proto-GL (and presumably after a Proto-East Nyanza phase). This Proto-Mara soon transitioned into a dialect continuum consisting of Proto-North and Proto-South varieties which maintained contact with each other. After a while these became distinct languages with their own internal dialect variation, and over time the dialects from North Mara and South Mara that were in continued contact with each other developed some additional common features that are not seen in all of the Mara languages.

Having taken a look at where the study is headed, I now explain some background information on Bantu historical-comparative linguistics that will set the context for the comparative nature of this research.

[6] No Suguti language variety was specifically studied for this thesis.

1.3 Bantu Language Classification and Historical-Comparative Linguistics

The Bantu languages are spoken by the majority of sub-Saharan Africans from Cameroon in the west to Kenya in the east and all the way south to South Africa; only small pockets of other families are interspersed amongst them within this vast area. The number of languages classified as "Bantu" ranges between 300 and 600, depending on how "language" is defined (Nurse & Philippson 2003a: 1-3; Hinnebusch 1989: 450-451). The size of the Bantu language family and its relative linguistic unity has drawn many researchers over the years to ask questions about the historical situation that gave rise to the current distribution.

1.3.1 Historical Overview

The history of Bantu language classification and historical-comparative linguistics has been covered in many excellent sources including Hinnebusch (1989) and Schadeberg (2003). I rely on their work for the following summary of that history, which provides a backdrop for comparative work in Mara Bantu.

Early comparative Bantu scholars included W.H.I. Bleek, who coined the term *Bantu* (Bleek 1862), and Carl Meinhof, who brought insights from the comparative method to reconstruct the phonology, some morphology and a number of lexical items for Proto-Bantu (Schadeberg 2003: 144). These early pioneers paved the way for the progress in historical-comparative Bantu linguistics that has come since the mid-twentieth century. A central figure of that era is Malcolm Guthrie who developed the geographically-based classification system of Bantu languages that is still widely used today (Schadeberg 2003: 144). Under this system, the Bantu area is divided into 15 geographical zones which are each represented by one of the following letters: A-H, K-N, P, R or S.[7] Within each zone, individual languages are

[7] The letters I, J, O, and Q were excluded from the original Guthrie classification, but I have never heard an explanation as to the reason for this.

referenced by a two or three digit number. The first digit of this number divides each region into subgroups and any following digits provide a unique number for each language within that sub-group. Guthrie's classification does not necessarily divide the languages into their genetic sub-groups and subsequent comparative work has shown that for many subgroups it is not reliable as a phylogenetic reference. Even so, it is still valuable as a point of reference to specifically identify many of the speech varieties within Bantu, especially if there are several names used for an individual speech variety (2003: 146; Nurse 1999: 1). Jouni Maho (2009) has worked more recently to update Guthrie's classification system, culminating in the *New Updated Guthrie List* (Maho 2009), which is referenced in this work to give a unique identifying code to each of the languages being studied.

A.E. Meeussen is another central figure in mid-twentieth century Bantu studies whose work as the head of linguistics at Tervuren in Belgium had an impact on Bantu classification. Those who followed in Meeussen's footsteps at Tervuren proposed the addition of zone J to Guthrie's classification based on the close relationship between many languages in the Great Lakes region of central and eastern Africa. All of the new zone J languages were originally part of Guthrie's zones D and E, which is why Maho's *NUGL* references these languages with two letters (JD or JE). The J indicates that the language is now considered part of zone J while the E (or D) indicates the zone of reference in Guthrie's original classification (Maho 2009: 7). The relevance of the work at Tervuren to this present study is that the Mara languages are currently classified as part of zone J. Since they were formerly considered part of Guthrie's zone E, all of the Mara languages start with the code JE. For instance, Kabwa is classified as JE405 (2009: 62).

The development of this classification system was one influential contribution of Guthrie's for Bantu linguistic studies; however, as mentioned earlier, the system has not proven

to be a reliable genetic classification of the Bantu languages. In spite of this, Guthrie did increase interest in Bantu historical-comparative studies. Understanding some of this history is relevant to my current study since it provides some context for the genetic classification of Great Lakes Bantu, which will be covered in depth in Section 1.5.2.

The highlight of Guthrie's career was the massive four volume *Comparative Bantu* (1967-71) that provided about 2500 "comparative series" of data, "each representing a set of words or bound morphemes from different languages connected by regular sound correspondences and identical meaning" (Schadeberg 2003: 144). Based on this data, Guthrie surmised that an initial Proto-Bantu[8] (PB) split into a western branch (PB-A) and an eastern branch (PB-B). Under PB-A were two coordinate sub-branches which split the languages of zones A, B, and C from the languages of zones H, R, K and L. There were also two sub-branches of PB-B. The first of these included the current Great Lakes languages (from zones D and E) with languages of zones M and S. The other sub-branch of PB-B was composed of languages from E50 as well as zones F, G, N and P (Hinnebusch 1989: 455-456). Several scholars in the early 1970's, including Henrici (1973), used lexicostatistics to present evidence against Guthrie's model of discrete splits. These new models proposed different versions of a right-branching descent from PB that tended to split Guthrie's zones A, B, and C from the rest at the highest branch in the tree (Hinnebusch 1989: 456; Mould 1976: 21-22). Additionally, some of the work at Tervuren showed that the Great Lakes languages (now zone J) are more closely linked, at least lexically, with the zone F languages (Bastin et al. 1983: 190; referenced in Schoenbrun 1990: 139-140).

[8] One thing that has been noted about the early reconstructions of PB is that they were skewed towards Eastern Bantu languages and that they did not include a sufficient sample of the languages in the northwest of the Bantu area, including Guthrie's zones A, B, and C (Mould 1976: 38-39).

As work proceeded into the early 1980's, a few studies began to compare phonological and morphological features of Bantu languages (i.e. Bastin 1983), but much of the early work in Bantu historical-comparative linguistics focused on the lexicon; and specifically on lexicostatistical comparison between various languages (Hinnebusch 1989: 456-461). This may have been a practical consideration based on the larger quantity of lexical data available from a broader sample of languages. Whatever the case, many Bantuists observed the limitations of relying uncritically on lexicostatistics for sub-grouping purposes. Even in the late 1980's, Hinnebusch realized the need for historical-comparative linguistic research in areas beyond lexicostatistics and he was optimistic about the possibility to make further progress. He stated that "[w]e can […] by considering lexicostatistical evidence, along with comparative studies that identify shared lexical, phonological, morphological and syntactic innovations, come up with a reliable classification" (1989: 460).

It is important to balance the results that come from lexical comparison with additional comparative data. Along these lines, Nurse and Philippson noted that:

> "[W]hile there is today widespread agreement that almost any linguistic feature or system can be transferred, vocabulary is the component of language that is most readily and quickly transferred, and over millennia it is possible for the vocabulary of any language to be so overwhelmed that its original core is hard to discern" (2003b: 166).

Lexicostatistics have a place in providing hypotheses for further reconstructive work (Dixon 1997: 37), but they must be taken with a grain of salt and corroborated by non-lexical comparisons. With that goal in mind, Nurse and Philippson were among the first to present a "comprehensive non-lexically based historical classification of the Bantu languages" (2003b: 164), but they also realized that their work was far from ideal since it lacked primary source data in many areas.

In addition to an increase in non-lexical historical-comparative work, more elegant ways to apply lexicostatistical data also emerged due to new applications from computer modeling. Holden and Gray (2006) pondered why a single unambiguous Bantu family tree has been so elusive, and they realized that the realities of the spread of Bantu languages resulted in rapid radiation, borrowing and dialect continua that tended to skew percentages of shared lexical items between languages. They employed computer software to better model these realities and to understand the probabilities of various scenarios. In this vein, they made use of Neighbor-Net (Huson & Bryant 2004) which is "an agglomerative method for constructing networks that selects taxa on the basis of similarity and groups them together" (Holden &Gray 2006: 25). The Neighbor-Net algorithm, when applied to percentages of shared lexical items between large numbers of languages, could be used to produce a graphic representation that more clearly pinpointed borrowing and similar phenomena. Application of these methods makes lexicostatistical data even more useful to comparative Bantuists. A more recent study by Roth (2011) triangulated results from application of the Neighbor-Net algorithm, historical/sociolinguistic data and the evidence from phonological/morphological innovations to discover a better classification of Wungu (F25, [wun]), which had previously been ambiguously affiliated within Bantu. As a secondary result of this, the contact phenomena which produced the skewing between Wungu and the nearby languages were understood much better. This type of research points not only to the validity of using better evaluative techniques for lexicostatistical data, but also to the need for interpreting quantitative methods in the light of insights from other linguistic and extralinguistic disciplines. For this present study, it is important to mention here that Schoenbrun (1990), the major classificatory work on Great Lakes (of which Mara is a sub-branch), has specifically focused on gathering results from a variety of linguistic and

non-linguistic domains. The breadth of his work is impressive and I will relate it to this current study in more detail in Section 1.5.2.

As a conclusion to this brief historical overview of Bantu historical-comparative linguistics, the following quote still seems particularly relevant in light of my current research. Hinnebusch urges for linguistics to go "beyond classifications based on lexicostatistics to classifications based on work that constructs the outlines of proto-intermediate groups, and the development of complementary evidence from other fields" (1989: 454-455). This thesis is aimed at going beyond the lexical component of historical-comparative work to look at an area of comparative morphosyntax that provides evidence to refine claims about "proto-intermediate groups" like Proto-GL and Proto-Mara.

1.3.2 Clarification of Historical-Comparative Linguistics

With the discussion of Bantu historical-comparative linguistics fresh in mind, a clarification needs to be made regarding the aims of this type of comparative work. In an article on comparative linguistic studies, Haspelmath argues for a view called "categorial particularism, that is, that grammatical categories cannot be equated across languages" (2010: 681). This view would appear to negate the possibility for cross-linguistic comparison of any type, but Haspelmath clarifies that carefully constructed "comparative concepts" allow for typological comparison (2010: 665-666). These "comparative concepts are typologists' constructs, not part of the structure of languages" (2010: 666). On the other side of the spectrum are descriptive linguists who work with the categories of a specific language and these must be created based on the facts of each language, so they are not comparable with the categories in other languages. The goals of descriptive linguistics and typological linguistics are distinct, which sometimes causes confusion between the two. Haspelmath concludes this theme by saying that "[t]he

analysis of particular languages and the comparison of languages are […] independent of each other as theoretical enterprises" (2010: 682).

It should be stated that this present research is not meant to be a descriptive linguistic analysis of any of the specific Mara language varieties. Since I gathered data from five different languages, a full description of all of them was beyond the scope of this thesis. This research is primarily a comparative analysis, and as such, some "comparative concepts" are employed to facilitate the search for similarities and differences between the Mara languages. Although this thesis also aims to provide preliminary descriptions of the TA systems in the Mara languages, it is important to note that these descriptions are analyzed in light of how they compare to the other TA systems. However, at this juncture it is also important to state openly that the categories used in this thesis have not been created for the purpose of typology, which is the goal of the "comparative concepts" in Haspelmath's reasoning.

A distinction needs to be made between historical-comparative linguistics and typological-comparative linguistics.[9] Historical-comparative linguistics focuses on sub-grouping related languages and reconstructing proto-languages as opposed to typological-comparative linguistics which is concerned with what is common and what is possible in the languages of the world. In fact, it is more likely that a typological-comparative linguist will aim to compare languages that are not historically related. These different goals mean that the idea of "comparative concepts" in historical-comparative linguistics must be different than the conception of them in typological-comparative studies.

For the purposes of this study, I have referred to previous research like Nurse 2008 (which is the focus of much of Section 1.4) to gain insight on TAM categories that are already

[9] The comments of Jamin Pelkey on this distinction were invaluable to my own thinking and I am indebted to him for this.

understood to be widely applicable to Bantu languages. It is still true that each language within the family is unique and has its own categories, but the similarities between the languages that stem from a common linguistic heritage provide a basis for comparison. Specifically, the purpose for establishing different comparative categories is to observe if there are any shared morphological realizations within a specific category. From a historical-comparative perspective, evidence for genetic similarities begins to build when there are morphophonologically related forms from at least two languages representing the same category, or even semantically related categories. The purpose of establishing TA categories in historical-comparative linguistics is more about providing a mental framework for understanding possible paths of semantic change from an earlier phase in a language's history to a later one. The measure of success for "comparative concepts" in historical-comparative linguistics is the extent to which they allow plausible hypotheses to be formed about language change.

Although the preceding paragraphs discuss my reasoning for focusing on comparative categories rather than descriptive language-internal categories, I realize that this approach is unsatisfying to descriptive linguists. In order to help bridge the gap to some extent, I have included introductory descriptions of how TAM functions within the individual systems of core Mara languages in Section 1.5.4. Additionally, it is my hope that this thesis will also encourage further research into the Mara languages that culminates in a full grammatical description of each one of them.

Having clarified the aims of this current research, I now move on to explain more about Bantu typology and specifically the predominant ways that Tense and Aspect are indicated as part of the verb.

1.4 The Bantu Languages and Study of Tense/Aspect

Even though the Bantu language family is vast, its internal typological cohesion is quite strong with the common exception being the northwest Bantu languages which "did not innovate or have lost some of the features" (Nurse & Philippson 2003a: 7). The following typological overview is dependent on Nurse and Philippson (2003a: 7-10).

From a PB system of seven vowels, most current languages either retain this or have reduced it to a five-vowel system. A phonological contrast between short and long vowels is not uncommon. The consonantal system of PB had either 3 or 4 places of articulation with voicing distinctions; "[r]emarkable features were clusters of nasal and homorganic stop […] and a general lack of fricatives" (2003a: 8). The modern seven vowel languages especially have kept many of these features. Most of the Bantu languages are tonal, distinguishing High and Low tone levels, with frequent application of tonal processes like downstep and spreading. Morphologically, the "Bantu languages are agglutinating" (2003a: 8) with many verbal affixes. Nouns are identified by a prefix that denotes a specific noun class; adjectives, other modifiers and verbs display agreement marking with the noun class. The typical word order is Subject-(Auxiliary)-Verb-Object, but focus on specific elements can be achieved through rearranging the canonical order or "by intraverb morphemes" (2003a: 9).

This overview is only meant to highlight some of the more prominent typological features in Bantu, but the following sections will expand on the typological elements of Bantu verbs that are particularly relevant to this study on TA marking. In Bantu studies, one major recent exploration of Tense and Aspect has drawn attention to the breadth of variety within the family as well as some of its implications for historical-comparative linguistics and reconstruction. Derek Nurse's *Tense and aspect in Bantu* (2008) is a seminal work on comparative TA in Bantu and has been inspirational for my current study.

1.4.1 Nurse's Conceptual Framework

The "conceptual framework" outlined by Nurse (2008: 10-14) forms the foundation for my current study so that this research can be seen as a micro-level contribution to his Bantu-wide work. This framework consists of eight main points, as detailed in Figure 1.

Figure 1: Nurse's Conceptual Framework for Analyzing TA in Bantu (2008: 10-14)

1. "Tense and aspect form a system."
2. "Tense and aspect systems are cognitively based, not direct representations of events in the real world."
3. "Tense and aspect form an interlocking system."
4. "A discrete verbal TA form has a specific and unique range of meaning."
5. "The system is not inflexible or unchanging."
6. "Any given (single) verb form can only have one tense."
7. "Every finite verb form has aspect."
8. "Most Bantu languages encode tense on the left and aspect to the right."

In light of this current study, many of these points require further explanation. The first three points all deal with the systematic aspects of TA within a specific language. Underlying the numerous morphosyntactic forms of TA is an interlocking system between the forms for both Tense and Aspect. In the system of an individual language, the nature of time has often been construed as linear with different degrees of past or future reference relating to successive time periods from the present. Indeed, this is the way that the temporal component is generally treated in Nurse (2008: 88-94) and I have followed his lead in my research, as is evident from Table 1 in

Section 3.1 or Table 2 in Section 3.3. This approach is unidimensional in relation to Tense, but more recent research by Botne (2013) and Botne & Kershner (2008) is revealing that a "multi-dimensional conceptualization of time and cognitive space" can unravel some of the perplexing elements of Bantu TA systems (2008: 146). Botne's approach is applied in Section 3.1.3 to help classify the Anterior (ANT) aspect within the Mara TA systems. This "multi-dimensional conceptualization" points to the complexity of the underlying systems but also reinforces the systematic relationships between forms in the same "dimension".

Along these lines, Nurse's framework also points to a system that is cognitively based, allowing learners to comprehend and reproduce it. "The various verbal categories do not directly reflect the events or objects of this world, but they rather reflect human organization, human categorization of these objects and events" (2008: 12). An outside researcher can find clues to the underlying cognitive system by looking for patterns in the "morphosyntactic forms, their place in the system, and their usage […]" (2008: 12). Henry Muzale, who studied under Nurse and wrote about TA amongst the Rutara[10] languages, mentions the importance of a diverse approach when he states that "a formal or purely functional approach […] fail[s] to address the intra-paradigmatic relationships that also contribute to holding the system together both synchronically and diachronically" (1998: 21). Studying the usage of morphosyntactic forms is important, since any given form takes on extended meaning as it is used in discourse. This is a complicating factor to the study of TA, sometimes resulting in polysemy, which is one of the driving forces behind language change. The discourse analysis of TA in this research is largely contained to one narrative in each language, which is inadequate to reveal widespread patterns; however, this is an important methodological step since it gives insight into how various TA forms are used in relationship to one another. Studying discourse allows not only the systematic

[10] This is another sub-branch of GL Bantu distantly related to the Mara languages.

function of a form to be ascertained but also aids in comprehending its extended usage, which can give clues to how the language is changing.

Even though a given TA form is used in a variety of contexts, there is a limit to its range of meaning. The fourth point of Nurse's conceptual framework states that "[a] discrete verbal TA form has a specific and unique range of meaning" (2008: 13). No two TA forms are used in all of the exact same situations, even if there is overlap between the ways that two different forms are used. On occasions in which two forms are used in exactly the same range of ways, one of them is prone to disappear in the near future. This principle has implications for my study since there are sometimes verb forms that seem to be used in identical situations. The most likely case is that they do not completely overlap in meaning and some nuance of the meaning or usage has not yet been discovered. An instance of superficial semantic equivalence between two forms can be observed in Examples (11) and (12) in Section 3.1.1; however, on deeper inspection the distinction between the two forms begins to become apparent.

Nurse's fifth point is that "the [TA] system is not inflexible or unchanging" (2008: 13). Language change springs from variation as new usages of certain TA forms cause a systematic shift to occur. A proto-language, with its proto-TA system, splits into distinctive linguistic groups which follow different historical paths; each of these groups make different linguistic choices based on the variation inherent in their system and from subsequent contact with other groups who introduce new variation (Bailey 1982). For my purposes though, the fact that TA systems are in flux means that what I present in this study cannot be as broad in scope as one might hope. The limits of time and resources only allowed for a single-speaker point of view in each of the five languages. A subsequent study on speaker variation within each of these languages is likely to enhance the findings of this current study.

Returning to Nurse's conceptual framework in Figure 1, points six through eight relate to the morphosyntactic realization of Tense and Aspect in Bantu. A verb form will always have at least one aspect (but may have more), it can only have one tense, and it tends to have tense marked farther to the left and aspect farther to the right. These concepts are clarified with a better understanding of how the verb is structured in Bantu languages, which is my next topic.

1.4.2 The Structure of the Bantu Verb

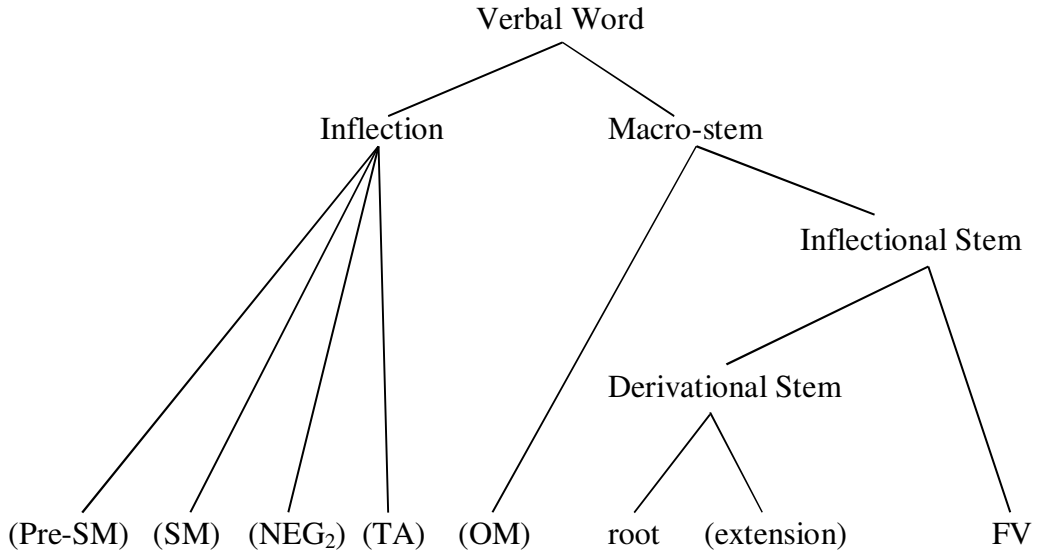
To introduce the topic of Bantu verbal structure, I am focusing on single-word inflectional constructions, which are capable of including both Tense and Aspect distinctions (Nurse 2008: 167-175). Multi-word compound constructions with auxiliaries also exist and these will be discussed further in Section 4.6. Both single-word and multi-word patterns follow the principles in points six through eight of Nurse's conceptual framework. In general, however, Bantu languages are agglutinating and much information can be stored on a single-word verb. If written with a linear template, the verbal word is formed from the slots presented in Figure 2.

Figure 2: The Slots on the Bantu Verb (adapted from Nurse 2008: 40)

(Pre-SM) + (SM) + (NEG_2) + (TA) + (OM) + root + (extension) + FV + (post-FV)

These slots will be explained more below, but first I must make clear that this linear template is a simplification of Bantu verbal structure. Based on the work of Bantu phonologists through the 1970's and 1980's, with specific attention to tonal patterning, the verbal word in Bantu came to be viewed as hierarchical and "not as a linear string" (Nurse 2008: 41). Figure 3 borrows and slightly modifies a visual model that Nurse uses to show the hierarchical structure of the verbal word.

Figure 3: The Hierarchical Structure of the Verbal Word (adapted from Nurse 2008: 42)



The bottom row of this figure corresponds to the slots in Figure 2, with the exception that the post-FV is not included. In this hierarchy, however, the slots relate to each other in mid-level and higher level groups. The main purpose of Figure 3 is to show that the concept of the verbal "stem" in Bantu is a tricky one. Any reference to a verbal stem must specify whether one is talking about the Derivational Stem (root and extensions), the Inflectional Stem (including the FV slot), or the Macro-Stem (with the addition of the OM slot).

The most basic Bantu verb form is simply an Inflectional Stem without any extensions on the verb root; this indicates that all the other slots besides root and FV are essentially optional.[11] In the Mara Bantu languages, this basic verb form indicates a Singular Imperative with *-a* in the FV (Final Vowel) slot. An example of this Singular Imperative for Simbiti is included in (1).

[11] This optionality is marked by parentheses around the slot labels in Figure 2 and Figure 3.

(1) *Tuk-a* *si-numbu!*
dig.IMP.SG-FV NC10.PL-potato
'Chimba viazi!'[12] / 'Dig up the potatoes (you, *sg.*)!'

To express any other categories on the verb, additional slots must be filled. I will explain the slots in Figure 3 starting from the left and working toward the right. The Pre-SM (or Pre-Subject Marker) slot is where some types of negation, relative objects, and newly grammaticalized morphemes (including tenses, aspects, conditionals and focus) are commonly marked (2008: 32, 40). In several of the core languages of this study (i.e. Ikoma and Simbiti), the most common morpheme in this slot is a copular nasal Focus (FOC) marker, which will be covered in Section 5.5. In addition to the Focus marking, Simbiti also marks most of its Main Clause (MC) negation at Pre-SM with the formative *te-* (see Section 5.6); the other core languages, however, have a preference for marking MC negation in the NEG_2 slot.

The SM (or Subject Marker) slot, both in wider Bantu and in the Mara languages, is reserved for the morpheme that specifies the Subject of the verb. With human Subjects the SM specifies person and number, while with non-human Subjects the SM specifies the particular noun class into which the Subject is grouped. Explicit lexical Subjects prior to the verbal word can be included for clarification, emphasis or other discourse purposes, but if the context is already understood, they are not necessary. In these cases the morpheme in the SM slot is the sole means of signifying the actor of a given sentence (Bearth 2003: 122).

The NEG_2 slot comes after the SM. In languages that express two types of negation (like Simbiti), this slot is where secondary negation (non-MC negation) is marked (Nurse 2008: 44). In these types of languages, marking at NEG_2 never co-occurs with the marking of primary negation (2008: 34). However, in the majority of the core languages for this study (except for

[12] This is the Swahili free translation, which was the original language of elicitation. All examples that were elicited through Swahili include this translation first followed by a forward slash (/) and then an English free translation.

Simbiti), this is the only slot where negation is marked morphologically. The formative *-ta-* is used to show negation at NEG_2 in the Mara languages. Negative verb forms are explained in more detail in Section 5.6.

The TA slot is one of the most important slots for this study since it is one of two primary slots where Tense/Aspect is marked in the Mara languages. In Example (2) from Ikizu, the formative *-ka-*, indicating a past tense, fills the TA slot following the SM *tɯ-* '1PL'.

(2) *Tɯ-**ka**-gaamb-an-a* *na-wɨ* *ɨkarɨ.*
1PL-P_2/P_3/P_4-speak-RECP-FV[13] CCONJ-3SG.PRON long.time.ago
'Tuliongea naye zamani.' / 'We spoke with him a long time ago.'

In broader Bantu it is possible for more than one morpheme to occur in the TA slot (2008: 34-36, 40), but synchronically, in the core languages of this study, this is not the case. As was mentioned in the previous section, if both Tense and Aspect are marked on the same verbal word, Tense will be found in the TA slot and Aspect at the FV slot.[14] An example of this from Simbiti in (3) shows that the marker of Present Tense, *-ra-*, fills the TA slot, while an aspectual marker, *-anga* 'HAB', is found at FV.

(3) *To-**ra**-tuk-**anga**.*
1PL-PRS-dig-HAB
'Tunachimba (tena na tena).' / 'We are digging (again and again).'

The slot following TA is the OM (or Object Marker) slot, which is generally only filled in two specific cases: when "the object denotes a specific human referent" or when "the referent of the object is already established as a discourse topic" (Bearth 2003: 123). In most cases, the OM slot is not filled, and this means that the formative in the TA slot is realized right before the verb root, as for *-ra-* 'PRS' in Example (3).

[13] The glossing notation P_2/P_3/P_4 shows that this formative can be used to indicate different degrees of past reference; the notation will be explained further in Section 3. This type of categorization is useful for comparing formatives across languages and is not meant to be a descriptive category for an individual language.
[14] However, it is possible for Aspect to be marked at both the TA and FV slots.

The "root" slot is reserved for the verb root and this slot is always filled, even for forms that have very little morphological marking, like the Singular Imperative in Example (1). The "extension" slot follows the "root" and consists mainly of "valency-changing derivational categories" (Nurse 2008: 37) like the causative, applicative, reflexive, reciprocal, stative and passive. Some of these categories can co-occur, and so more than one morpheme is possible in this slot.

Occasionally a verb root gets lexicalized with one of its extensions due to frequent usage and takes on a slightly unpredictable meaning. In these cases, it is no longer treated as a root with an extension; it has now become a new root morpheme that is even capable of taking new extensions. There are several places in the texts in Appendices B-F where such lexicalization is noted. Over a long period of time this can even lead to a situation where the original bare verb root (without the extension) is not found synchronically in the language. This is an area for further study amongst the Mara languages as there are some rich clues about past language change in lexicalized verb roots.

After the root and any extensions is the FV (or Final Vowel) slot. This is also a very important category for this study, since several TAM formatives are marked here. The most common formative that fills this slot is the vowel *-a* (usually indicating an indicative mood), which is why the slot is referred to as the Final Vowel. Another common formative is the subjunctive marked by *-e* or *-ɪ* (2008: 37, 40), which will be the topic of Section 5.1. In the core languages for this study, there are also several morphemes consisting of more than a single vowel that are quite common in this slot. I have already shown one of these formatives, *-anga* 'HAB', in Example (3), and it will in focus again in Section 4.3. Additionally, another longer formative found at FV is *-ire* 'ANT', which will be discussed more in Section 4.1.

Although not very productive in Mara, the post-FV slot is another place where newly grammaticalized material is often situated. In wider Bantu, this includes a 'plural imperative' morpheme, locatives, object pronouns, negatives and some TAM categories (2008: 39-40). For the core Mara languages, the only evidence of the post-FV slot being filled is with the formative *-hë*, which signals primary negation in Simbiti. Since this slot is not used widely in the Mara languages, it is not a central focus of my study; but it is nonetheless interesting for broader comparative work.

Up until now, I have been using the term "formative" without much explanation. As I showed in Example (3), when the Present Habitual is indicated in a language like Simbiti, it includes two formatives: *-ra-* (indicating Present Tense in the TA slot), and *-anga* (indicating Habitual aspect in the FV slot). This is repeated again as Example (4) with both of these formatives in bold for ease of reference.

(4) *To-**ra**-tuk-**anga**.*
1PL-PRS-dig-HAB
'Tunachimba (tena na tena).' / 'We are digging (again and again).'

Although each of the two formatives (*-ra-* and *-anga*) is an autonomous morpheme that can function as part of other verb forms as well, when taken together they indicate that this is a Present Habitual verb form. Any given "formative" can be used in a number of different verb "forms", which is why the terminological distinction is important. The way the formatives function together also helps to determine the underlying TA system in a given language.

Establishing the slots on the Bantu verb helps provide a descriptive framework that links the Mara languages to the rest of Bantu, but Nurse's study on *Tense and Aspect in Bantu* is not limited to typological considerations; it is also useful for its insight into historical change in Bantu.

1.4.3 Historical Implications of Nurse's TA Study

In Chapter 6 of *Tense and Aspect in Bantu*, Nurse compiles all his data to make proposals about categories and formatives that are traceable to PB (or at least to "early Bantu") for each of the slots on the verb. This culminates in a reconstruction of the basic PB TA system (2008: 279). Although some of the reconstructed formatives are more tentative then others, the overall picture is one of a coherent system that provides a strong hypothesis for further historical-comparative work. Chapter 7 then proceeds to analyze some of the more important "processes of change" over the course of historical development in Bantu TA systems in comparison with universally established grammaticalization paths for Tense and Aspect markers.

The historical insights from his work will be referenced throughout this thesis, but especially in Section 6, where I discuss the implications of the TA data for sub-grouping and gaining a better understanding of some of the historical relationships between different Mara speech varieties.

From comparative work at a Bantu-wide level, I now turn to the work that has been done amongst the Mara Bantu languages and provide further background on the language situation in the Mara region of Tanzania where the five core languages of this study are spoken.

1.5 Language Situation in the Mara Region of Tanzania

The Mara region of Tanzania is located just south of Kenya on the eastern side of Lake Victoria and is quite densely populated. The linguistic situation in the region is quite complex and deserves a fuller treatment, which will be covered in this section. First, I will discuss both non-genetic and genetic classifications into which the languages of the Mara region have been fit. This includes the way they have been classified by the *NUGL* in Section 1.5.1 and their genetic classification in 1.5.2. Following that, I give an overview of some of the previous literature that has been produced relating to the Mara languages (1.5.3). Then, a brief description of the

functioning of TAM systems in individual languages is provided (1.5.4). This is followed by Section 1.5.5, where I mention the usage of Swahili in Tanzania and how that influences gathering data in the country. In the final section (1.5.6), I pose some research questions that are worthwhile to consider when studying TA in the Mara languages.

1.5.1 Guthrie's Classification

Based on Guthrie's geographically oriented system of Bantu classification that has more recently been expanded by Maho (as explained in Section 1.3.1), JE40 is the group that includes all of the languages in this study. In addition to the seven languages covered in this thesis, there are a number of other speech varieties that fall into the JE40 referential group. The speech varieties that are listed by Maho as part of JE40 (2009: 62) are as follows: Ngoreme[15] (JE401, [ngq]), Ikizu (JE402, [ikz][16]), Suba[17] (JE403, [sxb][18]), Sizaki[19] (JE404, [ikz][20]), Kabwa (JE405, [cwa]), Singa/Cula[21] (JE406, [sgm]), Ware[22] (JE407[23]), Logooli (JE41, [rag]), Idaxo/Itoxo (JE411, [ida]), Isuxa/Isukha (JE412, [ida]), Tiriki (JE413, [ida]), Gusii[24] (JE42, [guz]), Kuria[25] (JE43, [kuj]), Simbiti (JE431, [ssc][26]), Hacha (JE432, [ssc]), Surwa (JE433, [ssc]), Sweta (JE434, [ssc]), Zanaki (JE44, [zak]), and Ikoma/Nata (JE45, [ntk]).

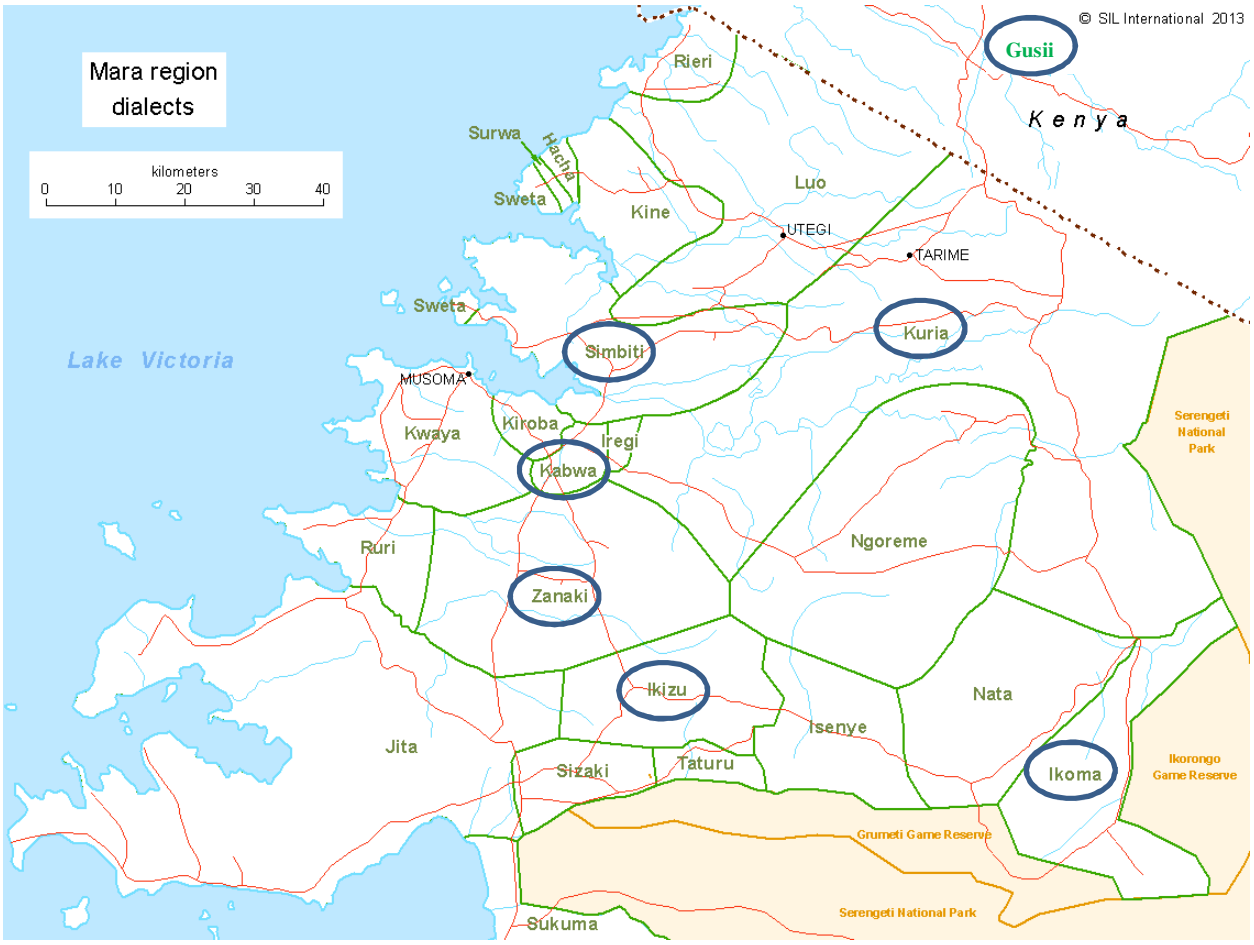
As noted above in the Introduction, the seven languages that are included in this study consist of five core languages from which I have personally gathered data and two additional

[15] Also referred to as Ngur(u)imi.
[16] This ISO 639-3 code now includes Sizaki as a dialect.
[17] This is the Suba spoken in Kenya and it includes the speech varieties Ivwang'ano, Gase, Kune, Muulu, Suuna, and Ngoe.
[18] The previous ISO 639-3 code was [suh].
[19] This is also referred to as Shashi.
[20] The previous ISO 639-3 code was [szk], but this has now been subsumed under [ikz].
[21] This is now extinct.
[22] This is now extinct.
[23] The ISO 639-3 code used by Maho is [wre] but this is not a current code.
[24] Also referred to as Kisii.
[25] Also referred to as Koria.
[26] The [ssc] ISO code is used for several speech varieties including Simbiti, Hacha, Surwa, and Sweta; this is explained further on page 29.

languages which already have published descriptions. The core languages are spoken entirely within the Mara region of Tanzania; these languages are Ikizu, Ikoma, Kabwa, Simbiti, and Zanaki. The two additional languages are spoken either wholly or partly outside of the Mara region. The first of these languages is Kuria, which is composed of a number of different clans that are located in an area stretching across the border between Kenya and Tanzania (Cammenga 2004: 15-19). The second additional language in this study is Gusii, which is spoken to the north of the Kuria area in Kenya. The map of the Mara region in Figure 4 (with highlighted speech varieties circled) gives a better idea of where these languages are located in relation to each other; this includes the addition of the approximate southern extent of the Gusii area across the border in Kenya.

Figure 4: Map of the Speech Varieties in the Mara Region of Tanzania

(© SIL International 2013)



One difficulty in classifying the speech varieties in the Mara region is determining which to treat as a distinct language and which to treat as a dialect. There is a lot of linguistic variation within the region and there are several closely related speech varieties that share either a *NUGL* classification number or an ISO 639-3 code. A number of these classification difficulties are important to this current work.

First, the three varieties Ikoma, Nata, and Isenye[27] are all grouped under the ISO code [ntk] and the lexicostatistic similarity has been measured in the 85-89% range between each of

[27] This is also referred to as Issenyi. Maho (2009) does not include Isenye at all.

the three varieties (Hill et al. 2007: 44).[28] Based on a dialect survey done amongst the groups, many non-lexical similarities have also been found (Smith et al. 2008). Nata and Ikoma appear to be more similar to each other than they are to Isenye, since both the vowel system and TA system in Isenye show more distinctions when compared to those of Nata and Ikoma (Higgins 2011: 5-6). All the data included for the [ntk] varieties in this thesis comes from Ikoma.

Second, Simbiti is one of several speech varieties that are included under the ISO code [ssc] for Suba-Simbiti, and it is worthwhile to mention that there is a cultural and linguistic link between Suba-Simbiti and Kuria (Cammenga 2004: 16-25). Several of the "Suba" varieties have distinct reference labels in Maho's *NUGL* classification including Hacha, Surwa and Sweta (2009: 62). In addition to these, the most recent Ethnologue lists Kine, Kiroba and Rieri as additional ethnic groups speaking "Suba" varieties (Lewis et al. 2013). This "Suba" grouping is distinct from the Suba language of Kenya, which only shares 40% lexical similarity to the "Suba" of the Mara region (Lewis et al. 2013). All data for [ssc] in this thesis is for Simbiti.

The third point to mention is that Kuria is composed of a number of different clans (Cammenga 2004: 19) and there is potential dialectal variation amongst them. Although the clans are grouped together culturally, a full survey is still necessary to determine how closely they are related linguistically.

Fourth, Ikizu and Sizaki (also called Shashi) are both included under the ISO code [ikz] (Lewis et al. 2013). A recent survey of the two varieties has shown that Sizaki is converging towards Ikizu, but previously the Sizaki claim that there was more linguistic separation between them (Mitterhofer & Sandeen 2012). All the data in this thesis for [ikz] is from Ikizu.

The other two core languages in this study, Zanaki and Kabwa, do not appear to have substantial linguistic variation. Although Zanaki is reported to have four dialects, mutual

[28] These lexicostatistical comparisons are based on a 304-item word list.

intelligibility amongst them is high and linguistic variation is minor (Mitterhofer et al. 2009). For Kabwa, which is a smaller language group, no dialectal variation has been observed (Walker 2010).

This section has highlighted some of the linguistic variation amongst the JE40 languages in the Mara region of Tanzania, and there are many areas for further research. Even so, there are linguistic reasons, and not just geographical reasons, for grouping the JE40 languages of the Mara region together. The genetic affiliation of these languages is the focus of the next section.

1.5.2 Genetic Affiliation

Most of the JE40 languages are claimed to be genetically related to each other, but the JE41 varieties are better grouped with the Luhya (sometimes spelled Luyia) languages of western Kenya (Mould 1976); these varieties include Logooli (JE41, [rag]), Idaxo/Itoxo (JE411, [ida]), Isuxa/Isukha (JE412, [ida]), and Tiriki (JE413, [ida]) (Nurse 1999: 8). Apart from the JE41 varieties, the rest of JE40 is proposed as the Mara sub-group of East Nyanza, which is a branch of Great Lakes (GL) Bantu (Nurse 1999:8; Schoenbrun 1990).

The usage of Mara for this sub-group should be distinguished from the name of the region since the Mara region of Tanzania contains not only the Mara languages, but also the Suguti languages which form the JE25 group in Maho's (2009) *NUGL*. These include Jita (JE25, [jit]), Kwaya (JE251, [kya]), Kara/Regi (JE252, [reg]), and Ruri[29] (JE253, [kya][30]). In addition to these main groupings, there is another GL Bantu language on a nearby island in Lake Victoria called Kerewe[31] (JE24, [ked]) and it is genetically grouped with the Rutara languages (Schoenbrun 1990; Muzale 1998; Nurse 1999:8), which are mainly found on the west and south sides of Lake Victoria. The Mara region also has several non-Bantu languages including Dholuo

[29] Also referred to as Rori (Maho 2009: 60).
[30] Ruri shares an ISO 639-3 code with Kwaya.
[31] Also referred to as Kerebe (Maho 2009: 60).

[luo] (western Nilotic), and a dialect of the southern Nilotic Datooga/Taturu [tcc] language (Hill et al. 2007; Lewis et al. 2013). The relative location of most of these language groups is included in the map in Figure 4.[32]

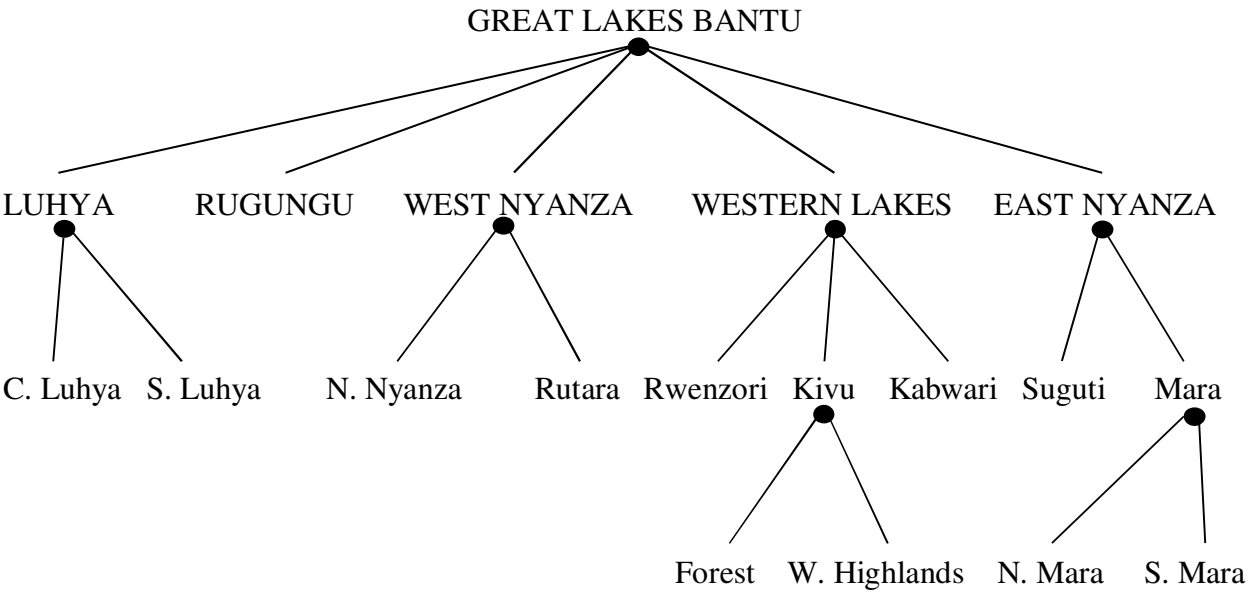
It has already been mentioned that the Mara languages are proposed as a sub-branch of Great Lakes (GL) Bantu, but up to this point, I have been using the term Great Lakes without defining it. This term generally denotes the proposed larger genetic grouping that has modern offshoots (including the Mara languages) in the "Great Lakes" area of Africa, especially the area around Lake Victoria. Schoenbrun defines the area where these languages are located as "the lands between and around lakes Rwitanzige, Rweru, Kivu, Victoria, Kyoga and the northern part of lake Tanganyika—to roughly where the Malagarasi river enters it" (1990: 9). Both the name Great Lakes Bantu (Schoenbrun 1990; Nurse & Muzale 1999) and the name Lacustrine (Mould 1976; Nurse & Philippson 1980) have been used to refer to the larger genetic linguistic grouping of Bantu languages in this area. However, since Schoenbrun's (1990) seminal work *Early history in eastern Africa's Great Lakes region* was published, the term Great Lakes has been used more frequently; it is also adopted for this thesis.

Schoenbrun (1990) makes the broadest proposal for which languages should be included in GL and how those languages are sub-grouped within GL. His research expands on several previous studies to posit five main branches of GL. He bases these divisions on lexicostatistic comparisons, shared lexical innovations, and shared borrowing from non-Bantu peoples in specific semantic domains that imply a common cultural vocabulary. The branches Schoenbrun proposes are Luhya,[33] Rugungu (a single language), West Nyanza, Western Lakes, and East

[32] The map excludes the locations of the Kara and Kerewe, since they dwell on islands to the west of the Jita peninsula in Lake Victoria.
[33] This is also sometimes spelled Luyia or Luhyia.

Nyanza (1990: 136). The chart in Figure 5 presents a schematic representation of the sub-grouping within GL Bantu.

Figure 5: Great Lakes Bantu and its Sub-groups (adapted from Schoenbrun 1990: 136)



In addition to lexically-based linguistic evidence for Great Lakes, Schoenbrun's research also gathers information from extra-linguistic sources like archaeology and palynology to paint a historical picture of the settlement of the Great Lakes area by Bantu speaking peoples. During the settlement of this region, the Bantu people came into contact with peoples from several other language families. Based on the interactions with these peoples, the Bantu people gained important cultural knowledge that helped them adjust their lifestyle to the ecological and economic realities in the Great Lakes area.

Though Schoenbrun's research is impressive in scope and includes evidence from a number of disciplines, he makes claims about the validity of Great Lakes Bantu that have been challenged. Regarding the linguistic evidence for Great Lakes, he states:

> "Supported by some thirteen innovations in the 200-word list of core vocabulary […] and several loan sets […], the genetic integrity of Great Lakes Bantu as the ancestral tongue of all Bantu languages spoken today in the Great Lakes region is almost beyond doubt" (Schoenbrun 1990: 150).

Nurse, for one, mentions that the linguistic data that has been gathered in support of Great Lakes Bantu is not "substantive" (1999: 9); this includes a lack of "nonlexical studies" covering the whole of the Great Lakes area (1999:9).[34] This state of affairs persists even though there is a large body of evidence from non-linguistic domains "suggest[ing] a unified and steady peopling of this area by a population likely Bantu-speaking, starting in the last half of the first millennium BC" (1999: 9). Nurse insists that GL is fruitful as a "working hypothesis" (1999: 7), but it should not be treated as a solid genetic grouping until more comprehensive reconstructive work can be done on the proto-language (1999: 9). Though this type of reconstruction is still an endeavor for the future, there is reason to be hopeful that it will yield beneficial results. Even if the entire set of current GL languages cannot be proven as a genetic unit, "[…] it seems likely that large parts of [GL] will ultimately cohere genetically" (Nurse & Muzale 1999: 518).

One of the languages that may eventually be excluded from GL is Gusii. It shares many things in common with the Central Kenya (CK)/Thagicu languages which are "impossible to explain by diffusion, as Gusii and CK have likely not been adjacent for some two millenia" (Nurse 1999: 28). One specific feature related to TA that connects Gusii with Central Kenya/Thagicu is explored further in Section 3.1.1. Gusii is of particular interest for this thesis because it is classified genetically as one of the Mara languages. However, due to the doubts raised regarding its genetic affiliation, it is treated as a marginal member of Mara for the purposes of this research.

[34] Schoenbrun expresses his intention to support the lexical evidence for Great Lakes with insight from "phonological and morphological innovation" (1990: 106), but actual data in these areas is lacking.

The preceding discussion emphasizes that the genetic linguistic unity of Great Lakes has yet to be proven and that work is still needed to confirm which languages are part of this grouping. However, many of the branches underneath Great Lakes (see Figure 4) are considered stronger genetic groupings by Schoenbrun based on group averaging of lexicostatistical data (1990: 114-115). Although East Nyanza is considered a "weak" grouping, Mara is considered "strong" since "its internal average exceeds all its external percentages" (1990: 115). Both North Mara and South Mara are considered "very strong" since "all [their] internal percentages exceed all [their] external percentages" (1990: 115). It is a goal of this thesis to provide evidence from TA that will either support or cast doubt on these sub-branches of GL, especially because there has not been much comparative work done amongst the Mara languages up until now.

Although this section provides some background on the genetic classification of the Mara languages within Great Lakes Bantu, I realize that not much has been said regaring research done within these languages. An overview of previous literature on the Mara languages is the topic of the next section.

1.5.3 Previous Literature on the Mara Languages

Until recently, most of the academic work on the Mara languages has been focused on Kuria and Gusii, which are spoken by a considerably larger number of people than the other Mara languages (Lewis et al. 2013). The initial work on Gusii, especially in relation to TA, was done by W.H. Whiteley (1960). Since that time, Cammenga (2002) has written more on the phonology of Gusii, which has included further study on some of the verbal morphology like TA marking. Cammenga's (2004) *Igikuria phonology and morphology* is also the main source for Kuria. Even though this also focuses on phonology, one of the final sections of the book features many prominent TA forms and their morphophonological realizations. Additionally, a verbal tone analysis of Kuria that relies on Cammenga's TA forms was done by Mwita in 2008.

Beyond these two languages, other Mara varieties have only received sporadic attention in the literature. More recently, though, the initiation of a language development project amongst several of these groups by SIL International has been spurring on research efforts. Each of the five core languages in this study has been the focus of orthography development and this work has produced some minor documents detailing many important phonological and morphological features of the languages. These "orthography sketches" are for Ikizu (Gray & Smith 2010a), Ikoma (Higgins 2010), Kabwa (Kiraka & Walker 2010), Simbiti (Compton & Walker 2010) and Zanaki (Gray & Smith 2010b). In addition to this, an overview of the grammars of Ikoma (Robinson 2013) and Zanaki (Futakamba et al. 2013) are in process; and documents detailing narrative discourse features have been produced in several of these languages including: Ikizu (Gray 2011a), Kabwa (Walker 2011), and Zanaki (Gray 2011b).

More detailed accounts of features in specific Mara languages include an MA thesis on Ikoma's vowel system (Higgins 2010), an article on nominal tone in Ikoma (Aunio 2010), another MA thesis describing Ikizu's locatives (Gray 2013), and a third thesis comparing discourse features in Ngoreme (JE401), Ikizu and Kabwa (Rundell 2012). Yet another work must be mentioned in regard to research amongst the Mara languages, although it is more of an anthropological and historical source than strictly a linguistic one. Jan Bender Shetler's (2003) *Telling our own stories: Local histories from south Mara, Tanzania* brings together ethnic histories from a variety of local sources and is an important academic source referring to the cultural dynamics in the Mara region of Tanzania.

In addition to published work on the Mara languages, the Uganda-Tanzania branch of SIL International keeps a computer database with a growing corpus of texts as well as a lexicon of around 2000-2500 items for each of the five core languages.

1.5.4 Brief Description of Individual TAM Systems in the Core Mara Languages

With the classification of the Mara languages in mind and a brief overview of the previous literature that has focused on them, it is important at this point to give a brief description of the TAM systems in each of the core languages. Although complete TAM descriptions are not available for the core Mara languages, it is important from a descriptive standpoint to get an idea of how each system functions, before explaining them in terms of comparative categories in Sections 3, 4, and 5.

The Zanaki language has a brief, unpublished description written on some of the TAM features in that language (Futakamba et al. 2013: 5-6). The primary thing to note about the Tense system in Zanaki is that there is a simple system of only two tenses distinguishing between Past (Section 3.1) and Non-Past (Section 3.3). Besides the tense distinction, there are one-word verb forms for Anterior (Section 4.1), Past Anterior (Section 5.7), Habitual (Section 4.3), Subjunctive (Section 5.1), Imperative (Section 1.4.2) and Conditional (Section 5.2), which is referred to as a "Simultaneous Subordinate" (2013: 5). In addition to these categories, there are also forms indicating Persistive (Section 4.4), Hypothetical (Section 5.3) and what appears to be Past Habitual (Section 3.1.1). These are discussed in this thesis, but are not well understood yet. There is one Present tense copula and three Past tense copular forms that can be used as auxiliaries to combine with a main verb forming three types of compound constructions. These compounds contrast between "duration", "process", and "non-duration" (2013: 5-6; see also Section 4.6). There is a Progressive verb form that can be included as part of compound verbs (Section 4.6).

Ikizu has a very similar system to the one found in Zanaki, but a fuller description is needed to understand where differences exist between the systems. Basically, the same two-way distinction between Past (Section 3.1.1) and Non-Past (Section 3.3.1) tenses holds for Ikizu. In

addition to the two tenses, there is a form not directly related to time that indicates a Vast Present (Section 3.3.4) as well as general categories like Anterior (Section 4.1), Persistive (Section 4.4), Hypothetical (Section 5.3), Habitual (Section 4.3), Subjunctive (Section 5.1), Conditional (Section 5.2), and Imperative (Section 1.4.2). For compound forms, Ikizu has at least two auxiliaries (see Table 16) indicating Non-Durative and Consecutive, although it is expected that further research would uncover a Durative category as well. Ikizu also has a Progressive form that can emerge on compound verbs (Section 4.6).

Ikoma also has an unpublished description of some of its grammatical features, including many of the prominent TAM distinctions in the language (Robinson 2013: 6-12). Ikoma has two basic Tense distinctions: Past (Section 3.1.1) and Non-Past (Section 3.3.1). In addition to the tense distinctions, Ikoma has single-word verbal categories (2013: 6-10) marking Anterior/"Perfect" (Section 4.1), Past Inceptive/"Incipient" (Section 4.5), Narrative (Section 3.2), Habitual (Section 4.3), Subjunctive (Section 5.1), Conditional (Section 5.2), and Imperative (Section 1.4.2). There is also a null-marked verbal form that is used for relating subsequent events to the TAM category established on a preceding verb (Section 3.3.4). There are three copular forms for marking the Past, Present and Future (2013: 10-11); the Past and Future forms can be used as auxiliaries in compound constructions to combine with main verbs marked in the Progressive/"Continuous", for example (2013: 8-10; see also Section 4.6 and Table 14).

The Kabwa language has a larger number of Tense distinctions than most of the other core languages. There are two pure Past tense forms: the first indicates an event that occurred yesterday (Section 3.1.2) and the second is for events that occurred before yesterday (Section 3.1.1). There are also two pure Future tenses: the first is for events that occur either later today or tomorrow (Section 3.3.2) and the second refers to events that occur after tomorrow (Section

3.3.1). In addition to these pure Tenses, there is an Anterior aspect being used to indicate the very near past (Sections 3.1.3 and 4.1) and a Future Inceptive form that can also be used as a very near future form (Sections 3.3.3 and 4.5). There is also a Vast Present form to indicate an event that is not directly associated with a temporal component (Section 3.3.4). Other types of single verb TAM forms are Habitual/Imperfective (Section 4.3), Persistive (Section 4.4), Subjunctive (Section 5.1), Imperative (Section 1.4.2), Hypothetical (Section 5.3), Progressive (Section 4.2) and Narrative (Section 3.2). Compound verbs are generally formed with one type of copular auxiliary in the Past tenses and another type of copular auxiliary in the Future (see Table 12). These compounds are used to indicate a non-Present Progressive or a non-Present Anterior (Section 4.6).

Simbiti has perhaps the most complex TAM system of the core Mara languages and deserves more detailed further study. As a general overview of the system, there are two pure Past tenses, three Future tenses and a Present tense (Section 3.3.4). The distinction between the two Past tenses is not yet well understood, but one Past appears to be used for events that took place earlier today and the other for events that occurred before today (Section 3.1). Additionally, there is a general Past tense that does not indicate a specific time (Section 3.2.2) and a Narrative tense (Section 3.2.1). Of the three Future tenses, the first form is used for events in the Immediate Future (Section 3.3.3), the second form is used for events that will take place later today (Section 3.3.2), and the third form indicates something will likely occur after today (Section 3.3.1). The other single-verb TAM forms that are found in Simbiti are Anterior (Sections 3.1.3 and 4.1), Habitual/Imperfective (Section 4.3), Persistive (Section 4.4), Hypothetical (Section 5.3), Subjunctive (Section 5.1), Imperative (Section 1.4.2), and Hortative (underneath Table 13). There are two different ways to construct compounds in Simbiti; one of

these is via the standard auxiliary-main verb ordering and the other has an auxiliary inversion so that it appears with the main verb first and then the auxiliary. Past Progressive/Anterior and Present Progressive are formed via the inverted auxiliary structure (Section 4.2), while Future Progressive/Anterior is formed with the standard auxiliary-main verb structure (Section 4.2).

1.5.5 The Influence of Swahili in Tanzania

One other thing to mention about the language situation in the Mara region of Tanzania is the influence of Swahili within the country as a whole. Swahili was promoted as the language of unification in the days when Tanzania was first establishing itself as a sovereign nation independent from British colonial rule. Swahili has a lot of prestige within Tanzania since it is frequently mentioned as one of the main reasons for the peaceful coexistence of people from over 120 different tribal groups.

Linguistically, the influence of Swahili is pervasive and it is used in interactions between people of different tribal groups. Since it is used as a prestigious second language by most Tanzanians, Swahili also has an increasing influence on the spoken local languages. Some of the implications of Swahili language usage in relation to the data in this thesis are discussed in more detail in Section 2.3.

1.5.6 Research Questions

In approaching TA research amongst the Mara Bantu languages, I have several questions that I want to explore; the preceding discussion on Bantu language classification provides important background for understanding these questions.

First of all, it is important to understand if Mara is a unified genetic grouping within GL Bantu. In other words, is there truly a Proto-Mara language phase that is the ancestor of all the modern Mara languages? In the search for features that can be traced to the hypothetical Proto-Mara phase, I treat Gusii as a marginal member for the reasons stated in Section 1.5.3. Although

TA research by itself cannot prove or disprove the existence of Proto-Mara, it can help identify common features that can be traced back to an earlier phase of language development. These in turn will serve as further evidence for work done in reconstruction amongst the Mara languages.

Secondly, based on Schoenbrun's (1990: 136) division of Mara into both North Mara and South Mara branches, it is important to see if these sub-branches are supported by innovations in the TAM systems. The question can be posed as: Are there innovations within the TAM systems of certain subsets of Mara languages (North Mara and South Mara) that point to a more recent shared inheritance between them? In order for both this question and the first question to be answered reliably, it is important to develop a comparative methodology that will help me analyze the data and pinpoint shared inheritance between languages. This topic will be explored in further detail in Section 2.

1.6 Overview of Research

In order to build a case for the conclusion that the current TAM systems in the Mara languages do not support the reconstruction of a Proto-Mara TAM system which is distinct from earlier GL systems, the following sections will explore the TAM systems of the Mara languages in detail. Section 2 explains the philosophy and methodology that underlie this study, including a description of the data collection procedures. In Section 3, the data on tenses in the Mara languages are presented and the implications of their significance are detailed. Following that, Section 4 turns to a description of aspectual categories and their relevance to understanding relationships between the Mara languages. The remainder of the verbal categories, such as mood/modality, negation and focus are the topic of Section 5. After looking at the data in Sections 3-5, Section 6 collates the insight gained from the research and assesses its implications for sub-grouping amongst the Mara languages. Some additional questions are also posed in that

Section to help guide future studies amongst the Mara languages. Finally, Section 7 presents a summary and conclusion to this study.

2 PHILOSOPHY AND METHODOLOGY

In order to appreciate the results of this thesis, it is important to understand how comparisons of the TAM systems from different languages are helpful for reconstructing a picture of the past. There has been much important work done in historical-comparative linguistics that establishes important terminology and provides philosophical insight to set a foundation for this present work. Some relevant insights are expounded in Section 2.1; these facilitate the search for individual-identifying innovations as the key to sub-grouping linguistic varieties into genetic groups. In Section 2.2, I summarize how the data for the present study was gathered with the goal of maximizing its value for comparative work. Finally, Section 2.3 discusses some of the advantages and limitations of the methodology I employed to gather data for this study.

2.1 Philosophical Underpinnings and Establishing Individual-Identifying Innovations

Various linguists have pointed out (e.g. Heine & Kuteva 2005) that there are a number of reasons why two languages could share similarities; and these "may be due to universal principles of linguistic discourse and historical development, to shared genetic relationship, to parallel development or drift, to language contact, or simply to chance" (2005: 2). Within a language family like Bantu, which has already been established as a coherent genetic grouping, it may be even more difficult to ascertain which type of development is responsible for a similarity between two languages. However, if two languages share a phonologically related morpheme or set of morphemes to mark a similar grammatical concept, the possibility that the two languages inherited it from a shared phase of language evolution should be considered.

As with other areas of a grammar, the morphophonological realization of TAM categories in a specific language could have come from a variety of sources. They can be borrowed from another language, they can develop following universal grammaticalization paths, or they can even develop by drift or chance. All these possibilities complicate any attempt at genetic subgrouping that is undertaken. It is therefore important to have some underlying philosophical principles that will guide comparative work. As I mentioned above in Section 1.4, sharing terminology and a conceptual framework with previous comparative TAM work in Bantu is already an advantage for historical-comparative research. One of the benefits of this is the ability to relate the data in this thesis to reconstructions (or at least comparative generalizations) of TAM systems at various phases in the spread and development of the Bantu languages.

I have mentioned contact-induced phenomena like borrowing as one of the major impediments to determining genetic similarities between languages, and this topic deserves further consideration. One of the main goals of historical-comparative linguistics is to determine genetic relatedness between the various linguistic varieties being studied and this often includes the drawing of linguistic trees to model which languages descended from the same proto-language. These trees imply that there are genetic relationships between various branches at higher and higher levels until the proto-language of an entire linguistic family is reached. The family tree model has been criticized since it "is not applicable everywhere and cannot explain every type of relationship between languages" (Dixon 1997: 29). Genetic relationships between languages are often obscured by language contact phenomena that have arisen as an individual language exists alongside other languages in a variety of relationships. LaPolla (2009: 228) mentions three types of influences in this regard: substratum, superstratum, and adstratum. In

some situations where languages come into contact, one language group is more dominant than the other sociolinguistically, and eventually all the speakers of the less dominant language end up shifting to the more dominant one. In these situations, however, there is usually a substratum of the language that has been subsumed within the language that remains, especially if there was a sizeable population speaking the less dominant language. A superstratum is the opposite type of scenario, where the language of the dominating people is lost because the common language is more popular; but many ideas from the language of domination seep into the common language and become the superstratum of that language. This is precisely what happened when the Norman French rulers shifted to English, which resulted in a strong French substratum in spoken English. The third type of situation occurs when two languages live side by side and influence each other in both directions, but neither becomes dominant. This is adstratum influence, which can greatly affect linguistic structures. Because my current study is concerned with closely related languages from the same family, language contact from within the family and influence from various strata are more difficult to discern since the languages often start out with a greater number of similarities.

In many linguistic situations, it is not just one language in contact with another, but instead a number of languages in a linguistic area that are influencing each other. In order to help explain areal contact situations, Dixon (1997) borrows a model from biology called "punctuated equilibrium" which claims that languages living in contact generally influence each other over a long period of equilibrium. Following this equilibrium period is a relatively brief period of punctuation which transforms the contact situation. This punctuation is caused by a relatively sudden change in conditions through, for example, a natural disaster, a military conquest, or the introduction of new technology (1997: 76-83). In brief, equilibrium involves the spreading of

areal features towards a "common prototype" (1997: 4), while punctuation involves the start of new languages. Since punctuation occurs in a relatively sudden manner, it can help us understand language change like the formation of Noun classes in PB, for instance. Dixon states:

> "It is unlikely that the ancestor language began with just one prefix, then added a second one a bit later, then another a bit later still, and so on, until eventually there were sixteen. I am not suggesting that this necessarily happened all at once (although it could have), but that there would not have been more than two or three stages involved" (1997: 54).

The concept of "punctuated equilibrium" emphasizes that genetic relationships are only valid in explaining part of the influence on a given language, and therefore it is unwise to treat family trees as the primary model for diagramming language relationships (1997: 4). Although this is true, it must be mentioned that family trees were never intended to model anything besides the genetic relationships between languages. For the purpose of historical-comparative linguistics then, as long as contact-induced changes are understood separately from inherited features, the family tree model is an appropriate way to diagram those genetic relationships. In recent work on Bantu, several scholars have proposed more robust models for diagramming relationships (see Holden & Gray 2006; Roth 2011), which I mentioned above in Section 1.3. Though Dixon is critical of family trees, he affirms that the concept of genetic relationships is still valid where these can be well established by comparative work; he mentions that these relationships can be established for families like Bantu (1997: 33-34).

In addition to contact-induced changes that can influence TAM marking in a given language, another possible type of change comes from grammaticalization. Essentially, the process of grammaticalization involves a shift "from lexical to grammatical and from grammatical to more grammatical forms" (Heine & Kuteva 2005: 14). This process contains three notable mechanisms that are occurring at approximately the same time: Desemanticization involving a loss of lexical meaning, decategorialization where a word becomes a clitic and shifts

towards a closed word-class rather than an open one, and the erosion of phonetic "substance" (Heine & Kuteva 2002a: 378-379). There has been a wealth of data on grammaticalization paths that are common in disparate language families all over the world (Heine & Kuteva 2002b; Bybee et al. 1994). Many times these grammaticalization paths are directly applicable to the development of TA categories in the languages of the world; and there is even research on grammaticalization paths for TA categories within Bantu (e.g. Nurse 2008). This previous research provides a basis for interpreting whether the morphemes representing a grammatical category in Bantu have arisen recently from grammaticalization processes in a specific language or whether the realization of the morphemes is not transparently the result of grammaticalization within the language. There are several places in this thesis, including parts of Sections 3.3.3 and 4.3, where grammaticalization paths are highlighted in relation to specific TA forms to provide a better understanding of the historical processes at work.

In relation to how grammaticalization may influence an understanding of genetic relatedness between languages, there are several things that must be remembered. If two genetically related languages inherit a TA formative with the same meaning, it may develop down different paths in each of the languages (or to different distances down the same path) causing a similar phonetic realization, but different usage in the two languages. Another possibility is for two languages to start off developing a morpheme in a similar direction, but then one of the languages has a new morpheme that begins to grammaticalize with the same meaning, and this new morpheme eventually replaces the old one. In cases like this, the languages are genetically related but the surface similarities have been obscured by later development in the languages.

Although there are universal paths of grammaticalization that are frequently evidenced in the languages of the world, the internal system of a specific language can also influence how that language develops over time (Dixon 1997: 13-14). Language typology dictates many of the processes of internal change. Dixon labels the three language types as isolating, agglutinative, and fusional. These can be pictured as hands on a typological clock with isolating at 4 o'clock, agglutinative at 8 o'clock and fusional at 12 o'clock (1997: 41-42). Any individual language is moving clockwise on the clock at a certain speed, which is not necessarily at the same pace as related languages. This typology can help in genetic classification and Dixon asserts that "[t]he ideal situation is to have a proto-language that is agglutinative and modern languages that vary from agglutinative to fusional" (1997: 41). This type of situation is more transparent for comparative work, since the isolating phase often replaces evidence from previous phases on the typological clock. Fortunately, this is very close to the situation for Bantu, which has an agglutinating PB phase and modern languages which are generally still agglutinative (Nurse 2008: 28-39). Since this thesis is studying languages of the Bantu family that have similar typological features (see Section 1.4), the internal drift of the languages is expected to be similar in many ways.

Based on the preceding discussion, some of the major impediments to understanding genetically shared linguistic traits are: contact-induced changes, grammaticalization influences, and internal drift. Additionally, there is always the possibility that two languages can share features purely by chance, and any study looking to establish a genetic link between languages should develop a methodology that can rule this out.

Interestingly, in the areas where external factors cause linguistic change, they often influence specific features in a grammar; whereas systematic shifts in the grammar are much

more likely to happen internally. Whole paradigms are much less likely to be borrowed than individual categories. In this regard, Johanna Nichols refers to the concept of an "individual-identifying" innovation, which means that "all the languages having it have acquired it, ultimately, from a single source" (1996: 50). For the case of grammatical innovations, the chance that a whole paradigm shared by two languages resulted from borrowing or grammaticalization does not have a high probability. This means that a historical-comparative linguist should look for paradigm-like evidence that is shared by two or more languages as a clue to a shared innovation that sub-groups them as a branch of related languages. Along these lines, Dixon notes that suppletions, morphological irregularities, and complete paradigms are very infrequently borrowed (1997: 22). This has relevance for my current study because the best type of evidence for genetic sub-grouping amongst the Mara languages would be a complete systematic equivalence of, for instance, all the past tenses (with related forms) between two or more languages. Unfortunately, the data is rarely this straightforward.

In relation to the search for paradigm-like "individual-identifying" innovations and how these apply to studies on TAM systems, there are several challenges that must be taken into account. First of all, it needs to be considered whether Tense, Aspect and Mood/Modality are really inter-related as a system in Bantu languages, or if they are separate categories marked at similar places on the verbal word. Second, the development of new categories influences the current system and eventually causing changes to the whole system, but the way this happens is not well understood. As the system begins to shift, it is frequent for certain verb forms to essentially have a "semantic shift" from their previous meanings. In a discussion of TA though, "semantic shift" is not well defined as it seems to be distinct from semantic shift in lexemes.

In addition to these broader challenges, how does one go about searching for genetic relatedness in the Mara Bantu languages, especially if there is a lack of paradigm-like "individual-identifying" innovations? Obviously, shared innovations are still the primary way to see that languages are genetically related. If paradigmaticity cannot be found, than the next best form of evidence is an unusual shared innovation. An unusual innovation is one that is not common and does not follow any typical grammaticalization path. If several of these are found between two languages, it provides even more evidence that they are genetically related. The collective force of a core of unusual shared innovations is much the same as a paradigmatic set of features in that it is not likely to have arisen in separate languages apart from a shared genetic source. If this type of evidence can be established for subsets of the Mara languages, then it adds evidence for sub-grouping those languages together.

2.2 Data Collection Procedures

The data for this study came from my interaction with one mother tongue speaker of each of the five core languages: Ikizu, Ikoma, Kabwa, Simbiti, and Zanaki (see Figure 4 for a map of the relative locations of these speech varieties). Additional data from other Mara varieties, such as Kuria and Gusii, was obtained through previously published sources (as listed in Section 1.5.3 above).

In working with the mother tongue speakers of the core languages, I initially obtained my data in two main ways. First of all, I created a list of 91 Swahili sentences (included in Appendix G), covering a wide variety of TAM distinctions based on categories highlighted in Nurse's (2008) research. Combining insight from Nurse with my personal knowledge of some of the TAM distinctions in the Mara languages from previous orthography development work with Kabwa and Simbiti, I created the list of Swahili sentences and translated them back into English. The sentences were specifically designed to show Tense distinctions between different temporal

frames, various Aspectual distinctions common in Bantu, and some additional distinctions between positive/negative clauses and hypothetical situations. Once the list had been created, all of these sentences were elicited orally from a translation of the Swahili into each of the core languages and then transferred to writing in the orthography[35] of each individual language with the help of literate and trained language consultants. The decision to transcribe in the orthographies was a practical one based on the amount of data that needed to be processed. As the sentences were elicited orally they were also recorded on a Zoom H4 digital recorder and downloaded for further reference. I focused on re-listening to any ambiguous verbal forms in the sentences to ascertain phonological and tonal distinctions that might have been obscured by the orthographic representation of the data. Much of the data obtained through the sentence elicitations is included in example sentences throughout the body of this thesis or in the Tables in Appendix A. Where reference is made in this thesis to a related form that is shared by several languages, an example sentence from only one of the languages is included.

The initial goal of this sentence elicitation was to elicit as many TAM forms as possible. As work progressed, however, a broader goal came into focus: to gain insight into the TAM system at work in each of the languages. Understanding the TAM system is a substantial part of understanding historical changes that may have taken place. Although gaining insight into the systematic functioning of TAM in each language was the ideal outcome of the sentence elicitation procedure, and much progress was made toward this goal, the scope necessary to completely understand the systems in five languages proved to be beyond the reach of the current project due to reasons that will be made clear in Section 2.3.

[35] The writing systems in each of these languages have been in development for the past six years, but a writing tradition is still being established and some minor changes to the orthographies can still be expected.

The second type of language data obtained for this project was the recording of a spoken discourse. In order to have a comparable text across all of the core languages, I showed each of the mother tongue language consultants the wordless video "The Pear Film".[36] This film originated from a large research project at the University of California at Berkeley in 1975 that culminated in the 1980 book, edited by Wallace Chafe, called "The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production." Following a viewing of the film (or two, if necessary), I asked each language consultant to recount to me, in his or her mother tongue, the story that they had seen while I recorded the narrative. Finally, I replayed the recording piece by piece with each language consultant, and obtained word-for-word and free translations of the entire text. I have included the transcriptions of each "Pear Story" as Appendices B through F, with morpheme-by-morpheme glosses and additional footnotes to highlight interesting linguistic elements. This body of data should be useful not only to my current study but also to any further studies looking at narrative discourse in the Mara languages.

It is important to mention a few things about the transcribed data itself before proceeding. Since I was working with a large body of transcribed material (91 sentences plus a transcribed copy of a three-to-five minute text) and since phonology is not the focus of this thesis, sentences are generally transcribed using the orthography of the language concerned rather than a phonological transcription. Thankfully, the phonologies of the Mara languages are fairly transparent from the orthographic form. In instances where it is important to highlight the specific phonological processes utilized in these languages, I have included footnotes that discuss those phenomena with references to additional sources where applicable. Tone is generally not marked in the orthographies of the Mara languages other than a few isolated

[36] The film can be accessed either on Youtube at http://www.youtube.com/watch?v=bRNSTxTpG7U, or at http://pearstories.org/pears_video.htm.

minimal pairs for grammatical tone in Ikoma, which are marked by special word initial symbols. When it is merely the tonal pattern differentiating one TAM formative from another, I make special note of the surface differences in prose. However, underlying tonal patterns cannot be ascertained until fuller descriptions of the verbal tone systems in these languages have been undertaken, which is beyond the scope of my research.

Before explaining more about the advantages and limitations of the methodology, one final note on data presentation is in order. Whenever an example sentence from a Mara language is included in this thesis, it is given both a Swahili and an English translation. This is deemed necessary to highlight the fact that the elicited sentence was originally translated directly from Swahili. The English is included because it is assumed that most readers will not be familiar with Swahili. As the following section points out, sentence elicitation has its drawbacks. I hope that a transparent presentation of the data will emphasize where literal translation from Swahili is potentially obscuring some of the systematic nature of TAM in a given Mara language.

2.3 Advantages and Limitations of the Methodology

As with any methodology, there are advantages and limitations that are inherent in the way the data was obtained and this project is no different.

The advantages of sentence elicitation are twofold. First, it is a relatively quick way to gather a broad range of TAM forms from languages that have not been well described. In order to gather an equivalent number of forms through natural texts would require a large corpus consisting of various genres and types. Even though a textual corpus is maintained by SIL for each of the core languages, it is not yet large enough or diverse enough to provide an adequate body of data for a broad study of TAM. Second, sentence elicitation allows for a quick comparison of a variety of forms in several languages. Even so, this statement requires a proviso. The TAM system in Swahili is not equivalent to the systems in the Mara languages (or any other

language) and so a one-to-one correspondence between categories cannot be assumed. When eliciting a specific verb form from Swahili, there is no guarantee that the verb root or the interpretation of its TAM marking will be uniform across the languages. There are several Swahili TAM forms which can be used in varying ways, depending on the speaker, his or her linguistic background, and the context they assume for the sentence. In fact, this turned out to be one of the difficulties associated with this method.

I discussed the impact that Swahili has on local language usage in Tanzania in Section 1.5.5. One drawback of eliciting through Swahili is exemplified in the discussion of the Anterior/Perfect in Section 3.1.3. This can be summarized here. In Swahili, the Anterior/Perfect formative *-me-* is used to cover any event that occurred in the very recent past, as Example (5) shows.

(5) *Tu-**me**-onge-a* *na-ye* *dakika* *kumi*
1PL-ANT-speak-FV CCONJ-3SG.PRON NC10.PL.minute ten

zi-li-zo-pit-a.
AGR10-PST-REL10-pass.by-FV
'We spoke with him ten minutes ago.'

In this instance, it is not natural to use the verb form *tuliongea* 'we spoke' for the Past Tense. When the language consultants in the Mara languages hear a sentence like this, they have a tendency to translate the Anterior/Perfect directly rather than using an Immediate Past tense, if it is available in their language. In this way, sentence elicitation has the potential to obscure categories in the languages into which the translation is being made. This means that further discussion beyond the sentence elicitation is necessary to understand the usage of the forms as much as possible; so through the elicitation process I did discuss the forms with the language consultants.

A related limitation imposed by sentence elicitation is that the speaker is translating directly from another language (in this case Swahili) rather than focusing on naturalness in their mother tongue, and so it is possible that borrowed Swahili forms or structures are used, which do not sound natural. This is seemingly the case for the Potential "can" forms, as is discussed further in Section 5.4.

In order to overcome some of the drawbacks of sentence elicitation, a method for gathering textual data was also employed, as discussed in Section 2.2. Using a single comparable text in each of the core languages is an achievable way to provide a more natural environment for observation of TAM forms than with elicited sentences. Texts more quickly reveal real-world variation of TAM forms, as well as providing a means of observing some TAM forms that are more difficult to elicit with a stand-alone sentence. One of the TAM forms that appears in the text is the Narrative (NAR); and this topic is covered in more detail in Section 3.2.

The main limitation to textual data, as was mentioned above, is that it often takes a large amount of data to be able to draw appropriate conclusions or to even have examples of a decent number of TAM forms in the language. But, since the textual data was meant to supplement the elicited sentences, it is used in this thesis as corroborating evidence for the semantic and pragmatic range of a specific TAM form.

Although the initial data was gathered through sentence elicitation and a narrative text, as I progressed with the analysis it became clearer to me that there are a lot of gaps in the data. One of the aims of understanding TAM in Bantu is to see how the different formatives function together as a system. If the TAM systems in the five core languages can be understood synchronically, there is a lot more that can be gleaned about the relationship between those systems and the possible diachronic influences that led to the present situation. In addition,

understanding systematic TAM data overcomes the drawback stated above; that sentence elicitation is capable of leading to varied interpretations of a specific TAM formative in Swahili. I therefore decided to include the reference charts for each of the languages in Appendix A, which systematize the TAM data in a visual way, and as much as possible give examples of the TAM formatives on the same verbs 'buy' or 'dance', which have a nearly identical form in all the core languages. The idea and format of these charts were influenced greatly by Nurse (2007). As much as possible the charts are filled out, with some reliance on two unpublished documents from SIL colleagues on the ground in Tanzania (Futakamba et al. 2013; Robinson 2013), but overall there are still gaps in the charts and further research will be needed to fully describe the TAM systems in these languages. In addition, the research itself was only conducted with one speaker from each of the core languages and many more samples of each of the languages will be needed to understand the range of synchronic variation in the system. Since this thesis is only intended to be a preliminary study, I leave both of these important topics for future research.

This leads to one other important point. There is evidence in Bantu that different classes of verbs (stative, active, durative, etc.) have a tendency to interact in unique ways with a specific TAM form (Nurse 2008: 13, 73). For this reason, a complete set of data would include each of the TAM forms on each class of verb. This too proved to be beyond the scope of this present study and is yet another reason why, to the fullest degree possible, the charts in Appendix A are included showing the main TA forms with the same active verbs 'buy' and/or 'dance'.

As noted above in Section 1.4.2, most TAM forms are constructed out of multiple formatives at the TA and FV slots (and sometimes the pre-SM slot as well). In the following sections it will be necessary to look at the morphemes present in each of these slots in order to determine which TAM form is being indicated. This is especially relevant to the tables below

where there are often two morphemes bolded to refer to a specific TAM form. With this reminder, I now turn my attention to the topic of Tense in the Mara Bantu languages.

3 TENSE

Although Tense is typically defined as "grammaticalised expression of location in time" (Comrie 1985: 9), this standard definition needs to be contextualized to the Bantu situation. As I mentioned earlier, the Bantu language family employs agglutinative verbal morphology, which means that the verb has a number of slots that are typically filled by morphemes from specific categories (refer to Section 1.4.2 for further discussion). For the majority of the Bantu languages (outside of the northwestern zone that covers Cameroon and Congo), the time of an action is expressed grammatically through inflection at the pre-stem TA slot (Nurse 2008: 80; a graphic representation of this is found in Figure 2 in Section 1.4.2). Example (6) from Zanaki shows the TA slot filled by the past tense marker *-ka-* prior to the verb root *gaamb* 'speak'.

(6) *Tu-**ka**-gaamb-an-a* *na-we* *ekare.*
1PL-P_2/P_3/P_4-speak-RECP-FV CCONJ-3SG.PRON long.time.ago
'Tuliongea naye zamani.' / 'We spoke with him a long time ago.'

In addition to the *-ka-* formative in the TA slot, the past tense form of the verb is indicated by *-a* in the FV slot.

In discussing Tense in Bantu, I will follow the lead of Nurse in making a distinction between two broad types of tense: absolute and relative (2008:120; based on Comrie 1985: 56). In addition, due to the nature of the Tense systems in several of the core languages in this study (Ikoma, Ikizu, and Zanaki) it is quite natural to distinguish between tenses that have past reference and those that have non-past reference. Based on these divisions, I begin my

exploration of Tense in Mara by observing how the Mara languages mark different degrees of absolute past tense in Section 3.1. Following this, I shift my attention to relative tenses, and most importantly the Narrative Tense, which relates subsequent events to a temporal frame that has already been established on the first verb in a sequence (3.2). Finally, I delve into the non-past absolute tenses relating to the future and/or the present (3.3).

3.1 Past Tenses

In Bantu languages, it is common to have various degrees of absolute past tense (Nurse 2008: 88-90). An "absolute" tense has the present as its reference point, so an absolute past tense would cover a certain period of time prior to the present. When considering Mara languages it is important to look at four of these: Remote Past (P_4, before yesterday), Hesternal Past (P_3, yesterday), Hodiernal Past (P_2, earlier today), and Immediate Past (P_1, within the last couple hours). The 'P_N' notation is based on Nurse (2008); however, there is a difference in how he uses his notation compared to how I am using it in this research. Nurse's use of the labels depends on the language he is analyzing. For instance, Nurse would only use P_1 and P_2 for a language that has two Past Tense distinctions, with P_1 representing the time frame nearer to the present and P_2 indicating the time frame that is further in the past. In this historical-comparative study, it was deemed necessary to have a cross-linguistic way of comparing the morphemes from various languages, so each label refers to a specific period of time. This four-way division is not merely ad hoc since it is recognized by Nurse in the Gusii system (2007:127-128).[37] In addition to that, the divisions of the Gusii system resemble the most common type of Bantu TA systems with four past tenses (Nurse 2008: 91). Amongst the languages with data included in this study, only Gusii

[37] Nurse's (2007: 128) time divisions for Gusii are as follows: P_1 - "last hour or two", P_2 - "today", P_3 "between 24 and 48 prior to reference point", and P_4 "beyond 2 days". Although these are essentially the same time periods as I am using for this research, they are sometimes expressed in number of hours rather than today, yesterday, and before yesterday.

has unique forms for all four degrees of past reference. Each of the other languages has less than four, but at least two forms related to the Past. This means that some of the degrees of reference in these languages are lumped together and represented by a single form. In Example (6) above for Zanaki, the morpheme gloss for the formative *-ka-* shows that it covers the P_2, P_3 and P_4 degrees of reference, even though it is being used with P_4 reference in that particular sentence. It is important to reiterate the relevance of the discussion in Section 1.3.2 to this topic; glossing morphemes in multiple categories like this is not meant to be descriptive of the system in that individual language, but merely to provide categories that can be compared based on their morphological realizations in genetically related languages.

In addition to the four degrees of past reference already mentioned, there is also an aspectual category in Bantu languages which is referred to as the Anterior (ANT) and as the name implies this can sometimes overlap with past reference. [38] Although the Anterior will be fully dealt with as an aspectual marker in Section 4.1, many of the Mara languages have extended the usage of the ANT to mark the P_1/P_2 temporal reference. For that reason it will also be covered in relation to Tense in Section 3.1.3 below.

Table 1 shows how each of the Mara languages marks the various degrees of past reference, including the Narrative tense (NAR), which will be discussed further in Section 3.2. These degrees of past tense were elicited with adverbials or temporal relative clauses denoting a specific distance from the present.[39] The Kuria and Gusii data in the chart are from published sources that mark underlying tone, while the data for the five core languages of this study are only marked with surface tone in instances where tone is one of the main differentiators of a minimal pair (otherwise no surface tone is marked). For this reason, the core languages cannot be

[38] The Anterior is also sometimes referred to as the Perfect (PRF) aspect.
[39] The Swahili sentences used for elicitation of past tense are included as numbers 11 to 19 in Appendix G.

compared tonally with Kuria and Gusii. Although tonal data is not comparable across all the languages, the tense formatives have been written using IPA for comparative purposes. In the table, R refers to the verb root and SM to the subject marker. All formatives that make up part of a tense form are italicized and bolded in their respective slots on either side of the root.

Table 1: Various Degrees of Past Reference in the Mara Languages

		Type of Past Reference				
		P_1	P_2	P_3	P_4	NAR/"Untimed"[40]
Language	Ikizu	SM-R-***iɾɪ***[41]	SM-***ka***-R-***a***[42]			
	Zanaki	SM-R-***iɾɪ***[43]	SM-***ka***-R-***a***			
	Ikoma	***n***-SM-R-***iɾi***		***n***-SM-***V***-R-***iɾi***[44]		1. SM-***Vka***-R-***a***[45] 2. SM-***ɾa***-R-***a***
	Kabwa	SM-R-***iɾi***		SM-***aa***-R-***iɾi***	SM-***aa***-Ŕ-***a***[46]	SM-***ka***-R-***a***
	Simbiti	***(n)***-SM-R-***iɾe***		***(n)***-SM-***aa***-R-***iɾe***[47]		1. SM-***aa***-R-***a*** 2. SM-***ka***-R-***a*** 3. SM-***ɾa***-R-***a***
	Kuria[48]	SM-***a***-R-***eɾe***	***ne***-SM-R-***eɾe***	***ne***-SM-***a***-R-***eɾe***		1. SM-***ká***-R-***a*** 2. SM-***a***-R-***a***
	Gusii[49]	SM-***á***-R-***a***[50]	***(n)***-SM-***á***-R-***ɛtɛ***[51]	SM-***aa***-R-***a***[52]	***(n)***-SM-***a***-R-***ɛ́tɛ***[53]	1. SM-***ɾaa***-R-***a***[54] 2. SM-***ka***-R-***a***[55]

[40] The "Untimed" category is based on Kuria research by Cammenga (2004: 287).
[41] The same SM-R-***iɾɪ*** form is also used with P_2 temporal reference.
[42] The form SM-***aa***-R-***iɾɪ*** also occurs with P_4 reference; this overlap in temporal reference is discussed further in Section 3.1.1.
[43] The same SM-R-***iɾɪ*** form is also used with P_2 temporal reference.
[44] The V indicates an unspecified vowel which surfaces as a copy of the vowel on the preceding SM.
[45] The V in this instance can cause lengthening of the SM, but this is not written in the orthography. In addition, the NAR is only distinguished from the PST.INCE (also referred to as the Incipient) by tone (Robinson 2013: 8).
[46] The P_4 has a high-tone on the vowel of the root. This contrasts tonally with the PRS in Kabwa, which has all low tones (SM-***aa***-R-***a***).
[47] There is also a similar form ***(n)***-SM-***a***-R-***iɾe*** with a short *a* before the root, but this was not elicited through the initial sentence translation and needs further investigation.
[48] The source for the [kuj] non-tonal data is Cammenga (2004), with underlying tones obtained from Mwita's (2008) analysis of Cammenga.
[49] The source for all the [guz] data is Nurse (2007: 127-129), which is based on Whiteley (1960); underlying tones are used for all the Gusii data.
[50] This category is defined by Nurse for Gusii as taking place in the "last hour or two" (2007: 128).
[51] This category is defined by Nurse for Gusii as taking place "today" (2007: 128).
[52] This category is defined by Nurse for Gusii as taking place "between 24 and 48 hours prior to reference point" (2007: 128).
[53] This category is defined by Nurse for Gusii as taking place "beyond two days" ago (2007: 128).
[54] This form indicates that "an action occurs after the one cited" (Nurse 2007, citing Whiteley 1960).
[55] This form indicates "distance from context, NAR" (Nurse 2007, citing Whiteley 1960).

It is important to note that Table 1 deals with the tense forms for Main Clause Affirmative (MCA) verb constructions. As I will mention later on in Sections 5.6 and 5.7, the variance in forms for relative clauses and negative verbs can be important for establishing a better picture of the historical situation since non-MCA verb constructions tend to be slower to take on innovations that have entered a language (Nurse & Muzale 1999: 521).

Another important point to mention about Table 1 is that the NAR/ "Untimed" column includes multiple forms in several of the languages. In addition, several footnotes in the Table mention that alternate forms are possible for the P_2 time frame in Zanaki and Ikizu, and for the P_4 time frame in Ikizu. All these forms were elicited through my research and point toward the variation that is inherent in any TA system. In fact, they could potentially signal ways that the system is changing. These concepts are encapsulated in points four and five of Nurse's conceptual framework from Figure 1 in Section 1.4.1. Since the areas where variation is found in the Mara languages are potential loci of language change, they are critical areas for further research.[56] I will discuss specific instances of Mara language variation in the appropriate sections below.

3.1.1 Remote Past (P_4)

The Remote Past, or P_4, includes any event that occurred before yesterday (pre-hesternal). This is a temporal distinction useful for comparing the Mara languages and is not meant to be extensible to other Bantu languages. Even so, in the broader comparison that Nurse undertook, the most common temporal ranges for Bantu languages with two and three degrees of past tense reference are comparable to the Mara languages (2008: 90-91).

[56] This research would be especially interested in clarifying semantic distinctions between forms that are used in similar ways.

There are two primary verb forms amongst the Mara languages for P_4: ***(n)***-SM-***a(a)***-R-***ire***[57] (or a variant of the *-VrV/-VtV* morpheme in the FV slot) in four languages and SM-***ka***-R-***a*** in two languages (Zanaki and Ikizu). The ***(n)***-SM-***a(a)***-R-***ire*** type is exemplified for Simbiti in Example (7).

(7) *N-kare,* *tw-**aa**-shumaash-**irë*** *na-we.*
FOC-long.time.ago 1PL-P_3/P_4-speak-ANT CCONJ-3SG.PRON
'Tuliongea naye zamani.' / 'We spoke with him a long time ago.'

The SM-***ka***-R-***a*** type of P_4 has already been seen in Example (6) for Zanaki; this is repeated as Example (8) for ease of reference.

(8) *Tu-**ka**-gaamb-an-**a*** *na-we* *ekare.*
1PL-P_2/P_3/P_4-speak-RECP-FV CCONJ-3SG.PRON long.time.ago
'Tuliongea naye zamani.' / 'We spoke with him a long time ago.'

Zanaki and Ikizu use the SM-***ka***-R-***a*** form to indicate multiple degrees of past reference. In fact, it is the general past tense (and Narrative) marker in these languages.[58]

Only Kabwa does not have either SM-***ka***-R-***a*** or ***(n)***-SM-***a(a)***-R-***ire*** for P_4 but instead has a variant of the latter of these two at P_3. The form Kabwa has for P_4 is SM-***aa***-Ŕ-***a***, and this is included in the sentence in Example (9).[59]

(9) *Tw-**a**-hair-**a***[60] *na-we* *akare.*
1PL-P_4-speak-FV CCONJ-3SG.PRON long.time.ago
'Tuliongea naye zamani.' / 'We spoke with him a long time ago.'

---

[57] This form was chosen as representative of Mara because it is most similar to the form reconstructed by Nurse for PB (or early Bantu) Past Anterior: *tʊ-**a**-gʊl-**i(le)*** 'we had bought' (2008: 279). This topic will be picked up again in Section 6.
[58] Section 3.1.3 is concerned with an aspectual form that indicates the near past in Zanaki and Ikizu.
[59] The Kabwa form is similar to the P_1 for Gusii (if the Kabwa high-tone is considered); however, it is unlikely that there is a historical connection between them since the Gusii form is most likely a recent innovation.
[60] The orthographical form <*twahaira*> 'we spoke' obscures the underlying long-vowel, which is evident in the first person singular form <*naahaira*> 'I spoke'.

Constructions with an *-a(a)-* formative at the TA slot (regardless of tone) together with *-a* in the FV slot are extremely common across Bantu. In fact, Nurse has this form connected to past reference in 78% of his matrix languages and claims it as reconstructible for PB (2008: 278). It is also highly likely that Proto-GL inherited this form with "Far Past/Middle Past" reference since it is found widely in the modern Great Lakes languages (Nurse & Muzale 1999: 533).

Four of the seven Mara languages under examination have a form like this for some type of past reference, but it is used slightly differently in these languages. Incidentally, all four languages that use this form are considered North Mara languages. Gusii has SM-***aa***-R-***a*** for P_3, while Kabwa has this same form for P_4.[61] Kuria has evidence of SM-***a***-R-***a*** indicating what is referred to as an "Untimed Past Anterior" (Mwita 2008: 36, originally from Cammenga 2004: 287). "Past" means that the action is prior to the speech act; "Untimed" that the action is not directly related to time; and "Anterior" that it is "related to some subsequent point or period of time, possibly the speech act" (Cammenga 2004: 285-286). In general then, the "Untimed Past Anterior" would refer to an action that began at an unspecified time in the past and has continued relevance for a later time (usually the present). The Simbiti form SM-***aa***-R-***a***, though including a long vowel at TA instead of a short one, appears to be related to the Kuria "Untimed Past Anterior". Example (10) shows an instance of this form taken from the narrative text in Appendix E, lines Ssc7a-b.

(10) *Bhoono, ha-ke igho n-**aa**-maah-**a** o-mo-ona*
now NC16-little only 1SG-PST-see-FV[62] AUG1-NC1.SG-child

u-wö-ndë.
AUG1-NC1.SG-certain
'Sasa, baada ya kidogo, nikaona mtoto fulani.' / 'Now, a little while later, I saw a certain child.'

[61] There is a high-tone that appears to surface on the penultimate mora for this form in Kabwa: [tʷaágʲa] 'we went (before yesterday)'; [tʷaagúra] 'we bought (before yesterday)'.
[62] For glossing purposes, the *-aa-* is treated as a general PST.

For the SM-***aa***-R-***a*** of Simbiti (as well as Gusii and Kabwa) to be treated as a corollary of the SM-***a***-R-***a*** found in Kuria there would need to be an explanation of the variance in the phonological shape of the morpheme (*-a-* vs. *-aa-*). Nurse (2008: 108; citing research in Nurse & Philippson 2006) mentions that, Bantu-wide, conditioned lengthening and tones from SMs can quite easily influence the length and tone of the following *-a-* formative in the TA slot. He goes on to say:

> "Thus over millennia *-a-* developed several tonal and length alloforms. Languages seized on this as a means of expressing new grammatical distinctions, and gradually the alloforms became phonologically and semantically distinctive" (2008: 108).

Since phonological transformations of this type appear to be common in Bantu and the North Mara variants of SM-***a(a)***-R-***a*** all relate to Past events, there does seem to be a connection between them.

The three South Mara varieties, on the other hand, behave differently. In Ikoma, there is no evidence for an SM-***a(a)***-R-***a*** form (nor the seemingly related SM-***V***-R-***a***) in its TA system. In the elicited data for Zanaki SM-***a(a)***-R-***a*** did not appear either; however, my language consultant told me that a similar form indicates a type of past habitual,[63] which would mean that, if related, the Zanaki form has experienced semantic shift from a pure marker of Past Tense. The Ikizu form is also a candidate for semantic shift since SM-***aa***-R-***a*** marks the Vast (or untimed) Present, but this is tonally and segmentally similar to the Vast Present (PRS) for Kabwa and so is most likely a different form (see Section 3.3.4 for further discussion of PRS). This means that none of the South Mara languages (Ikizu, Zanaki, and Ikoma) has SM-***a(a)***-R-***a*** indicating past tense, but the Zanaki Past Habitual could be related historically. This gives a brief overview of the topic,

[63] This form was not studied in detail, but it seems that the formative at TA is probably a long vowel, *-aa-*.

but the variants of *-a(a)-* in the TA slot and their implications for reconstruction will be in focus again in Table 9 of Section 6.1.1.

Returning to the discussion of the ***(n)***-SM-***a(a)***-R-***ire*** form that is common for P_4 in Mara, Ikizu, which generally uses SM-***ka***-R-***a*** at P_4, includes a variant of ***(n)***-SM-***a(a)***-R-***ire*** as an additional form referencing the P_4 time frame. These two forms can be seen in Examples (11) and (12) which were both elicited for the same phrase.

(11) *Tu-**ka**-gaamb-an-**a*** *na-wɨ* *ɨkarɨ.*
1PL-P_3/P_4-speak-RECP-FV CCONJ-3SG.PRON long.time.ago
'Tuliongea naye zamani.' / 'We spoke with him a long time ago.'

(12) *Tw-**aa**-gaamb-**ini*** [64] *na-wɨ* *ɨkarɨ.*
1PL-P_4-speak-RECP.ANT CCONJ-3SG.PRON long.time.ago
'Tuliongea naye zamani.' / 'We spoke with him a long time ago.'

The form in (11) is the basic P_4 form for Ikizu, but the form in (12) appears to be very close in usage.[65] In general, the two forms are distinguished by their usage in different clause types with SM-***ka***-R-***a*** employed in "Main Clause Affirmatives" (MCAs) and SM-***aa***-R-***irɪ*** generally found in Relative Clauses. I will explore this distinction more in Section 5.7.

One final thing should be mentioned in regard to the form ***(n)***-SM-***a(a)***-R-***ire*** in the Mara languages. Gusii has *-ɛtɛ* at FV for this form, which is quite distinctive from what is seen in the other languages in this study. In the wider context of Bantu, there are at least several dozen languages that have *-ite* (or a variant of it) as an allomorph of *-ire* (Nurse 2008: 267), but none of these is located in the Mara area. Gusii goes beyond mere allomorphy and shows evidence for a distinct *-ɛtɛ* formative, which is quite rare within Bantu. In fact, Nurse mentions that "the *-V(:)te* suffix as a form independent from *-ile* occurs only in Gusii, Sonjo, and E50 (not E56)" (2008:

[64] This morpheme is a combination of *-an* 'RECP' and the formative *-irɨ* 'ANT'.
[65] The semantic distinction between the two forms is not yet clear.

164). This is one of the pieces of evidence that has led researchers to posit a closer relationship between Gusii and the Central Kenya/Thagicu (E50) languages[66] (i.e. Nurse 1999: 28). In spite of this, Gusii is difficult to classify because it also has quite evident connections to some of the Mara languages, specifically Kuria. The proposed link between the ***(n)***-SM-***a***-R-***ɛ́tɛ*** form at P_4 in Gusii and the ***(n)***-SM-***a(a)***-R-***ire*** form in the majority of Mara languages is worthy of consideration in regard to the historical situation. The historical development of this Gusii form is taken up again in Sections 3.1.2 and 6.

Now that I have taken a look at some of the congruence and variation in the Remote Past (P_4) category, I turn my attention to more recent past events covered by the P_3 temporal category.

3.1.2 Hesternal Past (P_3)

The P_3 references an event that happened yesterday. Of the languages in focus for this study, only Kabwa and Gusii employ distinct forms for both P_3 and P_4. The other languages treat P_3/P_4 as a combined category marked by a single form. Example (13) displays SM-***aa***-R-***iri*** at P_3 in Kabwa.

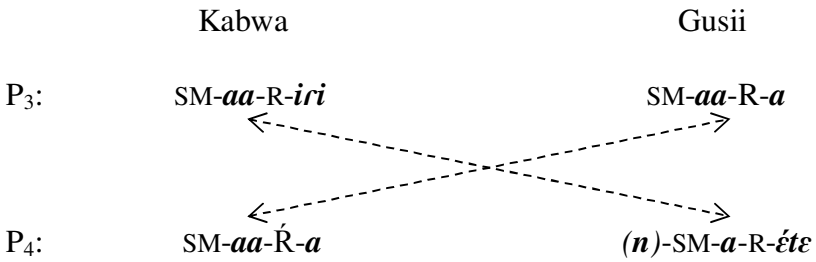
(13) *Tw-**a**-bhin-**iri***[67] *ejo.*
1PL-P_3-dance-ANT yesterday
'Tulicheza jana.' / 'We danced yesterday.'

An argument can be made that there is a relationship between the forms for P_3 and P_4 in Kabwa and Gusii, although the morphologically similar forms are in opposite grammatical categories. Where Kabwa has SM-***aa***-Ŕ-***a*** at P_4, Gusii has SM-***aa***-R-***a*** at P_3. Conversely, Kabwa has

[66] The E50 languages are generally called Central Kenya/Thagicu. Sonjo (E46) is also proposed to be genetically connected (Nurse & Rottland 1991/1992) and Gusii (JE42) has similarities that cannot be ignored (Nurse 1999).
[67] The underlying form of this is SM-***aa***-R-***iri*** but the long-vowel before the root is obscured by the orthography.

SM-***aa***-R-***iɾi*** at P_3 while Gusii has ***(n)***-SM-***a***-R-***ɛ́tɛ*** at P_4.[68] A schematic layout of the proposed relationship between opposite grammatical categories is seen in Figure 6.

Figure 6: Possible Relationships between P_3 and P_4 in Kabwa and Gusii[69]



The realizations of these forms are phonologically different, but the converse relationship in form/meaning pairings is interesting to note and begs further consideration. The possible connections between Gusii ***(n)***-SM-***a***-R-***ɛ́tɛ*** and Kabwa SM-***aa***-R-***iɾi*** are an extension of the discussion in Section 3.1.1 on the common Bantu phonological change between *-a-* and *-aa-*. I will discuss the *-a-/-aa-/-á-* distinction again in Section 6 where a visual display of the different forms in the Mara languages is presented in Table 9.

The P_4 form in Gusii, and specifically the relatively unique *-ɛtɛ* at FV, was also mentioned above in Section 3.1.1 as an innovation that links it with the Central Kenya/Thagicu (E50) languages. The *-V(:)te* suffix appears to be an allomorph of the standard *-ile/-ire* ANT formative and Nurse mentions that the two suffixes still have a traceable relationship in modern E50: *-eetɛ* "Anterior" and *-irɛ* "Perfective" (2008: 268). Gusii itself also has both of these suffixes, but they are used distinctively with *-ɛtɛ* as part of the P_1 and P_3 ("Perfective") forms and *-irɛ* as a "stative/ANT, often relative" (2007: 128). Sonjo (E46) includes its own unique

[68] The connection between *-ire/-iri* and *-ɛtɛ* at FV was discussed in Section 3.1.1.
[69] I give my regards to Jamin Pelkey for his suggestion of the visual representation in this Figure.

variations of these forms with *-atɛ* as a Progressive and *-irɛ* with PST and ANT uses (Nurse & Rottland 1991/1992: 180-181). Based on personal communication with P.R. Bennet, Nurse summarizes the possible development of this *-V(:)te* formative in Central Kenya/Thagicu (including Sonjo and Gusii) like this:

> "The original E50 *-VVte* suffix marked completed action with ongoing state (i.e. anterior, DN), the current meaning in E50 (Gikuyu, Kamba), from which both the Gusii Past and the Sonjo Present Progressive can be reasonably derived: dropping the completed action gives the Sonjo meaning, dropping the ongoing state gives the Gusii meaning" (2008: 165).

In connecting this with the Mara situation, the scenario outlined above at least allows for a connection between the P_4 ***(n)***-SM-***a***-R-***ɛ́tɛ*** of Gusii and the P_3 SM-***aa***-R-***iri*** of Kabwa. Obviously many more details need to be settled to flesh out this hypothesis, but I return to a possible historical scenario involving these forms in Section 6 after I have compared other features of the TAM systems in Mara. For now, I turn my attention to the more immediate past tenses, P_1 and P_2.

3.1.3 Hodiernal Past (P_2) and Immediate Past (P_1)

The P_2 refers to an event that occurred earlier today, but not within the last couple hours.[70] The P_1 is marked on events that happened within the last two hours, although distinct morphological marking between this and the P_2 is language specific[71]. Nurse mentions that an Immediate Past category is not widely found in Bantu and seems to be an innovation in the languages that do have it (2008: 100; see also the discussion later in this Section). In the five core Mara languages of this study, the temporal frame covering the Immediate Past (P_1) and the Hodiernal Past (P_2) is indicated by an Anterior/Perfect aspectual form, which is a variant of (***n***)-SM-R-***ire***. The variation

[70] Nurse's (2007: 127) division for P2 is "within the last 12 hours". This is interesting because his time divisions do not cover the 12-24 hours ago time frame.

[71] It must be remembered that the P_1 and P_2 are not descriptive categories for individual languages, but categories to aid in comparison. The fact that they are separate categories in several of the languages being studied means that they are important to distinguish in this study.

is mainly found in the last vowel of the *-iri* formative, where it also appears as *-ire* and *-iɪ*. The P_1 and P_2 time frames are generally indicated by adverbials or relative clauses as shown in examples (14) and (15) for Kabwa.

(14) *Tu-bhin-**iri*** *e-gi-dakika* *e-kumi* *gi-nú*
1PL-dance-ANT AUG10-NC10.PL-minute AGR10-ten AGR10-REL.PRON

gi-hit-iri.
SBJ10-pass.by-ANT
'Tumecheza dakika kumi zilizopita.' / 'We danced ten minutes ago.'

(15) *Tu-bhin-**iri*** *nyenkyo* *i-nu.*
1PL-dance-ANT NC9.SG.morning AGR9-PDEM
'Tulicheza asubuhi hii.' / 'We danced this morning.'

The usage of the Anterior/Perfect to cover the P_1/P_2 time frame is all the more interesting because other Mara languages like Kuria and Gusii have distinct non-Anterior forms for at least one of these two categories.

One difficulty in elicitation of the P_1 from Swahili is the fact that in Swahili the expression of an Immediate Past naturally takes an Anterior/Perfect formative *-me-* in the TA slot on the verb. Because of this association of the Swahili Anterior/Perfect with the Immediate Past, it is difficult to know whether the P_1 forms elicited in the core languages are best seen as Anterior (ANT) aspect, or as Immediate Past tenses. This ambiguity is not merely a result of the elicitation from Swahili since it is notoriously difficult to differentiate these categories in many Bantu languages (Nurse 2008: 94-99). As Nurse notes, it often takes close scrutiny of the system in a specific language to know whether a given form is ANT/PRF or whether it is a near past tense with perfective aspect, which is usually not marked in Bantu (2008: 94).

Nurse proposes several tests to which the language data can be subjected in order to determine the status of a form like (***n***)-SM-R-***ire*** in each of the individual Mara languages. For

the purposes of the present research, Nurse's second test is useful for making a decision. This test involves looking at compound constructions, consisting of an auxiliary and a main verb, and determining whether the form under consideration occurs in the auxiliary position or in the main verb position. Tense can only be marked on the auxiliary, which is the first word in the compound, and aspect can be marked on either the auxiliary or the main verb (2008: 96-97). In looking at the systematic usage of the (***n***)-SM-R-***ire*** form in Mara, it appears in the main verb position in compound constructions and so it is best considered an aspect rather than a tense. This type of usage is seen in another Kabwa sentence which is displayed in Example (16).

(16) *Tw-a*[72]*-ri-nga* *tu-bhin-**iri*** .
1PL-PST-COP-P_3/P_4 1PL-dance-ANT
'Tulikuwa tumecheza.' / 'We had danced (before today).'

Since the (***n***)-SM-R-***ire*** form functions aspectually in the Mara languages, I will use the term ANT to refer to it; more will be said about the ANT in Section 4.1. This, however, is a potential problem for interpretation because SM-R-***iri*** functions as a tense in Examples (14) and (15) in Kabwa (and similarly in the other core languages) since it refers to a specific period of time that cannot be referenced by any other past tense form. In addition, it has structural connections to the past tense system by differentiating the Hodiernal/Immediate Past from the Hesternal past that has the form SM-***aa***-R-***iri*** in Kabwa. The only difference between the two forms is that the Hesternal Past includes the *-aa-* formative in the TA slot.

For interpretational problems like this, where an aspectual form is being used functionally as part of the tense system, the research of Botne (2013) and Botne & Kershner (2008) is particularly relevant. This research, mentioned in Section 1.4.1, provides a multi-dimensional approach to understanding TA systems rather than assuming that all forms

[72] This is underlyingly a long-vowel [*aa*] but it is written short in the current orthography.

indicating a specific time range are best placed linearly along a single timescale (Botne 2013: 1). Botne explains that both "TENSE" and "TENOR" relationships are important features of most Bantu TA systems. "TENSE" relationships refer to the standard conceptualization of "a series of worlds arrayed linearly from the past into the future" (2013: 1). There are two common types of worlds: the P-domain and the D-domain. The P-domain is "the actual contemporal world in which speaker is interacting with hearer; that is, it is the privileged domain in which the utterance occurs" (2013: 1). Separate D-domains can occur in the past or the future, but each of these domains "denotes a disconnection of the discourse-narrated event from the speech context; that is, it is dissociated from the perceived 'realis' of the speaker's here-and-now at utterance time (UT)" (2013: 2). The "TENOR" relationships require that the timeline be construed "in terms of (adjoining) time regions" (2013: 3). This means that within a single world (like the P-domain) there can be more than one past or future region denoting both a Current Time Region and a Distal Time Region. An aspectual form like (***n***)-SM-R-***ire*** in the Mara languages functions in the Current Time Region of the contemporal P-domain allowing for other past forms to operate in the Distal Time Region of the P-domain or even in a dissociated Past D-domain. This multi-dimensional approach helps to explain overlapping TA systems like Ikizu where SM-R-***iri*** denotes the P_1/P_2 time frame, SM-***ka***-R-***a*** denotes the $P_2/P_3/P_4$ time frame, and SM-***aa***-R-***iri*** is also used for P_4. The SM-R-***iri*** form is in the Current Time Region of the P-domain (with Past reference) and the SM-***aa***-R-***iri*** form is in the Distal Time Region of the P-domain (with Past reference). The SM-***ka***-R-***a*** form, on the other hand, is dissociated from the contemporal world in the P-domain and is instead associated with a Past D-domain. This is an important area for further research amongst the Mara languages and is likely to provide richer explanations of language change.

Within the broader scope of the Great Lakes languages, the Rutara branch also shows the Perfect/Anterior being extended to near past usage (Muzale 1998: 18). Due to the connection of the near past to the ANT across the Bantu area and the widespread nature of an ANT with the formative *-ile/-ire* (or a variant) in the FV slot, Nurse has cast doubt on the antiquity of a near past tense. He suggests "that the immediate past/future category, where it exists, is a local innovation, together with the morphology instantiating it" (2008: 100). This would imply that the forms marking P_1 in Kuria and Gusii are more recent innovations. In addition, Nurse reconstructs *-ile* in the FV slot for PB, possibly already with the meaning ANT, although this is debatable (2008: 279). On this evidence, the *-ire/-irɪ/-iri* formative in Mara is very likely an inherited ANT aspectual marker. I will delve into this subject again in Section 4.1.

Before closing this section, there are two additional things to mention in relation to the near past tenses. The first is that Ikoma has an aspectual category called the Inceptive (INCE), which refers to actions that have just begun. This is conceptually related to the P_1 since both categories deal with events that started in the Immediate Past; I will cover it in detail in Section 4.5.

The second point to mention is that there are two forms that can be used to cover the P_2 time frame in Zanaki and Ikizu: SM-R-***irɪ***[73] and SM-***ka***-R-***a***. The form SM-***ka***-R-***a***, which is generally used in situations covering the more distant past (P_3/P_4) time frame, can also be used for the P_2 on at least some occasions. This overlap in reference can be explained by an aspectual difference between the ANT/PRF SM-R-***irɪ*** and the General Past SM-***ka***-R-***a*** (Futakamba et al. 2013). As I detailed for Ikizu above, Botne's multi-dimensional approach to TA systems handles the overlap in temporal reference in an elegant way.

[73] The formative at FV in Zanaki is phonetically [irɪ] but is orthographically <iri>.

Now that I have concluded this section on different degrees of absolute past reference, I will proceed to a discussion of relative tenses since they are also generally related to past reference.

3.2 Relative Tenses: The Narrative and Related Forms

Relative tenses are different than the absolute tenses I have focused on thus far. "Absolute tenses relate the situation to the present moment, while relative tenses relate it to some other time already established" (Nurse 2008: 120). The most frequently mentioned relative tense in Bantu is called the Narrative (NAR), which Nurse uses as a cover term for several related ideas like the consecutive/narrative, where the same subject is employed on sequential verbs, and the subsecutive/sequential, where a different subject is used on sequential verbs (2008: 121). The NAR relates subsequent events in a narrative to either the time of the first event or to a time that is already known to the speaker and listeners (2008: 120).

In addition to providing a broader context for the usage of many of the TAM forms from the elicited sentences in the core languages, one of the main purposes of gathering the discourse material found in Appendices B through F was to collect data on the usage of the NAR, which is difficult to observe through sentence elicitation alone. This proved to be a more difficult task in some languages because, as Table 1 shows in Section 3.1, languages like Ikoma and Simbiti have more than one form that could be interpreted as NAR. Since the NAR is a tense that surfaces in longer discourse, it is also potentially influenced more by individual speech styles; however, only a much broader discourse study would be able to unearth trends in usage. For my present study then, any conclusions on relative tenses must be taken as preliminary and in no way an indication of how a majority of speakers in a language community uses them.

3.2.1 Usage of the -(*V*)*ka*- Formative

In all seven of the Mara languages in this study, a variant of the SM-***(V)ka***-R-***a*** form is used in some type of Narrative function, with possible tonal variation on the -*ka*- morpheme.[74] Example (17), which is taken from Appendix D, lines Cwa3a and Cwa3b, shows a succession of verbs with -*ka*- formatives in Kabwa.

(17) *O-mw-ana* *wa-aye* *a-**ka**-aj-**a**,* *a-**ka**-hik-**a***
AUG1-NC1.SG-child AGR1-3SG.POSS 3SG-NAR-come-FV 3SG-NAR-arrive-FV

*a-**ka**-geg-**a*** *e-ki-mwi* *ku=Ø-bhaasikeri, …*
3SG-NAR-carry-FV AUG7-NC7.SG-one NC17=NC9.SG-bicycle
'Mtoto wake akaja, akafika, akabeba [kikapu] kimoja kwenye baiskeli yake,....' /
'His child came, [and] arrived, [and] carried one [basket] on his bike,....'

Ikoma is the only language that does not have SM-***ka***-R-***a*** for NAR and instead has an interesting variant, SM-***Vka***-R-***a***. Although this seems related, it remains to be seen how the Ikoma form developed an unspecified vowel in front of the -*ka*- formative.[75] In this regard, it is worthwhile to note that a survey has been done of the Ikoma, Nata, and Isenye language varieties, which are generally very similar (see Section 1.5.1 for more on this). Even amongst these three related varieties, there is variation noted between -*ka*- and -*Vka*- for the Narrative formative; however, the cause of this variation is unknown at the present (Smith et al. 2008). Additionally, in Botne's survey of future -*ka*- he mentions that in several parts of the Bantu area there are formatives that have -*ka*- following another morpheme (sometimes -*a*- or -*e*-) in the TA slot to indicate the future tense (Botne 1999: 490-492). This at least points to the potential variation between -*ka*- and -*Vka*- for a related formative in Bantu.

[74] Verification of this is dependent on a tonal analysis in each of the core languages. There is an apparent tonal distinction on this morpheme between the Kuria and Gusii reflexes of it (Mwita 2008; Nurse 2007).
[75] There are several other formatives in Ikoma that have an unspecified vowel; this usually occurs in places where other Mara languages have -*a*-.

Botne's research also shows the distribution of Future and Distal *-ka-* within Bantu. While not focusing on the Narrative *-ka-*, Botne does mention that the "Subsecutive" usage of *-ka-*, which is equivalent to what Nurse refers to as Narrative. This usage is widespread in the regions of the Bantu area where Distal *-ka-* is very rare (and vice versa), including the Great Lakes area (Botne 1999: 505). Nurse mentions all three of these usages of *-ka-* in his work on TA in Bantu, where he refers to Distal *-ka-* as an 'itive' and Future *-ka-* as a 'far future' (Nurse 2008: 240). The Distal/Itive usage is the most prevalent, occurring in at least a third of his matrix languages, while the Future usage occurs in 25-29 percent and the Narrative usage occurs in 29 percent of the matrix languages. Narrative *-ka-* occurs very frequently in the Great Lakes area as well as adjacent zones (2008: 242). One of the questions in the literature has been how these three relate to each other or whether they came from separate sources. There is, at the very least, semantic overlap between the Distal/Itive meaning of "location of an event away from the deictic center" and the Narrative meaning "A then B then C […]" (Nurse 2008: 242).

In addition, Nurse refers to another common usage of *-ka-* to mark the far past. This usage seems to be found more widely than just in GL Bantu since the distribution of this formative in eastern Bantu covers southern and southwestern Tanzania,[76] northern Zambia and northern Malawi (Nurse 2008: 243). A far past *-ka-* is semantically closely related to the Narrative *-ka-*, since NAR is generally used once past reference has been established. This state of affairs would imply that a far past tense (*-ka-*) could very easily become a NAR marker as well. Subsequently, the original far past reference could start to be marked by a new form and then only the NAR would remain with the original form. This is relevant because, as noted above

[76] This includes Gogo (G11), which shows some systematic evidence for being closely genetically related to other Great Lakes languages on the western side of Lake Victoria (Botne 1989/1990). However, the usage of *-ka-* as far past in Gogo could also be due to borrowing from other nearby languages which do not have a close link with Great Lakes.

in Table 1, Zanaki and Ikizu feature SM-***ka***-R-***a*** marking both the NAR and P_2/P_3/ P_4 categories; while in the other Mara languages SM-***ka***-R-***a*** marks NAR, but not P_2/P_3/ P_4. Based on a comparison of TA in GL Bantu, Nurse and Muzale see the *-ka-* formative as "a candidate for the original marker of Far Past in Great Lakes, kept in Rutara and part of [J]E.40, replaced elsewhere" (1999: 532-533). This would mean that Zanaki and Ikizu preserve the inherited *-ka-* for the far past, while the other Mara languages have more innovative systems that replaced this original form. Though this seems like a plausible scenario, it is not evident how the antiquity of a far past *-ka-* in Mara corresponds with the overwhelming existence of the form SM-***a(a)***-R-***ire***, which is currently represented as a Hesternal/Remote Past tense (P_3 and/or P_4) for MCAs in most of the Mara languages. Even where this form is not seen frequently in MCAs (i.e. especially Zanaki and to some extent Ikizu) it is still used in relative clauses (see Section 5.7). Another factor confusing the historical picture is that the LWC, Swahili, also uses *-ka-* to mark NAR and so the borrowed usage of this formative in the discourse of the Mara languages is a possibility. That being said, the fact that *-ka-* is used for NAR in all seven of the Mara languages strongly suggests that it should be reconstructed as such for Proto-Mara. I will return to this topic again in Section 6.1.1.

3.2.2 Other Formatives Used in Narrative Discourse

In both Ikoma and Simbiti, a formative *-ra-* is generally used in the TA slot on verb forms to indicate a PRS/PROG; this is detailed further in Section 4.2. However, in the narrative discourse that is included in the appendices for these languages, it is clear that this formative also functions similarly to the Narrative tense formative ***-(V)ka-***. Example (18), copied from the Ikoma narrative in Appendix C, lines Ntk4a-b, shows the discourse usage of the *-ra-* formative.

(18) *A-**ra**-tu,* *a-**ra**-tu,* *a-**ra**-tu,* *a-**ra**-ach-a,*
3SG-PROG-pick 3SG-PROG-pick 3SG-PROG-pick 3SG-PROG-come-FV

a-ri**[77]**-ichor-i *mo=bhe-tonga.*
3SG-PROG-fill-CAUS NC18=NC8.PL-basket
'Akachuma, akachuma, akachuma. Anakuja kujaza vikapu vyake.' / 'He picked [and] picked [and] picked [and then] went to fill his baskets.'

Similarly, Example (19), taken from Appendix E, lines Ssc24a-24b, shows this formative in Simbiti.

(19) *Së-haghwë* *si-yö* *se-**ra**-nyarag-an-a,*
NC10.PL-fruit AGR10-RDEM SBJ10-PROG-scatter-RECP-FV

*se-**ra**-shambok-a* *u-ku-y-a* *ahare* *i-yö.*
SBJ10-PROG-bounce-FV AUG15-INF-go-FV NC16.bicycle AGR23-RDEM
'Matunda yakaruka na kutawanyika kote kote.' / 'The fruit spilled out and scattered all over.'

It is conceivable that the usage of *-ra-* in discourse is extended from its Progressive usage and that it is slightly distinct from the NAR. One possibility is that *-ra-* indicates simultaneous or nearly simultaneous events rather than successive events. However, a fuller study of the discourse in these two languages is necessary to determine a more precise semantic distinction between the *-ka-* and *-ra-* formatives in the texts for these two languages. Another possibility is that the *-ra-* morpheme in Ikoma and Simbiti is related to the *-raa-* marker of NAR in Gusii. Nurse mentions that its usage in Gusii is distinctive from the typical usage of a NAR since it is placed first in a string of verbs and indicates that there is material following it (Nurse 2007: 128). Gusii is the only Mara language with a distinct narrative *-raa-* formative, which resembles a future formative *-raa-* in other North Mara languages (further discussion in Section 3.3.3) as well as the conditional formative *-raa-* that is prevalent throughout Mara (see specifically Section 5.2).

[77] This formative is *-ra-* underlyingly.

Another interesting form that behaves similarly to the Narrative *-ka-* in discourse is the Simbiti SM-***aa***-R-***a***, which appears to be related to the "Untimed Past Anterior" in Kuria (Cammenga 2004: 285-287), as I noted above in Section 3.1.1. The Simbiti and Kuria forms are evidently variants of SM-***a(a)***-R-***a*** (as further discussed in Sections 3.1.1 and 6.1.1), which is widely attested in Bantu as a marker of past tense (Nurse 2008: 82-83). This evidence strengthens the possibility of a historical connection between the Simbiti past tense system and the past tense system in Kabwa. This is noted here because these two languages seem to have very similar Tense systems in many regards, which is a topic that will be further developed in Section 6.1.

Having taken a look at the relative tense, which is generally related to past reference, I now shift my focus to the non-past tenses including the present and the future.

3.3 Non-Past Tenses

Any reference to a time other than the past is subsumed under the label non-past and this includes both the Present and the Future. The concept of "Present" is tricky to define since an action which is occurring at the present moment can be conceptualized in a variety of ways. This topic will be the focus of Section 3.3.4.

Reference to future events can be divided into various degrees depending on the temporal proximity of the action to the present moment. In order to keep a degree of symmetry with the references to past time divisions that were established in Section 3.1, the future time divisions cover similar temporal ranges; these include Immediate Future (F_1, within the next couple hours), Hodiernal Future (F_2, later today), Crastinal Future (F_3, tomorrow), and Remote Future (F_4, after tomorrow). Within the seven languages there are not as many distinctions in future reference as there are in past reference, which is typical in Bantu (Nurse 2008: 89). In fact, none of the Mara languages make a distinction between all four degrees of future reference; although

three of the languages have a distinction between three degrees of future reference. Ikizu, Zanaki and Ikoma don't even differentiate between the Future and the Present Tense (or Progressive Aspect). In spite of this, the establishment of four time divisions was deemed important to show the distinction between the usage of related forms in Kabwa and Simbiti; and this comes to light in Table 2.

Another important facet of future tense reference is its close relationship to degree of certainty, since the future has not yet occurred (Nurse 2008: 92). This is one of the reasons why there is semantic overlap between these two categories, which is observed in the Table below where the F_1 and Future Inceptive overlap. Morphological realizations of the various degrees of non-past reference in the Mara languages are illustrated in Table 2.

Table 2: Various Degrees of Non-Past Reference in the Mara Languages

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="5">Type of Non-Past Reference</td></tr>
<tr><td>F₁/FUT.INCE[78]</td><td>F₂</td><td>F₃</td><td>F₄</td><td>PRS</td></tr>
<tr><td rowspan="8">Language</td><td>Ikizu</td><td colspan="4">SM-ɾa-R-a</td><td>SM-aa-R-a</td></tr>
<tr><td>Zanaki</td><td colspan="5">SM-ɾa-R-a[79]</td></tr>
<tr><td>Ikoma</td><td colspan="4">n-SM-Vko-R-a</td><td>SM-Vká-R-a[80]</td></tr>
<tr><td>Kabwa</td><td>(n)-SM-ɾaa-R-ɛ</td><td colspan="2">SM-aka-R-ɛ</td><td>SM-ɾi-R-a</td><td>SM-aa-R-a[81]</td></tr>
<tr><td>Simbiti</td><td>(n)-SM-ɾaa-R-ɛ[82]</td><td>SM-aka-R-ɛ</td><td colspan="2">n-SM-ɾi-R-a</td><td>SM-ɾa-R-a</td></tr>
<tr><td>Kuria[83]</td><td>n(e)-SM-ɾaá-R-ɛ</td><td>SM-kaa-R-a</td><td colspan="2">SM-ɾee-R-a</td><td>SM-Ø-R-a</td></tr>
<tr><td>Gusii[84]</td><td colspan="4">(n)-SM-Ø-R-ɛ</td><td>SM-Ø-R-a</td></tr>
</table>

[78] The Immediate Future and Future Inceptive categories overlap in usage, which is why they are combined in this Table. This overlapping usage will be a focus for the discussions in Sections 3.3.3 and 4.5.
[79] A complex two-word form SM-***ra-az-a ku***-R-***a*** can be used for F_4 as well and may be lexicalizing as the Remote Future in Zanaki.
[80] This form is also used as a Past Inceptive in Ikoma indicating that an action has just begun (see Section 4.5).
[81] This form is distinct from the P_4 (SM-***aa***-Ŕ-***a***) in Kabwa because it lacks a high-tone on the verb root.
[82] The more common Simbiti form has the copular nasal in the Pre-SM slot, but the form without the nasal is also heard commonly in speech.
[83] The source for the [kuj] non-tonal data is Cammenga (2004), with underlying tones obtained from Mwita's (2008) analysis of Cammenga.
[84] The source for all the [guz] data is Nurse (2007: 127-129), which is based on Whiteley (1960); underlying tones are used for all the Gusii data.

3.3.1 Remote Future (F_4) and Crastinal Future (F_3)

The Remote Future (F_4) is a category that covers future events that are expected to occur after tomorrow, while the Crastinal Future (F_3) deals with events that are expected to occur tomorrow. In four of the Mara languages (Ikizu, Zanaki, Ikoma, and Gusii), the formative covering F_4 is also used to cover the $F_1/F_2/F_3$ degrees of future reference. In these systems it is better to refer to this combined category simply as Future (FUT). From a comparative perspective it is important to note that amongst these four languages there are three distinct forms that mark FUT. The first form is found in both Ikizu and Zanaki, which mark it as SM-***ra***-R-***a***. An example of its use for a Remote Future reference in Ikizu is found in (20).

(20) *Tʉ-**ra**-gaamb-an-**a*** *na-wɨ* *u-mw-aaka*
1PL-FUT-speak-RECP-FV CCONJ-3SG.PRON AUG3-NC3.SG-year

gʉ-nʉ *gu-ku-uz-a.*
AGR3-PDEM.REL SBJ3-NPST-come-FV
'Tutaongea naye mwaka ujao.' / 'We will speak with him next year.'

This is the same form that is seen in Example (21), where the reference is to the Immediate Future.

(21) *Tʉ-**ra**-gaamb-an-**a*** *na-wɨ* *i-baga* *i-suuhu*
1PL-FUT-speak-RECP-FV CCONJ-3SG.PRON NC5.SG-time AGR5-little

ri-nʉ *ri-ku-uz-a.*
AGR5-PDEM.REL SBJ5-NPST-come-FV
'Tutaongea naye baada ya muda mfupi.' / 'We will speak with him very soon.'

In addition to the uses of SM-***ra***-R-***a*** for FUT in Ikizu, it is also employed as a PROG marker which overlaps with Present Tense reference (see Section 4.2 for further discussion of this). Zanaki employs SM-***ra***-R-***a*** for both of the same categories where it is found in Ikizu (FUT and PROG) but also has extended it to cover the Vast Present (PRS), as is exemplified in Section 3.3.4 below. Since this form is extensible to both the FUT and PRS in Zanaki, the term Non-Past (NPST) is a better description of the Tense categories that SM-***ra***-R-***a*** covers in that language.

The second form marking general FUT in the Mara languages is the Ikoma ***n***-SM-***Vko***-R-***a***,[85] which is displayed in Example (22).

(22) ***N-to-ogho***[86]*-siiker-**a*** *na-we* *mo=mo-oka*
FOC-1PL-FUT-speak-FV CCONJ-3SG.PRON NC18=NC3.SG-year

u-no *o-ku*[87]*-uch-a.*
AGR3-PDEM.REL SBJ3-NPST-come-FV
'Tutaongea naye mwaka ujao.' / 'We will speak with him next year.'

Ikoma also uses its FUT form for PROG[88] as I showed for Ikizu and Zanaki; however, Ikoma does not extend this form to mark the Vast Present like Zanaki does. I will give an example of the PRS form in Ikoma in Section 3.3.4.

The third form for the FUT category, which is found in Gusii, is marked by ***(n)***-SM-***Ø***-R-***ɛ*** (Nurse 2007: 127-128). This is a widespread form for the subjunctive mood both in Mara and throughout Bantu (Nurse 2008: 259); therefore, it seems likely that the SBJV has been extended to FUT meaning in Gusii, replacing any previous formatives marking Future Tense. This type of extension does not occur in any of the other Mara languages and appears to be a local innovation in Gusii.

Amongst the other three Mara languages in this study (Kabwa, Simbiti, and Kuria) there are three distinct forms marking different degrees of Future reference. The F_4 is marked by a variant of ***(n)***-SM-***ri***-R-***a*** in all three languages. Example (23) shows this usage of the form in Simbiti.

[85] Along similar lines as the varation between *-ka-* and *-Vka-* in Ikoma/Nata/Isenye noted in Section 3.2, the unspecified vowel of *-Vko-* in the TA slot in Ikoma and Isenye shows variation with *-ko-* in Nata (Smith et al. 2008).
[86] Application of Dahl's Law causes the FUT formative to surface with a voiced *gh* rather than the voiceless *k*.
[87] This NPST formative is underlyingly *-oku-* (variant of *-Vko-*) but it surfaces as *-ku-* because there is a rule in Ikoma that disallows long vowels word-initially (Robinson: *pc*).
[88] For the Present Tense, but not for non-Present situations where there is a compound verbal form.

(23) *O-mo-oka* *ghu-yö* *ghu-ku-ush-a,*
AUG3-NC3.SG-year AGR3-RDEM.REL SBJ3-PRS-come-FV

n**-tu-**ri**-shumaash-**a *na-we.*
FOC-1PL-F_3/F_4-speak-FV CCONJ-3SG.PRON
'Tutaongea naye mwaka ujao.' / 'We will speak with him next year.'

Both Kuria and Simbiti use ***(n)***-SM-***ri***-R-***a*** for F_3 and F_4. Kabwa, on the other hand, employs a distinctive F_3, which is also used to cover the F_2. It is quite interesting that Kabwa lumps the F_2 and F_3 time frames together into a non-remote tense, whereas Simbiti lumps the F_3 and F_4 categories together into a remote tense. The F_3 in Kabwa is SM-***aka***-R-***e***, and it is exhibited in Example (24).

(24) *Tw-**aka**-hair-**e*** *na-we* *ejo.*
1PL-F_2/F_3-speak-SBJV CCONJ-3SG.PRON tomorrow
'Tutaongea naye kesho.' / 'We will speak with him tomorrow.'

Before closing this section, there is one other Remote Future form in Zanaki that needs to be discussed since it is a compound form that is quite distinct from the forms seen so far in this section. For my Zanaki language consultant, a compound form with a "come" auxiliary (SM-***ra-az-a ku***-R-***a***) was preferred over SM-***ra***-R-***a*** for actions in the very distant future. This compound F_4 form is displayed in (25).

(25) *Tu-**ra-az-a*** ***ku**-gaamb-an-**a*** *na-we* *o-mw-aaka*
1PL-FUT-come-FV INF-speak-RECP-FV CCONJ-3SG.PRON AUG3-NC3.SG-year

gu-no *gu-ku-uz-a.*
AGR3-PDEM.REL SBJ3-NPST-come-FV
'Tutaongea naye mwaka ujao.' / 'We will speak with him next year.'

It is important to note that SM-***ra***-R-***a*** can also be used for these situations, but according to my Zanaki language consultant the compound form is used more commonly in speech.

In relation to the Zanaki "come" compound being used for F_4, Ikizu also has a "come" compound. This compound is used quite distinctly from the Zanaki form to introduce a type of relative clause. There is a more detailed treatment of this phenomenon in Section 5.7 below.

Having considered the Remote Future (F_4) and the Crastinal Future (F_3) categories in the Mara Bantu languages, I now proceed to a discussion of a less distant future, namely the Hodiernal Future (F_2).

3.3.2 Hodiernal Future (F_2)

The Hodiernal Future is a time period occurring later today, but more than a few hours from the present. As we saw in the previous section, Kabwa, Simbiti and Kuria are the only three Mara languages in this study that have more than one future form. Example (24) above showed that Kabwa has the form SM-***aka***-R-***e*** for F_2/F_3. Simbiti has this same form, but only for F_2; its usage does not extend into F_3. The F_2 form for Simbiti is presented in Example (26).

(26) *Mo-ghorobha,*[89] ***n****-tw-****aka****-shumaash-****e*** *na-we.*
NC3.SG-evening FOC-1PL-F_2-speak-SBJV CCONJ-3SG.PRON
'Tutaongea naye jioni hii.' / 'We will speak with him this evening.'

A form of the shape SM-***aka***-R-***ɛ***[90] has a limited distribution within GL, so Nurse and Muzale treat it as a "post-proto-Great Lakes innovation" (1999: 530). Since the Suguti languages also have this form (1999: 529), and they are purportedly related to the Mara languages under the East Nyanza branch of Great Lakes, this is potentially a Proto-East Nyanza innovation. However, this interpretation is cast into doubt because the Luhya branch also has forms of this shape (1999: 529). The spread of the SM-***aka***-R-***ɛ*** form is interesting from an areal perspective as Mara, Suguti and Luhya are the three Great Lakes sub-groups along the eastern shores of Lake Victoria. In

[89] This word is placed in Noun Class 3 since it can take a standard plural in Noun Class 4; however, there are also indications that the singular form has become lexicalized as a temporal word without a noun class marker.
[90] Phonetically, [ɛ] is found at FV for this form in Kabwa and Simbiti. Both languages represent this orthographically as <e>.

broader Bantu perspective, Nurse found a future tense *-ka-* in the TA slot for 29% of the languages[91] he surveyed; and he mentions that sometimes this was as a "composite marker" with forms like [a+ka] or [ka+a] (2008: 85-86).

Of the seven languages in this study, it is interesting to see that only Kabwa and Simbiti have the SM-***aka***-R-***ɛ*** form with Future meaning. There is some surface similarity between this form and the Kuria F_2, which is SM-***kaa***-R-***a***. Although this Kuria form does have a variant of the *-ka-* formative in the TA slot, it also has *-a* at FV, which is structurally quite different than the Subjunctive *-ɛ* in the Kabwa and Simbiti forms. If a Future *-ka-* formative is even to be posited for Proto-North Mara, a plausible explanation for the existence of both the Kabwa/Simbiti form and the Kuria form should be established. Two possibilities present themselves.

The first is the possibility that Proto-North Mara had forms with both endings. This is based on Botne's discussion of Bantu languages that have Future tense *-ka-* formatives with both *-a* and *-e* final vowels. In these systems the SM-***ka***-R-***e*** would be a future subjunctive and the SM-***ka***-R-***a*** form would be a future indicative (1999:495). The only nearby languages that have both these forms are in the Central Kenya/Thagicu (E50) branch. Although there is evidence for a link between Gusii and Central Kenya/Thagicu (see Sections 3.1.1 and 3.1.2 for further explanation of this link), there is not yet enough evidence to posit a similar link between the rest of North Mara and Central Kenya/Thagicu. In addition, although subjunctive endings are posited as marking "nonfactuals" for the Future in GL (Nurse &Muzale 1999: 530), there is not any evidence I am aware of within GL for a system that includes both indicative and subjunctive futures. Since there is a lack of solid evidence for a system like this, it is rather fanciful to posit it for the Mara branch of Great Lakes; so, a split indicative/subjunctive future system cannot be supported for Proto-North Mara.

[91] With dubious cases excluded, the figure falls to 25%.

The second possibility for providing evidence that a *-ka-* future could have been inherited from Proto-North Mara comes from within the Kuria system. If it can be shown that the Kuria F_2 developed from a related form in that language, then the *-kaa-* formative does not need to be inherited from Proto-North Mara and the Kabwa/Simbiti form becomes the likely candidate for an inherited form. There does appear to be a formal similarity between SM-***kaa***-R-***a*** and the NAR in Kuria, which is SM-***ká***-R-***a***. Even so, it is difficult to posit a semantic connection between the two forms, since the Narrative is most frequently used with past tense reference instead of future reference. It therefore seems that this second possibility must also be discarded. It is still possible that the Kabwa/Simbiti SM-***aka***-R-***ɛ*** is an inherited post-Proto-GL (possibly Proto-East Nyanza) form that, outside this portion of North Mara, has only been maintained in Luhya and Suguti. However, the possibility that this distribution of the form is the result of areal diffusion cannot be ignored.

A related point of comparative interest is that the SM-***aka***-R-***ɛ*** form in Kabwa/Simbiti (as well as SM-***kaa***-R-***a*** in Kuria) shows some similarity to the Past Inceptive (PST.INCE) aspect of Ikoma (see Section 4.5 for further explanation of this category), which is realized as SM-***Vká***-R-***a***.[92] The PST.INCE in Ikoma has to do with events that have just begun or are beginning presently. There is a mismatch between the final vowels in the Kabwa/Simbiti and Ikoma forms, but since the PST.INCE category is not common amongst the Mara languages, its existence in Ikoma deserves an explanation. It seems semantically plausible that the meaning of this form could have arisen as an innovation in Ikoma from a reinterpreted future. This shift in meaning could have come about as Ikoma simplified its Tense system to a single Non-Past Tense and reanalyzed a previous Future tense form as the PST.INCE; this occurred as a shift from an event that will happen to one that is about to happen to one that is just starting. If we follow this

[92] This form is tonally distinct from the NAR in Ikoma, which is SM-***Vka***-R-***a***.

line of reasoning, then it is also plausible that in the process of this re-analysis the PST.INCE form (now without non-factual implication) was no longer thought of as related to the subjunctive and shifted its final vowel to the indicative *-a*. An alternative hypothesis is that the PST.INCE SM-***Vká***-R-***a*** was created on analogy to the now defunct Future form SM-***Vká***-R-***e*** based on an opposition between a non-factual subjunctive final vowel *-e* and the factual indicative *–a* (this is based on an argument by Nurse & Muzale 1999: 530). Subsequent to this development, the Future tense system was reduced in Ikoma and the original SM-***Vká***-R-***e*** was lost. I proffer both of these possible scenarios for further thought, but neither of them will be followed in this thesis since they are based on the existence of a lost form that is impossible to prove or disprove unless further comparative data becomes available to give it credence.

With the Hodiernal Future (F_2) category in mind, I now examine the Immediate Future (F_1) in the following section, since in some ways the Mara languages that have distinctions between the two categories show overlapping usage.

3.3.3 Immediate Future (F_1)

I mentioned the Inceptive aspect in the previous section and its meaning certainly has a relation to future events in the F_1/ F_2 time frame. Inceptive has an idea of 'beginning to do something' and will be dealt with as an aspectual marker in Section 4.5. F_1 certainly has a sense of 'beginning to' or 'about to' with it, but in Simbiti it indicates that an action will begin at a specific time in the near future rather than specifying the manner of the action. One case for this is the fact that elicitation of a specific Immediate Future time frame in Simbiti yields a distinct ***(n)***-SM-***raa***-R-***e***[93] form. This is exemplified in (27).

[93] Phonetically, this final vowel is [ɛ].

(27) *Ha-no* *se-raa-het-e* *se-dakika* *i-kömi,*
NC16-PDEM.REL SBJ10-F_1-pass.by-SBJV NC10.PL-minute AGR10-ten

n**-to-**raa**-shumaash-**e *na-we.*
FOC-1PL-F_1/F_2-speak-FV CCONJ-3SG.PRON
'Tutaongea naye baada ya dakika kumi.' / 'We will speak with him in ten minutes.'

Kuria has a variant of this form for F_1 that is realized as ***n(e)***-SM-***raá***-R-***ɛ***. In both Kuria and Simbiti, although the F_1 is generally realized in the Immediate Future (at the most a couple hours from the present), it also implies a relative amount of certainty that the event indeed will be performed (Mwita 2008: 47). The F_2, on the other hand, can be used for an event that is supposed to take place a couple of hours from now or later today, but it implies that there is less certainty whether the event will take place as planned.

Kabwa has a variant of the ***(n)***-SM-***raa***-R-***ɛ*** form as well, but its usage is slightly different. Although it can be used temporally for very near future events like an F_1 tense, this is not the most common usage. It is better thought of as an aspectual marker of the Inceptive with Future implications (FUT.INCE). The fact that the action is "about to" occur is more important than the exact time when it occurs and so its use as an F_1 Tense marker is secondary to its usage as a Future Inceptive marker. Obviously, the two categories overlap, which is why the chart in Table 2 includes F_1/FUT.INCE as a combined category. This topic will be explained in greater detail in Section 4.5.

On a final note before closing this section, Simbiti has a second way of expressing that an event occurred in the F_2 time frame. This compound form was not included in the previous section because it is more closely related in form to the F_1. Previously, in Example (25) from Section 3.3.1, we saw an F_4 compound form in Zanaki with the verb "come" as the auxiliary (first word in the compound). This "come" auxiliary was put into the NPST tense in Zanaki. Simbiti also has a compound verb employing "come" as an auxiliary, but it is based on the F_1

form (naturally Subjunctive) with the main verb in the Subjunctive mood. Example (28) displays this Simbiti compound form that indicates an action occurring in the F_2 time frame.

(28) ***N*-*to*-*raa*-*sh*-*e*** *tu-shumaash-**e*** *na-we* *mo-ghorobha.*
FOC-1PL-F_1-come-SBJV 1PL-speak-SBJV CCONJ-3SG.PRON NC3.SG-evening
'Tutaongea naye jioni hii.' / 'We will speak with him this evening.'

That compound verbs with "come" auxiliaries should shift to a future tense reference is not uncommon as a grammaticalization path (Heine & Kuteva 2002b: 331; Bybee et al. 1994); indeed, it is found in many other Bantu languages as well (Nurse 2008: 298). Nurse refers to a specific example from Sukuma-Kiiya (F21) where there is a near future (Hodiernal Future) compound form with a "come" auxiliary in the Subjunctive followed by a main verb in the Subjunctive (2008: 171). From a historical standpoint, a future tense (usually near future) based on a compound with a "come" auxiliary[94] has a "limited geographical distribution" in GL, so it is not reconstructible for Proto-GL but is likely a more recent innovation (Nurse & Muzale 1999: 529-530). Even if this is the case, the "come" verb to Future Tense grammaticalization path is common enough that precaution should be taken in positing it as an inherited form at any phase of development. Another possibility is that a similar form has been innovated at multiple times in multiple places within GL. Without some additional morphophonological clues to tie together historically related forms, it would be hard to determine this. A wider survey of the "come" Futures in GL is necessary to unravel this and could prove a fruitful avenue for future research.

Now that I have looked at the Future Tenses in the Mara languages, I will proceed to a discussion about the related yet somewhat ambiguous Present Tense category.

[94] In some cases the compound forms have reduced to a single verb form based on an original 'come' auxiliary (Nurse & Muzale 1999: 529).

3.3.4 Present Tense as the Vast Present (Timeless)

Present Tense is an elusive concept that includes the present moment within its semantic scope but usually also includes a period of time on either side of the very ephemeral present moment. Any type of "Present" will go beyond just a temporal component and will indicate some aspectual component as well (Nurse 2008: 116). The most basic Present Tense in Bantu is a widespread aspectual null form indicating a timeless (or "vast") Present Tense which is marked by SM-***Ø***-R-***a***. This is the form that Nurse reconstructs for Present Tense in PB, which could be indicated on the common Bantu verb *gʊ́l* 'buy' in the form *tʊ-**Ø**-gʊ́l-**a*** 'we buy' (2008: 279). This SM-***Ø***-R-***a*** form is also found in various branches of Great Lakes Bantu. It is a Present in the Rutara branch (Muzale 1998: 129), at Present/Habitual in the North Nyanza branch, and at Present in the Luhya branch (Mould 1976: 135-136). Nurse and Muzale find overwhelming evidence for positing this form for Proto-GL, although it appears in a distinct form in Suguti and has been replaced in Mara (Nurse & Muzale 1999: 522). To refer to this timeless present Nurse uses the term "Vast Present"[95] and this focuses on an action taking place in general, but does not necessarily associate any timing or habituality with the action (Nurse 2008: 118; Nurse & Muzale 1999: 522). For the purposes of this thesis, the Vast Present is what will be referred to as the Present Tense (PRS).

For the Mara languages, a locally innovative form SM-***ra***-R-***a*** with a meaning related to PRS appears to have replaced the inherited PB/Proto-GL SM-***Ø***-R-***a*** form in all but a few cases; these will be stated for individual Mara languages below. The *-ra-* formative generally indicates Present Tense and/or Progressive Aspect (see Section 4.2) in the Mara languages included in this study, but has even been extended to Future reference in Ikizu and Zanaki. Even though the form

[95] This Vast Present is distinct from the infinitive form (often ***ku***-R-***a***), which is often part of compound verbs in Bantu (or functions as a verbal noun) but tends not to stand by itself as a verb (Nurse 2008: 59-61).

has acquired this distinctive meaning in the Mara languages, a related *-ra-* formative is used in other parts of GL as well. Within the West Highlands subgroup of GL, a *-ra-* formative is used mainly for what is referred to as "Disjunctive Focus" (Harjula 2004: 98-101; Nurse 2008: 205). Disjunctive Focus will be explained more in Section 5.5, but in general it refers to verb focus as opposed to Conjunctive Focus, which places post-verbal elements in focus. Interestingly, the *-ra-* formative is also part of the Progressive form in the West Highlands language Ha (Harjula 2004: 107). Nurse notes that there is a connection between Focus and Progressive (based on Güldemann 2003), and posits a "chain of development: (disjunctive) focus > progressive > general present > future or non-past" (2008: 294). This relationship of Focus to Progressive/Present/Non-Past and its implications for the historical picture will be expounded upon not only in the remainder of this section, but also in Sections 4.2, 5.5, and 6. With this background in mind, then, I turn my attention to the realization of Present Tense (and related categories) in each of the individual Mara languages.

Kabwa has replaced the inherited PB/Proto-GL null form SM-***Ø***-R-***a***[96] with SM-***aa***-R-***a*** at PRS, while SM-***ra***-R-***a*** is used to show a Progressive. An example of the SM-***aa***-R-***a*** form is included in (29).

(29) *Tw-**a***[97]*-gy-**a*** *Darisaraamu.*
1PL-PRS-go-FV Dar.Es.Salaam
'Twaenda Dar Es Salaam.' / 'We go to Dar Es Salaam.'

In Kabwa, the Vast Present contrasts tonally with the P_4 form SM-***aa***-Ŕ-***a***, which is introduced in Section 3.1.1. Nurse mentions that the *-a(a)-* formative used to mark PRS has more than a local distribution in Bantu (2008: 117), but within Great Lakes this form is not widespread. On the

[96] Interestingly, Kabwa has a trace of this null form (with focus marking) in a specific construction: *Ntutura okusimba erituko* – 'Tunaweza tukachimba shimo'/'We could/might dig a hole.' Refer to Sections 5.4 and 5.5 for further explanation of this construction.
[97] The underlying long-vowel is obscured by the orthography here; the long-vowel is seen clearly in the first person singular form *naagya* 'I go'.

other hand, an *-a(a)-* formative of varying length and tone is quite common with mid to far past reference in Great Lakes (Nurse & Muzale 1999: 533). Of particular interest in this regard are systems like in Ha (West Highlands branch of GL) where a low toned *-a-* signifying "simple recent past" contrasts with high toned *-á-* for "simple remote past" (Harjula 2004: 95). These "simple" tenses are described as "anterior", meaning that a situation that began at a certain point in the past still had relevance at a later time. For the "simple recent past" this means that an event that began in today's past still has relevance in the present; whereas the "simple remote past […] usually indicates a generic situation" (2004: 95). It is not difficult to see how a system like this could be linked to the Kabwa system. If the past meaning associated with Simple Recent Past were removed, then present relevance would be the only remaining meaning, and the form would mark PRS. The Simple Remote Past, on the other hand, would still pertain to an event in the more distant past. These possible developmental links can be viewed schematically in Figure 7.

Figure 7: Possible Links Between *-a(a)-* and *-á-* Formatives in Ha and Kabwa

Ha		Kabwa
Simple Recent Past (SM-***a***-R-***a***)	<---------->	Vast Present (SM-***aa***-R-***a***)
Simple Remote Past (SM-***á***-R-***a***)	<---------->	Remote Past (SM-***aa***-Ŕ-***a***)

Of course, for this type of developmental path to have historical validity, more research is needed in both West Highlands and wider GL. This should clarify whether the Ha system is indicative of Proto-West Highlands and if it can be traced back even further to Proto-Great Lakes. On the other hand, it is also conceivable that the Kabwa system is older and that the Ha

system has innovated new uses for the formatives. However, since more systems in Great Lakes have past tenses with some variant of *-a(a)-* at TA (Nurse & Muzale 1999: 533), it is more likely that the PRS SM-***aa***-R-***a*** is the innovative usage at a post-Proto-GL phase of development. At the very least, the possible link is of interest to future comparative studies amongst Great Lakes languages.

Fortunately for comparative purposes, Kabwa is not the only Mara language with this type of PRS and this adds weight to positing this form as inherited from an earlier phase. Ikizu has also replaced the inherited SM-***Ø***-R-***a***[98] with SM-***aa***-R-***a*** at PRS. An example of the PRS in Ikizu is presented in (30).

(30) *Tw-**aa**-j-**a*** *Daresaraamu.*
1PL-PRS-go-FV Dar.Es.Salaam
'Twaenda Dar Es Salaam.' / 'We go to Dar Es Salaam.'

Since a connection was established between the SM-***ra***-R-***a*** form and the PRS/FUT/NPST in the Mara languages, it is important to make this connection explicit for Ikizu. Although SM-***ra***-R-***a*** is generally used for the PROG (see Section 4.2) it has also been extended to include any Future temporal frame in Ikizu as I showed in Examples (20) and (21) from Section 3.3.1.

Zanaki does not have a null form and if it had a SM-***aa***-R-***a*** form at PRS, this has also been replaced. Currently PRS is marked by the *-ra-* formative in the TA slot, which is most likely an extension of its meaning from PROG. This same SM-***ra***-R-***a*** form has been further extended to indicate all degrees of Future tense in Zanaki and is best treated as a NPST in that language. An example of the NPST translating a Swahili Vast Present is displayed in Example (31).

(31) *Tu-**ra**-gy-**a*** *Darisaraamu.*
1PL-NPST-go-FV Dar.Es.Salaam
'Twaenda Dar Es Salaam.' / 'We go to Dar Es Salaam.'

[98] There has been no evidence found for SM-***Ø***-R-***a*** in Ikizu's current TA system.

Simbiti still uses a null form with present meaning, in distinction to Kabwa, Ikizu and Zanaki. According to my language consultant this null form is used informally in speech with a vast present meaning, but the SM-***ra***-R-***a*** form is more standard in Simbiti for PRS. Example (32) displays this form.

(32) *To-**ra**-y-**a***[99] *Daresaramu.*
1PL-PRS-go-FV Dar.Es.Salaam
'Twaenda Dar Es Salaam.' / 'We go to Dar Es Salaam.'

Like Simbiti, Ikoma also has a null form that has a timeless, action-focused meaning associated with it. The usage of this form in discourse needs to be researched more, but it appears to function like a relative tense marker (see Section 3.2) once the initial Tense/Aspect has been set by the first verb. I will refer to it as a Non-Past Consecutive marker (NPST.CNS), since the Narrative tense for Ikoma could be considered more like a "Past Consecutive." An instance of this can be seen in Example (33) where the Habitual Aspect is marked on the first verb and then the null form is used to show that the habituality is carried on to the second action as well.

(33) ***N**-to-**haa**-yi* *mo=mu-tëëra* *to-**Ø**-ghor-**a***
FOC-1PL-HAB-go NC18=NC3.SG-market 1PL-NPST.CNS-buy-FV

cham-bori.
NC10.PL-goat
'Huwa tunaenda sokoni tunanunua mbuzi.' / 'We usually go to the market (and) we buy goats.'

However, this form does not appear to function on its own as a Vast Present. Instead, the form that was elicited from a sentence with the Swahili Vast Present is SM-***Vká***-R-***a***, which is exemplified in (34).[100]

[99] This is a variant of the more standard *toraghya* 'we go'.
[100] The final vowel *-a* does not appear in the example because the verb root *ye* 'go' already ends in a vowel.

(34) ^*To-**ogha**-ye Tarisaraamu.*[101]
1PL-PRS-go Dar.Es.Salaam
'Twaenda Dar Es Salaam.' / 'We go to Dar Es Salaam.'

This same from is used aspectually as the Inceptive in Ikoma (see Section 4.5) and its interpretation as a Vast Present is dependent on what type of verb it attaches to (Holly Robinson: *pc*). Research on the interaction of TAM with different verb types (active, stative, durative, etc.) is beyond the scope of my present research, but this is also an area that deserves more focused study in the future.

Another form related to the Present tense in Ikoma is the Present Progressive ***n***-SM-***Vko***-R-***a*** which marks all of the Future tenses as well. This basic division of Non-Past forms (Future and Present Progressive) from the Past is similar to Ikizu and Zanaki although the formatives used are distinct. An example of this form in Ikoma with reference to the Future was given in Example (22), but Example (35) shows that it can also be used with Present reference extending over a period of time.

(35) ***N**-to-**ogho**-tuk-**a*** *mw-ishe* *mo-gima.*
FOC-1PL-PRS.PROG-dig-FV NC3.SG-daytime AGR3-whole
'Tunachimba siku nzima.' / 'We are digging all day.'

Ikoma also has a PROG with the *-ra-* formative similar to the other Mara languages, but it is not used on its own as a Present Progressive; it is only seen in conjunction with other tense markers to show Future and Past Progressives. This is displayed in the Ikoma Chart in Table 14 of Appendix A and is discussed further in Section 4.2.

Both Kuria (Cammenga 2004: 286) and Gusii (Nurse 2007: 127) have null forms that represent a timeless quality already referred to as the Vast Present. From the data at hand, neither

[101] The carat sign (^) is used to distinguish this tense in the orthography from the NAR which is tonally distinct and written without the carat (Higgins 2010: 13).

of these languages has a *-ra-* form. The progressives in Kuria and Gusii resemble the compound form in Simbiti, which is very unique for Bantu and will be covered in detail in Section 4.2.

Based on all the information presented in this section, the various forms relating to the Present Tense are laid out in Table 3 for ease of comparison.

Table 3: Various Forms Related to the Present Tense in the Mara Languages

		Varieties of Forms				
		PRS	FUT	PRS.PROG	Other	Null form used?
Language	Ikizu	SM-***aa***-R-***a***	SM-***ɾa***-R-***a***		--	NO
	Zanaki	SM-***ɾa***-R-***a***			--	NO
	Ikoma	SM-***Vká***-R-***a***[102]	SM-***Vko***-R-***a***		SM-***ɾa***-R-***a***[103]	YES[104]
	Kabwa	SM-***aa***-R-***a***	Unrelated[105]	SM-***ɾa***-R-***a***	--	NO[106]
	Simbiti	1. SM-***ɾa***-R-***a*** 2. SM-***Ø***-R-***a***[107]	Unrelated[108]	***n-ko***-R-***a*** SM-**re**	--	YES
	Kuria[109]	SM-***Ø***-R-***a***	Unrelated[110]	***n-ko***-R-***a*** SM-**re**	--	YES
	Gusii[111]	SM-***Ø***-R-***a***	Unrelated[112]	***n-kó***-R-***a*** SM-**re**	--	YES

I now return to the discussion at the beginning of this section regarding the chain of development: focus > progressive > present > non-past (Nurse 2008: 209, 293-94). In Mara, there is evidence for this type of development with the SM-***ra***-R-***a*** form. If the situation in Ha (West Highlands) is taken as the initial state in Great Lakes then the *-ra-* formative started as a Disjunctive Focus marker that had already extended its usage as a Progressive by the time

[102] This is interpreted as a Vast Present on certain verb types; generally it is a Past Inceptive. The FUT/PRS.PROG form SM-***Vko***-R-***a*** can also be interpreted as PRS on certain verb types in Ikoma.
[103] This is used for non-present progressive compound forms and as a NAR tense.
[104] This form is glossed as a Non-Past Consecutive.
[105] There are multiple futures in Kabwa and they are not connected to the Present Tense; see Sections 3.3.1 to 3.3.3.
[106] There is still a frozen null form in one construction, but it is not currently productive.
[107] This form is not as common as the SM-***ɾa***-R-***a*** form for PRS.
[108] There are multiple futures in Simbiti and they are not connected to the Present Tense; see Sections 3.3.1 to 3.3.3.
[109] The source for the [kuj] non-tonal data is Cammenga (2004), with underlying tones obtained from Mwita's (2008) analysis of Cammenga.
[110] There are multiple futures in Kuria and they are not connected to the Present Tense; see Sections 3.3.1 to 3.3.3.
[111] The source for all the [guz] data is Nurse (2007: 127-129), which is based on Whiteley (1960); underlying tones are used for all the Gusii data.
[112] There is only one future in Gusii, but it is not connected to the Present Tense; see Section 3.3.1.

Proto-Great Lakes was splitting into daughter languages. The Mara languages do not show evidence of a *-ra-* marker for Disjunctive Focus; in fact, the only type of Focus marker in Mara is a homorganic nasal in the Pre-SM slot (see Section 5.5 for further discussion of Focus). All five of the core languages do, however, retain the progressive usage of *-ra-*, although it is restricted to compound forms in Ikoma and Simbiti. Neither Kuria nor Gusii have this form presently, but they are also the only two languages that have kept the null form as the only marker of PRS. Moving further along the chain of development, Zanaki and Simbiti have the Vast Present marked by a *-ra-* formative. Both Zanaki and Ikizu employ the *-ra-* formative to denote future reference in addition to PRS/PROG. It is interesting to see that only Zanaki currently has SM-***ra***-R-***a*** in all three of the later phases of development along the chain (Progressive, Vast Present, and Future). The value of the *-ra-* formative for subgrouping and reconstruction will be a focal topic in Section 6, but for now it is worthwhile to mention that the data does not lead to simple conclusions.

3.4 Section Summary and Conclusions

In Section 3, I discussed the realization of various tenses in seven Mara languages. As a basis for comparison between the various languages, four degrees of past reference were established based generally on categories found within wider Bantu and specifically on the categories found in Gusii. Similarly, four degrees of future reference were established as comparative categories to highlight differences in usage between similar forms in the Mara languages. Additionally, I covered the relative Narrative tense, and the Vast Present.

For past reference, the five core languages show similarities in that they have the ANT used to indicate the Immediate and Hodiernal Past time frames, whereas Gusii and Kuria do not. Ikizu, Zanaki, Ikoma and Simbiti only have one tense for more distant past, although Simbiti shows more complexity in Narrative marking with the addition of an “Untimed Past Anterior”

form; this is likely due to an extension of what was once a past tense to NAR. Kabwa and Gusii are the only languages that have both a P_3 and a P_4 form.

For non-past reference, Ikoma, Ikizu and Zanaki all have reduced systems that establish temporal distinctions between Past and Non-Past. Gusii is simplified as well since it only has a single Future tense referring to all degrees of temporal reference. Kabwa, Simbiti and Kuria have very similar future tense systems with a tripartite division. Kabwa, however, divides the time periods for the forms differently than Kuria and Simbiti.

Already at this point, it is evident that there are two types of systems functioning in the Mara languages. The South Mara languages (i.e. Ikoma, Ikizu and Zanaki) have tense systems that make distinctions between Past events and Non-Past events. The North Mara languages (i.e. Kabwa, Simbiti, Kuria and Gusii) have several distinct Past/Narrative tenses, and (with the exception of Gusii) their Future tense systems have multiple distinctions that are quite similar to each other. The similarities amongst the languages in the North Mara group and those amongst the languages in the South Mara group will be expounded upon in Section 6.

With this analysis of the tense systems in mind, I now move on to a description of aspect amongst the Mara Bantu languages.

4 ASPECT

In general terms, the difference between tense and aspect has been described well by Comrie: "Aspect is not concerned with relating the time of the situation to any other time-point, but rather with the internal temporal constituency of the one situation; […]" (1976: 5). Nurse takes his cue from Comrie and notes that the expression of "internal temporal constituency" is

grammaticalized on the verb in Bantu; and this most commonly occurs at the FV slot (2008: 128-131).

There are four important aspectual categories in focus for this thesis. The first is the Anterior/Perfect, which will be the focus of Section 4.1. Following that, I'll turn my attention to the Progressive/Continuous (4.2), the Habitual (4.3), and the Persistive (4.4). Intensive and/or Repetitive aspectual meaning is often marked by either complete or partial reduplication of the verb root in Mara, but this topic is not treated here since it is formally quite distinct from the other categories and deserves a study of its own. In Section 4.5, the focus shifts to the Inceptive; this category has elements of both Tense and Aspect in its meaning, making it somewhat distinct from other categories. Finally, in Section 4.6, I make some initial observations regarding the combination of Tense and Aspect in the Mara languages, both through the use of compound forms with auxiliaries and through the combination of Tense and Aspect marking on a single verb form.

4.1 Anterior/Perfect

The Anterior (or Perfect) "often involves the same situation at an earlier and a later time, chronologically ordered" (Nurse 2008: 95). The realization of this depends on whether the verb itself is stative or dynamic, but the Anterior (ANT) is generally quite likely to be connected with an event that happened in the recent past (2008: 95). As I mentioned previously in Section 3.1.3, the ANT often has extended usage referring to the recent past. In all five of the core languages, the ANT is used in MCAs to describe events that occurred in either the P_1 or P_2 time frame. Kuria and Gusii, on the other hand, have distinct non-Anterior forms for P_1.[113]

As for the morphological representation of ANT, all seven Mara languages have a variant of (***n***)-SM-R-***ire*** to represent this aspect. Example (36) shows the Anterior aspect in Simbiti

[113] It is interesting to note in Kuria that there is a focused form of the ANT indicating the P_2 time frame.

marked by *-irë* at FV without a specific temporal frame indicated by an adverbial or relative clause.

(36) *Tu-shumaash-**irë*** *na-we*.
1PL-speak-ANT CCONJ-3SG
'Tumeongea naye.' / 'We have spoken with him.'

If an adverbial or relative clause is added to refer to the P1 or P2 time frame, the Anterior aspect can still be used to refer to the action. Example (37) is a Simbiti sentence indicated as P2 by an adverbial.

(37) ***N**-tu-shumaash-**irë*** *na-we* *nyinkyö* *ha-ara*.
FOC-1PL-speak-ANT CCONJ-3SG NC9.SG.morning NC16-DDEM
'Tuliongea naye asubuhi hii.' / 'We spoke with him this morning.'

From a wider Bantu perspective, Nurse refers to a form with past reference that has a null morpheme in the TA slot and *-ile*/*-ire* at FV in 45% of his matrix languages (2008: 82); but more generally, the *-ile*/*-ire* formative is represented unambiguously in 66% of his languages (2008: 264). Nurse goes on to add that, based on the study of grammaticalization paths, the Anterior meaning is much more likely to be original with a later extension to indicate specific past time frames; although the extension from Anterior to Past reference is likely to have occurred multiple times in various languages (2008: 266). Based on the widespread nature of the *-ile* formative as a marker of ANT and the fact that is not found in broader Niger-Congo, Nurse reconstructs it as a probable innovation for ANT at the PB phase (2008: 266, 279). At the Proto-GL phase, *-ile* was probably already being used for both ANT and near past reference, although this polysemy is not unique to GL (Nurse & Muzale 1999: 533; Nurse 2008: 264-267). Since the ANT and near past usages of the *-ile* formative (realized in Mara as a variant of *-ire*) were inherited into the Mara languages, the systems in Kuria and Gusii are either more recent innovations from the

proto-language or are not genetically connected to Proto-Mara. In either case, their systems do appear to be innovations from Proto-GL, along with Luhya and North Nyanza which have systems with four degrees of past reference (1999: 531).

At this point in the discussion, the relevance of combinations of Tense with the Anterior should be mentioned. The core Mara languages in this study generally indicate Past or Future Anteriors with complex two-word verb forms. The copular verb is the root of the first word and it holds the tense marking, while the second word is the lexical verb which displays the Anterior marking. These combinations of Tense and Aspect are covered in greater detail in Section 4.6, but an example is included in (38) from Ikizu to show how it works.

(38) *Tw-**aa**-**ri*** *tu-bin-**iri***.
1PL-PST-COP 1PL-dance-ANT
'Tulikuwa tumecheza (ngoma).' / 'We had danced.'

Other examples of these compound Anteriors are included in the TA charts in Appendix A.

From a discussion of the Anterior aspect in the Mara languages, I now proceed to a description of the Progressive (or Continuous) marker and how it functions.

4.2 Progressive/Continuous

Although Nurse (2008: 137) makes a basic distinction between Progressive aspect (ongoing action over a brief period of time) and Continuous aspect (ongoing action over a longer period of time), this distinction is not evident in the Mara languages; and so I will only use the term Progressive in this thesis. The Progressive aspect refers to any action that is ongoing in relation to a specific temporal frame. When the Progressive aspect is in a time frame overlapping with the current moment, it is a Present Progressive (PRS.PROG, or simply PROG) and is generally unmarked for tense. In these cases it occurs on a single verb form. As noted above in Section 3.3.4, the formative *-ra-* is used to mark the PROG in the core languages in MCAs, although

Simbiti and Ikoma use it in more restricted settings. The simple Present Progressive can be seen for Kabwa in Example (39), but it is used similarly in Ikizu and Zanaki.

(39) *Tu-**ra**-simb-**a**.*
1PL-PROG-dig-FV
'Tunachimba.' / 'We are digging.'

Ikizu, Zanaki, and Kabwa also use *-ra-* for sentences indicating repetition of an action. This "Iterative" is marked adverbially instead of through another verbal formative, as is seen in (40) for Zanaki.

(40) *Tu-**ra**-tuk-**a*** *kweeki* *na* *kweeki.*
1PL-NPST[114]-dig-FV again CCONJ again
'Tunachimba tena na tena.' / 'We are digging over and over.'

Interestingly, for Kabwa there is another way of expressing the Iterative by using the formative *-anga* at FV. This same formative is seen as the main method to indicate repetition in Simbiti. It will be dealt with in more detail when I discuss Habitual aspect in Section 4.3.

I mentioned above that the Continuous aspect is not marked distinctly from the Progressive in the Mara languages, and this can be seen more clearly for Zanaki in (41) where *-ra-* is used to refer to an action performed over the course of an entire day.

(41) *Tu-**ra**-tuk-**a*** *o-ru-siku* *o-ru-gima.*
1PL-NPST-dig-FV AUG11-NC11-day AUG11-AGR11-whole
'Tunachimba siku nzima.' / 'We are digging all day long.'

Up to this point, I have shown some of the versatility in usage of the *-ra-* formative, but in Ikizu and Zanaki it has extended its usage even further. In these languages, the PROG *-ra-* is also employed more generally as a marker of Future events. Additionally, in Zanaki, *-ra-* is used

[114] The *-ra-* formative is used to indicate the PROG, the FUT and the PRS in Zanaki and is best considered a marker of NPST.

to indicate the Vast Present, so it is best treated as a marker of Non-Past in that language. This is discussed in Section 3.3.1.

In contrast to the languages that use *-ra-* for Present Progressive, Simbiti shares an innovative compound form with Kuria and Gusii to indicate the Present Progressive. The form consists of a focus-marked infinitival followed by a copular auxiliary (predominantly of the *-li/-ri*[115] variety), as seen for Simbiti in (42) with either a Progressive or Continuous meaning.

(42) (*U-ru-sikö* *u-ru-ghima,*) ***n-ku-tuk-a*** *tö-**rë***[116].
AUG11-NC11-day AUG11-AGR11-whole FOC-INF-dig-FV 1PL-COP
'Tunachimba (siku nzima).' / 'We are digging (all day long).'

Since the copular auxiliary generally appears as the first word in Bantu compounds (Nurse 2007: 128), this auxiliary inversion is an important innovation. In fact, within Eastern Bantu, the only other languages (outside of North Mara) that have formal correlates of an auxiliary inversion are Rangi (F33) and Mbugwe (F34), though the inversion is used for more than just Present Progressives (Nurse 2007: 128, 158). There are proposals that the Rangi/Mbugwe auxiliary inversion was influenced by contact with Southern Cushitic languages (Mous 2000, as referenced in Nurse 2007: 158); but even so, this type of innovation suggests similarities between North Mara and Rangi/Mbugwe that deserve further attention. If Southern Cushitic influenced the inversion in Rangi/Mbugwe, it may also have influenced a similar inversion in the North Mara languages which show other linguistic evidence of contact with Southern Cushitic (Schoenbrun 1990). In relation to this, another puzzle is why the auxiliary inversion is not seen in Kabwa and South Mara. The topic of auxiliary inversion cannot be given a thorough treatment in this research and deserves further attention.

[115] The copular *-li* is the standard form used as the historical pre-cursor in Eastern Bantu, but in the Mara languages this is evidenced as *-ri*, or with vowel variations, as in Simbiti *-rë*.
[116] Dahl's law is optional in Simbiti and so *nghutuka törë* is also an acceptable pronunciation.

Returning to the topic of Progressives in the Mara languages, Ikoma does not have a clear Present Progressive form since the lexical aspect of the verb seems to play a big role in whether or not the Past Inceptive (PST.INCE) form can be used to express the meaning of the Present Progressive. More generally, Ikoma has a general FUT form that can be used is situations where Present Progressive is indicated; and this form is marked by ***n***-SM-***Vko***-R-***a***. This PRS.PROG can be seen in Example (43).

(43) ***N***-*to*-***ogho***-*tuk*-***a*** *mw-ishe* *mo-ghima.*
FOC-1PL-PRS.PROG-dig-FV NC3.SG-day AGR3-whole
'Tunachimba siku nzima.' / 'We are digging all day long.'

In spite of the distinct realization of the Present Progressive in Ikoma, there are situations where this language uses the SM-***ra***-R-***a*** form. These are found specifically in compound Progressive constructions (i.e. Past and Future Progressives) where Progressive aspect is marked on the second (lexical) verb. An example of a Past Progressive for Ikoma is displayed in (44).

(44) ***N***-*to*-***o***-***re*** *to*-***ra***-*bhin*-***a***.
FOC-1PL-PST-COP 1PL-PROG-dance-FV
'Tulikuwa tunacheza.' / 'We were dancing.'

Simbiti also uses SM-***ra***-R-***a*** in certain types of compound Progressives, but these are reserved for Future time frames. The sentence in (45) shows a Future Progressive in the Remote Future (F_4) time frame.

(45) ***N***-*tu*-***ri***-***bh***-***anga*** *to*-***ra***-*bhin*-***a***.
FOC-1PL-F_3/F_4-COP 1PL-PROG-dance-FV
'Tutakuwa tunacheza.' / 'We will be dancing.'

The usage of SM-***ra***-R-***a*** to mark the Progressive aspect in compound Past or Future Progressives is not limited to Simbiti and Ikoma. The other Mara languages also use forms like

this with tense marked the on an auxiliary copula and Progressive aspect marked on the main verb. The use of auxiliaries will be covered more fully in Section 4.6 below.

Before leaving the topic of the *-ra-* formative, it is interesting to note that Kuria uses it on a form that Mwita calls the "Hortatory Imperative (3)" in contrast to two other hortatory imperatives (2008: 117). Based on its use in the sentences, it appears to be related to the Inceptive or Immediate Future categories in this study.

The discussion of various Progressive forms in Mara is condensed into Table 4 which shows the realizations of Present Progressives, (Remote) Past Progressives, and (Remote) Future Progressives in all seven of the Mara languages.

Table 4: Present, Past, and Future Progressives in the Mara Languages

		Type of Progressive		
		PRS.PROG	Past (P_4) Progressive	Future (F_4) Progressive
Language	Ikizu	SM-***ra***-R-***a***	SM-***aa-ri*** SM-***ra***-R-***a***	SM-***ra-b-a*** SM-***ra***-R-***a***
	Zanaki	SM-***ra***-R-***a***	SM-***aa-ri*** SM-***ra***-R-***a***	SM-***ra-bh-a*** SM-***ra***-R-***a***
	Ikoma	SM-***Vko***-R-***a***	***n***-SM-***V-re*** SM-***ra***-R-***a***	***n***-SM-***Vku-bh-a*** SM-***ra***-R-***a***
	Kabwa	SM-***ra***-R-***a***	SM-***aa-ri-nga*** SM-***ra***-R-***a***	SM-***ri-bh-anga*** SM-***ra***-R-***a***
	Simbiti	***n-ko***-R-***a*** SM-***re***	***(n)***-SM-***aa-re-nga ko***-R-***a***	***(n)***-SM-***ri-bh-anga*** SM-***ra***-R-***a***
	Kuria[117]	***n-ko***-R-***a*** SM-***re***	--[118]	--[119]
	Gusii[120]	***n-kó***-R-***a*** SM-***re***	P_3: SM-***a-are ko***-R-***a*** or ***n-ko***-R-***a*** SM-***a-are***[121]	--[122]

[117] The source for the [kuj] non-tonal data is Cammenga (2004), with underlying tones obtained from Mwita's (2008) analysis of Cammenga.
[118] This category is not specifically dealt with in Cammenga (2004).
[119] This category is not specifically dealt with in Cammenga (2004).
[120] The source for all the [guz] data is Nurse (2007: 127-129), which is based on Whiteley (1960); underlying tones are used for all the Gusii data.
[121] Underlying tones are not known for the Past Progressive forms.
[122] There is no data for this category included in Nurse (2007).

This chart clearly shows the ubiquitous nature of the SM-***ra***-R-***a*** form as a marker of Progressive amongst the core languages. The importance of this form from a historical-comparative standpoint is detailed in Section 6.

Now that I have looked at both the Progressive and Anterior aspectual categories, I continue on to a discussion of the Habitual, another important aspectual category amongst the Mara languages.

4.3 Habitual

The Habitual aspect (HAB) refers to an action that is performed on a regular basis. A closely related concept is the Iterative, which refers to an action that is performed repetitively. In 43% of the languages in Nurse's study of Bantu, the habitual was found to be "discrete from other imperfective categories [… and] overwhelmingly encoded by the inherited suffix *-ag-/-ang-/-aga/-anga*" (2008: 143-44). Of the Mara languages in this study, only Kabwa and Simbiti make use of this *-anga* morpheme at FV.[123] In addition, the usage of this morpheme appears to have extended beyond the Habitual, and can now be used as a more general imperfective, especially in Simbiti. Nurse and Muzale trace *-aga/-anga* to the Proto-GL phase, since it appears as a HAB marker across the GL area (1999: 524-525).

The formative *-anga* extends its Habitual meaning to the related Iterative concept. This can be seen for Simbiti in Example (46), but the same form is found in Kabwa as well.

(46) *To-**ra**-tuk-**anga***.
1PL-PRS-dig-HAB
'Tunachimba (tena na tena).' / 'We are digging (again and again).'

[123] The related formative *-aga* at FV is, however, attested in the nearby (and purportedly related) Suguti languages (i.e. Nurse & Muzale 1999: 524, 530).

In the three South Mara languages in this study (Ikoma, Ikizu and Zanaki), *-anga* is not used at all;[124] rather, a Pre-SM morpheme *-haa-* marks the Habitual. A sentence in Ikoma using this formative is seen in (47).

(47) *Mara* *kaaru,* *itu* *n-to-**haa**-ghi* *Tarisaraamu.*
NC10.PL.time many 1PL.PRON FOC-1PL-HAB-go Dar.Es.Salaam
'Mara kwa mara, sisi huenda Dar Es Salaam.' / 'We go to Dar Es Salaam frequently.'

The verb *-ha/-haana* 'give' in the South Mara languages is a potential source for the grammaticalization of the Habitual morpheme;[125] however, the grammaticalization path from 'give' to Habitual does not seem to be well attested in the languages of the world (Heine & Kuteva 2002b: 331; Bybee et al. 1994: 154-158). Additionally, Nurse does not mention a formative of similar shape to the *-haa-* Habitual within Bantu (2008: 143-145). All of these things together point toward *-haa-* being a unique development within the South Mara languages, which provides an interesting avenue for future research.

Based on the discussion in this section, there is a clear split between how the core languages mark the Habitual aspect. The North Mara languages keep the inherited *-anga* (from PB *-aga*) at FV (Nurse 2008: 279; Nurse & Muzale 1999: 525); the South Mara languages, on the other hand, have innovated the usage of the *-haa-* formative in the TA slot. I will review this evidence again in relation to sub-grouping the Mara languages in Section 6.2.

4.4 Persistive

Based on Nurse's definition, the Persistive (PER) is an aspect that "affirms that a situation has held continuously since an implicit or explicit point in the past up to the time of speaking" (2008:

[124] Isenye, a closely related speech variety to Ikoma, attests a Remote Past/Hesternal Past with the *-anga* formative at FV (Mitterhofer & Robinson 2012).
[125] Giha (JD66), another GL language (from the West Highlands branch), uses *-haaye* a Perfect form of the verb *-ha* 'give' as an auxiliary in progressive compound forms (Harjula 2004: 107).

145). In English this could be translated by the adverbials ‘still’ and ‘not yet’ depending on whether the polarity is positive or negative.

Nurse posits *-kí-* as the marker of PER for PB since it has broad distribution within the Bantu area (2008: 146). In Great Lakes a reflex of this has been combined with *-a(a)-*, for example in Haya (E22), to form a Persistive formative *-kiáá-* in the TA slot (2008: 147). This formative is found widely enough in GL Bantu, that it can be reconstructed for Proto-GL (Nurse & Muzale 1999: 526-527). Since the *-a(a)-* formative refers to past reference in the GL languages, the combined morpheme *-kiáá-* indicates that an event that started in the past is continuing into the present (1999: 526-527). Although *-kia(a)-/-kya(a)-* is reconstructible for PER in Proto-GL, there is sufficient occurrence of simply *-ki-*, or a related variant, to think that the system was a little more complex. It seems possible that there was variation between *-kia(a)-/-kya(a)-* and *-ki-* in the early GL system and that these two formatives were used to indicate different temporal relationships of Persistive situations.

There is variation of these two formatives in the Mara languages as well; several of the languages have both forms for the Persistive depending on its relationship to tense. In Zanaki, *-kyaa-* and *-ki-* are both seen depending on the tense to which the Persistive aspect is related. Example (48) shows that *-kyaa-* is used for a Persistive action that begins in the past and continues in the present.

(48) *Tu-**kyaa**-ri* *tu-ri*[126]*-ibhisiry-a* *na-we.*
1PL-PER-COP 1PL-PROG-speak-FV CCONJ-3SG
‘Bado tunaongea naye.’ / ‘We are still speaking with him.’

The *-kyaa-* formative is also used in situations where a situation did not hold in the past and still does not hold in the present. This type of usage can be seen in (49).

[126] This formative is underlyingly *-ra-*, but it surfaces as *-ri-* because there is vowel assimilation with the root-initial <i> (Gray & Smith 2010b: 15).

(49) *Tu-**kyaa**-ri* *kw-ibhisiry-a* *na-we.*
1PL-PER-COP INF-speak-FV CCONJ-3SG
'Bado hatujaongea naye.' / 'We still haven't spoken with him.'

However, there are situations where the *-ki-* formative is used in Zanaki instead of the *-kyaa-* formative. This can be observed in (50) where a situation that does not hold in the present is expected to come about in the future.

(50) *Tu-**ki**-ibhisiry-a* *na-we.*
1PL-FUT.PER-speak-FV CCONJ-3SG
'Bado tutaongea naye.' / 'We will still speak with him.'

In Ikizu, the formative *-k~~ii~~-* is used in most Persistive situations including those where *-kyaa-* is seen in Zanaki. The usage of Persistive *-k~~ii~~-* can be seen in Example (51) where a situation that held in the past continues to hold in the present.

(51) *T~~u~~-**k~~ii~~**-r~~i~~* *t~~u~~-ra-gaamb-an-a* *na-w~~i~~.*
1PL-PER-COP 1PL-PROG-speak-RECP-FV CCONJ-3SG
'Bado tunaongea naye.' / 'We are still speaking with him.'

However, in addition to the *-k~~ii~~-* formative, a *-cha-* formative is found in at least one Persistive construction as well. This construction is presented for Ikizu in (52) where a situation that does not hold currently is expected to occur in the future.

(52) *T~~u~~-**k~~ii~~**-r~~i~~* *t~~u~~-**cha**-az-a* *na-we.*
1PL-PER-COP 1PL-PER-come-FV CCONJ-3SG
'Bado tutaongea naye.' / 'We will still speak with him.'

In Ikizu, <ch> is known to have a palatalized /k^{j}/ underlyingly, so the *-cha-* formative appears to have originated from *-kya-* (Gray & Smith 2010a: 9).

Simbiti uses the formative *-këë-* (similar to *-k~~ii~~-* in Ikizu) to mark positive realizations of the PER, but negatives are marked with the formative *-raa-* that is also used on the F1. An example of a positive Persistive in Simbiti is included as Example (53).

(53) *N-ku-shumaash-a* *tö-**këë**-rë* *na-we.*
FOC-PROG-speak-FV 1PL-PER-COP CCONJ-3SG
'Bado tutaongea naye.' / 'We will still speak with him.'

This example also shows the auxiliary inversion that is discussed for Simbiti in Section 4.2.

Seemingly related to the Persistive in Simbiti, a *-ke-* formative (with focus at Pre-SM and *-ere* at FV) is also used in Kuria to mark a simultaneous aspect that includes reference to an action still going on (Mwita 2008: 46).

In comparison to the Mara languages already mentioned in this section, Kabwa and Ikoma have less variation in Persistive forms. However, the formatives these two languages use to express Persistive are different: Kabwa has *-kya-* while Ikoma has *-kee-*. The Persistive for Kabwa can be seen in Example (54).

(54) *Tu-**kya**-hair-a* *na-we.*
1PL-PER-speak-FV CCONJ-3SG
'Bado tunaongea naye.' / 'We are still speaking with him.'

Ikoma's *-kee-* Persistive in Example (55) includes the copular nasal Focus marker at Pre-SM, which is a frequent feature of MC verb forms in this language.

(55) *N-to-**kee**-re* *to-ra-siiker-a* *na-we.*
FOC-1PL-PER-COP 1PL-PROG-speak-FV CCONJ-3SG
'Bado tunaongea naye.' / 'We are still speaking with him.'

Based on the preceding examples and additional data that I elicited, Table 5 summarizes the various formatives related to Persistive aspect in the core Mara languages.

Table 5: Persistive Marking in the Core Mara Languages

		Types of Persistive Aspect		
		Situation holding in BOTH PAST AND PRESENT	Situation holding NEITHER IN PAST NOR IN PRESENT	Situation NOT holding at PRESENT BUT EXPECTED IN FUTURE
Language	Ikizu	SM-***kii-ri*** SM-***ra***-R-***a***	SM-***kii-ri ku***-R-***a***	SM-***kii-ri*** SM-***cha-az-a ku***-R-***a***
	Zanaki	SM-***kyaa-ri*** SM-***ra***-R-***a***	SM-***kyaa-ri ku***-R-***a***	SM-***ki***-R-***a***
	Ikoma	***n***-SM-***kee-re*** SM-***ra***-R-***a***	***n***-SM-***kee-re*** SM-***taa***-R-***a*** SM-***Ø***-R-***a***	***n***-SM-***kee-re n***-SM-***Vku-ch-a ko***-R-***a***
	Kabwa	SM-***kyaa***-R-***a***	SM-***kyaa-ri o-ku***-R-***a***	SM-***kya-aj-a o-ku***-R-***a***
	Simbiti	***n-ku***-R-***a*** SM-***këë-rë***	***te***-SM-***raa***-R-***a … hë***	***n***-SM-***raa-ash-e***[127] SM-***Ø***-R-***e***

More research is needed to understand the variation in Persistive forms amongst the Mara languages, but at this point it is clear that there are two main formatives employed in these constructions: *-kya(a)-/-cha(a)-* and *-kɪ(ɪ)-/-ki(i)-/-ke(e)-*.

With this background on the Persistive aspect, I conclude the sections focused pure aspectual distinctions in the Mara languages. However, there is one other category called the Inceptive that is closely associated with both Tense and Aspect; I turn to a discussion of that category now.

4.5 A Category Involving Tense and Aspect: The Inceptive

Nurse mentions that the Inceptive is a category in some Bantu languages that is related to the Anterior (2008: 161). The Inceptive (INCE) is either Future-oriented, dealing with the idea of being about to do something, or Past-oriented, referring to having just done something. These two concepts seem to be at odds with each other as one is referring to something that began in the immediate past and the other refers to a state that will begin in the immediate future. Nurse refers to these two types of Inceptive as the "state-just-before or at-its-inception" and the "state-after-its-coming-into-being" (2008: 162). The interesting thing to note about these two

[127] Underlyingly there are three vowels in a row in this word, but no more than two vowels in a row can surface in Simbiti.

types of Inceptive is that they have both temporal and aspectual properties associated with them. In other words, these categories blur the line between Tense and Aspect.

The "state-just-before or at-its-inception", or Future Inceptive (FUT.INCE) is realized as ***(n)***-SM-***raa***-R-***e*** in Kabwa and is shown in (57). This form is very similar to the F_1 in Simbiti, which was elicited in relation to a specific temporal frame, even though it also has some overlap with the F_2 in common use. An example of the F_1 for Simbiti with a specific Immediate Future time frame was shown in (27) above and is repeated in (56) for quick reference.

(56) *Ha-no* *se-raa-het-e* *se-dakika* *i-kömi,*
NC16-PDEM.REL SBJ10-F_1-pass.by-SBJV NC10.PL-minute AGR10-ten

n*-to-****raa****-shumaash-****e*** *na-we.*
FOC-1PL-F_1/F_2-speak-FV CCONJ-3SG.PRON
'Tutaongea naye baada ya dakika kumi.' / 'We will speak with him in ten minutes.'

Since the Simbiti form has a temporal reference, I will treat it as an Immediate Future tense (F_1) in this thesis. However, when the Swahili sentence in (56) is elicited for Kabwa, a distinctive temporal form SM-***aka***-R-***e*** is used. This indicates that ***(n)***-SM-***raa***-R-***e*** is used slightly differently in Kabwa.

The form ***(n)***-SM-***raa***-R-***e*** has more of an aspectual component in Kabwa indicating the immediacy of the action rather than the specific time it will take place. This is evident in Example (57).

(57) *Tu-****raa****-hair-****e*** *na-we* *bhw-angu* *ha-nu.*
1PL-FUT.INCE-speak-FV CCONJ-3SG.PRON NC14-quick NC16-PDEM
'Tumtaongea naye hivi karibuni.' / 'We will speak with him soon.'

This aspectual component to the Kabwa category signifies that it is better classified as a Future Inceptive rather than a pure Future tense. Even though there is a slight distinction between the

usage of ***(n)***-SM-***raa***-R-***e*** as an F_1 in Simbiti and as a FUT.INCE in Kabwa, the morphological similarity points toward a historical relationship between the two.

The "state-after-its-coming-into-being" Inceptive is a category that Ikoma employs; Robinson refers to this simply as the Inceptive (2013: 8). In this thesis, I refer to this form as a Past Inceptive (PST.INCE) in contradistinction to the Future Inceptive that is found in Kabwa. This form has some relationship to the temporal frame of an immediate past tense, but its aspectual component is more in focus. The PST.INCE indicates that an action has just started and carries current relevance. In Ikoma this PST.INCE is marked by the form SM-***Vká***-R-***a***, as seen in Example (58), which is adapted from Robinson (2013: 8).

(58) ^*Bh*[128]-***agha***-*tuk*-***a***.[129]
3PL-PST.INCE-dig-FV
'They have begun to dig.'

As I mentioned above in Section 3.3.4, the PST.INCE on certain verb types (active verbs like 'dig') also has relevance to the current moment and is thus used to translate the Vast Present (PRS).

From a broader comparative perspective, the formative -*Vká*- in Ikoma is of interest. Both the North Nyanza and Luhya branches of Great Lakes have -*aka*- as a marker of 'just past' (Nurse & Muzale 1999: 531). In the other Mara languages, there does not appear to be a similar formative marking past, except for the tonally distinct Narrative tense (refer to Section 3.2). Both Kabwa and Simbiti employ the formative -*aka*- for a near/middle future tense, but in these cases the final vowel is -*e* instead of -*a*.

Since, within Mara, the PST.INCE only occurs in Ikoma, the most immediate explanation is that it was innovated within the system of that language. The SM-***Vká***-R-***a*** form is definitely

[128] This morpheme is underlyingly *bha*-, but the second *a* is not written in the orthography (Robinson 2013: 8).
[129] The carat sign (^) at prior to the verb form is used to distinguish this tense in the orthography from the NAR which is tonally distinct and written without the carat (Higgins 2010: 13).

related to the tonally distinct SM-***Vka***-R-***a***, which is used for the NAR. Although there is often a link between the ANT and the INCE, there is less semantic similarity between the NAR[130] and the INCE. A verbal grammaticalization source for this form is also not readily apparent. Another explanation could be that the *-Vká-* formative is an inherited feature in Ikoma, thus explaining the similarity between this and the *-aka-* 'just past' formative in Luhya/North Nyanza. This is an interesting link and all the more puzzling since none of the other Mara languages have an obvious correlate. This formative deserves further study in Ikoma and the languages most closely related to it, but for now the origins of the PST.INCE in Ikoma remain difficult to ascertain.

With this overview of the two Inceptive categories in the Mara languages, I proceed to broaden the scope of my investigation by looking at constructions that clearly combine formatives for both Tense and Aspect or even for two different Aspectual categories.

4.6 Combinations of Tense and Aspect

When combining tense and aspect into a single verbal form, there are two widespread strategies across the main portion of the Bantu area: inflection and compounding. Even though an individual language generally prefers inflection over compounding, or vice versa, there is no known language that only allows one strategy at the expense of the other (Nurse 2008: 170). This is in large part due to the ongoing "grammaticalization cycle" whereby compound structures grammaticalize with concomitant morphophonological reduction to become inflectional forms over time. The inflectional forms can in turn be replaced by newer compound structures and the cycle begins over again (2008: 170-171).

Amongst the Mara languages, combinations of Tense and Aspect are frequently realized as compound constructions. In these compounds, tense is always marked first on an auxiliary

[130] This is quite possibly linked to a far past tense historically in Great Lakes (Nurse & Muzale 1999: 532-533).

verb and then the aspect is marked on the lexical verb, which comes second in the compound form. Example (59) shows this type of construction in Ikoma.

(59) ***N***-*to*-***oku***-*bh*-***a*** *to*-***ra***-*bhin*-***a***.
FOC-1PL-FUT-be-FV 1PL-PROG-dance- FV
'Tutakuwa tukicheza (ngoma).' / 'We will be dancing.'

Although compound forms are common across the Mara languages, the North Mara languages have a slightly stronger tendency for inflectional forms. This is seen especially in forms that make use of the suffix *-anga*. I already remarked in Section 4.2 that these forms represent Habitual aspect; however, Example (60) shows that in Kabwa this suffix can also be used as a more general marker of the Imperfective. This type of usage is found in Simbiti as well.

(60) *Tw*-***a***-*bhin*-***anga***.
1PL-P_3/P_4-dance-IPFV
'Tulikuwa tukicheza (ngoma).' / 'We were dancing.'

According to my language consultant, the verb form in (60) is very similar in meaning to the Kabwa compound form in Example (61).

(61) *Tw*-***a***-***ri***-***nga*** *tu*-***ra***-*bhin*-***a***.
1PL-PST-COP-P_3/P_4 1PL-PROG-dance- FV
'Tulikuwa tukicheza (ngoma).' / 'We were dancing.'

The inflectional form in (60) is a feature of the North Mara languages that links them more closely with the nearby Suguti languages, which often mark Tense/Aspect combinations inflectionally on a single-word verbal form in contradistinction to the majority of GL languages. Nurse and Muzale give an example of this in Ruri (adapted from 1999: 523), which is shown in (62).

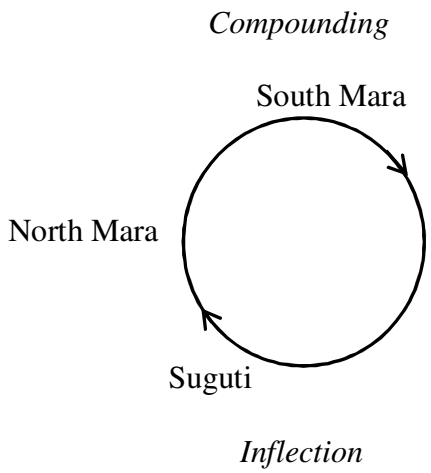
(62) a. ***e***-*ci*-***Ø***-*gur*-***aga*** 'we buy'

b. *ci*-***a***-*gur*-***aga*** 'we used to buy, were buying'

c. *ci*-***aka***-*gur*-***aga*** 'we will buy regularly'

Based on the grammaticalization cycle mentioned earlier in this section for combined Tense/Aspect forms in Bantu languages, a pictorial representation can be drawn to show whether a language (or group of languages) currently prefers inflection or compounding (or if the preference lies somewhere between the two). Figure 8 locates where Suguti, North Mara and South Mara are in relation to each other in this cycle.

Figure 8: Compounding/Inflection Cycle for Combined TA Forms



Previous examples have already shown that the Mara languages employ compounding strategies with copular auxiliaries. According to Botne, these copular auxiliaries are not "semantically empty" tense-holders, as was originally assumed; instead, they should be treated as shifters which "are *referential* in that they refer to a second locus of orientation and *indexical* in that they index an event from context (1986: 315-316). That is, the temporal marking on the auxiliary verb relates the timing of that event relative to another event by context (either explicit or implicit). The TA marking on the main (or lexical) verb, which is the second verb in the compound, can then be of two types: Tense or Aspect (1986: 310). The Mara languages are of the Aspect-type (like Swahili), which distinguishes them from the Tense-type West Highlands

languages Kinyarwanda and Kirundi (1986: 310, 312). In Examples (59) and (61) above, both Ikoma and Kabwa display the Aspect-type of compounding by marking the main verb with the PROG Aspect marker *-ra-*. If the main verb is marked for the Anterior instead, then the sentence is interpreted as a Past Perfect. This is displayed in Example (63).

(63) *Tw-**a**-**ri**-**nga** tu-bhin-**iri**.*
1PL-PST-COP-P_3/P_4 1PL-dance-ANT
'Tulikuwa tumecheza (ngoma).' / 'We had danced.'

In addition to distinctions between Tense and Aspect-type compounding languages, Botne also differentiates the general usage of two copular auxiliaries: *-ba*, which indicates "durative", and *-li*,[131] which refers to a "punctual" action (1986: 314). These two kinds of copular auxiliaries are evident in the Mara languages as well, as was revealed in Examples (59) and (61). The Zanaki data in (64) shows additional evidence that Botne's distinctions are valid in Mara (adapted from Futakamba et al. 2013: 5-6).

(64) a. *ni-**haa**-**bh**-**a** ni-**ra**-tuk-**a*** 'I was digging' (for a long time; "duration")

b. *ni-**ka**-**bh**-**a** ni-**ra**-tuk-**a*** 'I was digging' (after a process; "process")

c. *n-**aa**-**ri** ni-**ra**-tuk-**a*** 'I was digging' (for a short time; "non-duration")

The *-bha* copular form, when combined with two different TA markers, indicates duration and the *-ri* copular form indicates punctuality.

In addition to discussing the semantic difference between the two copular auxiliaries, it is important to elucidate their different lexical status. In all five of the Mara languages, the *-bha* copula comes from a lexical verb form meaning 'to be' or 'to become', whereas *-ri* is a defective copula that is only used in certain constructions. Example (65) shows the infinitival form of *-bha* in all five of the core languages.

[131] In the Mara languages there is no distinction between the sounds *l* and *r*, but *r*'s are used in the orthographies because they are a better representation of the underlying phoneme /r/. Therefore, *-li* is seen as *-ri* in the Mara languages.

(65) Kabwa: *o-ku-**bh-a*** ‘to be/become’

Simbiti: *o-ko-**bh-a*** ‘to be/become’

Ikizu: *k~~u~~-**b-a*** ‘to be/become’

Zanaki: *o-ku-**bh-a*** ‘to be/become’

Ikoma: *ku-**bh-a*** ‘to be/become’

Due to the fact that *-bha* is a regular verb root, it can hold any Future or Past tense marking, as well as aspectual marking like the ANT. The defective copula *-ri*, on the other hand, is only used with Past tense (as in (61) or (63) for Kabwa), or without a tense marker in locative constructions. The Zanaki example in (66), taken from lines Zak18a-b in Appendix F, exhibits the locative function of *-ri*.

(66) *…Waandibha a-ka-ror-a Kusaya a-**ri** i-guru bhu-kong’u.*
Wandiba 3SG-NAR-dance-FV Kusaya 3SG-COP NC23-high NC14-very
‘…Wandiba akaona Kusaya yuko juu sana.’ / ‘…Wandiba saw that Kusaya was really high.’

Before I close this section, it is important to mention that the copular auxiliaries are not the only auxiliary verbs used in compound constructions. We have already seen the use of the “come” auxiliary to denote the F_4 in Zanaki (Section 3.3.1) and the F_2 in Simbiti (Section 3.3.3). One other auxiliary found throughout the core languages is the verb *mara* ‘finish’, which indicates that an action is already completed. An instance of this is exemplified in (67) for Ikizu.

(67) *T~~u~~-mar-**iri** k~~u~~-gaamb-an-a na-w~~i~~.*
1PL-finish-ANT INF-speak-RECP-FV CCONJ-3SG.PRON
‘Tumeshaongea naye.’ / ‘We have already spoken with him.’

In addition to these, there are various other verbs used as auxiliaries for specific constructions in the Mara languages. Some of these auxiliaries are indicated in the tables for individual languages in Appendix A; however, the topic of auxiliation in Mara needs further research so that they can be studied in detail.

4.7 Section Summary and Conclusion

Throughout Section 4, I discussed various aspectual categories relevant to the Mara languages, and I observed how the morphological realizations of these categories relate to understanding historical relationships. The first category covered is the Anterior/Perfect, which is realized as a variant of the form (***n***)-SM-R-***ire*** in all the Mara languages and is reconstructible to an earlier language phase. The five core languages use this form to cover both the P_1 and P_2 time frames, so its usage overlaps with Past tense. Kuria and Gusii, however, have innovated new forms to refer to the P1 time frame.

In Section 4.2, I explained the realization of the Progressive aspect in more detail. For the five core Mara languages, compound Progressives (Past and Future Progressives) all employ the form SM-***ra***-R-***a*** in the main verb position. This form is also used for Present Progressives in Kabwa, Zanaki, and Ikizu; Simbiti and Ikoma use different forms to indicate a Present Progressive. In addition to this, Simbiti, Kuria and Gusii all have an auxiliary inversion for their Present Progressive that is very rare within Bantu.

The Habitual aspect is the topic of Section 4.3 and analysis of this category reveals a clear split between the North Mara and South Mara languages. The North Mara languages make use of the *-anga* morpheme at FV, in conjunction with other morphemes in the TA slot, to indicate a habitual action. The South Mara languages, on the other hand, have innovated the formative *-haa-* in the TA slot as a marker of habituality.

The Persistive aspect is the focus of Section 4.4, and a study of this category amongst the Mara languages shows that two morphemes are used widely: *-kya(a)-/-cha(a)-* and *-kɪ(ɪ)-/-ki(i)-/-ke(e)-*.

Section 4.5 introduces the Inceptive aspect which includes elements of both a tense and an aspect. In the Mara languages, the Inceptive is not a category included in enough of the

languages to reconstruct it for a hypothetical Proto-Mara phase. In Kabwa, a Future Inceptive overlaps in usage with actions in the F_1 time frame, but this category is focused more on the fact that the action is "about to" start then on the temporal component. Ikoma, on the other hand, makes use of a Past Inceptive that indicates that an action has just begun.

In Section 4.6, the focus is on forms that combine realizations of Tense and Aspect into complex forms. These combinations occur through either inflection or through compounding. The North Mara languages show a stronger preference for inflection, although they also employ compound forms. The South Mara languages tend to prefer compound forms.

The discussion of aspect in this section provides more evidence for a split between the North and South Mara languages. The evidence for a Proto-Mara language is scant and mainly based on categories that appear to be inherited from an even earlier phase of language evolution.

With this discussion in mind, Section 5 shifts the focus to additional verbal categories (besides Tense and Aspect) that can provide further support for genetic relationships between the Mara languages.

5 OTHER CATEGORIES: MODALITY, FOCUS, NEGATION, AND RELATIVE CLAUSES

There are a number of "other" categories that are commonly marked on the verb in Bantu including modality, focus, negation and relative clauses. Although many of these topics deserve a much wider treatment, they are not the focus of my present work and are mentioned in order to give a broader context for TAM in Mara Bantu. In addition, Negative polarity and Relative Clause (RC) constructions are of particular interest, as verbal marking in these categories tends to be more conservative, while innovative usage of TAM categories usually appears first in MCAs (Nurse & Muzale 1999: 520-521).

The modal categories in focus are: Subjunctive Mood (Section 5.1), Conditional (Section 5.2), Hypothetical (Section 5.3) and Potential (Section 5.4). Following that, Section 5.5 will deal with Focus. In Section 5.6, Negation will be the topic; and then Section 5.7 will proceed to a discussion of Relative Clauses. The concluding section will be a summary of Section 5.

5.1 Subjunctive Mood

The Subjunctive[132] mood (SBJV) indicates "speaker attitudes such as uncertainty, tentativeness, vagueness, suggestions, preferences, hopes, fears, wishes, and some commands, by contrast with the indicative, which represents situations as factual" (Nurse 2008: 317). The SBJV is overwhelmingly marked throughout the Bantu area by a mid-front vowel in the FV slot and has been reconstructed as *-é* for PB (2008: 276). The Mara languages have preserved this formative and show an overwhelming preference for *-e*/*-ɛ* in the FV position to mark the SBJV. These Subjunctives are commonly used to indicate a plural command. This can be seen in Example (68) for Ikoma.

(68) *Mo-**Ø**-tuk-**e*** *cha-numbu*!
2PL-IMP.PL-dig-SBJV NC10.PL-potato
'Chimbeni viazi!' / 'Dig up (pl.) potatoes!'

If the first person (i.e. 'we') is included as the subject of the command then the meaning can even be more of the expression of a wish or a desire, as observed in the Ikoma sentence in (69).

(69) *To-**Ø**-siiker-**e*** *na-we.*
1PL-IMP.PL-speak.with-SBJV CCONJ-3SG
'Tuongee naye.' / 'Let's speak with him.' or 'We should speak with him.'

Ikizu uses a slightly different vowel to mark Subjunctive than the other Mara languages, which is due to a vowel harmony process in that language (Gray & Smith 2010a: 13-14). The

[132] Optative is a closely related term that has been used instead of Subjunctive in some literature on the Mara languages (e.g. Gray 2013).

Ikizu Subjunctive appears as a higher front vowel, -***ɪ***,[133] if the vowel in the verb root is not a mid-vowel, as Example (70) shows.

(70) *Mʉ-**Ø**-sʉʉmb-**ɪ*** *zɪ-nʉʉmbʉ*!
2PL-IMP.PL-dig-SBJV NC10.PL-potato
'Chimbeni viazi!' / 'Dig up (pl.) potatoes!'

Across several of the Mara languages, and even further abroad within Bantu, a Subjunctive FV is used not only in Imperative forms but also in some of the future tense forms. Both Kabwa and Simbiti have it in their F_2 forms, which are variants of SM-***aka***-R-***ɛ***. Kabwa, Simbiti and Kuria all have a similar F_1/FUT.INCE form (a variant of ***(n)***-SM-***raa***-R-***ɛ***) that uses the Subjunctive FV. In Gusii, the only future tense, ***(n)***-SM-***Ø***-R-***ɛ***, is closely related to the Subjunctive form. This relationship between the Subjunctive and (near) Future seems to have something to do with the unplanned and unpredictable nature of events that have not yet occurred; therefore, it is plausible not to equate the Future with the Indicative mood which normally indicates "factual" situations. In fact, the Future/Subjunctive association occurs in more than merely a local distribution within Bantu (Nurse 2008: 261).

5.2 Conditional

The Conditional concept is encoded by the first clause in "If/when X, then Y" sentences or a negative formulation of these. This overlaps with the category that Nurse most frequently refers to as "concessive" which is marked most widely by the morpheme *-nga-* in Bantu (2008: 66, 238). In the Mara languages, *-nga-* is only seen as a Hypothetical (see Section 5.3 for more on this), and the Conditional is marked in other ways.

One of the difficulties in eliciting the Conditional (COND) through the use of Swahili is the fact that Swahili can explain this idea with or without the adverbial *kama* 'if'. When elicited with the adverbial, the sentences in the target languages also tend to appear with an adverbial

[133] Phonetically this sound is [ɪ].

followed by a verb form that is not necessarily distinctive from basic TA forms. For example, in Ikizu and Zanaki, the adverbial 'if' is followed by a PROG verb form SM-***ra***-R-***a***. Since we are mostly interested in discovering if a COND formative exists in the Mara languages, this is less useful for our present study.

The second way that the COND is marked in Swahili is using the formative *-ki-* at the TA slot without the use of an if-adverbial. This sentence type yields a clear formative *-raa-* at TA for all five of the Mara core languages. Generally, the COND form also has a SBJV at FV, as witnessed in Example (71) for Ikizu.

(71) *Tu-**raa**-buuk-**i*** *naangwiinu*, *tu-ra-keezer-a* *ku-hik-a*
1PL-COND-depart-SBJV now 1PL-FUT-be.late-FV INF-arrive-FV
'Tukiondoka sasa, tutachelewa kufika.' / 'If we depart now, we will arrive late.'

Ikoma, however, has the Indicative *-a* at FV, as shown in (72).

(72) *To-**raa**-bhook-**a*** *ibheere*, *n-to-oghu*[134]*-këëchër-a* *ko-hik-a*
1PL-COND-depart-FV now FOC-1PL-FUT-be.late-FV INF-arrive-FV
'Tukiondoka sasa, tutachelewa kufika.' / 'If we depart now, we will arrive late.'

From a historical perspective, it is still unclear why Ikoma does not have a SBJV ***-e*** in the FV position, as is seen in the other Mara languages. I already noted in Section 3.3.2 that Ikoma's Past Inceptive form SM-***Vká***-R-***a*** shows some formal similarities to the F_2 in Kabwa and Simbiti, which is realized as SM-***aka***-R-***ɛ***. At some point in the development of Ikoma's system, it appears that most of the distinctions between Subjunctive and Indicative forms were no longer needed and Ikoma began to favor the indicative final vowel. This hypothesis is based on minimal data, and future studies of Ikoma's TAM system will need to look at these forms in light of the rest of Ikoma's TAM system.

[134] This is an allomorph of the *-Vko-* formative for NPST in Ikoma. The *gh* is obtained by application of Dahl's Law.

Simbiti and Kabwa have a SBJV ***-ɛ***[135] at FV for the Conditional, but they also include a spatio-temporal demonstrative *hano*/*hanú* to introduce the clause. This can be observed in (73) for Simbiti.

(73) ***Ha-no*** *to-**raa**-tanor-**e*** *bhoono igho*,
NC16-PDEM 1PL-COND-depart-SBJV now this

n-to-raa-teghet-e *u-ku-hik-a.*
FOC-1PL-F_1-be.late-FV AUG15-INF-arrive-FV
'Tukiondoka sasa, tutachelewa kufika.' / 'If we depart now, we will arrive late.'

The spatio-temporal demonstrative *hano* 'here/when' (*hanú* in Kabwa) is also used to introduce temporal subordinate clauses. It makes sense that Simbiti and Kabwa would add a demonstrative to indicate a Conditional clause since the SM-***raa***-R-***ɛ*** form by itself is used for the F_1/FUT.INCE in these languages. The use of *hano*/*hanu* signals that the clause is subordinate, and therefore COND, rather than being read as an Immediate Future/Future Inceptive. The fact that Kabwa and Simbiti have both these uses for SM-***raa***-R-***ɛ*** has potential historical implications. There are two possibilities that readily present themselves. The first is that this form was inherited from Proto-GL (and possibly later through Proto-East Nyanza and then Proto-Mara) with both the COND and F_1 readings, and then subsequently the Future meaning dropped out of Ikizu/Zanaki/Ikoma at the Proto-South Mara phase. The second possibility is that the form was originally used only as a COND in the proto-language and was later extended to have an Immediate Future connotation in Kabwa and Simbiti after North Mara was distinct from South Mara. In regard to these two possibilities, it is important to note that Nurse and Muzale have reconstructed the *-laa-*/*-raa-* formative for Proto-GL with a Future Tense meaning (1999: 528-530). Since this is an inherited form, the second possibility can be ruled out in the Mara context. This means that Ikizu, Zanaki, and Ikoma all exhibit evidence of a shift away from the

[135] This vowel is orthographically an <e> in both languages.

-raa- formative carrying Future Tense meaning. This would have occurred at the point in time when the Future tense system in South Mara began to be reduced to its present form with only a PST/NPST distinction for tenses. This gives another clue that South Mara is a valid subgrouping.

Based on the examples of the *-raa-* formative in this section and in previous sections, the variance in usage in the Mara languages is presented in Table 6.

Table 6: The Distribution of the *-raa-* Formative in the Mara Languages

		Category		
		Conditional	F_1/FUT.INCE	Narrative
Language	Ikizu	SM-***raa***-R-***i***	--	--
	Zanaki	SM-***raa***-R-***e***	--	--
	Ikoma	SM-***raa***-R-***a***	--	--
	Kabwa	***hanú*** SM-***raa***-R-***e***	SM-***raa***-R-***e***	--
	Simbiti	***hano*** SM-***raa***-R-***e***	***(n)***-SM-***raa***-R-***e***	--
	Kuria	??[136]	***n(e)***-SM-***raá***-R-***ɛ***	--
	Gusii[137]	??[138]	--	1. ***(n)***-SM-***raa***-R-***a*** 2. SM-***raa***-R-***e***

Proceeding on from a discussion of the Conditional forms in this section, the following section explores the Hypothetical category in the Mara languages.

5.3 Hypothetical

The term Hypothetical (HYP) is not used by Nurse, but in this study it is translated by the concept of "would" in English. In other words, a Hypothetical is a verb form that indicates an irrealis event or state that is conceived in the speaker's mind, but has not been realized in

[136] Data not available from the sources at hand.
[137] The Gusii forms are obtained from Nurse (2007: 127, based on data from Whiteley 1960).
[138] Data not available from the sources at hand.

actuality and may not even be likely to occur. Generally, this first irrealis event or state influences some subsequent irrealis event or state. For my current research, I elicited the HYP through sentences translated as "If we would X, we would Y" or various negative formations of this structure.[139]

Ikoma, Ikizu and Zanaki (the South Mara languages) all have *-a(a)nga-* (surfacing as *-Vnga-* in Ikoma) as a formative for HYP in the TA slot. Additionally, these forms have the ANT formative *-iri*/*-iɾi* in the FV slot, as in Example (74) from Zanaki.

(74) *Tw-**aanga**-bhuuk-**iri** bhoono, tw-**aanga**-keezer-**i**[140] kuhika.*
1PL-HYP-depart-ANT now 1PL-HYP-be.late-FV INF-arrive-FV
'Tungeondoka sasa, tungechelewa kufika.' / 'If we would depart now, we would arrive late.'

Kabwa and Simbiti, however, have a different formative *-kaa-* at TA for HYP with ANT marked at FV. An example of this ***(n)***-SM-***kaa***-R-***irë*** form in Simbiti is exhibited in (75).

(75) *Ha-no to-**kaa**-tanööyë*[141] *bhoono igho,*
NC16-PDEM 1PL-HYP-depart.ANT now this

***n**-to-**kaa**-tëghët-**irë** u-ku-hik-a*
FOC-1PL-HYP-be.late-ANT AUG15-INF-arrive-FV
'Tungeondoka sasa, tungechelewa kufika.' / 'If we would depart now, we would arrive late.'

It is interesting to note the use of a *-kaa-* formative for the HYP form in Simbiti and Kabwa. Although I did not have access to data on the Hypothetical form in Kuria, the F_2 form in that language also has a *-kaa-* formative at TA and is realized as SM-***kaa***-R-***a***. There is a plausible

[139] The Swahili and English Hypothetical sentences are found in Appendix G, numbers 56 to 67.
[140] This is an allomorph of *-iri* that occurs after *r* final verb roots in Zanaki. The *-i* suffix has a significant enough geographical spread within Bantu to be reconstructible to PB and is probably even of earlier origin (Nurse 2008: 268, 273).
[141] The underlying root is *tanor* 'depart' and this is followed by the Anterior suffix *-irë*. In Simbiti there is an understudied process of Imbrication whereby *r*-final verb roots in conjunction with Anterior suffixes cause the *r* to elide and the final root-vowel to lengthen. The *r* in the Anterior suffix is pronounced as a *y*, and the final <ë> causes vowel raising of the lengthened root vowel.

semantic connection between a near/middle future and a Hypothetical since both of these categories have to do with events that have not yet been realized.

In the core Mara languages, North Mara realizes the Hypothetical as a variants of ***(n)***-SM-***kaa***-R-***ire*** while South Mara employs variants of the form SM-***a(a)nga***-R-***ire***. This split between the North and South languages gives further credence to separating these two as genetically distinct groups. In regards to this, one important question is whether these two formatives, *-a(a)nga-* and *-kaa-*, are inherited from an earlier phase (like Proto-GL or Proto-East Nyanza), or whether one of them is an innovation or due to contact induced language change. A related *-nga-* formative definitely has strong comparative support, especially in Eastern Bantu where there are many instances of a concessive/conditional/potential/irrealis with the formative *-nga-* (Nurse 2008: 251). This provides evidence for the antiquity of the *-nga-* formative and indicates that *-kaa-* is probably of more recent origin since it is not widely attested with an Irrealis meaning (Nurse 2008: 240-246). In fact, Nurse does not mention a long-vowel version of the formative *-ka-*; and in his description of the different meanings of the short-vowel formative *-ka-*, the closest meaning to Hypothetical is "if/when/conditional/participal/persistive" (2008: 243-244). From a comparative perspective, though, it is interesting that Ruri, a Suguti language, uses *-ka-* as a Conditional (2008: 243). These facts point to the potential for a link between North Mara and Suguti at the Proto-East Nyanza phase of development. However, since little broad-scale comparative work has been done within Great Lakes on Hypothetical (and the potentially related Conditional and Potential), it is difficult to make any definitive claims about historical relationships between different realizations of these categories.

As an additional note, in the Northwest portion of the Bantu area, a complex formative *-nga-*, probably originally composed of both *-n-* and *-ka-*, is used as a remote future (Botne 1999:

494). Although the formal similarity and semantic overlap of this remote future with the *-a(a)nga-* Hypothetical of Mara (and the related concessive/conditional/potential/irrealis *-nga-* of wider Eastern Bantu) is worthwhile to note, a historical connection cannot be assumed and more comparative work on this formative is needed in Bantu (see also Nurse 2008: 251-252).

5.4 Potential

The Potential relates to the ability or possibility for an action to occur. It is a related category to the Hypothetical, especially in Eastern Bantu, and there is sometimes formal overlap between these two categories, which both use the formative *-nga-* (Nurse 2008: 251-252). In the Mara languages, however, the Potential is expressed differently than HYP by using the verb 'be able to' in the position where an auxiliary would be found followed by the main verb, much as is done in Swahili. For example, in Kabwa the basic form (translated as "can" in English) has a morpheme-by-morpheme correspondence to its Swahili counterpart, with the exception that Kabwa includes augment vowels and Swahili does not. Example (76) displays this.

(76) *Tu-ra-tur-a* *o-ku-simb-a* *e-ri-tuko*
1PL-PROG-be.able-FV AUG15-INF-dig-FV AUG15-NC5.SG-hole
'Tu-na-wez-a Ø-ku-chimb-a Ø-Ø-shimo.'
'We can dig a hole.'

The Present Tense/Progressive Aspect in Swahili on the verb *weza* 'be able' followed by an infinitival form of the main verb is a perfect correspondence to the Non-Past tense on *tura* 'be able' followed by the infinitival main verb in Kabwa. All the other core languages translate this phrase in the same way and so the obvious influence from Swahili obscures any usefulness of this phrase for our current study.

There is another form of the Potential that translates as "could" in English. This is more helpful for tracing genetic similarities between the various branches of the Mara languages. Although the South Mara languages Ikoma, Ikizu and Zanaki again translate the Swahili in a

morpheme-by-morpheme fashion, Kabwa and Simbiti have a different form. The "could" form for Kabwa is shown in (77), but the Simbiti form has identical morpheme correspondence.

(77) ***N*-tu-*Ø*-tur-*a*** *o-ku-simb-a* *e-ri-tuko*
FOC-1PL-POT-be.able-FV AUG15-INF-dig-FV AUG5-NC5.SG-hole
'Tunaweza tukachimba shimo.' / 'We could/might dig a hole.'

A couple things that are especially interesting about this form in Kabwa are, first of all, that it includes a null morpheme at TA, which is an inherited form from GL, but has otherwise been lost in Kabwa (as mentioned in Section 3.3.4); and secondly, the appearance of a copular Focus marker, which is a remnant of Focus marking in Kabwa that has mostly been lost.

With this mention of Focus marking, it seems apropos to move on to a discussion of that topic in more detail in Section 5.5.

5.5 Focus

This section presents an introduction to how the concept of Focus relates to the Mara languages. As I mentioned above in Section 3.3.4, the formative *-ra-* was probably used for Disjunctive Focus in early GL and this usage has been retained in the West Highlands branch (Harjula 2004: 98-101; Nurse 2008: 205). This *-ra-* formative is included in Progressive forms in West Highlands as well (Harjula 2004: 107), and it seems likely that this double usage was already occurring at an early phase in GL since the ancestor of the Mara languages seized upon this Progressive usage of the *-ra-* formative at the expense of using it as a Focus marker.

There seems to have been another formative already in use at an early GL stage which employed the copular *ni* at the Pre-SM slot on the verb. The varied usage of this form in a number of modern GL languages makes it difficult to posit what its original meaning would have been. In several current West Highlands languages (Shubi, Hangaza) and in the geographically distant Luhya languages (Isuukha and Samia) the *ni-* formative is used to indicate if/when

clauses[142] (Nurse & Muzale 1999: 535). In Rutara and at least one Suguti speech variety (Regi/Kara), this same formative is used as a Progressive (1999: 534). In Mara (especially North Mara), however, this formative is used to mark Focus. This same type of Focus is also found outside of Great Lakes in Central Kenya/Thagicu (E50) and Kilimanjaro-Taita (E60) with general usage as a marker of "greater certainty on the part of the speaker about the validity of what is being said" (Nurse 2008: 207). One important thing to note about this copular Focus marking is that it does not contrast Disjunctive and Conjunctive like the *-ra-* Focus formative in West Highlands.

Having laid this foundation regarding the copular Focus, I now present several situations where it is realized on verbs[143] in the Mara languages. Both Kuria and Gusii have widespread usage of copular Focus on verbal forms, as can be seen in Table 1, Table 2, and Table 3 above. Nurse (2008: 208) presents data from Kuria that shows this distinction well. An unmarked form is [βaasɔmă] 'they have read'; this is contrasted with the Focus form [**m**baasɔ́ma] 'indeed they have read'.

Simbiti has a very similar Focus system to that of Kuria where a verbal form marked by an initial nasal represents emphasis on the verb itself. The Focus-marked forms are more frequent for MCAs than the unmarked forms. An example of a Focus-marked verb in Simbiti is presented in (78).

(78) ***N****-tw-aa-shumaash-irë* *na-we* *mo-oka*
FOC-1PL-P_3/P_4-speak-ANT CCONJ-3SG.PRON NC3.SG-year
'Tuliongea naye mwaka uliyopita.' / 'We spoke with him last year.'

[142] Interestingly, a very similar usage of this formative translating "variously as relative, conditional, and 'if/when'" is found outside Great Lakes in F21-22, F24, and F31-32 (Nurse 2008: 207, fn 47).
[143] Copular focus does also exist on some non-verbal constructions (especially cleft constructions) in at least a few of the Mara languages, but this area has not been adequately studied yet.

Emphasis can be shifted to a pre-verbal element by putting the copular nasal in front of that element. In cases like this, the verb form is not Focus-marked even though it is a Main Clause verb. This was seen in Example (7) above and is repeated as (79).

(79) ***N****-kare,* *tw-aa-shumaash-irë* *na-we.*
FOC-long.time.ago 1PL-P_3/P_4-speak-ANT CCONJ-3SG.PRON
'Tuliongea naye **zamani**.' / 'We spoke with him **a long time ago**.'

Additionally, in Simbiti, it is not possible to put Focus marking on a verb in a dependent clause. This is evident when the MCA verb ***n****twakashumaashe* 'we will speak' in Example (80) is compared with the RCA verb *bhaakateme* 'we will hit' in (81).

(80) *Mo-ghorobha,* ***n****-tw-****aka****-shumaash-****e*** *na-we.*
NC3.SG-evening FOC-1PL-F_2-speak-SBJV CCONJ-3SG.PRON
'Tutaongea naye jioni hii.' / 'We will speak with him this evening.'

(81) *Ha-no* *bha-****aka****-tem-****e*** *e-ke-tomwa*,
NC16-PDEM.REL 3PL-F_2-hit-SBJV AUG7-NC7.SG-drum

w-aka-nyoor-e *to-ra-bhin-a.*
2SG-F_2-find-SBJV 1PL-PROG-dance-FV
'Watakapopiga ngoma, tutakuwa tukicheza.' / 'When they (will) play the drum, we will be dancing. (lit. *you will find we are dancing*)'

Ikoma also has evidence of copular focus marking but it functions differently than in Simbiti. In Ikoma many MCA verb forms have a Pre-SM nasal, but it has become lexicalized on those forms and it does not indicate special emphasis on the verb like it does in Simbiti. In Example (82), a MCA Past Tense form includes a lexicalized Focus marker in Ikoma.

(82) ***N****-to-o-ghor-iri* *cham-bori* *kare.*
FOC-1PL-P_3/P_4-buy-ANT NC10.PL-goat long.time.ago
'Tulinunua mbuzi zamani.' / 'We bought goats a long time ago.'

However, similarly to what occurs in Simbiti, RCA forms like *bhaatëmiri* do not take the copular Focus marking. This can be seen in (83).

(83) *Ha-no* *bha-a-tëm-iri* *an-goma*,
NC16-PDEM.REL 3PL-P_2/P_3/P_4-hit-ANT NC9.SG-drum

n-to-o-re *to-ra-bhin-a.*
FOC-1PL-PST-PST.COP 1PL-PROG-dance-FV
'Walipopiga ngoma, tulikuwa tukicheza.' / 'When they played the drum, we were dancing.'

Although Simbiti has very active copular Focus marking and Ikoma preserves a number of lexicalized forms with the same marker, it is much less visible in the remainder of the Mara languages. Kabwa does have a few constructions, like the Potential (see Example (77) in Section 5.4 above) that show traces of a copular focus marker. Additionally, there is at least one form where it is optional without any evident change in meaning. The two variants of this form are included as Examples (84) and (85).

(84) ***N****-tu-raa-hair-e* *na-we* *bhwangu* *ha-nu.*
FOC-1PL-F_1-speak-SBJV CCONJ-3SG.PRON quickly NC16-PDEM
'Tutaongea naye hivi karibuni.' / 'We will speak with him soon.'

(85) *Tu-raa-hair-e* *na-we* *bhwangu* *ha-nu.*
1PL-F_1-speak-SBJV CCONJ-3SG.PRON quickly NC16-PDEM
'Tutaongea naye hivi karibuni.' / 'We will speak with him soon.'

Ikizu and Zanaki do not show any evidence of a copular Focus marker on verbal forms, which means they either lost it or never developed it in the first place. The case of losing the copular Focus marking makes more sense if we are to posit these languages as ancestors of GL (and genetically related to the rest of the Mara languages). In this regard it is interesting to note that Ikizu and Zanaki are also the two Mara languages that have developed the *-ra-* formative (previously a Focus marker) the furthest along the chain of development to a marker of Future/Non-Past. If the *-ra-* formative was always more prominent in the ancestors of these speech varieties, then it makes sense why another Focus marker (like the copular formative) never became prominent.

On a related topic to this, and to my discussion of Focus marking in GL at the beginning of this section, it is relevant that both Simbiti and Ikoma do not put a copular Focus marker on the SM-*ra*-R-*a* form. If both a *-ra-* formative at TA and a Pre-SM nasal were used as different types of Focus-marking strategies in early GL, which seems possible from their current distribution and usage, then they would not be expected to co-occur on the same form.

With this brief survey of Focus in Mara, I shift my attention to negation, which is another important verbal category that needs to be examined in order to better understand the entirety of the TAM systems.

5.6 Negation

Although this section on Negation and the one that follows on Relative Clauses are only an introduction to the systems at work in the Mara languages, they are important to gain a better picture of the historical situation. Up until this point, I have mainly looked at Main Clause Affirmatives (MCAs), which are generally the clause types with the richest variety of TAM formatives are found. In contrast, Main Clause Negatives (MCNs), Relative Clause Affirmatives (RCAs), and Relative Clause Negatives (RCNs), generally show less variation (Nurse & Muzale 1999: 520). In fact, it is suggested that "change tends to occur in MCAs rather than in the other structures so, if the forms of the MCAs differed from those in MCNs, RCAs, and RCNs, the latter may represent retention" (1999: 521). This also would imply that the most likely forms to show evidence of innovation are the MCAs. In this section, then, I will take a rudimentary look at MCNs to see if there are formatives in them that do not correspond with their affirmative counterparts.

The general negation pattern for the Mara languages is to use the formative ***-ta-*** in the NEG_2 slot on the verb. Sometimes the addition of this morpheme is the only distinction between

the MCA form and the MCN form. This is the case for the F_2/F_3 in Kabwa where the MCA form is SM-***aka***-R-***e*** and the MCN form is SM-***ta-aka***-R-***e***, as is evidenced in Example (86).

(86) *Tu-**ta-aka**-hik-**e*** *bhwangu*
1PL-NEG-F_2/F_3-arrive-SBJV early
'Hatutafika mapema.' / 'We will not arrive early.'

However, when the Present tense MCN is elicited, a different formative appears for Kabwa in the TA slot. Above in Section 3.3.4, we saw that Kabwa made use of the form SM-***aa***-R-***a*** to mark PRS, yet in the negative form it has SM-***ta-ku***-R-***a***, as seen in (87).

(87) *Tu-**ta-ku**-gy-**a***
1PL-NEG-PRS-go-FV
'Hatuendi.' / 'We are not going.'

Based on the proposal above, the formative *-ku-* represents a potentially older retention for PRS in Kabwa, which has now been replaced in MCAs with low-toned ***-aa-***. As I discussed in Section 4.2, the *-ku-* formative is found in all five of the Mara core languages as either PRS or with a connected aspectual meaning like PROG. The Negative Present in Kabwa links its use of the *-ku-* formative with the other core languages.

Both Ikizu and Zanaki have a distinction between a formative related to the present tense in positive and negative constructions as well. Ikizu has exactly the same forms as Kabwa with SM-***aa***-R-***a*** for the Vast Present and SM-***ta-ku***-R-***a*** for the Negative Non-Past form. Zanaki does not have SM-***aa***-R-***a***[144] indicating PRS currently, although it is quite possible that it has been replaced in the not too distant past with the SM-***ra***-R-***a*** form being extended from PROG. The SM-***ra***-R-***a*** form is used as a NPST marker now, but its negative form both in the present and future tenses is also SM-***ta-ku***-R-***a***. An example of a Negative Non-Past in Zanaki is observed in (88).

[144] My language consultant told me that SM-***aa***-R-***a*** indicates something like a Past Habitual in Zanaki.

(88) *Tu-**ta**-**ku**-hik-**a*** *bhwaangu.*
1PL-NEG-NPST-arrive-FV early
'Hatutafika mapema.' / 'We will not arrive early.'

In Ikoma, a seemingly related formative *-Vko-* is used as a NPST marker in both positive and negative constructions. The underlying negative form for NPST reference in Ikoma is SM-***ta-ako***-R-***a***. An allomorph of the NPST formative is exemplified in (89).

(89) *To-**ta-agho***[145]*-siiker-**a*** *na-we*
1PL-NEG-NPST-speak-FV CCONJ-3SG.PRON
'Hatuongei naye.' / 'We do not speak with him.'

The standard manner for marking negation in Simbiti is different than the rest of the core languages, although a formative *-ta-* in the NEG_2 slot is also found in certain negative forms giving credence to its antiquity. In Simbiti, and also in nearby Kuria, there is a clausal circumfix that is commonly used for negation (Compton & Walker 2010: 25-26; Mwita 2008: 178-180). The first morpheme of the circumfix is *te-* (with vowel harmony often affecting the height of the vowel) and it fills the Pre-SM slot. The second morpheme in the circumfix is *hë* and it is found clause finally, which means that it occurs after any post-verbal material like adverbs and explicit objects. An example of this clausal circumfix Negation can be seen in (90), which is borrowed from the Simbiti text in Appendix E, line Ssc39b.

(90) … ***te**-bh*[146]*-aa-töniibh-w-i* *na* *së-haghwë*
NEG=3PL-PST-concern-PASS-ANT CCONJ NC10.PL-fruit

se-era ***hë***.
AGR10-DDEM NEG
'…hawakuyajali yale matunda.' / '…they weren't concerned with that fruit.'

This type of clausal circumfix is extremely rare in Bantu languages and I am not aware that it occurs outside of Kuria and Simbiti (and possibly some closely related Suba speech varieties). This gives the negative clausal circumfix a higher degree of reliability as a shared

[145] The *gh* in *-agho-* surfaces through application of Dahl's law. (Higgins 2010: 23-24; Higgins 2011).
[146] This 3PL morpheme is underlyingly *bha* but three vowels cannot surface in a row in Simbiti.

innovation between these two languages and suggests that they both inherited it from a common ancestor.

In comparison to the other core Mara languages, it is also relevant to mention that the formative *-ko-/-ku-* (depending on vowel harmonization with the root) is used in Simbiti to show a Negative Present. This is distinct from the *-ra-* formative for PRS in positive sentences in Simbiti. An example of this Negative Present formative is shown in (91).

(91) ***Ti****-tu-****ku****-bhin-a* ***hë****.*
NEG-1PL-PRS-dance-FV NEG
'Hatuchezi (ngoma).' / 'We don't dance.'

Since all five of the core languages contain a formative *-ku-/-ko-/-Vko-* to indicate the Negative Present, this is potentially an inherited form. Indeed, Nurse and Muzale find a form like this throughout Great Lakes Bantu and think it is a likely reconstruction for the Progressive in Proto-GL (1999: 524). However, a word of caution is in order because the *-ku-/-ko-/-Vko-* formative is likely obtained through grammaticalization and reduction of "*-li-* + *-ku-* ('be + locative'" (1999: 524). A grammaticalization path like this seems to have occurred at different times and in different places across the Bantu area (1999: 524).

From the preceding discussion, Table 7 presents the morphological realization of the Negative Present in the core Mara languages and compares it to positive clauses with Present or Progressive verb forms.

Table 7: Comparison of Positive and Negative Presents and Progressives in core Mara

		Polarity and TA Type		
		Negative Present	Positive Present	Positive Present Progressive
Language	Ikizu	SM-***ta***-***ku***-R-***a***	SM-***aa***-R-***a***	SM-***ra***-R-***a***
	Zanaki	SM-***ta***-***ku***-R-***a***	SM-***ra***-R-***a***	SM-***ra***-R-***a***
	Ikoma	SM-***ta***-***ako***-R-***a***	SM-***Vká***-R-***a***[147]	SM-***Vko***-R-***a***
	Kabwa	SM-***ta***-***ku***-R-***a***	SM-***aa***-R-***a***	SM-***ra***-R-***a***
	Simbiti	***te***-SM-***ko***-R-***a hë***	1. SM-***ra***-R-***a*** 2. SM-***Ø***-R-***a***[148]	***n***-***ko***-R-***a*** SM-***re***

Table 5 shows clearly that the *-ku-/-ko-/-Vko-* formative is used in all five of the core Mara languages to indicate Negative Present. It is also clear that the forms for the Present Progressive in Ikoma and Simbiti make use of this same formative establishing a link between these two categories.

Before concluding this discussion of Negation in Mara, the Past Negative forms deserve further attention. I mentioned in Section 3.1.1 that Ikizu employs a form SM-***aa***-R-***irɪ*** that can be used to indicate the P_4 temporal frame, although SM-***ka***-R-***a*** is the more commonly seen form for Past tense in this language. The Negative form covering the Remote Past is SM-***ta***-***a***-R-***iri***, which is formally related to SM-***aa***-R-***iri***. This probably indicates that SM-***aa***-R-***irɪ*** is an older form in Ikizu. Example (92) displays this Negative Past form.

(92) *Tu-**ta**-**a**-gur-**iri**.*
1PL-NEG-PST-buy-ANT
'Hatukununua.' / 'We did not buy.'

Zanaki and Ikoma have variants of SM-***ta***-***a***-R-***ire*** for the Negative Past as well; and Kabwa has a similar form for Negative P_3. Simbiti again makes use of its clausal circumfix negation for past tenses, but inside the circumfix is a similar form SM-***aa***-R-***irë***.

[147] This is interpreted as a Vast Present on certain verb types; generally it is a Past Inceptive. The FUT/PRS.PROG form SM-***Vko***-R-***a*** can also be interpreted as PRS on active verbs in Ikoma.
[148] This form is not as common as the SM-***ra***-R-***a*** form for PRS.

This is a very basic look at negation in the Mara languages, but even so, it provides some interesting insights into changes that have occurred in the TAM systems of these languages. A much more thorough look at negation in these languages is necessary before broader conclusions can be drawn.

5.7 Relative clauses

Following the lines of reasoning in the previous section which highlighted some of the formatives preserved in MCNs, I now turn my attention to RCAs. Studying RCAs is another way to discover formatives that are no longer used in MCAs, but that likely preserve historical morphemes; these morphemes have been replaced in more frequently used areas of the grammar, like MCAs. This topic deserves a complementary study of RCNs, but that is beyond the scope of this thesis.

As already noted in Section 3.1, Ikizu and Zanaki have a Past Tense system that is distinctive from the other Mara languages since they make use of a SM-***ka***-R-***a*** form in the mid to distant past instead of using a variant of the ubiquitous ***(n)***-SM-***a(a)***-R-***ire*** form. Having said that though, it was also mentioned in Section 3.1.1 that Ikizu makes at least some usage of SM-***aa***-R-***iri*** (Ikizu's variant of the same ***(n)***-SM-***a(a)***-R-***ire*** form) at P_4. The semantic distinctions between the two forms for P_4 are not well understood at this point, but it is very interesting to see that SM-***aa***-R-***iri***[149] is used in Ikizu. I also showed in Section 5.6 that the Negative Past form SM-***ta***-***a***-R-***iri*** is related to the SM-***aa***-R-***iri*** form as well. It is intriguing, then, that the SM-***aa***-R-***iri*** form also shows up in relative clauses relating to the past, as can be observed in (93).

[149] This formative is [iɾɪ] phonetically.

(93) *Kira ha-nɯ b*[150]***-aa-****tɨm-****iri*** *en-goma,*
every NC16-PDEM.REL 3PL-PST-hit-ANT NC9.SG-drum

tɯ-ka-bin-a.
1PL-P_3/P_4-dance-FV
'Kila walipopiga ngoma, tulicheza.' / 'Every time they played the drum, we danced.'

Even in Zanaki, where a variant of ***(n)***-SM-***a(a)***-R-***ire*** is not evidenced in MCAs, such a form appears in RCAs relating to past events. An example of this is found in the text in Appendix F, lines Zak25a-Zak25c, which is repeated in (94) for ease of reference.

(94) *Ha-no ya*[151]***-a-****gw-****iiri,*[152] *bhasi o-mu-ukya wu-rya*
NC16- PDEM.REL 3SG-PST-fall-ANT then AUG1-NC1.SG-girl AGR1-DDEM

a-ke[153]*-enderer-a no o-ru-geendo*
3SG-NAR-continue-FV CCONJ.AUG11 AUG11-NC11.SG-journey

rw-aaye.
AGR11-3SG.POSS
'Kijana alipoanguka, binti yule akaendelea na safari yake.' / 'When the boy fell, that girl just continued on her journey.'

It is interesting to note that although SM-***a***-R-***iri***[154] is used with relatives to indicate past reference, SM-***ka***-R-***a*** can also be used for past reference with relatives.[155] Structurally, it would make sense for SM-***a***-R-***iri*** to indicate Past Anterior/Perfect since it is similar to the ANT/PRF form SM-R-***iri*** with the past tense *-a-* in the TA slot. In agreement with that claim, recent research into Zanaki grammar does mention that the SM-***a***-R-***iri*** form indicates Past Anterior (Futakamba et al. 2013: 5). In addition, based on structure, the SM-***ka***-R-***a*** form would relate to the PST/NAR *-ka-* and indicate a Past Consecutive form ("and then"). More research is needed to establish the

[150] The 3rd person plural morpheme is /ba/ underlyingly, but only two similar vowels can surface in a row in Ikizu.
[151] This *ya* allomorph of *a* appears word-initially before a long-vowel in Zanaki.
[152] The Anterior formative is written with a long-vowel as *-iiri* following any labialized root like *gw* 'fall'.
[153] This is the surface form of *-ka-* in this instance, since the vowel quality of the formative assimilates to the intial *e* of the verb root (Gray & Smith 2010b: 15).
[154] Just like in Ikizu, this formative is [iɾɪ] phonetically.
[155] See, for example, Appendix F, line Zak29a.

semantic distinctions between these two forms; however, all this leads to an expectation that SM-***a***-R-***iri*** can also be used in MCAs. In spite of semantic uncertainties, the previous two examples do give evidence that Ikizu and Zanaki also have variants of the ***(n)***-SM-***a(a)***-R-***ire*** form, which means that all five core languages have this form with Past reference. In positing a Proto-Mara language this form would have a high degree of reconstructibility. In addition, this tells us something about the shared origins of Ikizu and Zanaki since they shifted the usage of SM-***ka***-R-***a*** from a pure Narrative to a Past tense marker that generally replaced the ***(n)***-SM-***a(a)***-R-***ire*** form in MCAs. This was discussed in Section 3.2.1, and is brought up again in Section 6 in relation to the reconstruction of TA in Mara.

In relation to the previous section which discussed use of the *-ku-* formative on MCNs, all of the core languages except Kabwa have SM-***ku***-R-***a*** (or a related form) in RCAs to indicate something related to PRS/NPST/PROG. Kabwa uses its standard MCA verb forms with relative pronouns[156] to indicate RCAs, as Example (95) displays:

(95) *Kira ha-nú bha-**ra**-tem-**a** e-Ø-ng'oma,*
every NC16-REL.PRON 3PL-PROG-hit-FV AUG9-NC9.SG-drum

*tu-**ra**-bhin-**a**.*
1PL-PROG-dance-FV
'Kila wanpopiga ngoma, tunacheza.' / 'Whenever they play the drum, we are dancing.'

Both Ikizu and Zanaki use SM-***ku***-R-***a*** to show NPST in RCAs and this is a form that does not occur in MCAs in either of these languages. This usage in Zanaki is exemplified in (96).

[156] Relative pronouns are tonal variants of demonstrative pronouns: *bhanú* 'the ones who' vs. *bhanu* 'these (people)' (Kiraka & Walker 2010: 11).

(96) *Kira ha-no bha-**ku**-tem-**a** e-bhi-tomwe,*
Every NC16-PDEM.REL 3PL-NPST-hit-FV AUG8-NC8.PL-drum

*tu-**ra**-bhin-**a**.*
1PL-PROG-dance-FV
'Kila wanpopiga ngoma, tunacheza.' / 'Whenever they play the drums, we are dancing.'

However, there is an extension of this form in Zanaki to indicate aspectual meaning as well. Zanaki employs SM-***ku***-R-***a*** to mark the Progressive aspect for RCAs, while Ikizu uses SM-***ra***-R-***a*** to mark PROG, which is the same way these are marked in MCAs in this language.

The *-Vko-* formative in Ikoma appears to be related to the *-ku-* formative of Ikizu and Zanaki, which is found much more widely in GL and is even traceable to Proto-GL (Nurse & Muzale 1999: 522-524). In Ikoma, the formative appears as part of the SM-***Vko***-R-***a*** construction that is used to mark NPST in RCAs. This form can be seen in the first clause of Example (97).

(97) *Ha-no bha-**aghu***[157]*-tëm-**a** a-ngoma,*
NC16-PDEM.REL 3PL-NPST-hit-FV NC9.SG-drum

n-to-oku-bh-a to-ra-bhin-a.
FOC-1PL-NPST-be-FV 1PL-PROG-dance-FV
'Watakapopiga ngoma, tutakuwa tukicheza.' / 'When they (will) play the drum, we will be dancing.'

The usage of this form, however, is not a surprise in Ikoma since the MCA form for NPST is ***n***-SM-***Vko***-R-***a***. The only difference between the MCA form and the RCA form for NPST in Ikoma is the obligatory inclusion of Focus marking on the MCA form (as I mentioned above in Section 5.5). This contrast can be seen clearly in Example (98) when comparing the verb in the relative (first) clause with the main clause.

[157] An allomorph of *-Vko-* that is caused by vowel harmonization and the application of Dahl's Law.

(98) *Ha-no* *bha-**aghu**-tëm-**a*** *an-goma,*
NC16- PDEM.REL 3PL-NPST-hit-FV NC9.SG-drum

***n**-to-**oko**-bhin-**a**.*
FOC-1PL-NPST-dance-FV
'Wanpopiga ngoma, tunacheza.' / 'When they play the drums, we are dancing.'

Similarly, Simbiti uses focus marking for many TA forms on main clause verbs, but excludes it from relative forms. Based on this, it is clear that Focus-marking is restricted to non-relative clauses in Ikoma and Simbiti.

I have already noted that all three South Mara languages feature variants of SM-***ku***-R-***a*** in relative clauses to indicate NPST and that Kabwa does not have such a form. In the other North Mara languages, there is evidence of a related form; and it is realized in Simbiti as SM-***ko***-R-***a***, which marks PRS in RCAs. Example (99), taken from Appendix E, lines Ssc40a-Ssc40b, exhibits this.

(99) *Kë-mwë* *we* *a-ka-bh-a* *kya* *o-no* *ghwiki*
NC7-one 3SG.PRON 3SG-NAR-become-FV like NC1-PDEM.REL again

*a-**kw***[158]*-i-shaabh-a* *kö=mö-të* *haara,…*
3SG-PRS-REFL-hide-FV NC17=NC3.SG-tree NC16-DDEM
'Mara moja yeye akawa kama mtu anayejificha nyuma ya mti, …' / 'Immediately [the man] hid himself behind the tree ... (lit. *Immediately he was like someone who hides himself behind a tree*)'

In addition to its usage in Relative Clauses, the *-ko-* formative is also found in Main Clauses in Simbiti with a Progressive meaning that is connected to the Present tense. The most unique thing about this compound ***n-ko***-R-***a*** SM-***rë*** form is that auxiliary inversion has taken place. Instead of having the copular auxiliary first in the compound, it comes second. This is a rare phenomenon in Bantu and is explained in more detail in Section 4.2.

[158] The formative is *-ko-* underlyingly, but labialization takes place in front of the following vowel and it surfaces as *-kw-* (Compton &Walker 2010: 20).

As shown in Section 5.6, the *-ku-/-ko-/-Vko-* formative grammaticalized from the multimorphemic *-liku-*, which originated from the auxiliary *-li* followed by an infinitive (Nurse & Muzale 1999: 522-524). Based on this discussion of RCAs in the core languages, only Kabwa shows no evidence of a *-ku-* formative at TA in RCAs (or MCAs).[159] In Gusii, *-ko-* marks PROG (Nurse 2007: 128). I do not have data on RC forms in Kuria; otherwise a related morpheme does not seem to surface in MCAs. In summary then, a *-ku-* formative is inherited from Proto-GL *-liku-*, but this has now been precluded to RCs in most of Mara since it was replaced by a *-ra-* formative in MCAs.

A comparison of Present Tense Relative Clause verbs with Main Clause TA forms related to Present tense in the core Mara languages is included in Table 8.

Table 8: Comparison of MCA and RCA Presents and Progressives in core Mara

		Polarity and TA Type		
		RCA Present	MCA Present	MCA Present Progressive
Language	Ikizu	SM-***ku***-R-***a***	SM-***aa***-R-***a***	SM-***ra***-R-***a***
	Zanaki	SM-***ku***-R-***a***	SM-***ra***-R-***a***	SM-***ra***-R-***a***
	Ikoma	SM-***Vko***-R-***a***	SM-***Vká***-R-***a***[160]	SM-***Vko***-R-***a***
	Kabwa	SM-***ra***-R-***a***	SM-***aa***-R-***a***	SM-***ra***-R-***a***
	Simbiti	SM-***ko***-R-***a***	1. SM-***ra***-R-***a*** 2. SM-***Ø***-R-***a***[161]	***n-ko***-R-***a*** SM-***re***

Before I close this section, there is one other notable relative form that is more widespread in the Mara languages than is evident in the available data.[162] This compound form, composed of the "come" verb followed by a lexical verb, is found throughout the Ikizu text in

[159] Kabwa does use this formative in MCNs.
[160] This is interpreted as a Vast Present on certain verb types; generally it is a Past Inceptive. The FUT/PRS.PROG form SM-***Vko***-R-***a*** can also be interpreted as PRS on active verbs in Ikoma.
[161] This form is not as common as the SM-***ra***-R-***a*** form for PRS.
[162] I know personally from previous work on Kabwa discourse that this form is sometimes used in that language (i.e. Walker 2011).

Appendix B and generally appears as SM-***a-az-a ku***-R-***a***. Example (100) shows an instance of this from line Ikz6a.

(100) *A-**az**-a*[163] ***ku**-tiir-**a*** *kʉ=mʉ-ti,* *mʉ-ʉtʉ*
3SG.PST-come-FV INF-climb-FV NC17=NC3.SG-tree NC1.SG-person

a-ka-hʉt-a ***a-na*** *im-buri.*
3SG-NAR-pass.by-FV 3SG-with NC9.SG-goat
'Alipopanda kwenye mti, mtu mmoja akapita ana mbuzi.' / 'As he climbed up the tree, someone passed by with a goat.'

The construction *aaza* 'he came' followed by the infinitive is what signals that the clause should be interpreted as a temporal relative indicating "when/as/while". It appears that this construction has begun to grammaticalize and in the future could reduce phonologically to produce a new formative at TA based on the root *Vz* 'come'.

5.8 Section Summary and Conclusion

The first four subsections of Section 5 deal specifically with the realization of various modals. Section 5.1 presents examples of the Subjunctive Mood and shows that its main usage is to indicate plural imperatives. Within the Mara languages, as for Bantu in general, this mood is marked with *-e*/*-ɛ* at FV.

Section 5.2 analyzes the Conditional and shows that the formative *-raa-* is used in most of the Mara languages to indicate this category; however, there are additional elements involved in the realization of this category in several of the languages. Ikoma makes use of a different final vowel (*-a*) in the Conditional form, and both Simbiti and Kabwa include a spatio-temporal demonstrative in addition to the verb. Although the Conditional at an earlier language stage surely had a *-raa-* formative, it is difficult to reconstruct the morphological realization of the entire Conditional construction at this point.

[163] The underlying form of this word has three *a*'s in a row indicating 3SG, Past Tense, and the unspecified vowel of the verb root *-Vz* 'come' which takes its value of *a* from the preceding morpheme; however, no more than two vowels in a row can surface.

The third subsection displays the usage of the Hypothetical in the Mara languages and again shows a clear split between the North Mara and South Mara languages. The North Mara varieties of Kabwa and Simbiti make use of the form ***(n)***-SM-***kaa***-R-***ire*** for the Hypothetical while the South Mara languages use variants of SM-***a(a)nga***-R-***ire***.

In the fourth subsection, the Potential category is explained in greater detail. Although there is a lot of interference from Swahili when it comes to eliciting the Potential forms, the North Mara languages do display a unique form that translates the concept of "could" in English. This form is ***n***-SM-***Ø***-R-***a*** followed by an infinitive; it is one of the only places where focus is still marked in Kabwa.

Sections 5.5 to 5.7 survey the use of Focus marking, Negation, and Relative Clauses. Focus marking amongst the Mara languages is indicated by a copular nasal in the Pre-SM slot on the verb and is mainly seen in Ikoma and Simbiti, although there are remnants of it in Kabwa. Evidence from Negation in Main Clauses reveals some links between the formatives used in the Mara languages. This is especially useful for sub-grouping the South Mara languages by showing that each of the three languages has a historical Past/Non-Past system indicated by similar forms. Finally, Relative Clauses provide further evidence for historical similarities between the Mara languages.

These considerations from various verbal categories provide additional support for the conclusions that can be drawn from Tense and Aspect distinctions. Section 6 aims to bring together the evidence from various parts of the TAM systems of the Mara languages to better understand the historical relationships between them.

6 IMPLICATIONS FOR SUB-GROUPING AND RECONSTRUCTION

In the previous three sections I looked at Tense (Section 3), Aspect (Section 4), Mood/Modality (Sections 5.1 to 5.4), and other verbal categories related to TAM marking (Sections 5.5 to 5.7) in the Mara Bantu languages. Having laid the foundations for comparative TAM, I now turn my attention to the implications of this data for better understanding genetic relationships between the Mara languages. Although the ideal situation would be a reconstruction of the Proto-Mara TAM system, this has proven difficult for three reasons: 1) There are not a sufficient number of unusual shared innovations between all the Mara languages, even if I set Gusii aside as a marginal member; 2) The innovations do not function together in a paradigmatic fashion; 3) Many of the shared formatives that do exist cannot be unequivocally isolated to the Mara languages alone, but instead are traceable to an earlier Great Lakes phase, such as Proto-GL or a period like Proto-East Nyanza shortly following that. In addition, since the Mara languages have likely been coexisting in the same region for over a millennium (Schoenbrun 1990: 157), the influence of areal language contact and parallel grammaticalization processes on specific formatives cannot be overlooked.

When all three of these are taken together, it casts great doubt on whether the innovations can be considered “individual-identifying” (Nichols 1996: 50), which is the level of evidence I proposed as a basis for sub-grouping languages in Section 2.1. Where individual-identifying paradigm-like sets of innovations are lacking, there is not enough evidence to make claims about shared genetic heritage. This does not mean that Proto-Mara is a concept that should be abandoned, especially since there is evidence from other domains that the Mara languages do cohere as a group. It does mean, however, that the concept of Mara as a sub-group should be

approached with caution and its internal relationships may even require modification if the weight of linguistic evidence points elsewhere.

The following sections weigh and compare the TAM data in order to present a possible scenario for the historical development of the Mara languages. In Section 6.1, I give an overview of the comparative TAM data presented throughout this thesis and show that the evidence points against a reconstructible Proto-Mara system distinct from earlier phases like Proto-GL.[164] In Section 6.2, I show that there is substantial evidence from individual-identifying innovations to support the classification of both a North Mara and a South Mara sub-group of languages. Finally, in Section 6.3, I propose a list of areas that will need to be researched further in order to develop a more accurate picture of the historical situation in the Mara region.

6.1 The Evidence for Proto-Mara from TAM Marking

As detailed in Section 1.5.2, Schoenbrun (1990) gives evidence for a common genetic link between all the Mara languages based on both shared core vocabulary and on shared lexical innovations/borrowings. He lists 26 lexical items from a 200-word list as shared innovations or borrowings, although at least six of these are only found amongst either the North or South Mara languages (1990: 475-479). This suggests a strong link between the Mara languages and provides good evidence for a Proto-Mara phase of language evolution. However, comparison of these languages in other linguistic domains is difficult to interpret, giving the impression of a complex scenario of language development from Proto-Mara to the current Mara languages. For instance, Higgins (2011: 273-274) compares the current vowel systems of the seven-vowel languages and proposes a couple of scenarios on the developments that might have occurred as the modern languages evolved out of Proto-Mara. Based on the comparisons, she also questions whether "the

[164] The ideal situation would be to compare the hypothesized Proto-Mara system with the Proto-East Nyanza system, which is the proposed previous ancestor of both the Mara and Suguti languages (see Section 1.5.2). However, since Proto-East Nyanza has never been reconstructed, this is not possible.

7V languages of Mara [could] be split into two or more groups which have different historical origins" (2011: 271). She goes on to say that "there is enough variation in the region for this to be possible" (2011: 271). The TAM systems of the Mara languages also show enough variation that it is difficult to propose a historical scenario regarding their development into the modern Mara languages. Along these lines, the rest of this section overviews the variation in TAM forms that was presented in this thesis and concludes that reconstruction of a distinct Proto-Mara TAM system is not feasible.

6.1.1 Past Tense

In Section 3.1, I looked at Past Tense reference in the Mara Bantu languages. Related to this, there are three main points that need to be discussed.

First, the formatives *-á-/-aa-/-a-* in the TA slot on the verb (as part of forms with *-a* at FV), usually referring to some type of past reference, are quite common amongst the North Mara languages (Kabwa, Kuria, Simbiti, and even Gusii). However, they are less common in South Mara: in Zanaki with Past reference, in Ikizu with non-temporal reference, and not found at all in Ikoma. As I pointed out in 3.1.1, there are potential historical links between the short vowel *-a-* at TA and long vowel/high tone variants of the same formative (Nurse 2008: 108). Table 9 is included to help visualize the distribution of these potentially related formatives in the Mara languages.

Table 9: The Distribution of an *-a-/-aa-/-á-* Formative in the TA Slot with *-a* at Final Vowel

		Category			
		Vast Present (PRS)	P_4	"Untimed Past Anterior"/PST	Other Past
South Mara	Ikizu	SM-***aa***-R-***a***	--	--	--
	Zanaki	--	--	--	PST.HAB?: SM-***a(a)***-R-***a***
	Ikoma	--	--	--	--
North Mara	Kabwa	SM-***aa***-R-***a***	SM-***aa***-Ŕ-***a***[165]	--	--
	Simbiti	--	--	SM-***aa***-R-***a***	--
	Kuria	--	--	SM-***a***-R-***a***	--
	Gusii	--	--	--	P_1: SM-***á***-R-***a*** P_3: SM-***aa***-R-***a***

This distribution of formatives with a variety of meanings highlights the difficulty in reconstructing a likely form or even multiple forms for Proto-Mara. However, since five of the seven languages employ a variant of SM-***a(a)***-R-***a*** to confer some type of Past Tense meaning, this is likely inherited from an earlier phase of language evolution. Because the *-a(a)-* formative is so widespread in GL Bantu that it can be traced to an earlier GL phase as a marker of "Far" or "Middle Past" (Nurse & Muzale 1999: 533), it cannot be assigned as an innovation at a hypothetical Proto-Mara phase. It seems reasonable to posit *-a(a)-* as a marker of past tense that is inherited from Proto-GL (and subsequently from Proto-East Nyanza). Comparison of the usage of SM-***a(a)***-R-***a*** in the Suguti languages would help to determine how this form was being used at an earlier phase.

[165] The tonal realization of this form is distinct from the PRS form in Kabwa.

An additional difficulty in reconstructing a Proto-Mara form is positing whether the *-a(a)-* at TA was short or long and what tonal quality it had. This is also one of the issues that Nurse and Muzale faced in reconstructing the morpheme for Proto-GL (1999: 533). It is also possible that there could have been more than one length or tonal variant of *-a(a)-* at TA at the Proto-GL phase; a much more thorough survey of the *-a(a)-* variants throughout Great Lakes Bantu is necessary to shed light on this issue.

Along these lines, another interesting thing to note in Table 9 is that Kabwa and Gusii both have more than one variant of SM-***a(a)***-R-***a*** in their system. Since it is possible that this stems from multiple *-a(a)-* formatives in the proto-language, this topic needs to be considered in greater depth. Gusii has both of the *-a(a)-* variants referring to past tense, one of which is an Immediate Past. This type of pattern is extremely atypical for Bantu (Nurse 2008: 99-100). In fact, out of 100 "matrix" languages in Nurse's database only 5 of them "distinguished two past forms on the basis of a prosodic contrast (i.e. length and/or tone) between two different instances of *-a-* alone" (2008: 109). From this it seems extremely likely that the Immediate Past with *-á-* is an innovation in Gusii rather than an inherited form; this is also discussed in Section 3.1.3.

The multiple forms with an *-a(a)-* variant in Kabwa are different than those in Gusii since one refers to the Vast Present and one to the Past. A non-past *-a(a)-* formative is not traceable to PB, but it is widespread enough that Nurse proposes it could have been a focus marker in early Bantu (2008: 213). Interestingly, it is found as a focus marker added to past tense constructions in at least two other GL languages. In these constructions in the GL languages Ha and Haya the near past marker is *-a-*, so when the focus is included it is realized as a long vowel, *-aa-*, with varying tone (2008: 213). A grammaticalization path from a near past to a present reference is mentioned by Nurse (2008: 296). This type of *-a-* Focus, then, could be the source of the Vast

Present in Kabwa (and also Ikizu), but since it is lacking from the other five languages in this study it is difficult to reconstruct it for a hypothetical Proto-Mara. In this regard, refer to the discussion of possible developmental links between Kabwa and Ha in Figure 7 in Section 3.3.4.

The second point to discuss in relation to comparisons of Past Tense marking in Mara is the prevalence of two similar forms: ***(n)***-SM-***a(a)***-R-***ɪɾe*** (see Sections 3.1.1 and 3.1.2) and ***(n)***-SM-R-***ɪɾe*** (refer to Section 3.1.3). Their distribution amongst the Mara languages is included in Table 10.

Table 10: The Distribution of the *(n)*-SM-*a(a)*-R-*ɪɾe* and *(n)*-SM-R-*ɪɾe* Variants

		Category			
		ANT	P_1/P_2	P_3/P_4	Relative Clause
South Mara	Ikizu	SM-R-***ɪɾɪ***	SM-R-***ɪɾɪ***	P_4: SM-***aa***-R-***ɪɾɪ***	P_3/P_4: SM-***aa***-R-***ɪɾɪ***
	Zanaki	SM-R-***ɪɾɪ***	SM-R-***ɪɾɪ***	--	P_3/P_4: SM-***a***-R-***ɪɾɪ***
	Ikoma	***n***-SM-R-***ɪɾi***	***n***-SM-R-***ɪɾi***	***n***-SM-***V***-R-***ɪɾi***	$P_2/P_3/P_4$: ***n***-SM-***V***-R-***ɪɾi***
North Mara	Kabwa	SM-R-***ɪɾi***	SM-R-***ɪɾi***	P_3: SM-***aa***-R-***ɪɾi***	??[166]
	Simbiti	***(n)***-SM-R-***ɪɾe***	***(n)***-SM-R-***ɪɾe***	***(n)***-SM-***aa***-R-***ɪɾe***	P_3/P_4: ***(n)***-SM-***aa***-R-***ɪɾe***
	Kuria	***(ne)***-SM-R-***ɛɾɛ***	P_1: SM-***a***-R-***ɛɾɛ***[167] P_2: ***(ne)***-SM-R-***ɛɾɛ***	***ne***-SM-***a***-R-***ɛɾɛ***	??[168]
	Gusii	***n***-SM-R-***ɪɾe***[169]	P_2: ***(n)***-SM-***á***-R-***ɛtɛ***	P_4: ***(n)***-SM-***a***-R-***ɛ́tɛ***	??[170]

[166] This form is not evidenced in the data, but SM-***a***-R-***ɪɾi*** is expected to be the P_3 form for RCAs.
[167] There is a Simbiti form ***(n)***-SM-***a***-R-***ɪɾe*** with a short vowel in the TA slot that resembles this Kuria form, but its semantic scope is not well understood yet. This form was not elicited in the initial research, but is claimed by my language consultant to be distinct from ***(n)***-SM-***aa***-R-***ɪɾe***.
[168] No data available for Kuria RC forms.
[169] Nurse (2007: 127) puts a question mark next to this category as to whether it is Anterior.
[170] Nurse (2007: 129) comments that relatives are often distinguished tonally from their MCA counterparts.

The reconstructibility of these two forms is quite evident from the chart. All seven Mara languages have some variant of ***(n)***-SM-R-***ire*** to signify the Anterior. Additionally, all seven languages use a variant of ***(n)***-SM-***a(a)***-R-***ire*** to signify the past before today in their TA system. This is not surprising based on the antiquity of these forms in Bantu. Nurse posits a form SM-***a***-R-***i(le)*** as a Past Anterior for PB, or at least for early Bantu (2008: 279). Functioning systematically with this form is SM-R-***i(le)*** at Present Anterior in PB (2008: 279). At the Proto-GL phase, or soon afterwards (perhaps Proto-East Nyanza), the SM-R-***ile*** Anterior/Perfect meaning had already expanded to include near past reference; SM-***a***-R-***ile*** functioned systematically as a middle past in reference to it (Nurse & Muzale 1999: 526). This system was inherited into the Mara languages; and if Proto-Mara is taken as a valid phase of development, both forms can be reconstructed to that phase.

This insight for the Mara languages should be mentioned along with a couple of related points. First, the Zanaki form is only found in RCAs, which is its most common usage in Ikizu as well. As discussed in Section 5.7, the RCA form is more likely to be a conservative form when it differs from its MCA counterpart (Nurse & Muzale 1999: 520-521). Second, the Gusii formative *-ɛtɛ* at FV is distinct from the rest of Mara. As mentioned in Section 3.1.2 , there is a connection between Gusii and the Central Kenya/Thagicu (E50) languages, which is why Gusii is treated as a marginal member of Mara in this thesis. If Gusii is historically more closely related to Central Kenya/Thagicu, there is a possible scenario that presents itself (based on Nurse 2008: 165; as referenced above in Section 3.1.2). In this scenario, the Proto-Thagicu language had *-irɛ* initially meaning both Anterior and Perfective. As the Proto-Thagicu migrated to their present location they developed an allomorph *-itɛ*, which eventually took on an Anterior meaning; at the same time, the original *-irɛ* retained the Perfective meaning. As the Thagicu daughter languages split

up, Gusii remained in closer contact with the North Mara languages, which have *-ire*/*-ere* as ANT. This influenced its *-irɛ* form to shift more towards an ANT meaning while the *-itɛ* form (developing phonologically toward *-ɛtɛ*) dropped the "ongoing state" sense of meaning it previously contained and became a marker of Past tenses (2008: 165).

The third main point to discuss in relation to a comparative look at Past Tense reference in Mara relates to forms with *-ka-* at TA, which was previously covered in Sections 3.1.1 and 3.2. The *-ka-* formative (realized as *-Vka-* in Ikoma) used for the Narrative tense is ubiquitous in Mara since it occurs in all seven of the languages in this study (these forms are included in Table 1). This is reconstructible for Proto-Mara as a marker of the Narrative tense. In both Ikizu and Zanaki, this NAR has been extended to the Distant/Hesternal Past, which is a link between the systems of these two languages. Although this type of shift from Narrative to Past tense is potentially wider spread within the Bantu languages, Nurse does not find unambiguous evidence for it (2008: 304-305).

Nurse and Muzale do not find enough evidence to reconstruct *-ka-* to Proto-GL, but they note that it is at least found as a marker of "Far Past" in Rutara and South Mara (1999: 533). They make the claim that "*-ka-* is a candidate for original marker of Far Past in Great Lakes, kept in Rutara and parts of [J]E.40, replaced elsewhere" (1999: 532-533). However, since their research makes no mention of a Narrative tense, it is difficult to know how widespread this is in comparison to these few instances of "Far Past" in Great Lakes. When a broader range of data on the Narrative becomes available, this claim may need to be revised to state that *-ka-* marks the Narrative in Proto-GL and there are two instances within GL (Rutara and South Mara) where this later took on the meaning of "Far Past". It is my opinion that this latter case is more likely, but better data will be needed in Great Lakes to verify that. At this current time, then, it is unclear if

-ka- was a Narrative marker in Proto-GL or if it was grammaticalized from the Proto-GL "Far Past" marker at a post-Proto-GL stage like Proto-East Nyanza or Proto-Mara.

From the preceding discussion, if Proto-Mara is a valid proto-language, the reconstructions for Past Tense reference would be: SM-***a(a)***-R-***a*** 'General or Far Past', SM-***a(a)***-R-***ire*** 'Middle Past', SM-R-***ire*** 'Near Past/Anterior', and SM-***ka***-R-***a*** 'Narrative'.

6.1.2 Non-Past Tenses and the Progressive

For Non-Past Tenses, an initial statement should be made about the Future Tenses in the Mara languages. There is a clear split between the systems in the North Mara languages (Kabwa, Simbiti, and Kuria) and the South Mara languages (Ikoma, Zanaki, and Ikizu); since this fits nicely into the discussion in Section 6.2, further reference to the Future Tenses here is only for the purpose of establishing connections with forms related to the Present. With that in mind, there are three forms that need to be investigated for the purposes of comparison and reconstruction. The first is the null form SM-***Ø***-R-***a*** that has been reconstructed as a Vast Present both for PB (Nurse 2008: 279) and for Proto-GL (Nurse & Muzale 1999: 538). Within Mara, both Kuria and Gusii still use the null form at Vast Present. Simbiti uses it but not as its main marker of PRS. Kabwa, on the other hand, has completely replaced the null form at PRS, but some remnants of it are still found in specific constructions (i.e. its use in the Potential in Example (77) of Section 5.4). Ikoma has shifted the usage of the null form to a "Present Consecutive" (see Section 3.3.4); while in Zanaki and Ikizu there is not even a trace of the null form. The null form is reconstructible for a hypothetical Proto-Mara phase, but it has almost been lost from South Mara.

Secondly, in relation to the PRS, the SM-***ra***-R-***a*** form has replaced the null form in much of Mara and is one of the most likely candidates for an innovation distinguishing Proto-Mara

from earlier phases like Proto-GL and Proto-East Nyanza.[171] This form has already been discussed to great extent in Sections 3.3.4 and 4.2 with reference to the chain of development: focus > progressive > present > non-past (Nurse 2008: 209; 293-94). Ikizu and Zanaki have developed SM-***ra***-R-***a*** all the way to a Non-Past marker. Simbiti has developed the form to a Present marker with the somewhat similar usage in Kabwa as a Present Progressive. In Ikoma the SM-***ra***-R-***a*** form is only used as a Progressive marker in compound verbs. Neither Kuria nor Gusii has any evidence of this form. It is interesting to note in this regard that these two languages are also the only ones in this study of Mara that have retained the unrestricted usage of the null form at Vast Present. Since the SM-***ra***-R-***a*** form follows a known grammaticalization path in Bantu (Nurse 2008: 209; 293-94), the similarities in the function of SM-***ra***-R-***a*** between languages do not necessarily come from a more recent genetic link between them. Additionally, the usage of this form in one or more of the languages in the Mara contact zone could have influenced its usage in others.

The historical situation that could have brought about the present day distribution of this form is not straightforward. Treating Gusii as a marginal member of Mara, and Kuria as having long areal contact with Gusii, there is strong enough evidence from the other languages to reconstruct SM-***ra***-R-***a*** to Proto-Mara; however, the function of this form is more difficult to ascertain. It seems most likely that SM-***ra***-R-***a*** was a marker of Progressive in compounds (its current role in Ikoma) since all five core languages use it this way. Even so, its potential role in the hypothetical Proto-Mara system can only be understood in relation to another formative, which I turn to now.

[171] Precaution should be taken here since no comparison with the Suguti branch of East Nyanza has been undertaken.

The third formative related to Present is *-ku-/-ko-* in the TA slot to give a form SM-***ku***-R-***a***. The *-ku-* formative is phonologically reduced from *-liku-*, which appears to have originally been a compound form with *-li* as the auxiliary (which is still widely used in GL) and *ku-* as the infinitive form on the following main verb. This form is found beyond Mara in the Rutara and North Nyanza branches of GL usually with Progressive meaning (Nurse & Muzale 1999: 522-523). Nurse and Muzale do not find enough evidence to definitely posit it for Proto-GL, but they include it as a "possible category" at the GL phase (1999: 538). Within Mara, *-ku-/-ko-* is used in Present Progressive compounds in Gusii, Kuria and Simbiti, while it is not found at all in Kabwa. In Ikoma a variant (*-Vko-*) is used for Non-Past; and in Ikizu and Zanaki it is only found in RCAs, but with a Non-Past reference. From all this, the form is also reconstructible in Proto-Mara as a Progressive, but in distinction it would have been used only in non-compound (probably Present) Progressives.

When all these things are put together, the hypothetical Proto-Mara forms would have been: SM-***Ø***-R-***a*** 'PRS', SM-***ra***-R-***a*** 'PROG' (in non-Present compounds), and SM-***ku***-R-***a*** 'PRS.PROG'.

6.1.3 Comparing Other Categories

Moving on to aspectual categories, I have already mentioned the Anterior and Progressive aspects in the preceding Sections (6.1.1 and 6.1.2). As for the other aspectual categories, Habitual (from discussion in Section 4.3) shows a clear split between North Mara and South Mara languages; it is picked up again as a topic in Section 6.2. Persistive, on the other hand, is more complex with reflexes of both *-ki-* and *-kya-* found in the Mara languages (refer to Section 4.4). Since both of these variants are widespread in Great Lakes, they probably functioned together in some way in Proto-GL (Nurse & Muzale 1999: 526-527); neither of these is an

innovation in Mara. More focused study on the Persistive is needed before any historical-comparative conclusions can be drawn.

The Inceptive category is split into Past and Future Inceptives (see Section 4.5). The Future Inceptive in Kabwa is definitely related to the F_1 in Simbiti and since the Future Tense systems in these two languages are so similar, it is almost certain that they share a genetic connection (see Section 6.2 for further discussion of this). For the Past Inceptive, Ikoma is the only Mara language that employs this category and it makes use of the formative *-Vká-* (with tonal distinction from its NAR form). Nurse and Muzale also refer to an *-aka-* formative for "Just Past" mostly found in the North Nyanza and Luhya sub-branches of GL (1999: 533). These two formatives have a possible connection, which may explain it as an inherited form in Ikoma. In spite of this, though, it is not clear why a form of this type with past reference is not extant in any of the other Mara languages.

Of the remaining categories, the Subjunctive is the clearest one from a comparative perspective as all the languages have a form SM-***Ø***-R-***e*** or something related to this. In Gusii, this is now representative of any Future Tense reference, but in the other languages it is still used as an indicator of the Subjunctive mood and is reconstructible for Proto-Mara. This form is widespread in Bantu and is clearly inherited from PB or an even earlier phase (Nurse 2008: 261-262).

The Conditional is largely indicated by the form SM-***raa***-R-***e*** in Mara[172] with some variance at FV (see Section 5.2 for examples). Simbiti and Kabwa indicate the Conditional with the addition of a temporal relative pronoun before the verb; otherwise, the form is the same as the F_1/FUT.INCE in these languages. Nurse and Muzale find good evidence for reconstructing a *-laa-* (with *-raa-* being a normal variant) formative as a marker of "Near Future" in Proto-GL

[172] Data is lacking for a Kuria and Gusii Conditional form.

(1999: 538), so it seems an SM-***raa***-R-***e*** form is also hypothetically reconstructible for Proto-Mara as a marker of "Near Future" with extended usage as a marker of the Conditional as well.

For the Hypothetical there is once again a clear split between South Mara (SM-***a(a)nga***-R-***iri***) and North Mara (SM-***kaa***-R-***iri***[173]). The South Mara form is potentially a reanalyzed compound form from an auxiliary ending in the Proto-GL Habitual *-anga* followed by a main verb in the ANT aspect. There is no way to assign a potential form for the Proto-Mara Hypothetical since there is a clear split between the two forms.

For the Potential, there is interference from the Swahili elicitation; the forms were deemed to have an imposition from outside structures rather than being useful for historical-comparative purposes (see Section 5.4). The one unique form, the "could" Potential is an identical form in Kabwa and Simbiti (North Mara), adding credibility to the fact that these languages are closely related.

The marking of Focus in Mara is another area that shows cohesion amongst the languages in the North Mara group. Gusii, Kuria and Simbiti all have heavy use of a copular auxiliary to mark Focus. Kabwa does not have a productive Focus system currently, but there are a few frozen forms that resemble similar forms in Simbiti and so it seems that it was more productive in the past. Ikoma is the only South Mara language to show signs of verbal copular Focus marking, but it seems to behave differently than in Simbiti as it is required on certain verbal forms rather than optional.

When all the comparative data is lumped together, there are many morphemes that can be reconstructed for a putative Proto-Mara TAM system. However, when these morphemes are

[173] This shares some formal similarity to what Cammenga calls the "Hodiernal Past Progressive Anterior Focused" (2004: 289) in Kuria, which is realized as ***ne***-SM-***aka***-R-***ere***.

examined, only one of them, the *-ra-* Progressive (as part of compound Progressives) appears to be an innovation that could have arisen at the Proto-Mara phase of development. However, even the usage of *-ra-* in compound Progressives is found outside of the Mara languages in the West Highlands language Ha, for example (Harjula 2004: 107). This does not rule out a Proto-Mara phase, especially because the combination of morphemes proposed for the hypothetical Proto-Mara in Sections 6.1.1 through 6.1.3 forms a system that may indeed be distinct from other branches of Great Lakes. There are also gaps in the reconstructions for Proto-GL and this makes it difficult to know whether the proposed Proto-Mara forms go back all the way to Great Lakes or if they were innovated at a more recent phase of development, like Proto-East Nyanza. All this to say, though, a core of unusual TAM innovations has not been established as a verification of Proto-Mara. Specifically, there is no paradigmatic set of innovations that could be used as individual-identifying evidence for Proto-Mara. If there was a unique Proto-Mara TAM system it seems that it occurred at an early enough phase after Proto-Great Lakes split up that it did not differ significantly from the Proto-GL system. The split into North and South Mara must have occurred not too long after the hypothetical Proto-Mara phase; this scenario is discussed in Section 6.2.

6.2 Sub-grouping the Mara Languages

Schoenbrun (1990) identifies both a North Mara and a South Mara sub-branch of the Mara languages. Based on the evidence from TAM systems in this thesis, this North/South split seems to be much more strongly attested than their hypothetical Proto-Mara predecessor. One of the things that is lacking in the evidence for Proto-Mara is an "individual-identifying" paradigmatic set of innovations. Since a whole interlocking system, when shared by more than one language, is strong evidence for a common genetic heritage, this is the type of evidence that should be found before verifying that the languages in a sub-group are genetically related.

In North Mara, especially between Kabwa and Simbiti, there is a close relationship in the Tense systems providing good evidence for a shared genetic heritage. This includes nearly identical forms used to express three different degrees of Future reference in the two languages. The Past tense systems are very similar as well with the main distinction being the slightly different usage of the SM-***aa***-R-***a*** form as discussed above in Sections 3.1.1 and 6.1.1. In relation to this, Kuria has very similar F_1 and F_3/F_4 forms to Simbiti and Kabwa, while its F_2 form is distinct. In the Past tense system, Kuria has an innovative form in the Immediate Past and its Hodiernal Past is a Focus-marked version of the Anterior, but the rest of its system resembles Simbiti quite well. With all these systematic similarities there is confidence in positing what a Proto-North Mara Tense system looked like. A comparison of the Tense systems in Kabwa, Simbiti and Kuria, with a proposal of the Proto-North Mara Tense system is included in Table 11.

Table 11: Comparison of North Mara Tense Systems

		Tense System	
		Related to the Past	Related to the Future
	Kabwa	ANT/P_1/P_2: SM-R-***iɾi*** P_3: SM-***aa***-R-***iɾi*** P_4: SM-***aa***-Ŕ-***a*** NAR: SM-***ka***-R-***a***	FUT.INCE/F_1: ***(n)***-SM-***ɾaa***-R-***ɛ*** F_2/F_3: SM-***aka***-R-***ɛ*** F_4: SM-***ɾi***-R-***a***
	Simbiti	ANT/P_1/P_2: ***(n)***-SM-R-***iɾe*** P_3/P_4: ***(n)***-SM-***aa***-R-***iɾe***[174] "Untimed Past": SM-***aa***-R-***a*** NAR: SM-***ka***-R-***a***	F_1: ***(n)***-SM-***ɾaa***-R-***ɛ*** F_2: SM-***aka***-R-***ɛ*** F_3/F_4: ***n***-SM-***ɾi***-R-***a***
	Kuria[175]	ANT: SM-R-***eɾe*** P_1: SM-***a***-R-***eɾe*** P_2: ***ne***-SM-R-***eɾe*** P_3/P_4: ***ne***-SM-***a***-R-***eɾe*** "Untimed Past": SM-***a***-R-***a*** NAR: SM-***ká***-R-***a***	F_1: ***n(e)***-SM-***ɾaá***-R-***ɛ*** F_2: SM-***kaa***-R-***a*** F_3/F_4: SM-***ɾee***-R-***a***
	Proto-North Mara[176]	ANT/P_1/P_2: ***(n)***-SM-R-***iɾe*** P_3: ***(n)***-SM-***aa***-R-***iɾe*** P_4: SM-***aa***-Ŕ-***a*** NAR: SM-***ka***-R-***a***	F_1: ***(n)***-SM-***ɾaa***-R-***ɛ*** F_2: SM-***aka***-R-***ɛ*** F_3/F_4: SM-***ɾi***-R-***a***

In addition to the Tense system, there are many other formatives that Kabwa and Simbiti share (as detailed in 6.1.2 and 6.1.3). These include: use of the *-ra-* formative to indicate PRS/PROG, similar constructions for Subjunctive and Conditional, a Hypothetical form of the shape SM-***kaa***-R-***ire***, and the same construction for the "could" Potential.

Gusii does share some unusual features with other North Mara languages; in particular, the auxiliary inversion that was mentioned in Section 4.2. Although this type of innovation is very rare within Bantu, by itself it is not enough to secure Gusii as a North Mara language. Along these lines, it is interesting that Kabwa does not share in this auxiliary inversion.

[174] There is also a similar form ***(n)***-SM-***a***-R-***iɾe*** with a short *a* before the root, but this was not elicited through the initial sentence translation and needs further investigation.
[175] The source for the [kuj] non-tonal data is Cammenga (2004), with underlying tones obtained from Mwita's (2008) analysis of Cammenga.
[176] This is not meant to be a definitive reconstruction since there are many other North Mara language varieties that are not included in this.

Nurse (1999: 28) proposes a scenario for the relationship of Central Kenya/Thagicu to Gusii and possibly several of the other North Mara languages. Since Gusii shares the strongest link with Central Kenya/Thagicu, and the rest of North Mara does appear to be related to Great Lakes, it seems that there was a link between the two over a considerable time period in the past. They may have existed in somewhat of a dialect chain with Gusii as the link taking on aspects of both North Mara and Central Kenya/Thagicu.

The South Mara languages also have systematic similarities that provide good evidence for a paradigmatic individual-identifying set of innovations. The most notable of these is the reduction of the systems in Ikizu, Zanaki and Ikoma to a more basic Past/Non-Past system. This is not the system inherited from Great Lakes, which surely had more than one Past and Future Tense (Nurse & Muzale 1999: 538). The actual formatives used to express the Past and Non-Past Tenses in Zanaki and Ikizu are the same. The Past is expressed by SM-***ka***-R-***a*** and the Non-Past by SM-***ra***-R-***a***. In Ikoma the Past is expressed by ***n***-SM-***V***-R-***iri*** and the Non-Past by ***n***-SM-***Vko***-R-***a***. The system in Ikizu/Zanaki appears to be the more innovative one and forms related to the Ikoma PST/NPST appear in RCAs in Ikizu/Zanaki. These RCA forms in Ikizu and Zanaki are SM-***a(a)***-R-***iri*** for Past tense and SM-***ku***-R-***a*** for Non-Past. This points to a reconstructed Proto-South Mara PST/NPST distinction with the forms SM-***a(a)***-R-***iri*** **and** SM-***ku***-R-***a***, which were later replaced by new MCA forms in Ikizu/Zanaki. In Ikoma, phonological shifts in the language caused *a*'s in the TA slot to be reinterpreted as unspecified vowels.

Although this is already good evidence for a Proto-South Mara Tense system, there are other innovative forms that link the South Mara languages together. The most prominent of these is the Habitual form SM-***haa***-R-***a***, which is not found outside of South Mara. Additionally, the

Hypothetical forms are common within South Mara (as opposed to North Mara) with the form SM-***a(a)nga***-R-***iri***.

From this discussion, the North Mara and South Mara groups are strongly established by systematic similarities. Especially in the case of South Mara, the system is quite innovative from the Proto-GL system and the fact that there are similar formatives shared amongst the three languages is good evidence for a shared inheritance between them. Lexicostatistical data supports the split as well. Amongst the three South Mara languages in this study (Ikoma, Zanaki and Ikizu) there is 74-87% lexical similarity, while the averages of these three with the North Mara varieties (Kabwa and two dialects of Kuria[177]) are 60-68% (Hill et al. 2007: 44). Amongst the North Mara varieties there is 72-94% similarity (2007: 44). Even with this evidence, comparative work done on phonological innnovations (Hill et al. 2007: 44-46) suggests a more complicated scenario where Ikoma and Kuria share several innovations. Additionally, a brief survey of vowel systems in the Mara languages shows that the Ikizu vowel system is closer to the systems of Kuria and Simbit than it is to the systems in Ikoma and Zanaki (Higgins 2011: 274).

In spite of some challenges to historical-comparative work amongst the Mara languages, the TAM evidence establishes the North Mara and South Mara groups, with Gusii as a marginal North Mara member linked with the Central Kenya/Thagicu languages. The evidence for a Proto-Mara TAM system is more scant, on the other hand, and does not appear to differ significantly from what can be ascertained about the Proto-GL system. In addition, Simbiti/Kuria (North Mara) and Ikoma (South Mara) share some important features that connect them across the sub-grouping lines. Most notably this includes copular Focus marking.

[177] Simbiti was not included in the lexicostatistical study done by Hill et al. (2007).

I propose the following scenario for the development of the languages in the Mara region. As Proto-GL split up and diversified, Proto-East Nyanza was one of its offshoots and it formed into two branches: Proto-Suguti and Proto-Mara. Further work is needed to compare both of these branches before more can be said about this phase of development. Proto-Mara eventually developed into a dialect chain with the Proto-North Mara varieties in closer contact with pre-Gusii and the Proto-Central Kenya/Thagicu dialects.[178] The Proto-South Mara varieties were in closer contact with other Great Lakes varieties and possibly even Takama (Bantu zone F). As Proto-North and Proto-South became distinctive speech varieties they had already formed internal dialect differentiation. The pre-Ikoma dialect in the Proto-South Mara area and the pre-Simbiti/Kuria dialect in the Proto-North Mara area were in contact. Some of their features spread to each other causing them to look more similar, but both pre-Ikoma and pre-Simbiti retained systematic similarities in their TAM systems with the languages to which they are more closely genetically related.

6.3 Directions for Future Research

This research is only a first step and much more needs to be done to understand the TAM systems of the North and South Mara languages, as well as the historical processes at work in the Mara region leading to the present language situation. One type of research that is foundational for future historical-comparative work is detailed descriptions of the TAM systems of individual languages. With that foundation, the reconstructed TAM systems of both Proto-North Mara and Proto-South Mara could be the subject of their own studies with additional speech varieties

[178] I would also suggest that Proto-North Mara was in closer contact with Proto-Suguti, since the few areas where I have data on Suguti suggest some TAM similarities with North Mara. This claim awaits further research on both branches of East Nyanza (Mara and Suguti) to be verified.

included to expand the base for comparison.[179] This should help establish more solid reconstructions for the TAM systems in each of the proto-languages.

In addition, it will be important to understand how both North and South Mara relate to their nearby neighbors in Suguti (with whom they are grouped into the East Nyanza branch of GL). Along these lines, it will be necessary to have a specific study focused on the four main varieties of Suguti and to develop a reconstruction of the proto-TAM system in comparison with what is known about the systems in North and South Mara.

One of the interpretive issues in this thesis, and also in broader comparative studies of TA is the issue of the various phonological shapes of the *-a(a)-* formative (Nurse 2008: 108; Nurse & Muzale 1999: 531-533). The relationships between these forms, starting with closely related dialects, should be examined. Where there are differences between related varieties, they should be noted and an explanation of how the change was motivated should be undertaken. Once the changes at the micro-level are understood, it will give a better picture of how this might affect a larger scale effort in attempts to reconstruct the variants of *-a(a)-* within Great Lakes as a whole.

The study of TAM systems is not the only linguistic topic that needs attention within East Nyanza and Mara. There are various other linguistic topics, some directly related to my research, that will be important to study; these include how changes in vowel systems occur, how Focus-marking develops (and how Focus functions differently in related languages), and fuller descriptions of Negation and Relative Clauses.

An area that can further improve historical-comparative linguistic work in the Mara languages is an inclusion of insight from non-linguistic historical sources. The internal oral history of a tribe is one of the sources that should be considered. There are reasons to be cautious

[179] Other North Mara varieties that could be included are: Surwa, Sweta, Hacha, Kiroba, and Kine. Additionally, Kuria and Gusii's dialect differentiation need more study. Other South Mara varieties that will need to be studied are: Sizaki, Nata, Isenye, and Ngoreme.

about taking oral history as a literal description of past events, including the situation where a single clan dominates the sociopolitical/socioeconomic elements of a tribe and therefore develops the history according to their agenda (Mould 1976: 34-35). However, the ways that a people group chooses to talk about their history can give clues to contact they had with other people in the past and provide missing elements to reconstructing history. Additionally, there are written historical sources documenting events that have impacted the Mara region. Bender-Shetler mentions that tribal identities and histories have become more important in the Mara region since the colonial governments raised the status of the "tribe" in relation to the "descent group, region, rank, or age-set group" (2003: 3). A better grasp of these historical sources will need to be included in future historical-comparative linguistic work as well.

7 CONCLUSION

In this thesis, I provided evidence from the TAM systems of the Mara languages for sub-grouping North Mara and South Mara underneath the East Nyanza branch of Great Lakes Bantu. Past studies (Schoenbrun 1990) concluded that North Mara and South Mara should be immediate sub-branches of a unified Mara group; however, my present research did not find clear "individual-identifying" paradigmatic evidence (Nichols 1996: 50) that would indicate the Mara languages are a well-defined subgrouping within the East Nyanza branch of Great Lakes Bantu.

Though the primary purpose of this research was to search for TAM evidence to sub-group the Mara languages, a secondary goal was to provide a preliminary description of the TAM systems in five lesser described Mara languages: Kabwa, Ikizu, Ikoma, Simbiti and Zanaki. The data on these languages was gathered through sentence elicitation and the recording and transcription of an oral text. Appendices A through F, in conjunction with the example

sentences throughout the body of this paper, display the language data, which provides a foundation for future descriptive work amongst these and other Mara languages.

The first two sections of the thesis established the background for the research by providing information on Bantu historical-comparative linguistics, TAM studies, and previous classifications of the Mara languages. The TAM data from the Mara languages was presented in Sections 3 to 5 of this thesis, displaying data on Tense, Aspect, Mood/Modality, and other verbal categories (e.g. Negation and Relative Clauses). Section 6 then focused on an analysis of all the data and its implications for sub-grouping.

Based on the data, the only arguable innovation shared by all five core Mara languages in this study is the use of the formative *-ra-* with Present tense and/or Progressive aspectual meaning. Even the usefulness of this formative for sub-grouping is doubtful because its development in Mara follows a known grammaticalization path (Nurse 2008: 209; 293-94); it is therefore hard to know if the similarity between this morpheme in the core Mara languages is the result of genetic inheritance or other factors like parallel grammaticalization and areal diffusion. Regardless, one shared feature is not sufficient to establish a Proto-Mara TAM system since it leaves the reconstructed system looking much like what is proposed for earlier phases of language evolution like Proto-GL.

Despite the lack of individual-identifying evidence, the majority of Mara languages can trace elements of their TAM systems to an earlier language phase. These include the *-a(a)-* formative (or a similar variant) in the TA slot used to represent a Far Past or General Past, the extension of the Anterior SM-R-***ire*** to a Near Past, the usage of (***n***)-SM-***a(a)***-R-***ire*** with a Middle Past/Far Past reference and the formative *-ka-* indicating the Narrative Tense. However, each of

these can be traced as inherited features from Proto-Bantu or Proto-Great Lakes and are thus unhelpful for sub-grouping purposes.

Mould (1976) presents the historical picture that Gusii's proto-language (and supposedly the rest of Mara) originated on the west side of Lake Victoria, then moved around the south of the lake and ended up on the northeast side. Mould gives evidence from lexicostatistics, phonological comparisons, and morphosyntactic evidence to show that Gusii is more closely genetically affiliated to the West Highlands languages like Rwanda, Rundi and Ha than it is to its current neighbors in the Luhya (Luyia) branch of GL, which supposedly migrated around the north side of Lake Victoria. Schoenbrun (1990) proposes a similar migration for a Proto-East Nyanza group along the southern shores of Lake Victoria to their current location on the east side of the Lake.

There is one particular individual-identifying innovation that connects Gusii to the North Mara languages, Kuria and Simbiti, and that is the auxiliary inversion of infinitival compound constructions with copular *-ri/-li* auxiliaries (see Section 4.2). This construction is Focus-marked and indicates a Present Progressive in these three languages. This on its own is not strong enough evidence for sub-grouping, but it does suggest a link between Gusii and the North Mara languages that needs further consideration. It seems plausible that Gusii and the rest of the Mara languages existed as a dialect chain even as they started moving from the west side of the lake. This chain would have the predecessors of South Mara on one end, followed by a contact zone and then North Mara which is connected by Gusii to the E50 languages of Central Kenya. It is unknown how the Suguti languages (Jita, Kwaya, Ruri, etc.) would fit into this migration scenario since they are also proposed to be related to the Mara languages through a common East Nyanza ancestor (mother node above Mara on the tree).

What is clearer is that there are three main groupings within the Mara languages: Gusii, North Mara (Kuria, Simbiti, Kabwa), and South Mara (Ikizu, Zanaki and Ikoma). Kabwa is more closely aligned with North Mara, in that it shares a very similar Tense system with Simbiti (and, to some extent, Kuria), but it also did not go through the innovative auxiliary inversion, which is puzzling. South Mara is set apart by several innovations: the use of an unusual -*haa*- formative for the Habitual and the reduction from a more complex temporal system to a simpler PST/NPST distinction.

This thesis can be seen as the first step towards the reconstruction of the Proto-North Mara and Proto-South Mara TAM systems. As with any comparative work, more speech varieties will need to be included in future studies in order to get a fuller picture of the situation and work out some of the medial stages that led to the modern varieties. It will also be important to understand better how speech varieties borrow TAM morphemes from each other and to what degree this impacts the overall system within a specific variety. Pathways of change within Bantu languages are much better understood than they used to be, but there is still work to be done amongst the Mara languages to understand the processes that have shaped the modern speech varieties.

Appendix A: Individual Language Charts

Based on the research for this thesis, I have condensed the TA data into a simple chart for each language (displayed as Table 12 through Table 16) to show some of the systematic interaction between the TA forms. This is in line with points 1 and 3 of Nurse's conceptual framework (2008: 10-14) that is presented as Figure 1 in Section 1.4.1; and the format of these TA charts also follows the approach set out by Nurse in the appendices to his TA research (2007).

All of the TA forms in the charts are exemplified on the verbs 'buy' or 'dance', which are similar in all five of the core languages. Additionally, following each chart are a few short lists to capture more complex TAM forms, some of the Relative forms, and also any auxiliary verbs found in the data. All the Mara language data in these charts is transcribed orthographically with footnotes highlighting orthographical conventions that may obscure the underlying morphemes.

Table 12: Kabwa TA Chart

Tense	**Basic**	**Habitual/ Imperfective**[180]	**Progressive**	**Anterior/Perfect**
P_4 (before yesterday)	SM-***aa***[181]-R-***a*** *twagura*[182] 'we bought' Negative: SM-***ta-a***-R-***a*** *tutaagura* 'we did not buy'	SM-***aa***-R-***anga*** *twabhinanga* 'we (usually) danced'; *twaguranga* 'we (usually) bought'	SM-***aa-ri-nga*** SM-***ra***-R-***a***[183] *twaringa turagura* 'we were buying'	SM-***aa-ri-nga*** SM-R-***iri*** *twaringa tubhiniri* 'we had danced'; *twaringa tuguriri* 'we haad bought'
P_3 (yesterday)	SM-***aa***-R-***iri*** *twaguriri* 'we bought' Negative: SM-***ta-a***-R-***iri*** *tutaaguriri* 'we did not buy'	***Same as P_4 Habitual/ Imperfective***	***Same as P_4 Progressive***	***Same as P_4 Anterior/Perfect***

[180] The Habitual/Imperfective forms overlap in meaning with the Progressive forms.
[181] There is an underlying long-vowel, *-aa-*, in this and other Past forms in Kabwa (as well as the Vast Present). This is obscured orthographically following the SM *tw-* '1PL' after which a short-vowel <a> is found in writing.
[182] A high-tone surfaces on the vowel of the verb root in the Basic P_4 form.
[183] The form SM-***aa-bh-anga*** SM-***ra***-R-***a*** realized as *twabhanga turagura* 'we were buying' is also possible, but the distinction between this and the SM-***aa-ri-nga*** SM-***ra***-R-***a*** form is not well understood.

Tense	Basic	Habitual/ Imperfective	Progressive	Anterior/Perfect
P_1/P_2 (Anterior/ Perfect)	SM-R-***iri*** *tuguriri* 'we bought (today)' / 'we have bought' Negative: SM-***ta***-R-***iri*** *tutaguriri* 'we did not buy (today)'	***Not Possible for P_1/P_2:*** * *twaguranga* * *tuguranga*	Temporal: SM-***aa-ri*** SM-***ra***-R-***a*** *twari turagura* 'we were buying' Anterior (not temporal): SM-***bh-eeri*** SM-***ra***-R-***a*** *tubheeri turabhina*[184] 'we have been buying'	Temporal: SM-***aa-ri*** SM-R-***iri*** *twari tuguriri* 'we had bought'
PRS	SM-***aa***-R-***a*** *twagura*[185] 'we buy' Negative: SM-***ta-ku***-R-***a*** *tutakugura* 'we do not buy'	SM-***ra***-R-***anga*** *turaguranga* 'we (usually) buy'	SM-***ra***-R-***a*** *turagura* 'we are buying'	SM-R-***iri*** *tuguriri* 'we have bought'
F_1 (very soon)	***(n)***-SM-***raa***-R-***e*** *(n)turaagure*[186] 'we are about to buy' Negative: SM-***taa***-R-***e*** *tutaagure* 'we are not about to buy'	SM-***raa***-R-***enga*** *turaagurenga* 'we will be buying'	SM-***raa-bh-enga*** SM-***ra***-R-***a*** *turaabhenga turabhina* 'we will be dancing' *turaabhenga turagura* 'we will be buying'	SM-***raa-bh-e*** SM-R-***iri*** *turaabhe tuguriri* 'we will have bought'
F_2/ F_3 (today or tomorrow)	SM-***aka***-R-***e*** *twakagure* 'we will buy' Negative: SM-***ta-aka***-R-***e*** *tutaakagure* 'we will not buy'	SM-***aka***-R-***enga*** *twakagurenga* 'we will be buying'	SM-***aka-bh-enga*** SM-***ra***-R-***a*** *twakabhenga turabhina* 'we will be dancing'; *twakabhenga turagura* 'we will be buying'	SM-***aka-bh-e*** SM-R-***iri*** *twakabhe tuguriri* 'we will have bought'
F_4 (after tomorrow)	SM-***ri***-R-***a*** *turigura* 'we will buy' Negative: SM-***ta-ri***-R-***a*** *tutarigura* 'we will not buy'	SM-***ri***-R-***anga*** *turiguranga* 'we will be buying'	SM-***ri-bh-anga*** SM-***ra***-R-***a*** *turibhanga turagura* 'we will be buying'	SM-***ri-bh-a*** SM-R-***iri*** *turibha tuguriri* 'we will have bought'

Other TAM forms
Narrative/Consecutive: SM-***ka***-R-***a***
Persistive: SM-***kya***-R-***a***
Singular Imperative: ***Ø***-R-***a***
Subjunctive/Plural Imperative: SM-***Ø***-R-***e***
Conditional: SM-***raa***-R-***e***
Hypothetical: SM-***kaa***-R-***iri***
Negative Hypothetical: SM-***ta-kaa***-R-***iri***

[184] The form *twabha tubhiniri* also translates as 'we have been buying', but it is associated with a more permanent state of affairs. The form *tubheeri turabhina* is associated with a shorter time frame.
[185] This is distinct from the P_4 form since it is completely low-toned (i.e. no high-tone on the root like the P_4).
[186] When the copular nasal is included on this form, there is more doubt associated with the proposition.

Potential: ***n***-SM-***tur-a o-ku***-R-***a***

Relative/Subordinate Clause Forms[187]

Relating to Tense:
P_4: SM-***aa***-Ŕ-***a*** (same form as in MCAs)
PST COP: SM-***aa-ri-nga*** (same form as in MCAs)
F_2: SM-***aka***-R-***e*** (same form as in MCAs)

Relating to Aspect and Combined TA:
P_3/P_4 Progressive: SM-***aa-ri-nga*** SM-***ra***-R-***a*** (same form as in MCAs)
Present Progressive: SM-***ra***-R-***a*** (same form as in MCAs)
P_4 Anterior/Perfect: SM-***aa-ri-nga*** SM-R-***iri*** (same form as in MCAs)
ANT: SM-R-***iri*** (same form as in MCAs)

Relating to Non-TA categories:
Conditional Clause: SM-***raa***-R-***e*** (same form as F_1 in MCAs)
Negative Conditional: SM-***taa***-R-***e*** (same form as NEG F_1 in MCNs)

Auxiliary Verbs
-bha: Future/Anterior copula (from verb 'be/become')
-ri: Past/Progressive/Persistive/Hypothetical copula
-nyoora: Used in Habitual Progressive (from verb 'find')
-mara: Expressing the concept 'already' (from verb 'finish')
-tiga: Used to express negation in Conditionals and Hypotheticals (from verb 'quit')
-Vja: Used to express a Future Persistive (from verb 'come')
-enda: Used to express Future Inceptive 'about to'[188] (from verb 'want')
-tura: Used to express Potential (from verb 'be able to')

[187] This is not an exhaustive list of relative/subordinate clause forms, but merely lists the forms that are witnessed in the available data.
[188] However, this appears to be a direct translation from the Swahili rather than a natural Kabwa expression.

Table 13: Simbiti TA Chart

Tense	**Basic**	**Habitual/ Imperfective**[189]	**Progressive**	**Anterior**
P_4 (before yesterday)	***(n)***-SM-***aa***-R-***irë*** *ntwaabhinirë* 'we danced' Negative: *ti-twaabhinirë hë* 'we did not buy'	***(n)***-SM-***a***-R-***anga*** *ntwabhinanga* 'we (usually) danced'	***(n)***-SM-***aa-re-nga ko***-R-***a***[190] *ntwaarenga kubhina* 'we were dancing'	***(n)***-SM-***aa-re-nga*** SM-R-***irë*** *ntwaarenga tubhinirë* 'we had danced'
P_3 (yesterday)	***Same as P_4 Basic***	***Same as P_4 Habitual/ Imperfective***	***(n)***-SM-***aa-rë ko***-R-***a***[191] *ntwaarë kubhina* 'we were dancing'	***(n)***-SM-***aa-rë*** SM-R-***irë*** *ntwaarë tubhinirë* 'we had danced'
P_1/P_2 (Anterior/ Perfect)	***(n)***-SM-R-***irë***[192] *ntubhinirë* 'we danced' / 'we have danced'; *ntöghöörë* 'we bought' / 'we have bought' Negative: *ti-tubhinirë hë* 'we did not buy'	***No Distinct Form Found***	***(n)***-SM-***á-rë ko***-R-***a*** *ntwarë kubhina* 'we were dancing'	***(n)***-SM-***á-rë*** SM-R-***irë*** *ntwarë tubhinirë* 'we had danced'
PRS	SM-***ra***-R-***a*** *torabhina* 'we dance' Negative: ***te***-SM-***ko***-R-***a hë*** *ti-tukubhina hë* 'we do not dance'	Habitual Progressive: ***n-ko***-R-***anga*** SM-***rë*** *nkubhinanga törë* 'we are (usually) dancing' Habitual Present: SM-***ra***-R-***anga*** *torabhinanga* 'we (usually) dance'	***n-ko***-R-***a*** SM-***rë*** *nkubhina törë* 'we are dancing' Negative: SM-***taa-ko***-R-***a*** *totaakubhina* 'we are not dancing'	***(n)***-SM-R-***irë*** *tubhinirë* 'we have danced' *töghöörë* 'we have bought'
F_1 (very soon)	***(n)***-SM-***raa***-R-***e*** *ntoraabhine* 'we will dance' Negative: ***te***-SM-***V***-R-***e hë*** *ti-tuubhine hë* 'we will not dance'	(***n***)-SM-***raa***-R-***enga***[193] *ntoraabhinenga* 'we will be dancing'	***(n)***-SM-***raa-bh-enga*** SM-***ra***-R-***a*** *ntoraabhenga torabhina*[194] 'we will be dancing'	***n-o-raa-nyoor-e*** SM-R-***irë*** *noraanyoore tubhinirë* 'we will have danced (lit. you will find we have danced)'

[189] The Habitual/Imperfective forms overlap in usage with the Progressive forms.
[190] This form shows an act beginning and ending in the distant past.
[191] This form shows an act beginning either 'yesterday' (P_3) or the 'day before yesterday' (overlapping into P_4), but it also can show relevance continuing to the present moment.
[192] According to my language consultant there is another form that is used for P_2, although the initial sentence elicitation did not yield it. This P_2 is realized as ***(n)***-SM-***a***-R-***irë*** and is distinguished from the P_3/P_4 Basic form by both shorter vowel length on the pre-root TA morpheme and tonal pattern.
[193] This form overlaps with the F1 Progressive in meaning. A related form ***n***-SM-***Ø***-R-***anga*** (*ntubhinanga* 'we are/will be dancing') appears to relate the action itself to the imperfective without direct temporal reference.
[194] Swahili influence has increased the usage of this form in relation to the F1 Habitual/Imperfective.

Tense	Basic	Habitual/ Imperfective	Progressive	Anterior
F_2 (later today)	***(n)***-SM-***aka***-R-***e*** *ntwakabhine* 'we will dance' Negative: ***te***-SM-R-***a hë*** *ti-twakabhine hë* 'we will not dance'	(***n***)-SM-***aka***-R-***enga***[195] *ntwakabhinenga* 'we will be dancing'	***(n)***-SM-***aka-bh-enga*** SM-***ra***-R-***a***[196] *ntwakabhenga turabhina* 'we will be dancing'	***w-aka-nyoor-e*** SM-R-***irë*** *wakanyoore tubhinirë* 'we will have danced (lit. you will find we have danced)'
F_3/F_4 (tomorrow or after tomorrow)	***(n)***-SM-***ri***-R-***a*** *nturibhina* 'we will dance' Negative: ***te***-SM-***ri***-R-***a hë*** *ti-turibhina hë* 'we will not dance'	(***n***)-SM-***ri***-R-***anga*** *nturibhinanga* 'we will be dancing'	***(n)***-SM-***ri-bh-anga*** SM-***ra***-R-***a*** *nturibhanga turabhina* 'we will be dancing'	***u-ri-nyoor-a*** SM-R-***irë*** *urinyoora tubhinirë* 'we will have danced (lit. you will find we have danced)'

Other TAM forms

'Untimed Past Anterior'[197]: SM-***aa***-R-***a***
Narrative/Consecutive: SM-***ka***-R-***a***
Narrative (second form): SM-***ra***-R-***a***
General Future (less certainty): ***n***-SM-***Ø***-R-***a***
Negative Anterior: SM-***ta-raa***-R-***a***[198]
Persistive: ***n-ko***-R-***a*** SM-***kë-rë*** / SM-***ke***-R-***a***
Negative Persistive Anterior: ***te***-SM-***raa***-R-***a hë***
Future Persistive (or F_2): ***n***-SM-***raa-sh-e*** SM-R-***e***
Singular Imperative: ***Ø***-R-***a***
Subjunctive/Plural Imperative: SM-***Ø***-R-***e***
Hortative: SM-***ta***-R-***a***
Conditional: SM-***raa***-R-***e***
Hypothetical: ***n***-SM-***kaa***-R-***irë***
Negative Hypothetical: ***te***-SM-***kaa***-R-***irë hë***[199]
Past Hypothetical: ***singa*** SM-***aa-rë*** SM-***ta***-R-***irë***
Negative Past Hypothetical: ***te***-SM-***kaa***-R-***irë hë***
Potential: ***n***-SM-***tor-a o-ko***-R-***a***

Relative/Subordinate Clause Forms[200]

Relating to Tense:
P_3/P_4: SM-***aa***-R-***irë*** (in RCAs, same form as in MCAs but without focus marking)

[195] This form is used in spoken Simbiti, but was not considered proper usage by my language consultant.
[196] This form is used in spoken Simbiti, but was not considered proper usage by my language consultant.
[197] This terminology is borrowed from the analysis of a similar verb form in Kuria (Cammenga 2004: 285-287)
[198] This form is used in constructions indicating "before he/she had (R-ed)".
[199] There is another Negative Hypothetical form ***singa te***-SM-R-***irë hë*** which needs further research.
[200] This is not an exhaustive list of relative/subordinate clause forms, but merely lists the forms that are witnessed in the available data.

PRS: SM-***aa***-R-***a*** (in temporal RCAs, same form as "Untimed Past Anterior" marker in MCAs)
PRS: SM-***ko***-R-***a*** (in non-temporal RCAs, not used in MCAs)
F_1: SM-***raa***-R-***e*** (same form as in MCAs but without focus marking)
F_2: SM-***aka***-R-***e*** (same form as in MCAs but without focus marking)

Relating to Aspect and Combined TA:
Past Habitual: SM-***a***-R-***anga*** (same form as P_3/P_4 HAB in MCAs)
Anterior/P_1: SM-R-***irë*** (same form as in MCAs)

Relating to Non-TA categories:
Conditional Clause: SM-***raa***-R-***e*** (in RCAs, same form as F_1 in MCAs but without focus marking)
Negative Conditional: SM-***taa***-R-***e*** (in RCNs, distinct from NEG F_1 in MCNs)
Hypothetical: SM-***kaa***-R-***irë*** (in RCAs, same as HYP in MCAs but without focus marking)
Negative Hypothetical: SM-***ta-kaa***-R-***irë*** (in RCNs, distinct from NEG HYP in MCNs)

Auxiliary Verbs
-*bha*: Future/Anterior copula (from verb 'be/become')
-*rë*: Past/Progressive/Persistive/Hypothetical/Possessive/Locative copula (Habitual with *–nga* forms)
-*mara*: Expressing the concept 'already' (from verb 'finish')
-*Vsha*: Used to express some Futures together with a verb in the Subjunctive (from verb 'come')
-*tuna*: Used to express Future Inceptive 'about to'[201] (from verb 'want')
-*tora*: Used to express Potential (from verb 'be able to')
-*nyoora*: Used to express some Futures including Future Perfects (from verb 'find')

[201] However, this appears to be a direct translation from the Swahili rather than a natural Simbiti expression.

Table 14: Ikoma TA Chart

Tense	Basic	Progressive	Anterior/Perfect
P_3-P_4	***n***-SM-***V***-R-***iri*** *ntooghoriri* 'we bought (before today)'; *ntoobhiniri* 'we danced (before today)' Negative: SM-***ta***-***a***-R-***iri*** *totaaghoriri* 'we did not buy (before today)'	P_1/P_2/P_3/P_4 (temporal): ***n***-SM-***V-re*** SM-***ra***-R-***a*** *ntoore toraghora* 'we were buying' Narrative/Consecutive: SM-***Vka-bh-a*** SM-***ra***-R-***a*** *tookabha toraghora* '(and then) we were buying'	***n***-SM-***V-re*** SM-R-***iri*** *ntoore toghoriri* 'we had bought'
P_1-P_2 (Anterior/Perfect)	***n***-SM-R-***iri*** *ntoghoriri* 'we bought (earlier today)'/ 'we have bought'; *ntobhiniri* 'we danced (earlier today)'/ 'we have danced' Negative: SM-***ta***-R-***iri*** *totaghoriri* 'we did not buy (earlier today)'	Anterior (not temporal): ***n***-SM-***bh-eere*** SM-***ra***-R-***a*** *ntubheere toraghora* 'we have been buying'	***Not Possible:*** **ntubheere tobhiniri*
PRS	***n***-SM-***Vko***-R-***a*** *ntooghoghora* 'we buy' Negative: SM-***ta-ako***-R-***a*** *totaaghoghora* 'we do not buy'	***n***-SM-***Vko***-R-***a*** *ntooghoghora* 'we are buying' Negative: SM-***ta-ako***-R-***a*** *totaaghoghora* 'we are not buying'	***n***-SM-R-***iri*** *ntoghoriri* 'we have bought'; *ntobhiniri* 'we danced (earlier today)'
F_1- F_4	***n***-SM-***Vko***-R-***a*** *ntooghoghora* 'we will buy' Negative: SM-***ta-ako***-R-***a*** *totaaghoghora* 'we will not buy'	***n***-SM-***Vku-bh-a*** SM-***ra***-R-***a*** *ntookubha toraghora* 'we will be buying'	***n***-SM-***Vku-bh-a*** SM-R-***iri*** *ntookubha toghoriri* 'we will have bought'

Other TAM forms
Non-Past Consecutive: SM-***Ø***-R-***a***
Negative Consecutive?: SM-***taa***-R-***a***
Narrative/Past Consecutive: SM-***Vka***-R-***a***
Narrative (second form): SM-***ra***-R-***a***
Habitual: ***n***-SM-***haa***-R-***a***
Persistive: ***n***-SM-***kee***-R-***a***

Past Inceptive: SM-***Vká***-R-***a*** (high tone on -*Vká*- differentiates this from the Narrative)
Singular Imperative: ***Ø***-R-***a***
Subjunctive/Plural Imperative: SM-***Ø***-R-***e***
Conditional: SM-***raa***-R-***a***[202]
Hypothetical: ***n***-SM-***nga***-R-***iri***
Negative Hypothetical: SM-***ta***-***nga***-R-***iri***

Relative/Subordinate Clause Forms[203]

Relating to Tense:
P_2/P_3/P_4: SM-***V***-R-***iri*** (same form as in MCAs)
Past Copula: SM-***a***-***re***
Non-Past: SM-***Vko***-R-***a*** (same form as in MCAs)

Relating to Aspect and Combined TA:
Anterior: SM-R-***iri***
Persistive: SM-***kee***-***re***

Auxiliary Verbs
-*bha*: Future/Habitual/Anterior copula (from verb 'be/become')
-*nyi*: Present copula
-*re*: Past/Persistive copula
-*mara*: Expressing the concept 'already' (from verb 'finish')
-*Vcha*: Used to express a Future Persistive (from verb 'come')
-*enda*: Used to express Future Inceptive 'about to'[204] (from verb 'want')

[202] There appears to be overlap between this form and the SM-***ra***-R-***a*** form.
[203] This is not an exhaustive list of relative/subordinate clause forms, but merely lists the forms that are witnessed in the available data.
[204] However, this appears to be a direct translation from the Swahili rather than a natural Ikoma expression.

Table 15: Zanaki TA Chart

Tense	Basic	Progressive	Anterior/Perfect
P_2-P_4	SM-***ka***-R-***a*** *tukagura* 'we bought (earlier today or before today)' Negative: SM-***ta***-***a***-R-***iri*** *tutaaguriri* 'we did not buy (earlier today or before today)'	Narrative/Consecutive: SM-***ka-bh-a*** SM-***ra***-R-***a*** *tukabha turagura* 'we were buying'; *tukabha turabhina* 'we were dancing' Non-Durative[205]: SM-***aa-ri*** SM-***ra***-R-***a***[206] *twaari turabhina* 'we were dancing'	Narrative/Consecutive: SM-***ka-bh-a*** SM-R-***iri*** *tukabha tuguriri* 'we had bought'; *tukabha tubhiniri* 'we had danced' Non-Durative: SM-***aa-ri*** SM-R-***iri***[207] *twaari tubhiniri* 'we had danced'
P_1-P_2 (Anterior/Perfect)	SM-R-***irɪ*** *tuguriri* 'we bought (recently or earlier today)'/ 'we have bought'; *tubhiniri* 'we danced (recently or earlier today)'/ 'we have danced' Negative: SM-***ta***-R-***iri*** *tutaguriri* 'we have not bought (today)'	Anterior (not temporal): SM-***bh-eeri*** SM-***ra***-R-***a*** *tubheeri turabhina*[208] 'we have been dancing'	***Not Possible:*** **tubheeri tubhiniri*
PRS	SM-***ra***-Ŕ-***a*** *turagura* 'we buy'; *turabhina* 'we dance' Negative: SM-***ta-ku***-R-***a*** *tutakugura* 'we are not buying'	SM-***ra***-R-***a*** *turagura* 'we are buying'	SM-R-***iri*** *tuguriri* 'we have bought'
F_1- F_4	SM-***ra***-Ŕ-***a*** *turagura* 'we will buy' Negative: SM-***ta-ku***-R-***a*** *tutakugura* 'we will not buy'	SM-***ra-bh-a*** SM-***ra***-R-***a*** *turabha turagura* 'we will be buying'; *turabha turabhina* 'we will be dancing'	SM-***ra-bh-a*** SM-R-***iri*** *turabha tuguriri* 'we will have bought'; *turabha tubhiniri* 'we will have danced'

[205] This action is implied to be ongoing over a short period of time (Futakamba et al. 2013).
[206] There is an abbreviated version of this form that is only used in the 3rd person singular as ***ya-ara***-R-***a*** and the 3rd person plural as ***bha-bh-ara***-R-***a***. The 3rd person singular form can be seen in the Zanaki Pear Story (Appendix F), lines Zak10a and Zak12.
[207] There is an abbreviated version of this form: SM-***aa***-R-***iri***.
[208] This form was elicited in the initial session but was difficult to re-elicit. It is possibly a Swahili translation rather than a natural Zanaki form.

Other TAM forms
Narrative/Consecutive: SM-***ka***-R-***a***
Past Habitual?: SM-***a(a)***-R-***a***
Habitual: SM-***haa***-R-***a***
Past Persistive: SM-***ka-bh-a*** SM-***kyaa***-R-***a***[209]
Future Persistive: SM-***ki***-R-***a***
Persistive Progressive: SM-***kyaa-ri*** SM-***ra***-R-***a***
Negative Persistive Anterior: SM-***kyaa-ri ku***-R-***a***
Singular Imperative: ***Ø***-R-***a***
Subjunctive/Plural Imperative: SM-***Ø***-R-***e***
Conditional: SM-***raa***-R-***e***
Hypothetical: SM-***a-nga***-R-***iri*** / SM-***nga***-R-***iri***
Negative Hypothetical: SM-***ta-a-ri ku***-R-***a*** / SM-***ti-nga***-R-***iri***
Past Hypothetical: SM-***a-nga-ri***-R-***iri*** / SM-***nga-ri***-R-***iri***
Negative Past Hypothetical: SM-***ta-nga-ri***-R-***iri***

Relative/Subordinate Clause Forms[210]

Relating to Tense:
Past: SM-***ka***-R-***a*** (related to the PST/NAR in MCAs)
P_3/P_4: SM-***a***-R-***iri*** (in RCAs; no evidence of this form in MCAs)
Non-Past: SM-***ku***-R-***a*** (no evidence of this form in MCAs)

Relating to Aspect and Combined TA:
Past Progressive: SM-***aa-ri ku***-R-***a*** (in RCAs; no evidence of this form in MCAs)
Past Progressive: SM-***aa-ri*** SM-***ra***-R-***a*** (in conditional clauses; less frequent in MCAs)
Present Durative Progressive: SM-***ra-bh-a*** SM-***ra***-R-***a*** (same form as in MCAs)
Present/Progressive: SM-***ra***-R-***a*** (in conditional clauses; same form as in MCAs)
Anterior/Perfect: SM-R-***iri*** (same form as in MCAs)
Past Persistive: SM-***a-kya***-R-***a*** (in RCAs; no evidence of this form in MCAs)

Auxiliary Verbs
-bha: Durative copula indicating habituality or repeated action (from verb 'be/become')
-ri: Non-durative copula (used also with locatives and possessive 'be with' forms)
-mara: Expressing the concept 'already' (from verb 'finish')
-Vza: Expressing remote future (from verb 'come')
-enda: Used to express Negative Persistive and Future Inceptive 'about to'[211] (from verb 'want')
-tama: Used to express negation in some forms (from verb 'fail')
-nagya: Used to express ability (from verb 'be able to')

ni: Basic copula

[209] The formative *-kyaa-* 'PER' is written with a long vowel in the orthography because it follows a palatalized consonant. The actual phonetic length of the vowel has not been investigated.
[210] This is not an exhaustive list of relative/subordinate clause forms, but merely lists the forms that are witnessed in the available data.
[211] However, this appears to be a direct translation from the Swahili rather than a natural Zanaki expression.

Table 16: Ikizu TA Chart

Tense	**Basic**	**Progressive**	**Anterior/Perfect**
P_2-P_4	SM-***ka***-R-***a*** *tʉkabina* 'we danced (earlier today or before today)'; *tʉkagʉra* 'we bought (earlier today or before today)' P_4 Anterior: SM-***aa***-R-***iri*** *twaaguriri* 'we bought' Negative: SM-***ta***-***a***-R-***iri*** *tʉtaaguriri* 'we did not buy'	Narrative/Consecutive: SM-***ka***-***b***-***a*** SM-***ra***-R-***a*** *tʉkaba tʉrabina* 'we were dancing'; *tʉkaba tʉragʉra* 'we were buying' Non-Durative: SM-***aa***-***ri*** SM-***ra***-R-***a***[212] *twaari tʉrabina* 'we were dancing'; *twaari tʉragʉra* 'we were buying'	Narrative/Consecutive: SM-***ka***-***b***-***a*** SM-R-***iri*** *tʉkaba tuguriri* 'we had bought' Non-Durative: SM-***aa***-***ri*** SM-R-***iri*** *twaari tuguriri* 'we had bought'
P_1-P_2 (Anterior/Perfect)	SM-R-***iri*** *tʉguriri* 'we bought (recently or earlier today)' / 'we have bought' Negative: SM-***ta***-R-***iri*** *tʉtaguriri* 'we have not bought (today)'	Anterior (not temporal): SM-***b***-***iiri*** SM-***ra***-R-***a*** *tʉbiiri tʉrabina* 'we have been dancing'	***Not Possible:*** * *tʉbiiri tubhiniri*
PRS	SM-***aa***-R-***a*** *twaagʉra* 'we buy' Negative: SM-***ta***-***ku***-R-***a*** *tʉtakʉgʉra* 'we do not buy'	SM-***ra***-R-***a*** *tʉrabina* 'we are dancing'; *tʉragʉra* 'we are buying'	SM-R-***iri*** *tuguriri* 'we have bought'
F_1- F_4	SM-***ra***-Ŕ-***a*** *tʉragʉra* 'we will buy' Negative: SM-***ta***-***ku***-R-***a*** *tʉtakʉgʉra* 'we will not buy'	SM-***ra***-***b***-***a*** SM-***ra***-R-***a*** *tʉraba tʉrabina* 'we will be dancing'; *tʉraba tʉragʉra* 'we will be buying	SM-***ra***-***b***-***a*** SM-R-***iri*** *tʉraba tuguriri* 'we will have bought'

<u>Other TAM forms</u>
Narrative/Consecutive: SM-***ka***-R-***a***
Habitual: SM-***haa***-R-***a***
Progressive: SM-***ra***-R-***a***
Future Persistive: SM-***kii***-***ri*** SM-***cha***-***az***-***a*** ***ku***-R-***a***

[212] My language consultant told me that SM-***a***-***ri*** ***ku***-R-***a*** is used more frequently in speech as an abbreviated form with the same meaning.

Negative Future Persistive: SM-***kii-ri*** **na** SM-***ta-ku***-R-***a***
Persistive Progressive: SM-***kii-ri*** SM-***ra***-R-***a***
Negative Persistive Progressive: SM-***kii-ri*** SM-***ta-ku***-R-***a***
Negative Persistive Anterior: SM-***kii-ri ku***-R-***a***
Singular Imperative: ***Ø***-R-***a***
Subjunctive/Plural Imperative: SM-***Ø***-R-***i***
Conditional: SM-***raa***-R-***i***
Hypothetical: SM-***a-nga***-R-***iri***
Negative Hypothetical: SM-***ta-nga***-R-***iri***
Possessive: SM-***na***

Relative/Subordinate Clause Forms[213]

Relating to Tense:
Past: SM-***aa***-R-***iri*** (this form also used for P_4 in MCAs)
Past: SM-***aa-az-a ku***-R-***a*** (for temporal relatives; becoming lexicalized with no relative pronoun)
Non-Past: SM-***ku***-R-***a*** (no evidence of this form in MCAs)

Relating to Aspect and Combined TA:
Past Progressive: SM-***aa-ri*** SM-***ku***-R-***a*** (in RCAs; no evidence of this form in MCAs)
Past Progressive: SM-***aa-ri*** SM-***ra***-R-***a*** (in conditional clauses; same form as in MCAs)
Anterior Progressive: SM-***b-i-iri*** SM-***ra***-R-***a*** (same form as in MCAs)
Present/Progressive: SM-***ra***-R-***a*** (in RCAs and conditional clauses; same form as in MCAs)
Anterior/Perfect: SM-R-***iri*** (same form as in MCAs)

Auxiliary Verbs
-ba: Durative copula indicating habituality or repeated action (from verb 'be/become')
-ri: Non-durative copula (used also with Locatives and Persistive forms)
-mara: Expressing the concept 'already' (from verb 'finish')
-Vza: Expressing future with the Persistive (from verb 'come')
-enda: Used to express Future Inceptive 'about to'[214] (from verb 'want')
-tama: Used to express negation in some forms (from verb 'fail')
-tura: Used to express ability (from verb 'be able to')

[213] This is not an exhaustive list of relative/subordinate clause forms, but merely lists the forms that are witnessed in the available data.
[214] However, this appears to be a direct translation from the Swahili rather than a natural Ikizu expression.

Appendix B: "Pear Story" in Ikizu

Since the original text of the "Pear Story" in Ikizu was translated first into Swahili, rather than directly into English, the following transcription of the story is a six-row interlinear text which includes word-for-word and free translations into both Swahili and English, as well as a morpheme-by-morpheme English gloss. The first row of each interlinearized section has the Ikizu text of the story and is titled Ikz, followed by a sentence and line number (i.e. 1a, 1b, 2, etc.). These designations are used in the body of the thesis to refer to specific portions of this Ikizu text. The abbreviations for the other rows, in the order they appear, are: SWW (Swahili Word-for-Word), EWW (English Word-for-Word), MMG[215] (Morpheme-by-Morpheme Gloss), SFT (Swahili Free Translation) and EFT (English Free Translation). The latter two are included only on the last line of a sentence. All verbal words are bolded in the Ikizu text.

Ikizu Pear Story

Ikz1a:	*Rusiku*	*rumwi*	*ritabuuri,*	*musubi*	*wumwi*
SWW:	siku	moja	asubuhi	mwanaume	mmoja
EWW:	day	one	morning	man	one
MMG:	*ru-siku* NC11.SG-day	*ru-mwi* AGR11-one	*ri-tabuuri* NC5-morning	*mu-subi* NC1.SG-man	*wu-mwi* AGR1-one

Ikz1b:	***akabuuka,***	***akaja***	***kutwa***	*amapeasi.*
SWW:	akaamka	akaenda	kuchuma	mapeasi
EWW:	he woke up	he went	to pick	pears
MMG:	*a-ka-buuk-a* 3SG-NAR-wake.up-FV	*a-ka-j-a* 3SG-NAR-go-FV	*ku-tw-a* INF-pick-FV	*a-ma-peasi* AUG6-NC6.PL-pear
SFT:	'Siku moja asubuhi, mwanaume mmoja akaamka, akaenda kuchuma mapeasi.'			
EFT:	'One morning, a certain man woke up and went to pick pears.'			

[215] The MMG shows the underlying forms of the morphemes.

Ikz2a:	***Akahika,***	***akatiira***	*kʉmʉti*	*gwaazi,*
SWW:	akafika	akapanda	kwenye mti	wake
EWW:	he arrived	he climbed	at the tree	his
MMG:	*a-ka-hik-a* 3SG-NAR-arrive-FV	*a-ka-tiir-a* 3SG-NAR-climb-FV	*kʉ=mʉ-ti* NC17=NC3.SG-tree	*gʉ*[216]*-azi* AGR3-3SG.POSS

Ikz2b:	***akatanga***	***kutwa.***
SWW:	akaanza	kuchuma
EWW:	he began	to pick
MMG:	*a-ka-tang-a* 3SG-NAR-begin-FV	*ku-tw-a* INF-pick-FV
SFT:	'Akafika, akapanda kwenye mti wake, akaanza kuchuma.'	
EFT:	'He reached his tree, climbed up it and began to pick.'	

Ikz3a:	***Akatwa***	*ge*	*mbere,*	***akiika,***
SWW:	akachuma	ya	kwanza	akashuka
EWW:	he picked	of	first	he descended
MMG:	*a-ka-tw-a* 3SG-NAR-pick-FV	*g-e* AGR6.PL-AGR9.ASC	*m-bere* NC9.SG-first	*a-ka*[217]*-ik-a* 3SG-NAR-descend-FV

Ikz3b:	***akatʉʉra***	*mʉkihʉʉzʉ.*
SWW:	akaweka	mwenye kikapu (cha asili)
EWW:	he put	in the (traditional) basket
MMG:	*a-ka-tʉʉr-a* 3SG-NAR-put-FV	*mʉ=ki-hʉʉzʉ* NC18=NC7.SG-traditional.basket
SFT:	'Akachuma ya kwanza, akashuka, akaweka mwenye kikapu.'	
EFT:	'He picked the first ones, climbed down and put them in his basket.'	

Ikz4a:	***Akagarʉka***	*naatu*	***akatwa,***	***akatʉʉra***
SWW:	akarudi	tena	akachuma	akaweka
EWW:	he returned	again	he picked	he put
MMG:	*a-ka-garʉk-a* 3SG-NAR-return-FV		*a-ka-tw-a* 3SG-NAR-pick-FV	*a-ka-tʉʉr-a* 3SG-NAR-put-FV

[216] The *ʉ* in the prefix is labialized before a root-initial vowel, both here and other places in the text (Gray & Smith 2010a: 18). There is conditioned lengthening of the vowel following a labialized consonant, which is written as a long vowel in the orthography.

[217] The surface form of the *ka-* morpheme is *ki-*, both here and other places in this text, because of an assimilation process from the root vowel (Gray & Smith 2010a: 19).

Ikz4b:	*mʉkihʉʉzʉ.*
SWW:	mwenye kikapu (cha asili)
EWW:	in the (traditional) basket
MMG:	*mʉ=ki-hʉʉzʉ* NC18=NC7.SG-basket
SFT:	‘Akarudi tena, akachuma, akaweka mwenye kikapu.’
EFT:	‘He climbed up again, picked another and put it in the basket.’

Ikz5a:	***Akiizurya***	*mahʉʉzʉ*	*gaazi*	*abiri,*
SWW:	akajaza	vikapu (cha asili)	vyake	viwili
EWW:	he filled	(traditional) baskets	his	two
MMG:	*a-ka-izur-y-a* 3SG-NAR-be.full-CAUS-FV	*ma-hʉʉzʉ* NC6.PL-basket	*ga-azi* AGR6-3SG.POSS	*a-biri* AGR6-two

Ikz5b:	***akatiira***	*rwa*	*katatʉ*
SWW:	akapanda	(mara) ya	tatu
EWW:	he climbed	(time) of	third
MMG:	*a-ka-tiir-a* 3SG-NAR-climb-FV	*rʉ-a* AGR11.SG-ASC	*ka-tatʉ* NC12-third
SFT:	‘Akajaza vikapu vyake viwili, akapanda mara ya tatu.’		
EFT:	‘He filled his two baskets and climbed up a third time.’		

Ikz6a:	***Aaza***[218]	***kutiira***	*kʉmʉti,*	*mʉʉtʉ*
SWW:	alikuja	kupanda	kwenye mti	mtu
EWW:	(when) he came	to climb	on the tree	person
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-tiir-a* INF-climb-FV	*kʉ=mʉ-ti* NC17=NC3.SG-tree	*mʉ-ʉtʉ* NC1.SG-person

Ikz6b:	***akahita***	***ana***	*iimburi.*[219]
SWW:	akapita	ana	mbuzi
EWW:	he passed by	he is with	goat
MMG:	*a-ka-hit-a* 3SG-NAR-pass.by-FV	*a-na* 3SG-with	*im-buri* NC9.SG-goat
SFT:	‘Alipopanda kwenye mti, mtu mmoja akapita ana mbuzi.’		
EFT:	‘As he climbed up the tree, someone passed by with a goat.’		

[218] Underlyingly the form is *a-a-az-a* but there is a rule only allowing two similar vowels to surface in a row.
[219] The underlying form of the prefix is *im-*, but before a root-initial consonant it becomes a conditioned length environment and the vowel is written long as *iim-*.

Ikz7a:	*Iimburi*	*yaazi*	***ikaba***	***irama**:*
SWW:	mbuzi	wake	ikawa	inalia
EWW:	the goat	his	it was	it is bleating
MMG:	*im-buri* NC9.SG-goat	*ya-azi* AGR9-3SG.POSS	*i-ka-b-a* SBJ9-NAR-be-FV	*i-ra-m-a* SBJ9-PROG-bleat-FV

Ikz7b:	*"Mee.*	*Mee."*
SWW:	mee	mee
EWW:	mee	mee
MMG:		
SFT:	'Mbuzi wake akawa analia, "Mee. Mee." '	
EFT:	'His goat was bleating: "Mee. Mee." '	

Ikz8a:	*Wi*	*imburi*	***akahita**,*	***akaja***
SWW:	mwenye	mbuzi	akapita	akaenda
EWW:	owner	the goat	he passed by	he went
MMG:	*w-i* AGR1-AGR9.ASC	*im-buri* NC9.SG-goat	*a-ka-hit-a* 3SG-NAR-pass.by-FV	*a-ka-j-a* 3SG-NAR-go-FV

Ikz8b:	***araruta***	*iimburi*	*yaazi.*
SWW:	anavuta	mbuzi	wake
EWW:	he is pulling	goat	his
MMG:	*a-ra-rut-a* 3SG-PROG-pull-FV	*im-buri* NC9.SG-goat	*ya-azi* AGR9-3SG.POSS
SFT:	'Mwenye mbuzi akapita, akaenda anavuta mbuzi wake.'		
EFT:	'The owner of the goat passed by pulling his goat.'		

Ikz9a:	***Aaza***	***kuja**,*	*naatu*	*urubaara*
SWW:	alikuja	kwenda	tena	upande
EWW:	he came	to go	again	side
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-j-a* INF-go-FV		*u-ru-baara* AUG11-NC11.SG-side

Ikz9b:	***akarwiiyu***	*mwaana*
SWW:	akatokako	mtoto
EWW:	he came out from there	child
MMG:	*a-ka-rw-i*[220]*=yu* 3SG-NAR-come.out-FV=LOC23	*mu-ana* NC1.SG-child

Ikz9c:	***ari***	*kubasikiri.*
SWW:	yuko	kwenye baiskeli
EWW:	he is	on the bicycle
MMG:	*a-Ø-ri* 3SG-PRS-COP	*ku=Ø-basikiri* NC17=NC9.SG-bicycle
SFT:	'Alipokwenda, tena upande akatokea mtoto yuko kwenye baiskeli.'	
EFT:	'As he left, a child on a bicycle came from the side.'	

Ikz10:	*Umwaana*	*wuyo*	***akaaza***
SWW:	mtoto	huyo	akaja
EWW:	child	that	he came
MMG:	*u*[221]*-mu-ana* AUG1-NC1.SG-child	*wu-yo* AGR1-RDEM	a-ka-az-a 3SG-NAR-come-FV
SFT:	'Mtoto huyo akaja.'		
EFT:	'That child got closer.'		

Ikz11a:	***Aaza***	***kuhika***	*hayo*	*kumuti*
SWW:	alikuja	kufika	hapo	kwenye mti
EWW:	he came	to arrive	(right) there	at the tree
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-hik-a* INF-arrive-FV	*ha-yo* NC16-RDEM	*ku=mu-ti* NC17=NC3.SG-tree

Ikz11b:	*gwu*	*mukaruka*	*wuyo,*	***akiimiirira,***
SWW:	wa	mzee	huyo	akasimama
EWW:	of	old man	that	he stopped
MMG:	*gu-u* AGR3-AUG1.ASC	*mu-karuka* NC1.SG-old.man	*wu-yo* AGR1-RDEM	a-ka-imiirir-a 3SG-NAR-stop-FV

[220] The *-i* following the verb root may be a Subjunctive mood marker or it may be vowel harmonization of the mono-syllabic root to the locative post-clitic. This needs further investigation.
[221] There is an underlying *u* at AUG which is raised to *u* because of the labialization on the NC prefix preceding the vowel initial root (Gray & Smith 2010a: 18).

Ikz11c:	***akarora***	*amapeasi*	*garya*	*mwiihuuzu.*
SWW:	akaona	mapeasi	yale	mwenye kikapu (cha asili)
EWW:	he saw	pears	those	in the (traditional) basket
MMG:	a-ka-ror-a 3SG-NAR-see-FV	*a-ma-peasi* AUG6-NC6.PL-pear	*ga-rya* AGR6- DDEM	*mu=i*[222]*-huuzu* NC18=NC5.SG-basket
SFT:	'Alipofika hapo kwenye mti wa mzee huyo, akasimama, akaona mapeasi yale mwenye kikapu.'			
EFT:	'When he reached the old man's tree, he stopped and saw the pears in the basket.'			

Ikz12a:	***Akatuura***	*ibasikiri,*	***akayihindirya,***
SWW:	akaweka	baiskeli	akailaza
EWW:	he put	bicycle	he laid it down
MMG:	*a-ka-tuur-a* 3SG-NAR-put-FV	*i-basikiri* NC9.SG-bicycle	*a-ka-yi-hindir-y-a* 3SG-NAR-OBJ9-lay.down-CAUS-FV

Ikz12b:	***akiihinya***	*kwiihuuzu*	***kotoora***
SWW:	akainama	kwenye kikapu (cha asili)	kuokota
EWW:	he bent over	at the (traditional) basket	to pick up
MMG:	*a-ka-i-hiny-a* 3SG-NAR-REFL-bend.over-FV	*ku=i-huuzu* NC17=NC5.SG-basket	*ko-toor-a* INF-pick.up-FV

Ikz12c:	*ripeasi*	*rimwi.*
SWW:	peasi	moja
EWW:	pear	one
MMG:	*ri-peasi* NC5.SG-pear	*ri-mwi* AGR5-one
SFT:	'Akaweka baiskeli, akailaza, akainama kwenye kikapu kuokota peasi moja.'	
EFT:	'He laid his bike down and bent over the basket to pick up a pear.'	

Ikz13a:	***Akumara***[223]	***akamagamaga***	***akanunga***	*igunu*
SWW:	kisha	akaangaza	akainua	juu
EWW:	then	he stared	he lifted	up
MMG:	*a-ku-mar-a* 3SG-NPST-finish-FV	*a-ka-maga~mag-a* 3SG-NAR-stare~RDPL-FV	*a-ka-nung-a* 3SG-NAR-lift-FV	*i-guru* NC23-up

[222] An allomorph of *ri-* for the NC5 prefix when following a locative noun class prefix. Gray (2013: 17) traces the likely historical origin of this morpheme.

[223] This word is lexicalized as a temporal clause introducer, but it is still analyzable as a verbal form. When broken into verbal components, the morphemes are glossed as follows: *a-* '3SG', *-ku-* 'NPST.REL', *mar* 'finish', and *-a* 'FV'.

Ikz13b:	*umutwe*	*kʉmʉti,*	***akarora***	*ʉmʉkaruka*
SWW:	kichwa	kwenye mti	akaona	mzee
EWW:	head	at the tree	he saw	old man
MMG:	*u-mu-twe* AUG3-NC3.SG-head	*kʉ=mʉ-ti* NC17=NC3.SG-tree	*a-ka-ror-a* 3SG-NAR-see-FV	*ʉ-mʉ-karuka* AUG1-NC1.SG-old.man

Ikz13c:	*wurya*	***atakomorora***.
SWW:	yule	hamwoni
EWW:	that	he does not seem him
MMG:	*wu-rya* NC1.SG-DDEM	*a-ta-ko-mo-ror-a* 3SG-NEG-NPST-3SG-see-FV
SFT:	‘Kisha akaangaza, akainua juu kichwa kwenye mti, akaona mzee yule hamwoni.’	
EFT:	‘Then he lifted up his head and glanced in the tree. He saw that the old man didn’t see him.’	

Ikz14a:	***Akiihinya,***	***akatoora***	*rihʉʉzʉ*
SWW:	akainama	akaokota	kikapu (cha asili)
EWW:	he bent over	he picked up	the (traditional) basket
MMG:	*a-ka-i-hiny-a* 3SG-NAR-REFL-bend.over-FV	*a-ka-toor-a* 3SG-NAR-pick.up-FV	*ri-hʉʉzʉ* NC5.SG-basket

Ikz14b:	*ryʉʉsi,*	***akabuucha***	*ɨbasikɨri,*	***akagega***
SWW:	chote	akainua	baiskeli	akabeba
EWW:	all	he lifted up	bicycle	he carried
MMG:	*ri*[224]*-ʉsi* AGR5-all	*a-ka-bʉʉk-y*[225]*-a* 3SG-NAR-arise-CAUS-FV	*ɨ-basikɨri* NC9.SG-bicycle	*a-ka-geg-a* 3SG-NAR-carry-FV

Ikz14c:	*rihʉʉzʉ*	*ryaazɨ,*	***akatʉʉra***
SWW:	kikapu (cha asili)	chake	akaweka
EWW:	the (traditional) basket	his	he put
MMG:	*ri-hʉʉzʉ* NC5.SG-basket	*ri-azɨ* AGR5-3SG.POSS	*a-ka-tʉʉr-a* 3SG-NAR-put-FV

[224] The *i* in the prefix is palatalized before a root-initial vowel, both here and other places in the text (Gray & Smith 2010a: 17-18).
[225] An underlying *ky* surfaces as *ch* in Ikizu (Gray & Smith 2010a: 9). Also, the causative extension causes raising of the root vowels from underlying *ʉ* /ʊ/ to *u* /u/ (14-15).

Ikz14d:	*kʉbasikɨri,*	***akaja***.
SWW:	kwenye baiskeli	akaenda
EWW:	on the bicycle	he went
MMG:	*kʉ=Ø-basikɨri* NC17=NC9.SG-bicycle	*a-ka-j-a* 3SG-NAR-go-FV
SFT:	‘Akainama, akaokota kikapu chote, akainua baiskeli, akabeba kikapu chake, akaweka kwenye baiskeli, akaenda.’	
EFT:	‘He bent over, took a whole basket, lifted up the bike, carried his basket, put it on the bike and left.’	

Ikz15a:	***Aaza***	***kuja***	*mʉnzɨra,*	*kwe*
SWW:	alikuja	kwenda	njiani	kwa
EWW:	he came	to go	on the path	(at the place) of
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-j-a* INF-go-FV	*mʉ=n-zɨra* NC18=NC9.SG-path	*kʉ-e* NC17-AGR9.ASC

Ikz15b:	*mbere*	*muucha*	*wʉmwɨ*	*wʉʉsi*	***akaba***
SWW:	mbele	binti	mmoja	naye	akawa
EWW:	front	girl	one	also	she was
MMG:	*m-bere* NC9.SG-front	*mu-ucha* NC1-girl	*wʉ-mwɨ* AGR1-one	*wʉ-ʉsi* AGR1-all	*a-ka-b-a* 3SG-NAR-be-FV

Ikz15c:	***araaza***	*nɨ*	*basikɨri.*
SWW:	anakuja	na	baiskeli
EWW:	she is coming	with	bicycle
MMG:	*a-ra-az-a* 3SG-PROG-come-FV	*nɨ* CCONJ.AUG9	*Ø-basikɨri* NC9.SG-bicycle
SFT:	‘Alipokwenda njiani, kwa mbele binti mmoja naye akawa anakuja na baiskeli.’		
EFT:	‘As he went down the path, in front of him a girl was also coming on a bike.’		

Ikz16a:	*Umuucha*	*wuyo*	***aaza***	***kuhika,***	*hanʉ*
SWW:	msichana	huyo	alikuja	kufika	hapa
EWW:	girl	that	she came	to arrive	here
MMG:	*u-mu-ucha* AUG1-NC1-girl	*wu-yo* AGR1-RDEM	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-hik-a* INF-arrive-FV	*ha-nʉ* NC16-PDEM

Ikz16b:	***bagiiri***	***kuhitirana***.
SWW:	wameenda	kupishana
EWW:	they have gone	to pass by each other
MMG:	*ba-gi-iri* 3PL-go-ANT	*ku-hit-ir-an-a* INF-pass.by-APPL-RECP-FV
SFT:	'Binti huyo alipofika, walipishana.'	
EFT:	'When she arrived, they passed by each other.'	

Ikz17a:	*Umuucha*	*wurya*	***akahuna***	*ingu̱hiira*
SWW:	msichana	yule	akagonga	kofia
EWW:	girl	that	she hit	hat
MMG:	*u-mu-ucha* AUG1-NC1.SG-girl	*wu-rya* AGR1-DDEM	*a-ka-hun-a* 3SG-NAR-hit-FV	*in-gu̱hiira* NC9.SG-hat

Ikz17b:	*yaazi,*	*kumutwi,*	*ingu̱hiira*	*ikagwa.*
SWW:	yake	kichwani	kofia	ikaanguka
EWW:	his	on head	hat	it fell
MMG:	*ya-azi* AGR9-3SG.POSS	*ku=mu-twi* NC17=NC3.SG.head	*in-gu̱hiira* NC9.SG-hat	*i-ka-gw-a* SBJ9-NAR-fall-FV
SFT:	'Binti yule akagonga kofia yake kichwani, kofia ikaanguka.'			
EFT:	'The girl hit the hat on his head and it fell off.'			

Ikz18a:	***Yaaza***	***kugwa,***	*umwaana*	*wurya*
SWW:	ilikuja	kuanguka	mtoto	yule
EWW:	it came	to fall	child	that
MMG:	*i-a-az-a* SBJ9-PST.REL-come-FV	*ku-gw-a* INF-fall-FV	*u-mu̱-ana* AUG1-NC1.SG-child	*wu-rya* AGR1-DDEM

Ikz18b:	***kwiitacha***	***kwiichu̱ra***.
SWW:	kushtuka	kugeuka
EWW:	to be surprised (himself)	to turn (himself)
MMG:	*ku̱-i-tach-a* INF-REFL-surprise-FV	*ku̱-i-chu̱r-a* INF-REFL-turn-FV
SFT:	'Ilipoanguka, yule mtoto akashtuka, akageuka.'	
EFT:	'When it fell, the boy was surprised and turned around.'	

Ikz19a:	***Kwiirahuta***	*inyuma,*	***akiihuna***	*kwiibwɨ*
SWW:	kugeuka	nyuma	akajigonga	kwenye jiwe
EWW:	to turn (himself)	back	he hit (himself)	at rock
MMG:	*kʉ-i-rahut-a* INF-REFL-turn-FV	*i-nyuma* NC9-back	*a-ka-i-hun-a* 3SG-NAR-REFL-hit-FV	*ku=i-bwɨ* NC17=NC5.SG-rock

Ikz19b:	*nɨ*	*basikɨri*	*yaazɨ,*
SWW:	na	baiskeli	yake
EWW:	and	bicycle	his
MMG:	*nɨ* CCONJ.AUG9	*Ø-basikɨri* NC9.SG-bicycle	*ya-azɨ* AGR9-3SG.POSS

Ikz19c:	***akiitema***	*haasɨ:*	*“pu!”*
SWW:	akaanguka	chini	pu
EWW:	it fell (itself)	down	pu
MMG:	*a-ka-i-tem-a* 3SG-NAR-REFL-hit-FV	*ha-asɨ* NC16-ground[226]	
SFT:	‘Alipogeuka nyuma, akajigonga kwenye jiwe na baiskeli yake akaanguka chini: “pu!” ’		
EFT:	‘When he turned around, the bicycle hit a rock and fell over: “pu!” ’		

Ikz20:	*Na*	*mapeasi*	*gaazɨ*	***gakanyaragana****.*
SWW:	na	mapeasi	yake	yakatawanyika
EWW:	and	pears	his	they were scattered
MMG:	*na* CCONJ.AUG6	*ma-peasi* NC6-pear	*ga-azɨ* AGR6-3SG.POSS	*ga-ka-nyaragan-a* SBJ6-NAR-be.scattered-FV
SFT:	‘Na mapeasi yake yakatawanyika.’			
EFT:	‘And his pears were scattered all over.’			

Ikz21a:	***Gaaza***	***kʉnyaragana****,*	***akabʉʉka****,*
SWW:	yalikuja	kutawanyika	akaamka
EWW:	they came	to be scattered	he arose
MMG:	*ga-a-az-a* SBJ6-PST.REL-come-FV	*ku-nyaragan-a* INF-be.scattered-FV	*a-ka-bʉʉk-a* 3SG-NAR-arise-FV

[226] Compare with NC9 word *ɨsɨ* ‘earth/ground’ and see also Gray (2013: 15).

Ikz21b:	***akiikong'ererya,***	***akabuuka.***
SWW:	akajikaza	akaamka
EWW:	he mustered courage	he arose
MMG:	*a-ka-i-kong'erery*[227]*-a* 3SG-NAR-REFL-muster.courage-FV	*a-ka-buuk-a* 3SG-NAR-arise-FV
SFT:	'Yalipotawanyika, akajikaza, akaamka.'	
EFT:	'When they scattered, he gathered himself together and got up.'	

Ikz22a:	***Aaza***	***kubuuka,***	*nawe*	*kwiiki*	*kwe*
SWW:	alikuja	kuamka	lakini	tena	kwa
EWW:	he came	he arose	but	again	(at the place) of
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-buuk-a* INF-arise-FV			*ku-e* NC17-AGR9.ASC

Ikz22b:	*mbere*	*waazi,*	***baari***	***bareho***	*baana*
SWW:	mbele	yake	walikuwa	wapo	watoto
EWW:	front	his	they were	they are (there)	children
MMG:	*mbere* NC9.SG-front	*wa-azi* NC1-3SG.POSS	*ba-aa-ri* 3PL-NAR-COP	*ba-re*[228]*=ho* 3PL-COP=LOC16	*ba-ana* NC2.PL-child

Ikz22c:	*ba*	*kisubi*	*batatu*	***barabaraana.***
SWW:	wa	kiume	watatu	wanacheza
EWW:	of	male	three	they are playing
MMG:	*b-a* AGR2-ASC	*ki-subi* NC7- male	*ba-tatu* AGR2-three	*ba-ra-baraan-a* 3PL-PROG-play-FV
SFT:	'Alipoamka, lakini tena kwa mbele yake, walikuwepo watoto wakiume watatu wakicheza.'			
EFT:	'As he got up, there were three boys playing way off in front of him.'			

Ikz23a:	*abaana*	*barya*	***bakamorora,***	***bakaryaara***
SWW:	watoto	wale	wakamwona	wakakimbia
EWW:	children	those	they saw him	they ran
MMG:	*a-ba-ana* AUG2-NC2.PL-child	*ba-rya* AGR2-DDEM	*ba-ka-mo-ror-a* 3PL-NAR-3SG-see-FV	*ba-ka-ryaar-a* 3PL-NAR-run-FV

[227] This word appears to be derivable at an even more basic level from *kong'u* 'strong' with the applicative (*–er*) and causative (*-y*) extensions (Gray: pc). However, derivations of this type have not been adequately studied yet and so the gloss in the text is kept simple.

[228] This is an allomorph of the Copula *-ri* that is affected by the locative post-clitic (Gray 2013: 50, footnote 8).

Ikz23b:	***kuuza***	***kʉmʉsakirya,***	***bakamubuucha,***
SWW:	kuja	kumsaidia	wakamwamsha
EWW:	to come	to help him	they got him up
MMG:	*ku-uz-a* INF-come-FV	*kʉ-mʉ-sakiry-a* INF-3SG.OBJ-help-FV	*ba-ka-mu-buuk-y-a* 3PL-NAR-3SG.OBJ-arise-CAUS-FV

Ikz23c:	***bakamʉkʉng'ʉta,***	***bakamwaarwa***	***kotoora***
SWW:	wakamfuta	wakamsaidia	kuokota
EWW:	they dusted him off	they helped him	to pick up
MMG:	*ba-ka-mʉ-kʉng'ʉt-a* 3PL-NAR-3SG.OBJ-dust.off-FV	*ba-ka-mu-arw-a* 3PL-NAR-3SG.OBJ-help-FV	*ko-toor-a* INF-pick.up-FV

Ikz23d:	*amapeasi*	*gaazi,*	***bakatʉʉra***	*mwiihʉʉzʉ.*
SWW:	mapeasi	yake	wakaweka	mwenye kikapu
EWW:	pears	his	they put	in the basket
MMG:	*a-ma-peasi* AUG6-NC6-pear	*ga-azi* AGR6- 3SG.POSS	*ba-ka-tʉʉr-a* 3PL-NAR-put-FV	*mʉ=i-hʉʉzʉ* NC18=NC5.SG-basket
SFT:	'Watoto wale wakamwona, wakakimbia kwenda kumsaidia, wakamwamsha, wakamfuta, wakamsaidia kuokota mapeasi yake, wakaweka mwenye kikapu.'			
EFT:	'Those boys saw him, ran to help him, lifted him up, dusted him off, gathered up his pears and put them in the basket.'			

Ikz24a:	*Rihʉʉzʉ*	***rikiizʉra,***	***bakamwaarwa***	***kogega,***
SWW:	kikapu (cha asili)	kikajaa	wakamsaidia	kubeba
EWW:	(traditional)	it filled up	they helped him	to carry
MMG:	*ri-hʉʉzʉ* NC5.SG-basket	*ri-ka-izʉr-a* SBJ5-NAR-fill.up-FV	*ba-ka-mu-arw-a* 3PL-NAR-3SG.OBJ-help-FV	*ko-geg-a* INF-carry-FV

Ikz24b:	***bakatʉʉra***	*kʉbasikiri.*
SWW:	wakaweka	kwenye baiskeli
EWW:	they put	on the bicycle
MMG:	*ba-ka-tʉʉr-a* 3PL-NAR-put-FV	*kʉ=Ø-basikiri* NC17=NC9.SG-bicycle
SFT:	'Kikapu kikajaa, wakamsaidia kubeba, akaweka kwenye baiskeli.'	
EFT:	'The basket filled up and they helped him to carry it and put it on the bike.'	

Ikz25:	*Umwaana*	*wurya*	***akatanʉra***	***kuja***.
SWW:	mtoto	yule	akaondoka	kwenda
EWW:	child	that	he departed	to go
MMG:	*ʉ-mʉ-ana* AUG1-NC1.SG-child	*wu-rya* AGR1-DDEM	*a-ka-tanʉr-a* 3SG-NAR-depart-FV	*ku-j-a* INF-go-FV
SFT:	‘Mtoto yule akaondoka kwenda.’			
EFT:	‘That boy started on his way.’			

Ikz26a:	***Aaza***	***kʉtanʉra***	***kuja***,	*banʉ*	*bʉʉsi*
SWW:	alikuja	kuondoka	kwenda	hawa	nao
EWW:	he came	to depart	to go	these	also
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*kʉ-tanʉr-a* INF-depart-FV	*ku-j-a* INF-go-FV	*ba-nʉ* NC2-PDEM	*bʉ-ʉsi* AGR2-all

Ikz26b:	***bakeenderera***	*no*	*rogeendo*	***kogeenda***
SWW:	wakaendelea	na	safari	kutembea
EWW:	they continued	with	journey	to walk
MMG:	*ba-ka-enderer-a* 3PL-NAR-continue-FV	*no* CCONJ.AUG11	*ro-geendo* NC11.SG-journey	*ko-geend-a* INF-walk-FV

Ikz26c:	*hasuuhu,*[229]	***bakarumana***	*nʉ*	*ngʉhiira*	*yaazʉ,*
SWW:	kidogo	wakaonana	na	kofia	yake
EWW:	little (bit)	they found	with	hat	his
MMG:	*ha-suuhu* NC16-little	*ba-ka-rum-an-a* 3PL-NAR-find-RECP-FV	*nʉ* CCONJ.AUG9	*n-gʉhiira* NC9.SG-hat	*ya-azʉ* AGR9- 3SG.POSS

Ikz26d:	***akaba***	***atigiri,***	***bakatoora***	*ingʉhiira.*
SWW:	akawa	ameacha	wakaokota	kofia
EWW:	he was	he has left	they picked up	hat
MMG:	*a-ka-b-a* 3SG-NAR-be-FV	*a-tig-irʉ* 3SG-leave-ANT	*ba-ka-toor-a* 3PL-NAR-pick.up-FV	*ʉn-gʉhiira* NC9.SG-hat
SFT:	‘Alipoondoka kwenda, hawa nao wakaendelea na safari kutembea kidogo, wakaiona kofia yake, alikuwa ameacha, wakaokota kofia.’			
EFT:	‘When he departed, the boys also continued a little further on their way. They found the hat he had left and picked it up.’			

[229] This is a lexicalized adverb derived from a locative noun class originally (Gray 2013: 23).

Ikz27:	*~~U~~w~~u~~mw~~i~~*	***akarurya***.
SWW:	mmoja	akapiga mluzi
EWW:	one (of them)	he whistled
MMG:	*~~u~~-w~~u~~-mw~~i~~* AUG1-NC1-child	*a-ka-rury-a* 3SG-NAR-whistle-FV
SFT:	‘Mmoja wao akapiga mluzi.’	
EFT:	‘One of them whistled.’	

Ikz28a	***Aaza***	***kururya***,	*wurya*	***akiirahuta***
SWW:	alikuja	kupiga mluzi	yule	akageuka
EWW:	he came	to whistle	that (boy)	it turned
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-rury-a* INF-whistle-FV	*wu-rya* NC1.SG-DDEM	*a-ka-i-rahut-a* 3SG-NAR-REFL-fill.up-FV

Ikz28b:	*inyuma.*
SWW:	nyuma
EWW:	back
MMG:	*i-nyuma* NC9.back
SFT:	‘Alipopiga mluzi, yule akageuka nyuma.’
EFT:	‘When he whistled, the other boy turned back.’

Ikz29a:	***Aaza***	***kwiirahuta***	*inyuma,*	***akarora***
SWW:	alikuja	kugeuka	nyuma	akaona
EWW:	he came	to turn	back	he saw
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-i-rahut-a* INF-REFL-fill.up-FV	*i-nyuma* NC9-back	*a-ka-ror-a* 3SG-NAR-see-FV

Ikz29b:	*~~i~~ng~~u~~hiira*	*yaz~~i~~,*	***akiimiirira***.
SWW:	kofia	yake	akasimama
EWW:	hat	his	he stopped
MMG:	*~~i~~n-g~~u~~hiira* NC9.SG-hat	*ya-az~~i~~* AGR9-3SG.POSS	*a-ka-imi~~i~~r~~i~~r-a* 3SG-NAR-stop-FV
SFT:	‘Alipogeuka nyuma, akaona kofia yake, akasimama.’		
EFT:	‘When he turned around, he saw his hat and stopped.’		

Ikz30a:	*Umwaana*	*~~u~~wu~~u~~mw~~i~~*	***akageenda***
SWW:	mtoto	mmoja	akatembea
EWW:	child	one (of them)	he walked
MMG:	*~~u~~-mu~~u~~-ana* AUG1-NC1.SG-child	*~~u~~-wu~~u~~-mw~~i~~* AUG1-AGR1-one	*a-ka-geend-a* 3SG-NAR-walk-FV

Ikz30b:	***akamuhirira***	*~~i~~ngu~~u~~hiira.*
SWW:	akampelekea	kofia
EWW:	he took him	hat
MMG:	*a-ka-mu-hir-~~i~~r-a* 3SG-NAR-3SG.OBJ-take-APPL-FV	*~~i~~n-gu~~u~~hiira* NC9.SG-hat
SFT:	‘Mtoto mmoja akatembea, akampelekea kofia.’	
EFT:	‘One child walked to him and gave him the hat.’	

Ikz31a:	***Aaza***	***kumuhikirya***	*~~i~~ngu~~u~~hiira,*
SWW:	alikuja	kumfikishia	kofia
EWW:	he came	to give back to him	hat
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-mu-hik-ir-y-a* INF-3SG.OBJ-arrive-APPL-CAUS-FV	*~~i~~n-gu~~u~~hiira* NC9.SG-hat

Ikz31b	*wurya*	*hang’u*	***akamoha***	*mapeasi*	*atatu~~u~~.*
SWW:	yule	pia	akampa	mapeasi	tatu
EWW:	that	also	he gave him	pears	three
MMG:	*wu-rya* NC1-DDEM		*a-ka-mo-h-a* 3SG-NAR-3SG.OBJ-give-FV	*ma-peasi* NC6.PL-pear	*a-tatu~~u~~* AGR6-three
SFT:	‘Alipomfikishia kofia, yule pia akampa mapeasi matatu.’				
EFT:	‘When he gave that other boy back the hat, the other boy gave him three pears.’				

Ikz32a:	*Nawe*	*hanu~~u~~*	***baari***	***bamwaarirw~~i~~***
SWW:	lakini	ambapo	walikuwa	wamemsaidia
EWW:	but	while	they were (there)	they have helped him
MMG:		*ha-nu~~u~~* NC16- PDEM	*ba-aa-r~~i~~* 3PL-PST-COP	*ba-mu-arirw~~i~~*[230] 3PL-3SG.OBJ-help.ANT

[230] The root for ‘help’ is *arw*, which appears to be a lexicalized passive form of the root *ar* ‘spread’. The passive suffix *-w* is infixed into the ANT *-iri* when both morphemes occur together.

Ikz32b:	***kotoora,***	***ataabahiiri***	*chuchuusi,*	*naangu*
SWW:	kuokota	hakuwapa	chochote	sasa
EWW:	to pick up	he did not give him	anything	now
MMG:	*ko-toor-a* INF-pick.up-FV	*a-ta-a-ba-h-iiri*[231] 3SG-NEG-PST-3PL-give-ANT	*ch-u-ch-uusi* NC7-any-NC7-any	

Ikz32c:	***akamoha***	*amapeasi*	*atatu.*
SWW:	akampa	mapeasi	tatu
EWW:	he gave him	pears	three
MMG:	*a-ka-mo-h-a* 3SG-NAR-3SG.OBJ-give-FV	*a-ma-peasi* AUG6-NC6.PL-pear	*a-tatu* AGR6-three
SFT:	'Lakini wakati walikuwa walimsaidia kuokota, hakuwapa chochote, sasa akampa mapeasi matatu.'		
EFT:	'Though while they had been helping him pick up the pears, he didn't give them anything, now he gave them three pears.'		

Ikz33:	*Umwaana*	*wuyo*	***akagaruka***	*nagu.*
SWW:	mtoto	huyo	akarudi	nayo
EWW:	child	that (one)	he returned	with them
MMG:	*u-mu-ana* AUG1-NC1.SG-child	*wu-yo* AGR1-RDEM	*a-ka-garuk-a* 3SG-NAR-return-F	*na-gu* CCONJ-AGR6
SFT:	'Mtoto huyo akarudi nayo.'			
EFT:	'That boy went back with the pears.'			

Ikz34a:	***Aaza***	***kuhika,***	***akasoondera***
SWW:	alikuja	kufika	akagawia
EWW:	he came	to arrive	he divided
MMG:	*a-a-az-a* 3SG-PST.REL-come-FV	*ku-hik-a* INF-arrive-FV	*a-ka-soonder-a* 3SG-NAR-divide-FV

Ikz34b:	*abakiindi*	*chazi.*
SWW:	wenzake	wake
EWW:	companions	his
MMG:	*a-ba-kiindi* AUG2-NC2.PL-companion	*ch-azi* NC7-3SG.POSS
SFT:	'Alipofika, akawagawia wenzake (kitu) chake.'	
EFT:	'When he reached his companions, he gave them each a pear.'	

[231] This is an allomorph of the Anterior formative that occurs after monosyllabic roots.

Ikz35a:	***Bakataanga***	***kurya***	***baraja,***	***bakageenda***
SWW:	wakaanza	kula	wanaenda	wakatembea
EWW:	they began	to eat	they are going	they walked
MMG:	*ba-ka-taang-a* 3PL-NAR-begin-FV	*ku-ry-a* INF-eat-FV	*ba-ra-j-a* 3PL-PROG-go-FV	*ba-ka-geend-a* 3PL-NAR-walk-FV

Ikz35b:	***kuhika***	*eembere,*	***bakabona***	*~~u~~mu~~u~~karuka*
SWW:	kufika	mbele	wakakuta	mzee
EWW:	to arrive	front	they encountered	old man
MMG:	*ku-hik-a* INF-arrive-FV	*eem-bere* NC9.SG-front	*ba-ka-bon-a* 3PL-NAR-encounter-FV	~~u~~-mu~~u~~-karuka AUG1-NC1.SG-old.man

Ikz35c:	*wuyo*	*ku~~u~~mu~~u~~ti*	*harya*
SWW:	huyo	kwenye mti	pale
EWW:	that (one)	at the tree	there
MMG:	*wu-yo* AGR1-RDEM	*ku~~u~~=mu~~u~~-ti* NC17=NC3.SG-tree	*ha-rya* NC16-DDEM
SFT:	‘Wakaanza kula wanaenda, wakatembea kufika mbele, wakakuta mzee huyo kwenye mti pale.’		
EFT:	‘They ate them as them as they went on walking and came to the old man at the tree.’		

Ikz36:	*Harya*	*~~u~~mu~~u~~ti*	*wurya*	***aari.***
SWW:	pale	mti	yule	alikuwa
EWW:	there	tree	that (man)	he was
MMG:	*ha-rya* NC16-DDEM	*~~u~~-mu~~u~~-ti* AUG3-NC3.SG-tree	*wu-rya* AGR1-DDEM	*a-aa-ri* 3SG-PST-COP
SFT:	‘Alikuwepo kwenye mti ule.’			
EFT:	‘He was there at the tree.’			

Ikz37a:	*Umusubi*	*wunu*	*wuusi*	***akaba***
SWW:	mwanaume	huyu	naye	akawa
EWW:	man	this	also	he was
MMG:	*u-mu-subi* AUG1-NC1.SG-male	*wu-nu* AGR1-PDEM	*wu-usi* AGR1-all	*a-ka-b-a* 3SG-NAR-be-FV

Ikz37b:	***iikiri***	*kumuti.*
SWW:	ameshuka	kwenye mti
EWW:	he has descended	on the tree
MMG:	*a*[232]*-ik-iri* 3SG-descend-ANT	*ku=mu-ti* NC17=NC3.SG-tree
SFT:	'Huyo mzee naye, akawa ameshuka kwenye mti.'	
EFT:	'The man had come down from the tree.'	

Ikz38a:	*Naangu,*	***aaza***	***kwiika****,*
SWW:	sasa	alikuja	kushuka
EWW:	now	he came	to descend
MMG:		*a-a-az-a* 3SG-PST.REL-come-FV	*ku-ik-a* INF-descend-FV

Ikz38b:	***akataanga***	*kumagamaga.*
SWW:	akaanza	kuangaza
EWW:	he began	to stare
MMG:	*a-ka-taang-a* 3SG-NAR-begin-FV	*ku-maga~mag-a* INF-stare~RDPL-FV
SFT:	'Sasa, aliposhuka, akaanza kuangaza.'	
EFT:	'Now, when he came down, he began to stare.'	

Ikz39a:	*"Rihuuzu*	*ryaani*	***rerehe****,"*	*kina*
SWW:	kikapu (cha asili)	changu	kiko wapi	kama
EWW:	(traditional) basket	my	it is where	like
MMG:	*ri-huuzu* NC5.SG-basket	*ri-ani* AGR5-1SG.POSS	*re-re=he* SBJ5-COP=where	

Ikz39b:	***akwiibuurya****.*
SWW:	anajiuliza
EWW:	he is asking himself.
MMG:	*a-ku-i-buury-a* 3SG-NPST-REFL-ask-FV
SFT:	"Kikapu changu kiko wapi," kama anajiuliza.'
EFT:	"Where did my basket go," is what he asks himself.'

[232] The *a-* 3rd person singular marker surfaces as an *i-* here because it assimilates to the root vowel (Gray & Smith 2010a: 19).

Ikz40:	*Nawe*	*rihʉʉzʉ*	***ritarʉrikini*.**
SWW:	lakini	kikapu (cha asili)	hakikuonekana
EWW:	but	the (traditional) basket	it was not visible
MMG:		*ri-hʉʉzʉ* NC5.SG-basket	*rɨ-ta-rʉrɨkɨnɨ* SBJ5-NEG-be.visible.ANT
SFT:	'Lakini kikapu hakikuonekana.'		
EFT:	'But the basket was nowhere to be found.'		

Ikz41a:	*Abaana*	*bayo*	***bakahɨta***	***bararya***
SWW:	watoto	hao	wakapita	wanakula
EWW:	children	those	they passed by	they are eating
MMG:	*a-ba-ana* AUG2-NC2.PL-child	*ba-yo* AGR2-RDEM	*ba-ka-hɨt-a* 3PL-NAR-pass.by-FV	*ba-ra-ry-a* 3PL-PROG-eat-FV

Ikz41b:	*amapeasi*	*gaabʉ*	***baraja*.**
SWW:	mapeasi	yao	wanaenda
EWW:	pears	their	they are going
MMG:	*a-ma-peasi* AUG6-NC6.PL-	*ga-abʉ* AGR6-3PL.POSS	*ba-ra-j-a* 3PL-PROG-go-FV
SFT:	'Watoto hao wakapita wanakula mapeasi yao.'		
EFT:	'Those boys passed by eating their pears.'		

Ikz42a:	*Ʉmʉkaruka*	*wuyo*	***akatanga***	***kʉbarora***
SWW:	mzee	huyo	akaanza	kuwaona
EWW:	old man	that	he began	to see them
MMG:	*ʉ-mʉ-karuka* AUG1-NC1.SG-old.man	*wu-yo* AGR1-RDEM	*a-ka-tang-a* 3SG-NAR-begin-FV	*kʉ-ba-ror-a* INF-3PL.OBJ-see-FV

Ikz42b:	***akarʉgʉʉrabu***	*chiɨmbu*	***bakurya***
SWW:	akashangaa tu	jinsi	wakila
EWW:	he was simply surprised	how	they eat
MMG:	*a-ka-rʉgʉʉr-a=bu* 3SG-NAR-be.surprised-FV=only		*ba-ku-ry-a* 3SG-NPST-eat-FV

Ikz42c:	*amapeasi.*
SWW:	mapeasi
EWW:	pears
MMG:	*a-ma-peasi* AUG6-NC6.PL-pear
SFT:	'Na mzee huyo akaanza kuwaona, akawashangaa tu jinsi wanavyokula mapeasi.'
EFT:	'And the old man saw them and was simply shocked that they were eating the pears.'

Ikz43:	*Rigan~~u~~*	***~~ri~~kari~~i~~nga***	*hayo.*
SWW:	hadithi	ikakomea	hapo
EWW:	story	it finished	there
MMG:	*ri-gan~~u~~* NC5.SG-story	*~~ri~~-ka-~~ri~~ng-a* SBJ5-NAR-finish-FV	*ha-yo* NC16-RDEM
SFT:	'Hadithi ikakomea hapo.'		
EFT:	'The end.'		

Appendix C: "Pear Story" in Ikoma

Since the original text of the "Pear Story" in Ikoma was translated first into Swahili, rather than directly into English, the following transcription of the story is a six-row interlinear text which includes word-for-word and free translations into both Swahili and English, as well as a morpheme-by-morpheme English gloss. The first row of each interlinearized section has the Ikoma text of the story and is titled Ntk, followed by a sentence and line number (i.e. 1a, 1b, 2, etc.). These designations are used in the body of the thesis to refer to specific portions of this Ikoma text. The abbreviations for the other rows, in the order they appear, are: SWW (Swahili Word-for-Word), EWW (English Word-for-Word), MMG[233] (Morpheme-by-Morpheme Gloss), SFT (Swahili Free Translation) and EFT (English Free Translation). The latter two are included only on the last line of a sentence. All verbal words are bolded in the Ikoma text.

Ikoma Pear Story

Ntk1a:	*Ibheere,*	*orosiko*	*rumu*	*atabhoori,*
SWW:	sasa	siku	moja	asubuhi
EWW:	now	day	one	morning
MMG:		*o-ro-siko* AUG11-NC11.SG-day	*ru-mu* AGR11-one	*a-tabhoori* NC9.SG-morning

Ntk1b:	*omooto*	*umu*	***akaru***	*awaache,*
SWW:	mtu	mmoja	akatoka	kwake
EWW:	person	one	he left	his (place)
MMG:	*o-mo-oto* AUG1-NC1.SG-person	*u-mu* AGR1-one	*a-aka-ru*[234] 3SG-NAR-leave	*awa-ache* NC23?[235]-3SG.POSS

[233] The MMG shows the underlying forms of the morphemes.
[234] There is a rule in Ikoma that prohibts long vowels word-initially, so this word surfaces as *akaru* instead of *aakaru* (Higgins 2011: 81). There are many instances of this rule being applied throughout the remainder of this text.
[235] In nearby Ikizu, "home base" semantics are found in NC23 and possessive pronominal agreement is indicated by the *wa-* prefix (Gray 2013: 33). More research is needed to establish a specific noun class for this form and the similar form found in line Ntk23d.

Ntk1c:	***aghaye***	*mo-mote*	*umu*	*o*
SWW:	akaenda	mtini	mmoja	wa
EWW:	he went	to the tree	one	of
MMG:	*a-agha-ye* 3SG-NAR-go	*mo=mo-te* NC18=NC3.SG-tree	*u-mu* AGR3-one	*o* AGR3.ASC

Ntk1d:	*manyëmbë gha maghuta*[236]	***ghuye***
SWW:	maembe ya mafuta	kwenda
EWW:	mangos of oil (i.e. pears)	going
MMG:	*ma-nyëmbë gha ma-ghuta* NC6.PL-mango AGR6.ASC NC6-oil	*ghu-ye* INF-go

Ntk1e:	***ghotu***	*amanyëmbë*	*ghayo.*
SWW:	kuchuma	maembe (ya mafuta)	hayo
EWW:	to pick	pears	those
MMG:	*gho-tu* INF-pick	*a-ma-nyëmbë* AUG6-NC6.PL-pear	*gha-yo* AGR6-RDEM
SFT:	'Siku moja asubuhi, mtu mmoja akatoka kwake, akaenda mtini kuchuma maembe (ya mafuta).'		
EFT:	'One morning, a certain man left his house and went to a tree to pick pears.'		

Ntk2a:	*Omooto*	*uyo*	***akabha***	***asambekiri***
SWW:	mtu	huyo	akawa	amebeba begani
EWW:	person	that	he was	he has put on shoulders
MMG:	*o-mo-oto* AUG1-NC1.SG-person	*u-yo* AGR1-RDEM	*a-aka-bh-a* 3SG-NAR-be-FV	*a-sambek-iri* 3SG-carry.on.shoulder-ANT

Ntk2b:	*oghosuku*	*ghokoro,*	***akiimoki***	*ne*
SWW:	mfuko	mkubwa	akachukua	na
EWW:	bag	big	he took	with
MMG:	*o-gho-suku* AUG20-NC20.SG-bag	*gho-koro* AGR20-big	*a-aka*[237]*-imoki* 3SG-NAR-take	*ne*[238] CCONJ.AUG8

[236] Here *manyembe ga maguta* 'pears' identifies the fruit, but following references abbreviate to *manyembe* 'mangos', which in context is understood to still be referring to 'pears'.

[237] In Ikoma, when an *a* before a morpheme break adjoins with an *i* in the following morpheme, the *a* assimilates to the *i* and they become a long *ii* (Higgins 2010: 21). This also occurs on other morphemes in this text (i.e. *ra* becomes *ri* before *i* in line Ntk14b).

[238] The final vowel in the conjunction *na* elides in speech when in front of a vowel initial word. This form is written in the orthography with the first vowel of the following word, both here and other places in the text (Higgins 2010: 29).

Ntk2c:	*ebhetonga*	*bhyaru,*	***aghaye***	*nabhyo.*
SWW:	vikapu	vingi	akaenda	navyo
EWW:	baskets	many	he went	with them
MMG:	*e-bhe-tonga* AUG8-NC8.PL-basket	*bhi*[239]*-aru* AGR8-many	*a-agha-ye* 3SG-NAR-go	*na-bhyo* CCONJ-NC8
SFT:	'Mtu huyo akawa amebeba begani mfuko mkubwa, akachukua vikapu vingi, akaenda navyo.'			
EFT:	'That man had carried a big bag on his shoulders with several baskets in it.'			

Ntk3a:	*Hano*	***ahikiri,***	***naare***	*na*
SWW:	hapa	amefika	alikuwa	na
EWW:	here	he has arrived	he was	with
MMG:	*ha-no* NC16-PDEM.REL	*a-hik-iri* 3SG-arrive-ANT	*n*[240]*-a-a-re* FOC-3SG-PST-PST.COP	*na* CCONJ

Ntk3b:	*angachi*	***arayetoora,***	***aratiira***
SWW:	ngazi	anaiweka	anapanda
EWW:	ladder	he is putting it	he is climbing
MMG:	*an-gachi* NC9.SG-ladder	*a-ra-ye-toor-a* 3SG-PROG-OBJ9-climb-FV	*a-ra-tiir-a* 3SG-PST-climb-FV

Ntk3c:	*amote,*	***arasëëma***	***ghotu***
SWW:	kwenye mti	anaanza	kuchuma
EWW:	on the tree	he is beginning	to pick
MMG:	*a*[241]*=mo-te* NC16=NC3.SG-tree	*a-ra-sëëm-a* 3SG-PROG-begin-FV	*gho-tu* INF-pick

[239] The *i* in the prefix is palatalized before a root-initial vowel, both here and other places in the text (Higgins 2010: 20-21).
[240] The Focus occurs here in a main/independent clause, but not on the related relative/dependent clause form *are* in Ntk5a. This distinction is explained more in Section 5.5.
[241] This appears to be a "remnant of noun class 16" that has been reduced from *ha* (Holly Robinson: *pc*).

Ntk3d:	*amanyëmbë*	*ghayo.*
SWW:	maembe (ya mafuta)	hayo
EWW:	pears	those
MMG:	*a-ma-nyëmbë* AUG6-NC6.PL-pear	*gha-yo* AGR6-RDEM
SFT:	'Alipofika, alikuwa na ngazi, akaiweka, akapanda mti, akaanza kuchuma maembe (ya mafuta) hayo.'	
EFT:	'When he arrived, he had a ladder, so he placed it, climbed the tree and began to pick pears.'	

Ntk4a:	***Aratu,***	***aratu,***	***aratu,***	***araacha,***
SWW:	anachuma	anachuma	anachuma	anakuja
EWW:	he is picking	he is picking	he is picking	he is coming
MMG:	*a-ra-tu* 3SG-PROG-pick	*a-ra-tu* 3SG-PROG-pick	*a-ra-tu* 3SG-PROG-pick	*a-ra-ach-a* 3SG-PROG-come-FV

Ntk4b:	***ariichori***	*mo-bhetonga.*
SWW:	anajaza	katika vikapu
EWW:	he is filling	in the baskets
MMG:	*a-ra-ichor-i* 3SG-PROG-fill-CAUS	*mo=bhe-tonga* NC18=NC8.PL-basket
SFT:	'Akachuma, akachuma, akachuma. Anakuja kujaza vikapu vyake.'	
EFT:	'He picked [and] picked [and] picked [and then] went to fill his baskets.'	

Ntk5a:	*Hano*	*wiiki*	***are***	***akeere***
SWW:	hapa	tena	alikuwa	bado
EWW:	here	again	he was	still
MMG:	*ha-no* NC16-PDEM.REL		*a-a-re*[242] 3SG-PST-PST.COP	*a-kee-re* 3SG-PER-PST.COP

Ntk5b:	***areenderera***	***ghotu,***	***araacha***	*omooto*
SWW:	anaendelea	kuchuma	anakuja	mtu
EWW:	he is continuing	to pick	he is coming	person
MMG:	*a-ra-enderer-a* 3SG-PROG-continue-FV	*gho-tu* INF-pick	*a-ra-ach-a* 3SG-PROG-come-FV	*o-mo-oto* AUG1-NC1.SG-person

[242] This is another instance of the word-initial long vowel prohibition in Ikoma that was remarked on in the first footnote in line Ntk2b. This prohibition makes the relative clause form *are* even more distinct from the main clause form *naare* in line Ntk3a.

Ntk5c:	***areenderera***	***anyi***	*na*	*ambori.*
SWW:	anaendelea	akiwa	na	mbuzi
EWW:	he is continuing	he is	with	goat
MMG:	*a-ra-enderer-a* 3SG-PROG-continue-FV	*a-nyi* 3SG-PRS.COP	*na* CCONJ	*am-bori* NC9.SG-goat
SFT:	'Aliporudi tena kuendelea kuchuma, mtu mwingine akaja na mbuzi.'			
EFT:	'As he continued to pick fruit, another person came by with a goat.'			

Ntk6a:	*Ibheere*	*hano*	***bhahikiri***	*hayo*	***bhating'iri,***
SWW:	sasa	hapa	wamefika	hapo	wamekaribia
EWW:	now	here	they have	this place	they have come close
MMG:		*ha-no* NC16-PDEM.REL	*bha-hik-iri* 3PL-arrive-ANT	*ha-yo* NC16-RDEM	*bha-ting'-iri* 3PL-approach-ANT

Ntk6b:	*ambori*	*iyo*	***eriighomba***	*kure*
SWW:	mbuzi	hiyo	anatamani	mbali
EWW:	goat	that	it is wanting	far
MMG:	*am-bori* NC9.SG-goat	*i-yo* AGR9-RDEM	*e-ra-ighomb-a* SBJ9-PROG-want-FV	*ku-re* NC17-DDEM

Ntk6c:	*amanyëmbë*	*ghayo,*	*amare*	*omwene*
SWW:	maembe (ya mafuta)	hayo	lakini	mwenye
EWW:	pears	those	but	owner
MMG:	*a-ma-nyëmbë* AUG6-NC6.PL-pear	*gha-yo* AGR6-RDEM		*o-mu*[243]*-ene* AUG1-NC1.SG-owner

Ntk6d:	*ambori*	***arayighööta***	***arayeruta***
SWW:	mbuzi	anaishika	anaivuta
EWW:	goat	he is grabbing it	he is pulling it
MMG:	*am-bori* NC9.SG-goat	*a-ra-yi-ghööt-a* 3SG-PROG-OBJ9-grab-FV	*a-ra-ye-rut-a* 3SG-PROG-OBJ9-grab-FV

[243] The *u* in the prefix is labialized before a root-initial vowel, both here and other places in the text (Higgins 2010: 20-21).

Ntk6e:	***araye***	***koyibhöha***	*yiire.*
SWW:	anaenda	kuifunga	kule
EWW:	he is going	to tie it up	there
MMG:	*a-ra-ye* 3SG-PROG-go	*ko-yi-bhöh-a* INF-OBJ9-tie.up-FV	*yii-re* NC23[244]-DDEM
SFT:	'Walipokaribia mtini, mbuzi akatamani kula maembe (ya mafuta) hayo, lakini mwenye mbuzi akamshika, akamvuta, akaenda kumfunga kwake.'		
EFT:	'When they got close to the tree, the goat wanted to eat some pears, but the owner took hold of it, pulled it away, and led it off to tie it up at home.'		

Ntk7a:	*Omugharuka*	*uyo*	*we*
SWW:	mzee	huyo	yeye
EWW:	old man	that	he
MMG:	*o-mu-gharuka* AUG1-NC1.SG-old.man	*u-yo* AGR1-RDEM	*we* 3SG.PRON

Ntk7b:	***areenderera***	***ghotu***	*amanyëmbë.*
SWW:	anaendelea	kuchuma	maembe
EWW:	he is continuing	to pick	pears
MMG:	*a-ra-enderer-a* 3SG-PROG-continue-FV	*gho-tu* INF-pick	*a-ma-nyëmbë* AUG6-NC6.PL-pear
SFT:	'Mzee huyo anaendelea kuchuma maembe (ya mafuta).'		
EFT:	'The old man just continued picking pears.'		

Ntk8a:	*Mbe,*	***aratu***	*amanyëmbë,*	*antenero*
SWW:	basi	anachuma	maembe (ya mafuta)	hatimaye
EWW:	so	he is picking	pears	at last
MMG:		*a-ra-tu* 3SG-PROG-pick	*a-ma-nyëmbë* AUG6-NC6.PL-pear	

[244] The form of this marker is Noun Class 23, which has semantics referring to the "home base" in nearby Ikizu (Gray 2013: 16, 34). In Ikoma, however, my language consultant said that "home base" is not implied.

Ntk8b:	*omwereeti*	*woose*	***araacha***.
SWW:	mwanaye	naye	anakuja
EWW:	his son	also	he is coming
MMG:	*o-mu-ereeti* AUG1-NC1.SG-son	*wo-ose* AGR1-all	*a-ra-ach-a* 3SG-PROG-come-FV
SFT:	'Mbe, akachuma maembe (ya mafuta), hatimaye mwanaye naye akaja.'		
EFT:	'So, he picked pears, and eventually his son came along.'		

Ntk9a:	***Araacha***	***anyi***	*na*	*abhayisikeeri,*
SWW:	anakuja	akiwa	na	baiskeli
EWW:	he is coming	he is	with	bicycle
MMG:	*a-ra-ach-a* 3SG-PROG-come-FV	*a-nyi* 3SG-PRS.COP	*na* CCONJ	*a-bhayisikeeri* NC9.SG-bicycle

Ntk9b:	***araacha***	***ghwichera***	*amanyëmbë*	*ghayo.*
SWW:	anakuja	kufuata	maembe (ya mafuta)	hayo
EWW:	he is coming	to follow	pears	those
MMG:	*a-ra-ach-a* 3SG-PROG-come-FV	*ghu-icher-a* INF-follow-FV	*a-ma-nyëmbë* AUG6-NC6.PL-pear	*gha-yo* AGR6-RDEM
SFT:	'Akaja na baiskeli akitafuta maembe (ya mafuta) hayo.'			
EFT:	'He came on a bike in search of those pears.'			

Ntk10a:	*Hano*	***ahikiri,***	***aragharamereri***
SWW:	hapa	amefika	anaangalia
EWW:	here	he has arrived	he is looking
MMG:	*ha-no* NC16-PDEM.REL	*a-hik-iri* 3SG-arrive-ANT	*a-ra-gharamereri* 3SG-PROG-look

Ntk10b:	*ighoro*	***arakeeri***	*ishe.*
SWW:	juu	anasalimia	baba yake
EWW:	up	he is greeting	his father
MMG:	*i-ghoro* NC23-up	*a-ra-keeri* 3SG-PROG-greet	*Ø-ishe* NC1.SG-father
SFT:	'Alipofika, akaangalia juu, akasalimia baba yake.'		
EFT:	'When he arrived, he looked up and greeted his father.'		

Ntk11a:	*Hano*	***amariri***	***komukeeri,***	***ariimoki***
SWW:	hapa	amemaliza	kumsalimia	anachukua
EWW:	here	he has finished	to greet him	he is taking
MMG:	*ha-no* NC16-PDEM.REL	*a-mar-iri* 3SG-finish-ANT	*ko-mu-keeri* INF-3SG.OBJ-greet	*a-ra-imoki* 3SG-PROG-take

Ntk11b:	*eghetonga*	*kimu,*	***araghetiiri***	*mo-bhayisikeeri*
SWW:	kikapu	kimoja	anakiweka	kwenye baiskeli
EWW:	basket	one	he is putting it	on bicycle
MMG:	*e-ghe-tonga* AUG7-NC7.SG-basket	*ki-mu* AGR7-one	*a-ra-ghe-tiiri* 3SG-PROG-OBJ7-put	*mo=Ø-bhayisikeeri* NC18=NC9.SG-bicycle

Ntk11c:	*yaache,*	***arasëëma***	*orughëndö*	*ro*
SWW:	yake	anaanza	safari	ya
EWW:	his	he is starting	journey	of
MMG:	*ya-ache* AGR9-3SG.POSS	*a-ra-sëëm-a* 3SG-PROG-begin-FV	*o-ru-ghëndö* AUG11-NC11-journey	*ro* NC11.ASC

Ntk11d:	***ghugharuka***	*aka.*
SWW:	kurudi	nyumbani
EWW:	to return	to home
MMG:	*ghu-gharuk-a* INF-return-FV	*a-aka*[245] NC9.SG-home
SFT:	‘Alipomaliza kumsalimia baba, akachukua kikapu kimoja, akakiweka kwenye baiskeli yake, akaanza safari ya kurudi nyumbani.’	
EFT:	‘When he had greeted his father, he took a basket, put it on his bike, and began biking back home.’	

Ntk12a:	*Ibheere*	*hano*	***are***	*mo-nchera,*
SWW:	sasa	hapa	alikuwa	njiani
EWW:	now	here	he was	on the way
MMG:		*ha-no* NC16-PDEM.REL	*a-a-re* 3SG-PST-PST.COP	*mo=n-chera* NC18=NC9.SG-way

[245] This word is typically found in NC23 in the Mara languages (i.e. *iika* in Kabwa; *yɨika* in Ikizu). In Ikoma, it appears to be in NC9 and even takes a regular NC10 plural *chaaka.*

Ntk12b:	***aratëmëërana***	*no*	*omoghiikoro*	*umu.*
SWW:	anakutana	na	mwanamke	mmoja
EWW:	he is meeting	with	woman	one
MMG:	*a-ra-tëmëër-an-a* 3SG-PROG-meet-RECP-FV	*no* CCONJ.AUG1	*o-mo-ghiikoro* AUG1-NC1.SG-woman	*u-mu* AGR1-one
SFT:	'Sasa alipokuwepo njiani, akakutana na mwanamke mmoja.'			
EFT:	'Now, as he went down the path, he came upon a girl.'			

Ntk13a:	*Woose*	***naare***	*na*	*abhayisikeeri,*	*amare*
SWW:	naye	alikuwa	na	baiskeli	lakini
EWW:	also	she was	with	bicycle	but
MMG:	*wo-ose* AGR1-all	*n-a-a-re* FOC-3SG-PST-PST.COP	*na* CCONJ	*a-bhayisikeeri* NC9.SG-bicycle	

Ntk13b:	*nawe*	***akabha***	***araacha***
SWW:	naye	akawa	anakuja
EWW:	she also	she was	she is coming
MMG:	*na-we* CCONJ-3SG.PRON	*a-aka-bh-a* 3SG-NAR-be-FV	*a-ra-ach-a* 3SG-PROG-come-FV

Ntk13c:	***naakohombahomba***	*mo-nchera.*
SWW:	akiyumbayumba	njiani
EWW:	she is swerving	on the way
MMG:	*n-a-ako-homba~homb-a* FOC-3SG-SMLT-swerve~RDPL-FV	*mo=n-chera* NC18=NC9.SG-way
SFT:	'Naye akawa na baiskeli, lakini akaja akiyumbayumba.'	
EFT:	'She was also on a bicycle, but she was swerving as she came towards him.'	

Ntk14a:	*Hano*	***amariri***	***komoheteera,***
SWW:	hapa	amemaliza	kumpita
EWW:	here	she has finished	to pass by him
MMG:	*ha-no* NC16-PDEM.REL	*a-mar-iri* 3SG-finish-ANT	*ko-mo-hete-er-a* INF-3SG.OBJ-pass.by-APPL-FV

Ntk14b:	*omoona*	*uyo*	***ariishoorori,***
SWW:	mtoto	huyo	akageuka
EWW:	child	that	he is turning
MMG:	*o-mo-ona* AUG1-NC1.SG-child	*u-yo* AGR1-RDEM	*a-ra-i-shooror-i* 3SG-PROG-REFL-turn-CAUS

Ntk14c:	***ariikara***	***komumööhi***	*abhayisikeeri,*	***araye***
SWW:	anakaa	kumtazama	baiskeli	anaenda
EWW:	he is sitting	to gaze	bicycle	he is going
MMG:	*a-ra-ikar-a* 3SG-PROG-sit-FV	*ko-mu-möömi* INF-3SG.OBJ-gaze	*a-bhayisikeeri* NC9.SG-bicycle	*a-ra-ye* 3SG-PROG-go

Ntk14d:	***arataara***	*riibhohe.*
SWW:	anagonga	jiwe
EWW:	he is hitting	rock
MMG:	*a-ra-taar-a* 3SG-PROG-hit-FV	*rii-bhohe* NC5-rock
SFT:	‘Alipomaliza kumpita, mtoto huyo akageuka kumtazama, baiskeli ikagonga jiwe.’	
EFT:	‘When she had already passed by, the boy turned to gaze at her and the bicycle hit a rock.’	

Ntk15a:	*Hano*	***itaariri***	*riibhohe,*	*omoona*
SWW:	hapa	imegonga	jiwe	mtoto
EWW:	here	it has hit	rock	child
MMG:	*ha-no* NC16-PDEM.REL	*i-taar-iri* SBJ9-hit-ANT	*rii-bhohe* NC5.SG-rock	*o-mo-ona* AUG1-NC1.SG-child

Ntk15b:	*uyo*	***aragho***	*haase*	*na*	*amanyëmbë*
SWW:	huyo	anaanguka	chini	na	maembe (ya mafuta)
EWW:	that	he is falling	down	and	pears
MMG:	*u-yo* AGR1-RDEM	*a-ra-gho* 3SG-PROG-fall	*ha-ase* NC16-down	*na* CCONJ	*a-ma-nyëmbë* AUG6-NC6.PL-pear

Ntk15c:	*ghayo*	***gharahomoka,***	***gharasancha***
SWW:	hayo	yaramwagika	yakasambaa
EWW:	those	they are spilling	they are spreading out
MMG:	*gha-yo* AGR6-RDEM	*gha-ra-homok-a* SBJ6-PROG-spill-FV	*gha-ra-sanch-a* SBJ6-PROG-spread.out-FV

Ntk15d:	*haase*	*hayo.*
SWW:	chini	hapo
EWW:	down	there
MMG:	*ha-ase* NC16-down	*ha-yo* NC16-RDEM
SFT:	'Baiskeli ilipogonga jiwe, mtoto huyo akaanguka chini na maembe (ya mafuta) yakamwagika, yakasambaa njiani.'	
EFT:	'When the bicycle hit the rock, the boy fell down and the pears spilled out of the basket all over the path.'	

Ntk16a:	*Amare*	*bhaabha okumangi*[246]	***mbaareho***
SWW:	lakini	bahati nzuri	walikuwepo
EWW:	but	good luck	they were there
MMG:			*m-ba-a-re=ho* FOC-3PL-PST-PST.COP=LOC16

Ntk16b:	*abhaato*	*bhatato.*
SWW:	watu	watatu
EWW:	people	three
MMG:	*a-bha-ato* AUG2-NC2.PL-person	*bha-tato* AGR2-three
SFT:	'Lakini, kwa bahati nzuri, walikuwepo watu watatu.'	
EFT:	'But fortunately, there were three people there.'	

Ntk17:	***Bhaare***	*ambere*	*yaache.*
SWW:	walikuwa	mbele	yake
EWW:	they were (there)	front	his
MMG:	*bha-a-re* 3PL-PST-PST.COP	*am-bere* NC9.SG-front	*ya-ache* AGR9-3SG.POSS
SFT:	'Walikuwa mbele yake.'		
EFT:	'They were there in front of him.'		

Ntk18a:	*Hano*	***bhahikiri,***	***bharaacha,***	***bharamotoori,***
SWW:	hapa	wamefika	wanakuja	wanamsaidia
EWW:	here	they have	they are coming	they are helping him
MMG:	*ha-no* NC16-PDEM.REL	*bha-hik-iri* 3PL-arrive-ANT	*bha-ra-ach-a* 3PL-PROG-come-FV	*bha-ra-mo-toori* 3PL-PROG-3SG.OBJ-help

[246] This phrase is idiomatic and is difficult to parse.

Ntk18b:	***bharamughööta***	***bharamwimeereri,***
SWW:	wanamshika	wanamsimamisha
EWW:	they are taking hold of him	they are standing him up
MMG:	*bha-ra-mu-ghööt-a* 3PL-PROG-3SG.OBJ-take.hold-FV	*bha-ra-mu-imeerer-i* 3PL-PROG-3SG.OBJ-stand-CAUS

Ntk18c:	***bharamotoori***	***ghughëgha***	*amanyëmbë*	*ghayo*
SWW:	wanamsaidia	kubeba	maembe (ya mafuta)	hayo
EWW:	they are helping him	to carry	pears	those
MMG:	*bha-ra-mo-toori* 3PL-PROG-3SG.OBJ-help	*ghu-ghëgh-a* INF-carry-FV	*a-ma-nyëmbë* AUG6-NC6.PL-pear	*gha-yo* AGR6-RDEM

Ntk18d:	*na*	***ghoghatoora***	*mo-ghetonga,*	***bharamusakeri***
SWW:	na	kuyaweka	kwenye kikapu	wanamsaidia
EWW:	and	to place them	in the basket	they are helping him
MMG:	*na* CCONJ	*gho-gha-toor-a* INF-OBJ6-carry-FV	*mo=ghe-tonga* NC18=NC7.SG-basket	*bha-ra-mu-sakeri* 3PL-PROG-3SG.OBJ-help

Ntk18e:	***ghwikari***	*mo-bhayisikeeri.*
SWW:	kukaliza	kwenye baiskeli
EWW:	to make sit down	on the bicycle
MMG:	*ghu-ikar-i* INF-sit-CAUS	*mo=Ø-bhayisikeeri* NC18=NC9.SG-bicycle
SFT:	‘Walipofika, wakamshika, wakamsimamisha, wakamsaidia kubeba maembe (ya mafuta) hayo na kuyaweka kwenye kikapu, wakamkaliza kwenye baiskeli.’	
EFT:	‘When they arrived, they picked him up, helped him carry the pears and put them in the basket, and sat him down on his bike.’	

Ntk19a:	*Ibheere*	*omobhureni*	*uyo*	***araheta,***
SWW:	sasa	kijana	huyo	anapita
EWW:	now	young male	this	he is passing by
MMG:		*o-mo-bhureni* AUG1-NC1.SG-young.male	*u-yo* AGR1-RDEM	*a-ra-het-a* 3SG-PROG-pass.by-FV

Ntk19b:	***araghe***	*aka,*	*amare*	*eriibhagha*	*riyo*
SWW:	araenda	nyumbani	lakini	wakati	huo
EWW:	he is going	home	but	time (period)	that
MMG:	*a-ra-ghe* 3SG-PROG-go	*a-aka* NC9.SG-home		*e-rii-bhagha* AUG5-NC5-time.period	*ri-yo* AGR5-RDEM

Ntk19c:	***naare***	***anyahaarekiri***	*oghoghoro,*
SWW:	alikuwa	ameumia	mguu
EWW:	she was	he has injured	foot
MMG:	*n-a-a-re* FOC-3SG-PST-PST.COP	*a-nyahaarek-iri* 3SG-be.injured-ANT	*o-gho-ghoro* AUG15-NC15-foot

Ntk19d:	***akabha***	***araghe***	***naaghosoghora.***
SWW:	akawa	anaenda	anachechemea
EWW:	he was	he is going	he is limping
MMG:	*a-aka-bh-a* 3SG-NAR-be-FV	*a-ra-ghe* 3SG-PROG-go	*n-a-aghosoghor-a* FOC-3SG-SMLT-limp-FV
SFT:	'Sasa, kijana huyo akapita kwenda nyumbani, lakini wakati huo alikuwa ameumia mguu, akaenda akichechemea.'		
EFT:	'Now, the boy continued home, but he had injured his foot so he was limping.'		

Ntk20a:	*Hano*	***ishukiri***	*hasuuhu,*	***bharatööra***
SWW:	hapa	amesogea	kidogo	wanaokota
EWW:	here	he has moved	little	they are finding
MMG:	*ha-no* NC16-PDEM	*a*[247]*-ishuk-iri* 3SG-move-ANT	*ha-suuhu* NC16-little	*bha-ra-töör-a* 3PL-PROG-find-FV

Ntk20b:	*angösiira*	*abhabhureni*	*bhare,*
SWW:	kofia	vijana	wale
EWW:	hat	young males	those
MMG:	*an-gösiira* NC9.SG-hat	*a-bha-bhureni* AUG2-NC2.PL-young.male	*bha-re* AGR2-DDEM

Ntk20c:	***bharamutëmëra***	*omorori,*	*omobhureni*
SWW:	wanampigia	mluzi	kijana
EWW:	they are hitting for him	whistle	young male
MMG:	*bha-ra-mu-tëm-ër-a* 3PL-PROG-3SG.OBJ-hit-APPL-FV	*o-mo-rori* AUG3-NC3.SG-whistle	*o-mo-bhureni* AUG1-NC1.SG-young.male

Ntk20d:	*uyo*	***ariimeerera,***	***aramohirera***
SWW:	huyo	anasimama	anampelekea
EWW:	that	he is stopping	he is taking to him
MMG:	*u-yo* AGR1-RDEM	*a-ra-imeerer-a* 3SG-PROG-stop-FV	*a-ra-mo-hir-er-a* 3SG-PROG-3SG.OBJ-take-APPL-FV

[247] The 3SG morpheme does not surface here because the *a* assimilates to the *i* and then the long vowel only surfaces as a short vowel since it is in word-initial position. See note on line Ntk1b for more on this.

Ntk20e:	*angösiira*	*iyo.*
SWW:	kofia	hiyo
EWW:	hat	that
MMG:	*an-gösiira* NC9.SG-hat	*i-yo* AGR9-RDEM
SFT:	'Aliposogea kidogo, vijana wale wengine wakaokota kofia yake, wakampigia mluzi, akasimama, na mmojawapo akampelekea kofia.'	
EFT:	'When he had gone a little ways, the three boys found his hat and whistled to him. Then he stopped and one of them brought it to him.'	

Ntk21a:	*Hano*	***ahikiri,***	***aramusabha***
SWW:	hapa	amefika	anamwomba
EWW:	here	he has arrived	he is asking him
MMG:	*ha-no* NC16-PDEM.REL	*a-hik-iri* 3SG-arrive-ANT	*a-ra-mu-sabh-a* 3SG-PROG-3SG.OBJ-ask-FV

Ntk21b:	***amuhë***	*amanyëmbë*	*ghayo*	*gha*
SWW:	ampe	maembe	hayo	ya
EWW:	he should give him	pears	those	of
MMG:	*a-mu-h-ë* 3SG-3SG.OBJ-give-SBJV	*a-ma-nyëmbë* AUG6-NC6.PL-pear	*gha-yo* AGR6-RDEM	*gha* AGR6.ASC

Ntk21c:	*amaghuta,*	*omobhureni*	***aramuha***
SWW:	mafuta	kijana	anampa
EWW:	oil	young male	he is giving him
MMG:	*a-ma-ghuta* AUG6-NC6-oil	*o-mo-bhureni* AUG1-NC1.SG-young.male	*a-ra-mu-h-a* 3SG-PROG-3SG.OBJ-give-FV

Ntk21d:	*amanyëmbë*	*atato.*
SWW:	maembe (ya mafuta)	matatu
EWW:	pears	three
MMG:	*a-ma-nyëmbë* AUG6-NC6.PL-pear	*a-tato* AGR6-three
SFT:	'Alipofika (na kofia), akamwomba ampe maembe ya mafuta. Kijana huyo akampa maembe (ya mafuta) matatu.'	
EFT:	'When the boy (with the hat) arrived, he asked for some pears. The other boy gave him three pears.'	

Ntk22a	*Hano*	***ahikiri,***	***arahööhi***	*kira*	*mooto*
SWW:	hapa	amefika	anagawia	kila	mtu
EWW:	here	he has arrived	he is dividing up	each	person
MMG:	*ha-no* NC16-PDEM.REL	*a-hik-iri* 3SG-arrive-ANT	*a-ra-höö̈hi* 3SG-PROG-divide		*mo-oto* NC1.SG-person

Ntk22b:	*erinyëmbë*	*kerimu*	*rimu.*
SWW:	embe (la mafuta)	moja	moja
EWW:	pear	one	one
MMG:	*e-ri-nyëmbë* AUG5-NC5.SG-pear	*ke=ri-mu* NC7[248]=AGR5-one	*ri-mu* AGR5-one
SFT:	‘Aliporudi kwa wenzake, akagawia kila mtu embe (la mafuta) moja moja.’		
EFT:	‘When he returned to his friends, he handed out one pear to each person.’		

Ntk23a:	*Ibheere*	*mooto*	*woose*	***aghatanura,***
SWW:	sasa	mtu	wote	akaondoka
EWW:	now	person	all	he left
MMG:		*mo-oto* NC1.SG-person	*wo-ose* AGR1-all	*a-agha-tanur-a* 3SG-NAR-left-FV

Ntk23b:	*abhabhureni*	*bhayo*	***bhaaghaye***	*no*
SWW:	vijana	hao	wakaondoka	na
EWW:	young males	those	they left	and
MMG:	*a-bha-bhureni* AUG2-NC2.PL-young.male	*bha-yo* AGR2-RDEM	*bha-agha-yë* 3PL-NAR-go	*no* CCONJ.AUG1

Ntk23c:	*omoona*	*ure*	***akeenderera***	***ghuye***
SWW:	mtoto	yule	akaendelea	kwenda
EWW:	child	that	he continued	to go
MMG:	*o-mo-ona* AUG1-NC1.SG-child	*u-re* AGR1-DDEM	*a-aka-enderer-a* 3SG-NAR-continue-FV	*ghu-ye* INF-go

[248] This usage of class 7 refers to the fact that each person (*kira mooto*) is getting one pear a piece.

Ntk23d:	*awaabho.*
SWW:	kwao
EWW:	to their (place)
MMG:	*awa-abho* NC23[249]-3PL.POSS
SFT:	'Sasa, kila mtu akaondoka: vijana hao wakaondoka na mtoto huyo akaendelea kwenda nyumbani.'
EFT:	'Now everyone left: the three youths went on their way and the other boy continued towards his place.'

Ntk24:	*Eriighano*	***rerasirera***	*hayo.*
SWW:	hadithi	inaishia	hapo
EWW:	story	it is ending	here
MMG:	*e-rii-ghano* AUG5-NC5.SG-story	*re-ra-sirer-a* SBJ5-PROG-end-APPL-FV	*ha-yo* NC16-RDEM
SFT:	'Hadithi inaishia hapo.'		
EFT:	'The end.'		

[249] See note on this noun class marker in footnote for *awaache* in line Ntk1b.

Appendix D: "Pear Story" in Kabwa

Since the original text of the "Pear Story" in Kabwa was translated first into Swahili, rather than directly into English, the following transcription of the story is a six-row interlinear text which includes word-for-word and free translations into both Swahili and English, as well as a morpheme-by-morpheme English gloss. The first row of each interlinearized section has the Kabwa text of the story and is titled Cwa, followed by a sentence and line number (i.e. 1a, 1b, 2, etc.). These designations are used in the body of the thesis to refer to specific portions of this Kabwa text. The abbreviations for the other rows, in the order they appear, are: SWW (Swahili Word-for-Word), EWW (English Word-for-Word), MMG[250] (Morpheme-by-Morpheme Gloss), SFT (Swahili Free Translation) and EFT (English Free Translation). The latter two are included only on the last line of a sentence. All verbal words are bolded in the Kabwa text.

Kabwa Pear Story

Cwa1a:	*Erigano*	*eryeng'i*	*rya*	*Ekikabhwa*
SWW:	hadithi	fupi	ya	Kikabwa
EWW:	story	short	of	Kabwa
MMG:	*e-ri-gano* AUG5-NC5.SG-story	*e-ri*[251]*-eng'i* AUG5-AGR5-short	*ry-a* AGR5-ASC	*e-ki-kabhwa* AUG7-NC7-Kabwa

Cwa1b:	*eguru*	*ya*	*omuntu*	*umwi*
SWW:	juu	ya	mtu	mmoja
EWW:	about	of	person	one
MMG:	*e-guru* NC9-about	*i-a* AGR9-ASC	*o-mu-ntu* AUG1-NC1.SG-person	*u-mu*[252]*-i* AUG1-AGR1-one

[250] The MMG shows the underlying forms of the morphemes.
[251] The underlying form is *ri*, but palatalization occurs before a root-initial vowel and it surfaces as *ry*, which is the orthographic form (Kiraka & Walker 2010: 12). Palatalization occurs in various places in the text.
[252] The underlying form is *mu*, but labialization occurs before a root-initial vowel and it surfaces as *mw,* which is the orthographic form (Kiraka & Walker 2010: 12). Labialization occurs in various places in the text.

Cwa1c:	*unú*	***yaaringa***	*mumugunda*	*gwaye*
SWW:	ambaye	alikuwa	shambani	mwake
EWW:	who	he was	on the farm	his
MMG:	*u-nú* AGR1-REL.PRON	*a-aa*[253]*-ri-nga* 3SG-PST-COP-P_3/P_4	*mu=mu-gunda* NC18=NC3.SG-farm	*gu-aye* AGR3-3SG-POSS

Cwa1d:	***aratwa***	*amanyembe ga amaguta.*
SWW:	anachuma	maembe ya mafuta
EWW:	he is picking	pears
MMG:	*a-ra-twa* 3SG-PROG-pick	*a-ma-nyembe g-a amaguta* AUG6-NC6.PL-pear AGR6-ASC AUG6-NC6-oil
SFT:	'Hadithi fupi ya Kikabwa juu ya mtu mmoja aliyeenda shambani mwake kuchuma maembe ya mafuta.'	
EFT:	'A short story in Kabwa about a certain man who went to his farm to pick pears.'	

Cwa2a:	*Omuntu*	*uyo*	***yaatwanga***
SWW:	mtu	huyo	alikuwa akichuma
EWW:	person	that	he was picking
MMG:	*o-mu-ntu* AUG1-NC1.SG-person	*u-yo* AGR1-RDEM	*a-aa-tw-anga* 3SG-P_3/P_4-pick-IPFV

Cwa2b:	*amanyembe ga amaguta,*	***akatwa,***
SWW:	maembe ya mafuta	akachuma
EWW:	pears	he picked
MMG:	*a-ma-nyembe g-a amaguta* AUG6-NC6.PL-pear AGR6-ASC AUG6-NC6-oil	*a-ka-tw-a* 3SG-NAR-pick-FV

Cwa2c:	***akeejurya***	*ebhitonga*	*bhibhiri.*
SWW:	akajaza	vikapu	viwili
EWW:	he filled	baskets	two
MMG:	*a-ka*[254]*-ijur-y-a* 3SG-NAR-be.filled-CAUS-FV	*e-bhi-tonga* AUG8-NC8.PL-basket	*bhi-bhiri* AGR8-two
SFT:	'Mtu huyo alikuwa akichuma maembe ya mafuta, akajaza vikapu viwili.'		
EFT:	'That man was picking pears and filled up two baskets.'		

[253] An allomorph of *a* for 3rd person singular is *y*, which surfaces in order to prevent a three vowel sequence word-initially. This process needs to be further researched in Kabwa.
[254] When *a* and *i* come together at a morpheme boundary (*ka* 'NAR' + *ijur* 'be filled') partial assimilation takes place and a long *e* is formed (Kiraka & Walker 2010: 12).

Cwa3a:	*Omwana*	*waaye*	***akaaja,***	***akahika,***
SWW:	mtoto	wake	akaja	akafika
EWW:	child	his	he came	he arrived
MMG:	*o-mu-ana* AUG1-NC1.SG-child	*wa-aye* AGR1-3SG.POSS	*a-ka-aj*[255]*-a* 3SG-NAR-come-FV	*a-ka-hik-a* 3SG-NAR-arrive-FV

Cwa3b:	***akagega***	*ekimwi*	*kubhaasikeri,*	***akagya***
SWW:	akabeba	kimoja	kwenye baiskeli	akaenda
EWW:	he carried	one (basket)	on bicycle	he went
MMG:	*a-ka-geg-a* 3SG-NAR-carry-FV	*e-ki-mwi* AUG7-NC7.SG-one	*ku=Ø-bhaasikeri* NC17=NC9.SG-bicycle	*a-ka-gy-a* 3SG-NAR-go-FV

Cwa3c:	***okuhira***	*iika.*
SWW:	kupeleka	nyumbani
EWW:	to take	to home
MMG:	*o-ku-hir-a* AUG15-INF-take-FV	*ii-ka* NC23-home
SFT:	‘Mtoto wake akaja, akafika, akabeba [kikapu] kimoja kwenye baiskeli yake, akaenda kukipeleka nyumbani.’	
EFT:	‘His son came, arrived, put one [basket] on his bike, and went to take it home.’	

Cwa4a:	*Kwa*	***okubha***	***yaaringa***	*mumasaka,*
SWW:	kwa	kuwa	alikuwa	porini
EWW:	for	being	he was	in the bush
MMG:		*o-ku-bh-a* AUG15-INF-be-FV	*a-aa-ri-nga* 3SG-PST-COP-P_3/P_4	*mu=ma-saka* NC18=NC6.PL-bush

Cwa4b:	*enjira*	***yaaringa***	*n-nyerere.*
SWW:	njia	ilikuwa	ni nyembamba
EWW:	path	it was	is narrow
MMG:	*e-n-jira* AUG9-NC9.SG-path	*e-aa-ri-nga* SBJ9-PST-COP-P_3/P_4	*n=ny-erere* COP=AGR9-narrow
SFT:	‘Kwa kuwa alikuwa porini, njia ilikuwa nyembamba.’		
EFT:	‘Since he was in the bush, the path was narrow.’		

[255] This root is technically *Vj* ‘come’ and the unspecified V takes on the value of the vowel preceding it: *okuuja* ‘to come’.

Cwa5a:	*Hanu*	***yaagyanga,***	***yaasikana***	*na*
SWW:	hapa	alikuwa akienda	alikutana	na
EWW:	here	he was going	he met	with
MMG:	*ha-nu* NC16-PDEM	*a-aa-gy-anga* 3SG-P_3/P_4-go-IPFV	*a-aa-sik-an-a* 3SG-P_4-meet-RECP-FV	*na* CCONJ

Cwa5b:	*omwana*	*oumwi*	*wa*	*ekihara*
SWW:	mtoto	mwingine	wa	kike
EWW:	child	another	of	female
MMG:	*o-mu-ana* AUG1-NC1.SG-child	*o-u-mwi* AUG1-AGR1-one	*u-a* AGR1-ASC	*e-ki-hara* AUG7-NC7-female

Cwa5c:	*unú*	***yaajanga***	*wonse*	*kunjira*
SWW:	ambaye	alikuwa akija	naye	katika njia
EWW:	who	was coming	also	on the path
MMG:	*u-nú* AGR1-REL.PRON	*a-aa-j-anga* 3SG-P_3/P_4-come-IPFV	*u-onse* AGR1-also	*ku=n-jira* NC17=NC9.SG-path

Cwa5d:	*iyo.*
SWW:	hiyo
EWW:	that
MMG:	*i-yo* AGR9-RDEM
SFT:	‘Alipokuwa akienda, alikutana na mtoto mwingine wa kike aliyekuwa akija naye katika njia hiyo.’
EFT:	‘As he was going, he met a girl who was coming down the same path.’

Cwa6:	*Hanu*	***yaahika,***	***bhakeetemana.***
SWW:	hapa	alifika	wakagongana
EWW:	here	she arrived	they ran into each other
MMG:	*ha-nu* NC16-PDEM	*a-aa-hik-a* 3SG-P_4-arrive-FV	*bha-ka-i-tem-an-a* 3PL-NAR-REFL-run.into-RECP-FV
SFT:	‘Alipofika, wakagongana.’		
EFT:	‘When she reached him, they ran into each other.’		

Cwa7a:	*Omwana*	*unú*	***yaaringa***	*na*
SWW:	mtoto	ambaye	alikuwa	na
EWW:	child	who	he was	with
MMG:	*o-mu-ana* AUG1-NC1.SG-child	*u-nú* AGR1-REL.PRON	*a-aa-ri-nga* 3SG-PST-COP-P_3/P_4	*na* CCONJ

Cwa7b:	*amanyembe ga amaguta*	***akagwa***	*hansi,*
SWW:	maembe ya mafuta	akaanguka	chini
EWW:	pears	he fell	down
MMG:	*a-ma-nyembe g-a a-ma-guta* AUG6-NC6.PL-pear AGR6-ASC AUG6-NC6-oil	*a-ka-gw-a* 3SG-NAR-fall-FV	*ha-nsi* NC16-down

Cwa7c:	*amanyembe*	***gakeetika.***
SWW:	maembe (ya mafuta)	yakamwagika
EWW:	pears	they spilled out
MMG:	*a-ma-nyembe* AUG6-NC6.PL-pear	*ga-ka-itik-a* SBJ6-NAR-be.spilled-FV
SFT:	'Mtoto aliyekuwa na maembe ya mafuta akaanguka chini, maembe (ya mafuta) yakamwagika.'	
EFT:	'They child who had the pears fell down and the pear spilled out.'	

Cwa8a:	*Hanu*	***geetika,***	*hanu*
SWW:	hapa	yalimwagika	hapa
EWW:	here	they spilled out	here
MMG:	*ha-nu* NC16-PDEM	*ga-aa*[256]*-itik-a* SBJ6-P_4-be.spilled-FV	*ha-nu* NC16-PDEM

Cwa8b:	***yeemuka***	***okugabhooja,***	*abhaana*
SWW:	aliamka	kuyaokota	watoto
EWW:	he arose	to pick them up	children
MMG:	*a-aa-imuk-a* 3SG-P_4-arise-FV	*o-ku-ga-bhooj-a* AUG15-INF-OBJ6-pick.up-FV	*a-bha-ana* AUG2-NC2.PL-child

[256] The underlying sequence of four vowels is only written as two vowels, both here and other places in the text. The actual phonetic length of a three or four-vowel sequences, in comparison with regular long vowels, is an area for further research.

Cwa8c:	*abhandi*	*bhonse*[257]	***bhaaringa***	***bharaaja***
SWW:	wengine	nao	walikuwa	wanakuja
EWW:	others	also	they were	they are coming
MMG:	*a-bha-ndi* AUG2-AGR2-other	*bha-onse* AGR2-also	*bha-aa-ri-nga* 3PL-PST-COP-P_3/P_4	*bha-ra-aj-a* 3PL-PROG-come-FV

Cwa8d:	*kunjira*	*kuyo.*
SWW:	katika njia	hiyo
EWW:	on the path	that
MMG:	*ku=n-jira* NC17=NC9.SG-path	*ku-yo* AGR17-RDEM
SFT:	'Yalipomwagika na alipoamka kuyaokota, watatu wengine walikuwa wakija njiani.'	
EFT:	'When they spilled out and the boy got up to gather them, there were other children coming down the path.'	

Cwa9a:	*Hanu*	***bhaahika,***	***bhakamubheera***
SWW:	hapa	walifika	wakamsaidia
EWW:	here	they arrived	they helped him
MMG:	*ha-nu* NC16-PDEM	*bha-aa-hik-a* 3PL-P_4-arrive-FV	*bha-ka-mu-bheer-a* 3PL-NAR-3SG.OBJ-help-FV

Cwa9b:	***okugabhooja.***
SWW:	kuyaokota
EWW:	to pick them up
MMG:	*o-ku-ga-bhooj-a* AUG15-INF-OBJ6-pick.up-FV
SFT:	'Walipofika, wakamsaidia kuyaokota.'
EFT:	'When they arrived, they helped him pick up the pears.'

Cwa10a:	***Bhakabhooja***	*gonsego,*	***akatuura***	*mukitonga,*
SWW:	wakaokota	yote	akaweka	katika kikapu
EWW:	they picked up	all (the pears)	he put	in the basket
MMG:	*bha-ka-bhooj-a* 3PL-NAR-pick.up-FV	*ga-onsego* AGR6-all	*a-ka-tuur-a* 3SG-NAR-put-FV	*mu=ki-tonga* NC18=NC7.SG-basket

[257] Atlhough vowel lengthening takes place before a pre-nasalized consonant, it is not phonemic and is written as a single vowel in the orthography (Kiraka & Walker 2010: 6-7). That rule is generally broken when there are underlyingly two vowels across a morpheme boundary followed by a pre-nasalized consonant (7); however, common conjunctive words like *bhonse* are exceptions due to their unambiguous nature and high frequency.

Cwa10b:	***akarin'ya***	*kubhasikeeri,*	***akagya.***
SWW:	akapandisha	kwenye baiskeli	akaenda
EWW:	he got on	on the bicycle	he went
MMG:	*a-ka-rin-y-a* 3SG-NAR-go.up-CAUS-FV	*ku=Ø-bhasikeeri* NC17=NC9.SG-bicycle	*a-ka-gy-a* 3SG-NAR-go-FV
SFT:	'Wakaokota maembe (ya mafuta) hayo yote, kijana akaweka kwenye kikapu, akapandisha baiskeli, akaenda.'		
EFT:	'They gathered all the pears and the boy put them in the basket. Then he got on his bike and left.'		

Cwa11a:	*Nawe*	***yaaringa***	***anyaharikiri***
SWW:	lakini	alikuwa	ameumia
EWW:	but	he was	he has injured
MMG:		*a-aa-ri-nga* 3SG-PST-COP-P_3/P_4	*a-nyahar-ik-iri* 3SG-injure-STAT-ANT

Cwa11b:	*okuguru.*
SWW:	mguu
EWW:	foot
MMG:	*o-ku-guru* AUG15-NC15-foot
SFT:	'Lakini alikuwa ameumia mguu.'
EFT:	'But he had injured his foot.'

Cwa12:	***Yaagendanga***	***arahegera.***
SWW:	alikuwa akitembea	anachechemea
EWW:	he was walking	he is limping
MMG:	*a-aa-gend-anga* 3SG-P_3/P_4-walk-IPFV	*a-ra-heger-a* 3SG-PROG-limp-FV
SFT:	'Alikuwa anatembea akichechemea.'	
EFT:	'He was walking with a limp.'	

Cwa13a:	*Hanu*	***yaahika***	*hare-ko,*	*abhaana*
SWW:	hapa	alifika	mbali kidogo	watoto
EWW:	here	he arrived	far (there)	children
MMG:	*ha-nu* NC16-PDEM	*a-aa-hik-a* 3SG-P_4-arrive-FV	*ha-re=ko* NC16-far=LOC17	*a-bha-ana* AUG2-NC2.PL-child

Cwa13b:	*bharya*	***bhakabhooja***	*engofiira*	*yaaye*
SWW:	wale	wakaokota	kofia	yake
EWW:	those	they picked up	hat	his
MMG:	*bha-rya* AGR2-DDEM	*bha-ka-bhooj-a* 3PL-NAR-pick.up-FV	*e-n-gofiira* AUG9-NC9.SG-hat	*ya-aye* AGR9-3SG.POSS

Cwa13c:	*inú*	***yaaringa***	***agwisirye.***
SWW:	ambayo	alikuwa	amedondosha
EWW:	which	he was	he has dropped
MMG:	*i-nú* AGR9-REL.PRON	*a-aa-ri-nga* 3SG-PST-COP-P_3/P_4	*a-gw-isirye* 3SG-fall-CAUS.ANT
SFT:	‘Alipofika mbali kidogo, wale wale watoto waliokota kofia yake aliyokuwa amedondosha.’		
EFT:	‘When he had gone a little ways, those same boys picked up the hat that he had dropped.’		

Cwa14a:	*Oumwi*	*unú*	***yaaringa***
SWW:	mmoja	ambaye	alikuwa
EWW:	one (of them)	who	he was
MMG:	*o-u-mwi* AUG1-NC1.SG-one	*u-nú* AGR1-REL.PRON	*a-aa-ri-nga* 3SG-PST-COP-P_3/P_4

Cwa14b:	***arahooya***	*omupiira*	*kukibhaau*
SWW:	anacheza	mpira	kwenye kiubau
EWW:	he is playing	ball	on board
MMG:	*a-ra-hooy-a* 3SG-PROG-play-FV	*o-mu-piira* AUG3-NC3-ball	*ku=ki-bhaau* NC17-NC7.SG-board

Cwa14c:	***akamutemera***	*ekinyu,*
SWW:	akampigia	mluzi
EWW:	he hit for him	whistle
MMG:	*a-ka-mu-tem-er-a* 3SG-NAR-3SG.OBJ-hit-APPL-FV	*e-ki-nyu* AUG7-NC7-whistle

Cwa14d:	***akamuhirira***	*engofiira.*
SWW:	akampelekea	kofia
EWW:	he took to him	hat
MMG:	*a-ka-mu-hir-ir-a* 3SG-NAR-3SG.OBJ-take-APPL-FV	*e-n-gofiira* AUG9-NC9.SG-hat
SFT:	'Mmojawapo aliyekuwa anacheza mpira katika kiubao akampigia mluzi, akampelekea kofia.'	
EFT:	'One of them, who was playing with a paddleball, whistled to the boy and took his hat to him.'	

Cwa15a:	*Hanu*	***yaahika,***	*omwana*	*urya*
SWW:	hapa	alifika	mtoto	yule
EWW:	here	he arrived	child	that
MMG:	*ha-nu* NC16-PDEM	*a-aa-hik-a* 3SG-P_4-arrive-FV	*o-mu-ana* AUG1-NC1.SG-child	*u-rya* AGR1-DDEM

Cwa15b:	*wonse,*	*kwa*	*obhujomerwa,*	*kya*	*ego*	***bhaaringa***
SWW:	naye	kwa	furaha	kama	hivyo	walikuwa
EWW:	also	with	joy	like	that way	they were
MMG:	*u-onse* AGR1-also		*o-bhu-jomerwa* AUG14-NC14-joy			*bha-aa-ri-nga* 3PL-PST-COP-P_3/P_4

Cwa15c:	***bhamusakirya,***	***akabhahaana***
SWW:	wamemsaidia	akawapa
EWW:	they have helped him	he gave them
MMG:	*bha-mu-sakiry*[258]*-a* 3PL-3SG.OBJ-help-ANT	*a-ka-bha-haan-a* 3SG-NAR-3PL.OBJ-give-FV

Cwa15d:	*amanyembe maguta*	*gasatu,*	*kwa*	***okubha***
SWW:	maembe (ya mafuta)	matatu	kwa	kuwa
EWW:	pears	three	for	being
MMG:	*a-ma-nyembe ma-guta* AUG6-NC6.PL-mango NC6-oil	*ga-satu* AGR6-three		*o-ku-bh-a* AUG15-INF-be-FV

[258] *Okusakirya* is a lexicalized verb meaning 'to help'; however, the etymology of this word is interesting since Causative *-y* and Applicative *-ir* suffixes are added to the root *okusaka* 'to snatch up'. The literal meaning of *okusakirya* is 'to cause to snatch up for'.

Cwa15e:	***bhaaringa***	*bhasatu.*
SWW:	walikuwa	watatu
EWW:	they were	three
MMG:	*bha-aa-ri-nga* 3PL-PST-COP-P_3/P_4	*bha-satu* AGR2-three
SFT:	'Alipofika, mtoto yule naye, kwa furaha jinsi walivyokuwa wamemsaidia, akawapa maembe ya mafuta matatu, kwa kuwa walikuwa watatu.'	
EFT:	'When he arrived, the other boy, with joy because they had helped, gave them three pears, since there were three of them.'	

Cwa16:	*Bhonse*	***bhakamuhaana***	*engofiira*	*yaaye.*
SWW:	nao	wakampa	kofia	yake
EWW:	they also	they gave him	hat	his
MMG:	*bha-onse* AGR2-also	*bha-ka-mu-haan-a* 3PL-NAR-3SG.OBJ-give-FV	*e-n-gofiira* AUG9-NC9.SG-hat	*ya-aye* AGR9-3SG.POSS
SFT:	'Nao wakampa kofia yake.'			
EFT:	'And they gave him his hat.'			

Cwa17a:	*Kwa*	*ego,*	*omuntu*	*urya*	***akagendererya***
SWW:	kwa	hivyo	mtu	yule	akaendelea
EWW:	for	that way	person	that	he continued
MMG:			*o-mu-ntu* AUG1-NC1.SG-person	*u-rya* AGR1-DDEM	*a-ka-genderery-a* 3PL-NAR-continue-FV

Cwa17b:	***okutwa***	*amanyembe maguta*	*gaaye,*
SWW:	kuchuma	maembe (ya mafuta)	yake
EWW:	to pick	pears	his
MMG:	*o-ku-tw-a* AUG15-INF-pick-FV	*a-ma-nyembe ma-guta* AUG6-NC6.PL-mango NC6-oil	*ga-aye* AGR6-3SG.POSS

Cwa17c:	***akamara.***
SWW:	akamaliza
EWW:	he finished
MMG:	*a-ka-mar-a* 3PL-NAR-finish-FV
SFT:	'Kwa hivyo, mtu urya akaendelea kuchuma maembe ya mafuta yake, akamaliza.'
EFT:	'Meanwhile, the man (at the tree) continued picking pears until he finished.'

Cwa18a:	*Hanu*	***yaamara,***	***akaamba***	*obhutaratibhu*
SWW:	hapa	alimaliza	akaanza	taratibu
EWW:	here	he finished	he started	process
MMG:	*ha-nu* NC16-PDEM	*a-aa-mar-a* 3SG-P_4-finish-FV	*a-ka-amb-a* 3PL-NAR-start-FV	*o-bhu-taratibhu* AUG14-NC14-process

Cwa18b:	*bhwa*	***okuhinda***	*iika.*
SWW:	ya	kurudi	nyumbani
EWW:	of	returning	home
MMG:	*bhu-a* AGR14-ASC	*o-ku-hind-a* AUG15-INF-return-FV	*ii-ka* NC23-home
SFT:	'Alipomaliza, akaanza taratibu ya kurudi nyumbani.'		
EFT:	'When he finished, he started to pack up to return home.'		

Cwa19:	*Erigano*	*rinu*	***rihwereri***	*hayo.*
SWW:	hadithi	hii	imeishia	hapo
EWW:	story	this	it has ended	here
MMG:	*e-ri-gano* AUG5-NC5.SG-story	*ri-nu* AGR5-PDEM	*ri-hwer-eri* SBJ5-end-APPL.ANT	*ha-yo* NC16-RDEM
SFT:	'Hadithi hii imeishia hapo.'			
EFT:	'The end.'			

Appendix E: "Pear Story" in Simbiti

Since the original text of the "Pear Story" in Simbiti was translated first into Swahili, rather than directly into English, the following transcription of the story is a six-row interlinear text which includes word-for-word and free translations into both Swahili and English, as well as a morpheme-by-morpheme English gloss. The first row of each interlinearized section has the Simbiti text of the story and is titled Ssc, followed by a sentence and line number (i.e. 1a, 1b, 2, etc.). These designations are used in the body of the thesis to refer to specific portions of this Simbiti text. The abbreviations for the other rows, in the order they appear, are: SWW (Swahili Word-for-Word), EWW (English Word-for-Word), MMG[259] (Morpheme-by-Morpheme Gloss), SFT (Swahili Free Translation) and EFT (English Free Translation). The latter two are included only on the last line of a sentence. All verbal words are bolded in the Simbiti text.

Simbiti Pear Story

Ssc1a:	***Ndareeta,***	***ndatuna***	***mbatëëbi***	*bhono*
SWW:	ninaleta	ninataka	niwaambie	jinsi
EWW:	I bring	I want	I should tell you (pl.)	how
MMG:	*n-ra*[260]*-reet-a* 1SG-PRS-bring-FV	*n-ra-tuna-a* 1SG-PRS-want-FV	*m-ba-tëëbi* 1SG-3PL.OBJ-tell.SBJV	*bho-no* AGR14-PDEM

Ssc1b:	***naarë***	***ndööshë***	*ipicha*	*iyëndë*
SWW:	nilikuwa	nimeona	picha	nyingine
EWW:	I was	I have seen	film	another
MMG:	*n-aa-rë* 1SG-PST-COP	*n-röösh*ë 1SG-see.ANT	*i-picha* NC9.SG-film	*iy-ëndë* AGR9-another

[259] The MMG shows the underlying forms of the morphemes.
[260] An underlying *r-* surfaces as *d-* when it follows a nasal, both here and other places in the text (Compton & Walker 2010: 10).

Ssc1c:	*mobhasö*	*hano.*
SWW:	mchana	hapa
EWW:	afternoon	here
MMG:	*mo-bhasö* NC3.SG-afternoon	*ha-no* NC16-RDEM
SFT:	'Ninataka niwaambie kwamba niliona picha nyingine mchana huo.'	
EFT:	'I would like to tell you all that I saw another film this afternoon.'	

Ssc2a:	*Ku-picha*	*iyö*	*eno*	***yaarë***
SWW:	Katika picha	hiyo	ambayo	ilikuwa
EWW:	in film	that	which	it was
MMG:	*ku=Ø-picha* NC17=NC9.SG-film	*i-yö* AGR9-RDEM	*e-no* AGR9-PDEM.REL	*e*[261]*-aa-rë* SBJ9-PST-COP

Ssc2b:	*ko-kompyuta*	*iya*	*mosaani*	*uwaane,*
SWW:	kwenye kompyuta	ya	rafiki	yangu
EWW:	on computer	of	friend	mine
MMG:	*ko=Ø-kompyuta* NC17=NC9.SG-computer	*iy-a* AGR9-ASC	*mo-saani* NC1-friend	*u-wa-ane* AUG1-NC1-1SG.POSS

Ssc2c:	***naarë***	***ndööshë***	*nyinkyö,*	*ahasë*
SWW:	nilikuwa	nimeona	asubuhi	mahali
EWW:	I was	I have seen	morning	place
MMG:	*n-aa-rë* 1SG-PST-COP	*n-röösh*ë 1SG-see.ANT	*nyinkyö* NC9.SG.morning	*a-ha-së* AUG16-NC16-place

Ssc2d:	***haarë***	***habhaayë***	*kya*	*nyinkyö*	*igho.*
SWW:	palikuwa	pamekuwa	kama	asubuhi	hivi
EWW:	it was	it has become	like	morning	like this
MMG:	*ha-aa*[262]*-rë* AGR16-PST-COP	*ha-bha-irë*[263] SBJ16-be-ANT		*nyinkyö* NC9.SG.morning	
SFT:	'Katika picha hiyo iliyokuwa kwenye kompyuta ya rafiki yangu, niliona mahali palikuwa kama asubuhi.'				
EFT:	'In that film, which was on my friend's computer, I saw a place at morning time.'				

[261] This surfaces as <*y*> before the low-vowel <*a*>.
[262] There is a rule preventing a triple length vowel, so it is realized as a long-vowel.
[263] A process of imbrication causes the Anterior morpheme to surface as *-ayë*, but this process needs to be better studied and described for Simbiti. Similar processes occur elsewhere in this text without further comment.

Ssc3a:	***Naighwa***	*sengoko*	***seraraghora,***	***naamaaha***
SWW:	nikasikia	kuku	wakiwika	nikaona
EWW:	I heard	chickens	they are crowing	I saw
MMG:	*n-aa-ighw-a* 1SG-PST-hear-FV	*sen-goko* NC10.PL-chicken	*se-ra-raghor-a* SBJ10-PROG-crow-FV	*n-aa-maah-a* 1SG-PST-see-FV

Ssc3b:	*ömöntö*	*kya*	*ono*	***atuubhaini***
SWW:	mtu	kama	ambaye	amefanana
EWW:	person	like	who	he resembles
MMG:	*ö-mö-ntö* AUG1-NC1.SG-person		*o-no* AGR1-PDEM.REL	*a-tuubha-ini* 3SG-resemble-ANT.RECP

Ssc3c:	*umwibhi*	*handë*	*ono*	***amutirë***
SWW:	mwizi	au	ambaye	amewahi
EWW:	thief	or	who	he is early
MMG:	*u-mu*[264]*-ibhi* AUG1-NC1.SG-thief		*o-no* AGR1-PDEM.REL	*a-mut-irë* 3SG-be.early-ANT

Ssc3d:	***ukuya***	*kö-möghöndö*	*ughwaye.*
SWW:	kwenda	kwenye shamba	lake
EWW:	to go	to farm	his
MMG:	*u-ku-ya* AUG15-INF-go	*kö=mö-ghöndö* NC17=NC3.SG-farm	*u-ghu-aye* AUG3-AGR3-3SG.POSS
SFT:	'Nikasikia majogoo wakiwika, nikaona mtu aliyekuwa amefanana na mwizi au kama ambaye amewahi kwenda shambani kwake.'		
EFT:	'I heard roosters crowing and I saw a man who looked kind of like a thief or like someone who had gone out early to his farm.'		

Ssc4a:	***Aratwa***	*sëhaghwë*	*handë*	*seno*	***tokobherekera***
SWW:	anachuma	matunda	au	ambayo	tunayaita
EWW:	he is picking	fruits	or	which	we call
MMG:	*a-ra-tw-a* 3SG-PROG-pick-FV	*së-haghwë* NC10.PL-fruit		*se-no* AGR10-PDEM.REL	*to-ko-bhereker-a* 1PL-PRS-call-FV

[264] The *u* in the prefix is labialized before a root-initial vowel, both here and other places in the text (Compton & Walker 2010: 20).

Ssc4b:	*igha*	*amatunda.*
SWW:	kwamba	matunda
EWW:	that	fruits
MMG:		*a-ma-tunda* AUG6-NC6-fruit
SFT:	‘Alikuwa akichuma matunda.’	
EFT:	‘He was picking fruit.’	

Ssc5:	***Yaarë***	*na*	*amatonga*	*atatö.*
SWW:	alikuwa	na	matenga	matatu
EWW:	he was	with	baskets	three
MMG:	*a*[265]*-aa-rë* 3SG-PST-COP	*na* CCONJ	*a-ma-tonga* AUG6-NC6.PL-basket	*a-tatö* AGR1-three
SFT:	‘Alikuwa na matenga matatu.’			
EFT:	‘He had three baskets.’			

Ssc6a:	*Urwa*	*kwanza*	*ruyö,*	***naarë***	***mmörööshë,***
SWW:	mara ya	kwanza	hiyo	nilikuwa	nimemwona
EWW:	time of	first	that	I was	I have seen him
MMG:	*u-ru-a* AUG11-NC11-ASC		*ru-yö* AGR11-RDEM	*n-aa-rë* 1SG-PST-COP	*m-mö-röö-shë* 1SG-3SG.OBJ-see.ANT

Ssc6b:	***yaarë***	***aishööyi***	*ritonga*	*imwë.*
SWW:	alikuwa	amejaza	tenga	moja
EWW:	he was	he has filled	basket	one
MMG:	*a-aa-rë* 3SG-PST-COP	*a-ishööyi* 3SG-be.full.CAUS.ANT	*ri-tonga* NC5.SG-basket	*i-mwë* AGR5-one
SFT:	‘Mara hiyo ya kwanza, niliona alikuwa amejaza tenga moja.’			
EFT:	‘That first time, I saw he had filled one basket.’			

Ssc7a:	*Bhoono*	*hake*	*igho*	***naamaaha***	*omoona*
SWW:	sasa	kidogo	tu	nikaona	mtoto
EWW:	now	little	only	I saw	child
MMG:		*ha-ke* NC16-little		*n-aa-maah-a* 1SG-PST-see-FV	*o-mo-ona* AUG1-NC1.SG-child

[265] The third person singular morpheme surfaces as <*y*> when it comes before a long vowel word-initially, both here and other places in this text.

Ssc7b:	*uwöndë.*
SWW:	fulani
EWW:	certain
MMG:	*u-wö-ndë* AUG1-NC1.SG-certain
SFT:	'Sasa, baada ya kidogo, nikaona mtoto fulani.'
EFT:	'Now, a little while later, I saw a certain child.'

Ssc8a:	*Ghwiki*	*hake*	*igho*	***naighwa***
SWW:	tena	kidogo	tu	nikasikia
EWW:	again	little	only	I heard
MMG:		*ha-ke* NC16-little		*n-aa-ighw-a* 1SG-PST-hear-FV

Ssc8b:	*ëmböri*	***eraana***
SWW:	mbuzi	analia
EWW:	goat	it is crying
MMG:	*ëm-böri* NC9.SG-goat	*e-ra-an-a* SBJ9-PROG-cry-FV
SFT:	'Tena baada ya kidogo, nikasikia mbuzi analia.'	
EFT:	'After a little while more, I heard a goat crying.'	

Ssc9a:	*Ëmböri*	*eera*	*yonswe,*	*ono*	***yaarë***
SWW:	mbuzi	yule	nayo	huyu	alikuwa
EWW:	goat	that	also	this (one)	he was
MMG:	*ëm-böri* NC9.SG-goat	*e-era* AGR9-DDEM	*y-onswe* AGR9-all	*o-no* NC1-PDEM	*a-aa-rë* 3SG-PST-COP

Ssc9b:	*nayo*	***yaatora***	*nayo,*	***naamaaha***
SWW:	nayo	alitokea	nayo	nikaona
EWW:	with it	he came out	with it	I saw
MMG:	*na-yo* CCONJ-AGR9	*a-aa-tor-a* 3SG-PST-come.out-FV	*na-yo* CCONJ-AGR9	*n-aa-maah-a* 1SG-PST-see-FV

Ssc9c:	***yaaheta***	*haang'ë*	*na*	*amatonga*	*agha*
SWW:	alipita	karibu	na	matenga	ya
EWW:	he passed	near	with	baskets	of
MMG:	*a-aa-het-a* 3SG-PST-pass-FV	*ha-ang'ë* NC16-near	*na* CCONJ	*a-ma-tonga* AUG6-NC6.PL-basket	*a-gh-a* AUG6-NC6-ASC

Ssc9d:	*ömöntö*	*uyö.*
SWW:	mtu	huyo
EWW:	person	that
MMG:	*ö-mö-ntö* AUG1-NC1.SG-person	*u-yö* AGR1-RDEM
SFT:	'Aliyekuwa na mbuzi akatokea, nikamwona anapita karibu na matenga ya mtu huyo.'	
EFT:	'The owner of the goat came out and I saw him passing close to the other man's baskets.'	

Ssc10a:	*Ömöntö*	*uyö*	*hano*	***yaarë***
SWW:	mtu	huyo	hapa	alikuwa
EWW:	person	that	here	he was
MMG:	*ö-mö-ntö* AUG1-NC1.SG-person	*u-yö* AGR1-RDEM.REL	*ha-no* NC16-PDEM	*a-aa-rë* 3SG-PST-COP

Ssc10b:	***amörööshë,***	***yaimaana***	*kö-mötë.*
SWW:	amemwona	akajibanza	kwenye mti
EWW:	he has seen him	he pressed himself	to the tree
MMG:	*a-mö-röösh*ë 3SG-3SG.OBJ-see.ANT	*a-aa-i-maan-a* 3SG-PST-REFL-press-FV	*kö=mö-të* NC17=NC3.SG-tree
SFT:	'Mtu huyo aliyekuwa akichuma matunda, alipomwona yule kijana, akajibanza kwenye mti.'		
EFT:	'The man who was picking fruit, when he saw this man, pressed himself up against the tree.'		

Ssc11a:	*Nyamböri*	*uyö*	*wonswe*	***yaaheta***
SWW:	mwenye mbuzi	huyo	naye	alipita
EWW:	goat owner	that	also	he passed by
MMG:	*nya=m-böri* POSS.NC1=NC9.SG-goat	*u-yö* AGR1-RDEM	*w-onswe* AGR1-all	*a-aa-het-a* 3SG-PST-pass-FV

Ssc11b:	***ti-yaarë***	***atöniibhwi***	*na*	*amatonga*
SWW:	hakuwa	amejali	na	matenga
EWW:	he did not	he has been concerned	with	baskets
MMG:	*ti=a-aa-rë* NEG=3SG-PST-COP	*a-töniibh-w-i* 3SG-concern-PASS-ANT	*na* CCONJ	*a-ma-tonga* AUG6-NC6.PL-basket

Ssc11c:	*gaara*	***hë.***
SWW:	yale	NEG
EWW:	those	NEG
MMG:	*ga-ara* NC6-DDEM	*hë* NEG
SFT:	'Na huyo aliyekuwa na mbuzi, naye alipita bila kujali yale matenga.'	
EFT:	'The man with the goat passed by without looking at the baskets.'	

Ssc12a:	*Ëmböri*	*eera*	***yaarë***	***ököörökya***
SWW:	mbuzi	yule	alikuwa	kuonyesha
EWW:	goat	that	it was	to show
MMG:	*ëm-böri* NC9.SG-goat	*e-era* AGR9-DDEM	*a-aa-rë* SBJ9-PST-COP	*ö-kö-öröky-a* AUG15-INF-show-FV

Ssc12b:	*igha*	***okanyoora***	*wonswe,*	*na*	*ahasë*
SWW:	kwamba	ukakuta	naye	na	mahali
EWW:	that	you found	he also	with	place
MMG:		*o-ka-nyoor-a* 2SG-NAR-find-FV	*w-onswe* AGR1-all	*na* CCONJ	*a-ha-së* AUG16-NC16-place

Ssc12c:	***aghirutirë***	***ukurwa,***	***naghëkööbhööyë.***
SWW:	amemvuta	kutoka	amemwiba
EWW:	he has pulled it	from	he has stolen it
MMG:	*a-ghi-rut-irë* 3SG-NC9-pull-ANT	*u-ku-rw-a* AUG15-INF-out.of-FV	*n-a-ghë-kööbhör-irë* FOC-3SG-NC9-steal-ANT
SFT:	'Yule mbuzi alionekana kwamba huenda mtu huyo amemwiba kutoka sehemu fulani.'		
EFT:	'That goat was looking around like the man had stolen it from somewhere.'		

Ssc13a:	*Na*	*hano*	***yaarë***	***amaarë***	***okoheta***
SWW:	na	hapa	alikuwa	amemaliza	kupita
EWW:	and	here	he was	he has finished	to pass by
MMG:	*na* CCONJ	*ha-no* NC16-RDEM.REL	*a-aa-rë* 3SG-PST-COP	*a-mar-irë* 3SG-finish-ANT	*o-ko-het-a* AUG15-INF-pass.by-FV

Ssc13b:	*ömöntö*	*uyö,*	*ghwiki*	*ömöntö*	*uyö*
SWW:	mtu	huyo	tena	mtu	huyo
EWW:	person	that	again	person	that
MMG:	*ö-mö-ntö* AUG1-NC1.SG-person	*u-yö* AGR1-RDEM		*ö-mö-ntö* AUG1-NC1.SG-person	*u-yö* AGR1-RDEM

Ssc13c:	***yaarë***	*mu-risisi*	*muyö*	***aratwa***
SWW:	alikuwa	porini	humo	anachuma
EWW:	he was	in the bush	in there	he is picking
MMG:	*a-aa-rë* 3SG-PST-COP	*mu=ri-sisi* NC18=NC5.SG-bush	*mu-yö* NC18-RDEM	*a-ra-tw-a* 3SG-PROG-pick-FV

Ssc13d:	*sëhaghwë,*	***akariina***	*ghwiki*	***ukutwa***	*sëhaghwë.*
SWW:	matunda	akapanda	tena	kuchuma	matunda
EWW:	fruits	he climbed	again	to pick	fruits
MMG:	*së-haghwë* NC10.PL-fruit	*a-ka-riin-a* 3SG-NAR-climb-FV		*u-ku-tw-a* AUG15-INF-pick-FV	*së-haghwë* NC10.PL-fruit
SFT:	‘Na yule mtu aliyekuwa na mbuzi alipokwisha kupita, yule mchuma matunda shambani akapanda tena, akaendelea kuchuma.’				
EFT:	‘When the man with the goat had passed by, the fruit picker climbed up (the tree) again and continued picking.’				

Ssc14a	*Hake*	*igho*	***naamaaha***	*ghwiki*	*omoona*
SWW:	kidogo	tu	nikaona	tena	mtoto
EWW:	little	only	I saw	again	child
MMG:	*ha-ke* NC16-little		*n-aa-maah-a* 1SG-PST-see-FV		*o-mo-ona* AUG1-NC1.SG-child

Ssc14b:	***araasha***	***arakënyia***	*ëbhaisikëëri.*
SWW:	anakuja	anakimbiza	baiskeli
EWW:	he is coming	he is running	bicycle
MMG:	*a-ra-ash-a* 3SG-PROG-come-FV	*a-ra-këny-i-a* 3SG-PROG-run-CAUS-FV	*ë-bhaisikëëri* NC9.SG-bicycle
SFT:	‘Baada ya kidogo, nikamwona kijana akiendesha baiskeli anakuja kupitia eneo hilo.’		
EFT:	‘A little while later, I saw a boy pushing a bicycle toward the tree.’		

Ssc15a	***Yaakenya***	*këmwëkëmwë,*	***yaasha,***	***yaaheta***
SWW:	akakimbia	harakaharaka	akaja	akapita
EWW:	he ran	quickly	he came	he passed by
MMG:	*a-aa-keny-a* 3SG-PST-run-FV	*kë-mwë~këmwë* NC7-one~RDPL	*a-aa-ash-a* 3SG-PST-come-FV	*a-aa-het-a* 3SG-PST-pass.by-FV

Ssc15b:	*haang'ë*	*na*	*amatonga*	*ghayö.*
SWW:	karibu	na	matenga	hayo
EWW:	near	with	baskets	those
MMG:	*ha-ang'ë* NC16-near	*na* CCONJ	*a-ma-tonga* AUG6-NC6.PL-basket	*gha-yö* AGR6-RDEM
SFT:	'Akaendesha harakaharaka, akapita karibu na matenga hayo.'			
EFT:	'He drove quickly and passed close to where the baskets were.'			

Ssc16a:	*Bhuyö*	***yaarë***	***aghëëyë***	***okomaaha,***
SWW:	huo	alikuwa	ameenda	kuona
EWW:	at that time	he was	he has gone	to see
MMG:	*bhu-yö* AGR14-RDEM.REL	*a-aa-rë* 3SG-PST-COP	*a-ghe-irë* 3SG-go-ANT	*o-ko-maah-a* AUG15-INF-see-FV

Ssc16b:	*amatonga*	*ghayö*	*abhërë*	***ghaishööyë***	*sëhaghwë.*
SWW:	matenga	hayo	mawili	yamejaa	matunda
EWW:	baskets	those	two	they are full	fruits
MMG:	*a-ma-tonga* AUG6-NC6.PL-basket	*gha-yö* AGR6-RDEM	*a-bhërë* AGR6-two	*gha-ishor-irë* SBJ6-be.full-ANT	*së-haghwë* NC10.PL-fruit
SFT:	'Mara alipoona hayo matenga, mawili yamejaa matunda.'				
EFT:	'At the time he went to look, two baskets had already been filled.'				

Ssc17a:	*Wonswe*	***naamaaha***	***araghaiseeheera,***
SWW:	naye	nikaona	anayafikiria
EWW:	he also	I saw	he is thinking about them
MMG:	*w-onswe* AGR1-all	*n-aa-maah-a* 1SG-PST-see-FV	*a-ra-gha-iseeheer-a* 3SG-PROG-OBJ6-think.about-FV

Ssc17b:	***yaamagamaga***	*eno*	*na*	*eno,*	***yaagharamera***
SWW:	akaangaza	huku	na	huku	akainua macho
EWW:	he stared	here	and	here	he looked up
MMG:	*a-aa-maga~mag-a* 3SG-PST-stare~RDPL-FV	*e-no* NC23-PDEM	*na* CCONJ	*e-no* NC23-PDEM	*a-aa-gharamer-a* 3SG-PST-look.up-FV

Ssc17c:	*ighörö,*	***yaamaaha***	*ömöntö*	*oora*
SWW:	juu	akaona	mtu	yule
EWW:	up	he saw	person	that
MMG:	*i-ghörö* NC23-up	*a-aa-maah-a* 3SG-PST-see-FV	*ö-mö-ntö* AUG1-NC1.SG-person	*o-ora* AGR1-DDEM.REL

Ssc17d:	***akutwa***	*sëhaghwë*	*seera*
SWW:	anayechuma	matunda	yale
EWW:	that is picking	fruits	those
MMG:	*a-ku-tw-a* 3SG-PRS-pick-FV	*së-haghwë* NC10.PL-fruit	*se-era* AGR10-DDEM

Ssc17e:	***ta-komomaaha***	***hë.***
SWW:	hamwoni	NEG
EWW:	he does not see him	NEG
MMG:	*t-a-ko-mo-maah-a* NEG-3SG-PRS-3SG.OBJ-see-FV	*hë* NEG
SFT:	'Nikamwona ameshawishika kuichukua, akaangalia upande na upande, akaona mwenye kuchuma hamwoni.'	
EFT:	'I saw him thinking about taking something, so he looked all around and saw that the man picking fruit doesn't see him.'	

Ssc18:	***Ta-ana***	*agha*	*ikyaro*	***hë.***
SWW:	hana	ya	nchi	NEG
EWW:	he does not have	of	country	NEG
MMG:	*ta=a-na* NEG=3SG-be.with	*a-gh-a* AUG6-AGR6-ASC	*i-ki-aro* AUG7-NC7.SG-country	*hë* NEG
SFT:	'Hana habari.'			
EFT:	'He has no idea.'			

Ssc19a:	***Yaihiinya,***	***yaaghegha***	*ritonga*	*rimwë*
SWW:	akainama	akachukua	tenga	moja
EWW:	he bent over	he took	basket	one
MMG:	*a-aa-ihiny-a* 3SG-PST-bend.over-FV	*a-aa-ghegh-a* 3SG-PST-take-FV	*ri-tonga* NC5.SG-basket	*ri-mwë* AGR5-one

Ssc19b:	*irya*	*sëhaghwë,*	***yaatoora***	*ko-bhaisikëëri*
SWW:	la	matunda	akaweka	kwenye baiskeli
EWW:	of	fruits	he put	on bicycle
MMG:	*i-ri-a* AUG5-AGR5-ASC	*së-haghwë* NC10.PL-fruit	*a-aa-toor-a* 3SG-PST-put-FV	*ko=Ø-bhaisikëëri* NC17=NC9.SG-bicycle

Ssc19c:	*iyaaye.*
SWW:	yake
EWW:	his
MMG:	*i-ya-aye* AUG9-AGR9-3SG.POSS
SFT:	'Akainama, akachukua tenga moja la matunda, akaweka juu ya baiskeli yake.'
EFT:	'He bent over, took a basket of fruit and put it on his bicycle.'

Ssc20a:	***Yaaghörrökya***	*ëbhaisikëëri,*	***yaatoora-ko***
SWW:	akainua	baiskeli	akaweka juu yake
EWW:	he lifted up	bicycle	he put in on bicycle
MMG:	*a-aa-ghörröky-a* 3SG-PST-lift.up-FV	*ë-bhaisikëëri* NC9.SG-bicycle	*a-aa-toor-a=ko* 3SG-PST-put-FV=LOC17

Ssc20b:	*sëhaghwë*	*siyö,*	***yaakenya***	*këmwë.*
SWW:	matunda	hayo	akakimbia	mara moja
EWW:	fruits	those	he ran	at once
MMG:	*së-haghwë* NC10.PL-fruit	*si-yö* AGR10-RDEM	*a-aa-keny-a* 3SG-PST-run-FV	*kë-mwë* NC7-one
SFT:	'Akainua baiskeli, akaweka matunda juu yake, akaendesha mara moja.'			
EFT:	'He lifted up the bike, put the fruit on it and rode off quickly.'			

Ssc21a:	*Bhuyö*	***yaarë***	***aghëëyë***	***okokenya***	*igho,*
SWW:	huo	alikuwa	ameenda	kukimbia	tu
EWW:	at that time	he was	he has gone	to run	only
MMG:	*bhu-yö* AGR14-RDEM.REL	*a-aa-rë* 3SG-PST-COP	*a-ghe-irë* 3SG-go-ANT	*o-ko-keny-a* AUG15-INF-run-FV	

Ssc21b:	***bhaaya***	*ghwiki,*	***bhaasikana***	*na*	*uwöndë,*
SWW:	walienda	tena	wakakutana	na	mwingine
EWW:	they went	again	they met each other	with	another (person)
MMG:	*bha-aa-y-a* 3PL-PST-go-FV		*bha-aa-sik-an-a* 3PL-PST-meet-RECP-FV	*na* CCONJ	*u-wö-ndë* AUG1-NC1.SG-other

Ssc21c:	*wonswe*	***arëngë***	*ko-bhaisikëëri.*
SWW:	naye	akiwa	kwenye baiskeli
EWW:	also	she is (there)	on bicycle
MMG:	*w-onswe* AGR1-all	*a-rë-ngë* 3SG-COP-LOC	*ko=Ø-bhaisikëëri* NC17=NC9.SG-bicycle
SFT:	'Mara alipoanza kukimbia tu, akakutana na mwingine akiwa kwenye baiskeli.'		
EFT:	'Soon after he started out, he met someone else who was riding a bicycle.'		

Ssc22:	***Naamaaha***	***aramwihëënërya.***
SWW:	nikaona	anampisha
EWW:	I saw	he is letting her pass (him)
MMG:	*n-aa-maah-a* 1SG-PST-see-FV	*a-ra-mu-i-hëënëry-a* 3SG-PROG-3SG.OBJ-REFL-remove.APPL-FV[266]
SFT:	'Nikaona anampisha.'	
EFT:	'I saw her passing him.'	

Ssc23a:	*Bhuyö*	***yaarë***	***amwihëënëëyi***
SWW:	huo	alikuwa	anampisha
EWW:	at that time	she was	she is passing him
MMG:	*bhu-yö* AGR14-RDEM.REL	*a-aa-rë* 3SG-PST-COP	*a-mu-i-hënëëyi* 3SG-3SG.OBJ-REFL-remove.APPL.ANT

Ssc23b:	*igho,*	***ariituuratuura***	*këmwë*	*ku-righena,*
SWW:	tu	anajikwaa	mara moja	kwenye jiwe
EWW:	only	he stumbles (himself)	at once	to a stone
MMG:		*a-ra*[267]*-i-tuura~tuur-a* 3SG-PROG-REFL-stumble~RDPL-FV	*kë-mwë* NC7-one	*ku=ri-ghena* NC17=NC5.SG.stone

Ssc23c:	***araghwa***	*hansë.*
SWW:	anaanguka	chini
EWW:	he is falling	down
MMG:	*a-ra-ghw-a* 3SG-PROG-fall-FV	*ha-nsë* NC16-down
SFT:	'Walipopishana, kijana akagonga jiwe, akaanguka chini.'	
EFT:	'As they passed by each other, the boy hit a stone and fell down.'	

[266] The sense of this construction is that the boy is removing himself from the place where the girl is passing. More simply he is letting her pass.

[267] Although an *a* final morpheme does not normally assimilate to the place of an *i-* in the following morpheme (Compton & Walker 2010: 19-20), the REFL morpheme is an exception and does cause assimilation. This exception has not been previously described.

Ssc24a:	*Sëhaghwë*	*siyö*	***seranyaragana,***
SWW:	matunda	hayo	yanatawanyika
EWW:	fruits	those	they are scattered
MMG:	*së-haghwë* NC10.PL-fruit	*si-yö* AGR10-RDEM	*se-ra-nyarag-an-a* SBJ10-PROG-scatter-RECP-FV

Ssc24b:	***serashamboka***	***ukuya***	*ahare*	*iyö.*
SWW:	yanaruka	kwenda	mbali	huko
EWW:	they are bouncing	to go	far	there
MMG:	*se-ra-shambok-a* SBJ10-PROG-bounce-FV	*u-ku-y-a* AUG15-INF-go-FV		*i-yö* AGR23-RDEM
SFT:	'Matunda yakaruka na kutawanyika kote kote.'			
EFT:	'The fruit spilled out and scattered all over.'			

Ssc25a:	*Murikyaye*	*oora*	*we*
SWW:	mwenzake	yule	yeye
EWW:	companion	that	she
MMG:	*mu-riky-aye* NC1.SG-companion-3SG.POSS	*o-ora* AGR1-DDEM	*we* 3SG.PRON

Ssc25b:	***akaheteerania***	*këmwë.*
SWW:	akapitiliza	moja kwa moja
EWW:	she passed (by him)	straight (away)
MMG:	*a-ka-het-eer-an-i-a* 3SG-NAR-pass.by-APPL-RECP-CAUS-FV	*kë-mwë* NC7-one
SFT:	'Yule mwingine akaendelea na safari.'	
EFT:	'The girl just passed by and left.'	

Ssc26a:	*Rakini*	*ghwiki*	***naamaaha***	*abhandë*
SWW:	lakini	tena	nikaona	wengine
EWW:	but	again	I saw	other (people)
MMG:			*n-aa-maah-a* 1SG-PST-see-FV	*a-bha-ndë* AUG2-NC2.PL-other

Ssc26b:	***bharoorekerrania,***	***bharoorokania***	*igha:*
SWW:	wanaelekezana	wanaonyeshana	kwamba
EWW:	they are headed toward each other	they are showing each other	that
MMG:	*bha-ra-orekerr-an-i-a* 3PL-PROG-head.toward-RECP-CAUS-FV	*bha-ra-orok-an-i-a* 3PL-PROG-show-RECP-CAUS-FV	

Ssc26c:	*"**Tamaaha***	*ömöntö*	***aghööyë***	*hansë*	*haara.*
SWW:	haoni	mtu	ameanguka	chini	pale
EWW:	he does not see	person	he has fallen	down	there
MMG:	*t-a-maah-a* NEG-3SG-see-FV	*ö-mö-ntö* AUG1-NC1.SG-person	*a-gho-irë* 3SG-fall-ANT	*ha-nsë* NC16-down	*ha-ara* NC16-DDEM
SFT:	'Lakini tena nikaona watu wengine wakielekezana, wakisema kwamba, "Angalia, kuna mtu ameanguka chini!".'				
EFT:	'But then I saw some others heading toward the boy, saying, "Did you see that? Somebody fell down over there!'				

Ssc27a:	*Abhantö*	*bhatatö*	***bhaasha***	*këmwë,*
SWW:	watu	watatu	wakaja	mara moja
EWW:	people	three	they came	at once
MMG:	*a-bha-ntö* AUG2-NC2.PL-person	*bha-tatö* AGR2-three	*bha-aa-ash-a* 3PL-PST-come-FV	*kë-mwë* NC7-one

Ssc27b:	*bhöngöbhöngö*	***bhaamötöörya***	***ukushoora***
SWW:	harakaharaka	wakamsaidia	kuokota
EWW:	very quickly	they helped him	to pick up
MMG:	*bhö-ngö~bhöngö* NC14-quick~RDPL	*bha-aa-mö-töörу-a* 3PL.PST-3SG.OBJ-help-FV	*u-ku-shoor-a* AUG15-INF-pick.up-FV

Ssc27c:	*sëhaghwë*	*siyö,*	*bhayö*	***bhaaghya***	*na*
SWW:	matunda	hayo	hao	wakaenda	na
EWW:	fruits	those	those	they left	and
MMG:	*së-haghwë* NC10.PL-fruit	*si-yö* AGR10-RDEM	*bha-yö* AGR2-RDEM	*bha-aa-ghy-a* 3PL-PST-go-FV	*na* CCONJ

Ssc27d:	***bhaatoora***	*mu-ritonga*	*iryaye,*
SWW:	wakaweka	tengani	mwake
EWW:	they put	in basket	his
MMG:	*bha-aa-toor-a* 3PL-PST-put-FV	*mu=ri-tonga* NC18= NC5.SG-basket	*i-ri-aye* AUG5-NC5-3SG.POSS

Ssc27e:	***bhaamöghörrökya,***	***bhaamweya***	*urutu.*
SWW:	wakamwinua	wakamfuta	vumbi
EWW:	they lifted him up	they wiped him off	dust
MMG:	*bha-aa-mö-ghörröky-a* 3PL-PST-3SG.OBJ-lift.up-FV	*bha-aa-mu-ey-a* 3PL-PST-3SG.OBJ-wipe-FV	*u-ru-tu* AUG11-NC11-dust
SFT:	'Watu watatu wakafika mara moja, wakamsaidia kuokota haraka haraka yale matunda, wakaweka ndani ya tenga lake, wakamwinua na kumfuta vumbi.'		
EFT:	'Three boys came, quickly helped him pick up the fruit and put it in the basket. They also helped him up and dusted him.'		

Ssc28a:	*Na*	***yaariina***	*ko-bhaisikëëri*	*iyaaye,*
SWW:	na	akapanda	kwenye baiskeli	yake
EWW:	and	he climbed	on bicycle	his
MMG:	*na* CCONJ	*a-aa-riin-a* 3SG-PST-climb-FV	*ko=Ø-bhaisikëëri* NC17=NC9.SG-bicycle	*i-ya-aye* AUG9-AGR9-3SG.POSS

Ssc28b:	*ghwiki*	***yaawanza***	***okokenya.***
SWW:	tena	akaanza	kukimbia
EWW:	again	he started	to run
MMG:		*a-aa-wanz-a* 3SG-PST-start-FV	*o-ko-keny-a* AUG15-INF-run-FV
SFT:	'Naye akapanda juu baiskeli yake, akaanza kuendesha.'		
EFT:	'He got on his bike and started to ride off.'		

Ssc29:	***Ataraahika***	*ahare*	*igho,*	***bhakamobherekera.***
SWW:	kabla hajafika	mbali	tu	wakamwita
EWW:	he will not arrive	far	only	they called him
MMG:	*a-ta-raa-hik-a* 3SG-NEG-F_1-arrive-FV			*bha-ka-mo-bherek-er-a* 3PL-NAR-3SG.OBJ-call-APPL-FV
SFT:	'Kabla hajaenda mbali, wakamwita.'			
EFT:	'Before he got far, they called to him.'			

Ssc30a:	*Ghwiki*	*uwöndë*	***akamotemera***
SWW:	tena	mwingine	akampigia
EWW:	again	another	he hit for him
MMG:		*u-wö-ndë* AUG1-NC1.SG-other	*a-ka-mo-tem-er-a* 3SG-NAR-3SG.OBJ-hit-APPL-FV

Ssc30b:	*ömöröryö,*	*'Fi*	*fyuuuu.'*
SWW:	mbinja	fi	fyuuuu
EWW:	whistle	fi	fyuuuu
MMG:	*ö-mö-röryö* AUG3-NC3.SG-whistle		
SFT:	'Mwingine akampigia mbinja, "Fi fyuuuu."'		
EFT:	'One of them whistled to him, "Fi fyuuuu."'		

Ssc31:	*Këmwë*	***akaisyörya***	*nyuma.*
SWW:	mara moja	akageuka	nyuma
EWW:	at once	he turned (himself)	back
MMG:	*kë-mwë* NC7-one	*a-ka-i-syöry-a* 3SG-NAR-REFL-turn-FV	
SFT:	'Mara moja akageuka nyuma.'		
EFT:	'He turned around immediately.'		

Ssc32a:	*Bhuyö*	***yaaghëëyë***	***kwisyörya***	*igho,*
SWW:	hivyo	alienda	kugeuka	tu
EWW:	like that	he went	to turn (himself)	only
MMG:	*bhu-yö* AGR14-RDEM.REL	*a-aa-ghe-irë* 3SG-P_3/P_4-go-ANT	*ku-i-syöry-a* INF-REFL-turn-FV	

Ssc32b:	***akamaaha***	***bharamöörökya***	*ëngöfiira,*
SWW:	akaona	wanamwonyesha	kofia
EWW:	he saw	they are showing him	hat
MMG:	*a-ka-maah-a* 3SG-NAR-see-FV	*bha-ra-mö-öröky-a* 3PL-PROG-3SG.OBJ-show-FV	*ën-göfiira* NC9.SG-hat

Ssc32c:	***akaimeera.***
SWW:	akasimama
EWW:	he stopped (himself)
MMG:	*a-ka-emeer-a* 3SG-NAR-stop-FV
SFT:	'Alivyogeuka tu, akaona wanamwonyesha kofia, akasimama.'
EFT:	'Just as he turned around, he saw them showing him the hat, so he stopped.'

Ssc33a:	*Uwöndë*	*ono*	***akamoheera***
SWW:	mwingine	huyu	akampelekea
EWW:	another	this	he took to him
MMG:	*u-wö-ndë* AUG1-NC1.SG-other	*o-no* NC1-PDEM	*a-ka-mo-heer-a* 3SG-NAR-3SG.OBJ-take.to-FV

Ssc33b:	*ëngöfiira*	*bhöngöbhöngö.*
SWW:	kofia	harakaharaka
EWW:	hat	very quickly
MMG:	*ën-göfiira* NC9.SG-hat	*bhö-ngö~bhöngö* NC14-quick~RDPL
SFT:	‘Mmoja wao akampelekea kofia kwa haraka.’	
EFT:	‘One of them took the hat to him quickly.’	

Ssc34a:	*Bhuyö*	***yaaghëëyë,***	***yaahëkirë***	*haara,*
SWW:	hivyo	alienda	alifika	pale
EWW:	like that	he went	he has arrived	there
MMG:	*bhu-yö* AGR14-RDEM	*a-aa-ghe-irë* 3SG-P_3/P_4-go-ANT	*y-aa-hëk-irë* 3SG-P_3/P_4-arrive-ANT	*ha-ara* NC16-DDEM

Ssc34b:	***akamohaana.***
SWW:	akampa
EWW:	he gave him
MMG:	*a-ka-mo-haan-a* 3SG-NAR-3SG.OBJ-give-FV
SFT:	‘Akaenda, akamkabidhi.’
EFT:	‘He went and gave the hat to him.’

Ssc35a:	*Na*	*wonswe*	***akamohaana***	*ghwiki*	*sëhaghwë*
SWW:	na	naye	akampa	tena	matunda
EWW:	with	him also	he gave him	again	fruits
MMG:	*na* CCONJ	*w-onswe* AGR1-all	*a-ka-mo-haan-a* 3SG-NAR-3SG.OBJ-give-FV		*së-haghwë* NC10.PL-fruit

Ssc35b:	*kama*	*isatö.*
SWW:	kama	matatu
EWW:	like	three
MMG:		*i-satö* AGR10-three
SFT:	‘Naye huyo kijana akampa kiasi cha matunda kama matatu.’	
EFT:	‘And that boy gave him three pears in turn.’	

Ssc36a:	*Na*	*sëhaghwë*	*isatö*	*siyö*	***akaya,***
SWW:	na	matunda	matatu	hayo	akaenda
EWW:	with	fruits	three	those	he went
MMG:	*na* CCONJ	*së-haghwë* NC10.PL-fruit	*i-satö* AGR10-three	*si-yö* AGR10-RDEM	*a-ka-y-a* 3SG-NAR-go-FV

Ssc36b:	***akatwenania***	*bharikyaye*	*bhaara*
SWW:	akagawia	wenzake	wale
EWW:	he divided up	his companions	those
MMG:	*a-ka-twen-an-i-a* 3SG-NAR-divide-RECP-CAUS-FV	*bha-riky-aye* NC2.PL-companion-3SG.POSS	*bha-ara* AGR2-DDEM

Ssc36c:	*bhöngöbhöngö,*	***bhakarya.***
SWW:	harakaharaka	wakala
EWW:	very quickly	they ate
MMG:	*bhö-ngö~bhöngö* NC14-quick~RDPL	*bha-ka-ry-a* 3PL-NAR-eat-FV
SFT:	‘Na hayo matunda matatu akayagawa kwa wenzake, mara wakaanza kula.’	
EFT:	‘And the three pieces of fruit he gave to his friends, and they immediately ate them.’	

Ssc37a:	*Na*	*abhatatö*	*bhayö*	***bhakaheta***	*ghwiki*
SWW:	na	watatu	hao	wakapita	tena
EWW:	with	three	those	they passed by	again
MMG:	*na* CCONJ	*a-bha-tatö* AUG2-NC2.PL-three	*bha-yö* AGR2-RDEM	*bha-ka-het-a* 3PL-NAR-pass.by-FV	

Ssc37b:	*haang’ë*	*na*	*amatonga*	*ghaara.*
SWW:	karibu	na	matenga	yale
EWW:	near	with	baskets	those
MMG:	*ha-ang’ë* NC16-near	*na* CCONJ	*a-ma-tonga* AUG6-NC6.PL-basket	*gha-ara* AGR6-DDEM
SFT:	‘Na hao watu watatu wakapita karibu na yale matenga.’			
EFT:	‘And later on, those three boys walked close to the baskets (by the tree).’			

Ssc38a:	*Na*	***okanyoora***	*ömöntö*	*oora*
SWW:	na	ukakuta	mtu	yule
EWW:	with	you found	person	that
MMG:	*na* CCONJ	*o-ka-nyoor-a* 2SG-NAR-find-FV	*ö-mö-ntö* AUG1-NC1.SG-person	*o-ora* AGR1-DDEM.REL

Ssc38b:	***yaarë***	***ukutwa***	*sëhaghwë*	*seera*
SWW:	alikuwa	kuchuma	matunda	yale
EWW:	he was	to pick	fruits	those
MMG:	*a-aa-rë* 3SG-PST-COP	*u-ku-tw-a* AUG15-INF-pick-FV	*së-haghwë* NC10.PL-fruit	*se-era* AGR10-DDEM

Ssc38c:	*mu-risisi*	*moora*	***aituumirë***	***akamaaha***
SWW:	porini	mle	ameshuka	akaona
EWW:	in the bush	in there	he has gone down	he saw
MMG:	*mu=ri-sisi* NC18=NC5.SG-bush	*mo-ora* AGR18-DDEM	*a-i-tuum-irë* 3SG-REFL-go.down-ANT	*a-ka-maah-a* 3SG-NAR-see-FV

Ssc38d:	*n-kya*	*ono*	***aghëghëëywë***
SWW:	ni kama	ambaye	amechukuliwa
EWW:	it is like	who	he has been taken from
MMG:	*n=kya* COP=like	*o-no* AGR1-PDEM.REL	*a-ghëgh-ëëywë*[268] 3SG-carry-PASS.ANT

Ssc38e:	*sëhaghwë*	*isyaye.*
SWW:	matunda	yake
EWW:	fruits	his
MMG:	*së-haghwë* NC10.PL-fruit	*i-si-aye* AUG10-AGR10-3SG.POSS
SFT:	'Mtu yule aliyekuwa akichuma matunda akawa ameshuka na kushangaa kuona matunda yamepungua.'	
EFT:	'The man who had been picking the fruit came down and was shocked to see some of his fruit had disappeared.'	

Ssc39a:	*Na*	*bhano*	*hano*	***bhahëtirë***	*haara*
SWW:	na	hawa	hapa	wamepita	pale
EWW:	with	these	here	they have passed by	there
MMG:	*na* CCONJ	*bha-no* NC2-PDEM	*ha-no* NC16-PDEM.REL	*bha-hët-irë* 3PL-pass.by-ANT	*ha-ara* NC16-DDEM

[268] This consists of the Passive morpheme *-w* and the Anterior morpheme *-irë*. Additionally, a process of Imbrication has been applied.

Ssc39b:	***te-bhaatöniibhwi***	*na*	*sëhaghwë*	*seera*	***hë.***
SWW:	hawakuyajali	na	matunda	yale	NEG
EWW:	they were not concerned	with	fruits	those	NEG
MMG:	*te=bha-aa-töniibh-w-i* NEG=3PL-PST-concern-PASS-ANT	*na* CCONJ	*së-haghwë* NC10.PL-fruit	*se-era* AGR10-DDEM	*hë* NEG
SFT:	'Na hao watu watatu walipopita karibu na yale matunda, hawakuyajali.'				
EFT:	'When the three boys had passed by the baskets, they hadn't touched them.'				

Ssc40a:	*Këmwë*	*we*	***akabha***	*kya*	*ono*	*ghwiki*
SWW:	mara moja	yeye	akawa	kama	ambaye	tena
EWW:	at once	he	he became	like	who	again
MMG:	*kë-mwë* NC7-one	*we* 3SG.PRON	*a-ka-bh-a* 3SG-NAR-become-FV		*o-no* NC1-PDEM.REL	

Ssc40b:	***akwishaabha***	*kö-mötë*	*haara,*	*bhano*
SWW:	anayejificha	kwenye mti	pale	hawa
EWW:	who is hiding himself	at the tree	there	those (people)
MMG:	*a-ku-i-shaabh-a* 3SG-PRS-REFL-hide-FV	*kö=mö-të* NC17=NC3.SG-tree	*ha-ara* NC16-DDEM	*bha-no* NC2-PDEM

Ssc40c:	***bhakaheta.***
SWW:	wakapita
EWW:	they passed by
MMG:	*bha-ka-het-a* 3PL-NAR-pass.by-FV
SFT:	'Mara moja yeye akawa kama mtu anayejificha nyuma ya mti, hao watu wakapita.'
EFT:	'Immediately the man hid himself behind the tree and the boys passed by.'

Ssc41a:	*Bhuyö*	***bhaarë***	***bhaghëëyë***	***okoheta***
SWW:	huo	walikuwa	wameenda	kupita
EWW:	at that time	they were	they have gone	to pass by
MMG:	*bhu-yö* AGR14-RDEM.REL	*bha-aa-rë* 3PL-PST-COP	*bha-ghe-irë* 3PL-go-ANT	*o-ko-het-a* AUG15-INF-pass.by-FV

Ssc41b:	*igho*	*këmwë,*	*we*	***yaawanza***	*ghwiki,*	***yaariina***
SWW:	tu	mara moja	yeye	alianza	tena	alipanda
EWW:	only	at once	he	he started	again	he climbed
MMG:		*kë-mwë* NC7-one	*we* 3SG.PRON	*a-aa-wanz-a* 3SG-PST-start-FV		*a-aa-riin-a* 3SG-PST-climb-FV

Ssc41c:	*kö-mötë*	*koora,*	***yaatwa***	*sëhaghwë*
SWW:	kwenye mti	kule	alichuma	matunda
EWW:	to the tree	there	he picked	fruits
MMG:	*kö=mö-të* NC17=NC3.SG-tree	*ko-ora* AGR17-DDEM	*a-aa-tw-a* 3SG-PST-pick-FV	*së-haghwë* NC10.PL-fruit

Ssc41d:	*seera,*	*ghwiki*	***yaatoora***	*mo-matonga*
SWW:	yale	tena	aliweka	matengani
EWW:	those	again	he put	in the baskets
MMG:	*se-era* AGR10-DDEM		*a-aa-toor-a* 3SG-PST-put-FV	*mo=ma-tonga* NC18=NC6.PL-basket

Ssc41e:	*moora.*
SWW:	mle
EWW:	in there
MMG:	*mo-ora* AGR18-DDEM
SFT:	'Walipokwisha kupita kwenye eneo hilo, yule mtu akapanda tena kuchuma yale matunda na kuyaweka kwenye matenga yake.'
EFT:	'When they had left the area, the man climbed up his tree again, picked fruit, and put it in his baskets.'

Ssc42a:	*Bhoono,*	*hake*	*igho,*	***naamaaha***	*abhantö*
SWW:	sasa	kidogo	tu	nikaona	watu
EWW:	now	little	only	I saw	people
MMG:		*ha-ke* NC16-little		*n-aa-maah-a* 1SG-PST-see-FV	*a-bha-ntö* AUG2-NC2.PL-person

Ssc42b:	*bhaara,*	*eeee,*	***bhahëtirë***	***bharaya.***
SWW:	wale	eeee	wamepita	wanaenda
EWW:	those	eeee	they have passed	they are going
MMG:	*bha-ara* AGR2-DDEM		*bha-hët-irë* 3PL-pass.by-ANT	*bha-ra-y-a* 3PL-PROG-go-FV
SFT:	'Baada ya kidogo, nikaona wale watu wamepita wanaenda zao.'			
EFT:	'After a little while, I saw those boys had passed on to their homes.'			

Ssc43a:	*Na*	*këmwë*	***naanga***	***okomanya***	*keno*
SWW:	na	mara moja	nilishindwa	kujua	ambacho
EWW:	and	at once	I failed	to know	which
MMG:	*na* CCONJ	*kë-mwë* NC7-one	*n-aa-ang-a* 1SG-PST-fail-FV	*o-ko-many-a* AUG15-INF-know-FV	*ke-no* NC7-PDEM

Ssc43b:	***këkögëndëërya,***	*na*	***bhwarë***	***bhökëëyë.***
SWW:	kinaendelea	na	kulikuwa	kumekucha
EWW:	it is continuing	and	it was	night had fallen
MMG:	*kë-kö-gëndëëry-a* SBJ7-PRS-continue-FV	*na* CCONJ	*bho-a-rë* SBJ14-P_1/P_2-COP	*bho-ke-irë* SBJ14-be.nighttime-ANT
SFT:	‘Mara nikashindwa kuelewa jambo linaloendelea na kulikuwa kumekucha.’			
EFT:	‘Then I couldn’t understand what happened after that and night came.’			

Ssc44:	*Bhoono,*	***tigha***	***ndenge***	*hayö.*
SWW:	sasa	acha	niishie	hapo
EWW:	now	let	I should finish	here
MMG:		*Ø-tigh-a* IMP.SG-leave-FV	*n-reng-e* 1SG-finish-SBJV	*ha-yö* NC16-RDEM
SFT:	‘Ebu sasa niishie hapo.’			
EFT:	‘Let me finish now.’			

Appendix F: "Pear Story" in Zanaki

Since the original text of the "Pear Story" in Zanaki was translated first into Swahili, rather than directly into English, the following transcription of the story is a six-row interlinear text which includes word-for-word and free translations into both Swahili and English, as well as a morpheme-by-morpheme English gloss. The first row of each interlinearized section has the Zanaki text of the story and is titled Zak, followed by a sentence and line number (i.e. 1a, 1b, 2, etc.). These designations are used in the body of the thesis to refer to specific portions of this Zanaki text. The abbreviations for the other rows, in the order they appear, are: SWW (Swahili Word-for-Word), EWW (English Word-for-Word), MMG[269] (Morpheme-by-Morpheme Gloss), SFT (Swahili Free Translation) and EFT (English Free Translation). The latter two are included only on the last line of a sentence. All verbal words are bolded in the Zanaki text.

Zanaki Pear Story

Zak1a:	*Reero*	***ni***	*orusiku*	*rwe*	*etarehe*
SWW:	leo	ni	siku	ya	tarehe
EWW:	today	is	day	of	date
MMG:		*ni* COP	*o-ru-siku* AUG11-NC11.SG-day	*ru*[270]*-e* AGR11-AUG9.ASC	*e-Ø-tarehe* AUG9-NC9.SG-date

Zak1b:	*miroongo ebhiri*	*na*	*muhuungati.*
SWW:	ishirini	na	saba
EWW:	twenty	and	seven
MMG:	*mi-roongo e-bhiri* NC4.PL-ten AGR4-two	*na* CCONJ	
SFT:	'Leo ni tarehe ishirini na saba.'		
EFT:	'Today is the 27th (of the month).'		

[269] The MMG shows the underlying forms of the morphemes.
[270] The *u* (phonetically *ʊ*) in the prefix is labialized before a vowel-initial following morpheme, both here and other places in the text (Gray & Smith 2010b: 14).

Zak2:	*Rino*	***ni***	*erigano*	*rye*	*Ekizanaki.*
SWW:	hii	ni	hadithi	ya	Kizanaki
EWW:	this	is	story	of	Zanaki
MMG:	*ri-no* AGR5-PDEM	*ni* COP	*e-ri-gano* AUG5-NC5.SG-story	*ri*[271]*-e* AGR5-AUG7.ASC	*e-ki-zanaki* AUG7-NC7-zanaki
SFT:	'Hii ni hadithi ya Kizanaki.'				
EFT:	'This is a story in Zanaki.'				

Zak3a:	*Erigano*	*rino,*	***ndarora,***	***rirahusu***
SWW:	hadithi	hii	naona	
EWW:	story	this	I see	
MMG:	*e-ri-gano* AUG5-NC5.SG-story	*ri-no* AGR5-PDEM	*n-ra*[272]*-ror-a* 1SG-PROG-see-FV	*ri-ra-husu* SBJ5-PROG-be.about

Zak3b:	*abhaatu*	*bhatatu,*	*na*	*abhaatu*	*bhatatu*
SWW:	watu	watatu	na	watu	watatu
EWW:	people	three	and	people	three
MMG:	*a-bha-atu* AUG2-NC2.PL-person	*bha-tatu* AGR2-three	*na* CCONJ	*a-bha-atu* AUG2-NC2.PL-person	*bha-tatu* AGR2-three

Zak3c:	*bhano,*	*owumwe*	***arabhirikirwa***	*Waambura,*
SWW:	hawa	mmoja	anaitwa	Wambura
EWW:	these	one	he is called	Wambura
MMG:	*bha-no* AGR2-PDEM	*o-wu-mwe* AUG1-NC1.SG-one	*a-ra-bhirikir-w-a* 3SG-PROG-call-PASS-FV	

Zak3d:	*owuundi*	***arabhirikirwa***	*Kusaya,*	*na*
SWW:	mwingine	anaitwa	Kusaya	na
EWW:	another	he is called	Kusaya	and
MMG:	*o-wu-ndi* AUG1-NC1.SG-another	*a-ra-bhirikir-w-a* 3SG-PROG-call-PASS-FV		*na* CCONJ

[271] The *i* (phonetically *ɪ*) in the prefix is palatalized before a vowel-initial following morpheme, both here and other places in the text (Gray & Smith 2010b: 14).
[272] An underlying *r-* surfaces as *d-* when it follows a nasal, both here and other places in the text (Gray & Smith 2010b: 8).

Zak3e:	*owuundi*	*Waandibha.*
SWW:	mwingine	Wandiba
EWW:	another	Wandiba
MMG:	*o-wu-ndi* AUG1-NC1.SG-another	
SFT:	'Hadithi hii inahusu watu watatu: Wambura, Kusaya na Wandiba.'	
EFT:	'This story is about three people: Wambura, Kusaya and Wandiba.'	

Zak4a:	*Bhoono*	*Kusaya*	***yaari***	***agiiri***	*mwiisiku*
SWW:	sasa	Kusaya	alikuwa	ameenda	uani
EWW:	now	Kusaya	he was	he has gone	into the property
MMG:			*a-aa-ri*[273] 3SG-PST-COP	*a-gi-iri* 3SG-go-ANT	*mu=i-siku* NC18=NC5.SG-property

Zak4b:	*ryaye*	*ryo*	*omuguundu.*
SWW:	mwake	mwa	shamba
EWW:	his	of	farm
MMG:	*ri-aye* AGR5-3SG.POSS	*ri-o* AGR5-AUG3.ASC	*o-mu-guundu* AUG3-NC3.SG-farm
SFT:	'Sasa, Kusaya alikuwa ameenda shambani mwake.'		
EFT:	'Now, Kusaya had gone out to his farm.'		

Zak5a:	***Agiiri***	***kutwa***	*amachuungwa*	*mumuguundu*
SWW:	ameenda	kuchuma	machungwa	shambani
EWW:	he has gone	to pick	oranges	on the farm
MMG:	*a-gi-iri* 3SG-go-ANT	*ku-tw-a* INF-pick-FV	*a-ma-chuungwa* AUG6-NC6.PL-orange	*mu=mu-guundu* NC18=NC3.SG-farm

Zak5b:	*gwaaye.*
SWW:	mwake
EWW:	his
MMG:	*gu-aye* AGR3-3SG.POSS
SFT:	'Alikuwa ameenda kuchuma machungwa.'
EFT:	'He had gone to pick oranges.'

[273] A three-vowel sequence is disallowed and is shortened to a two-vowel sequence. When this long-vowel appears word-initially a *y* is inserted before the long vowel. This process needs to be further researched in Zanaki.

Zak6a:	*Ewe*	***yaari***	*mwiisiku*	*kwa sababu*	*hayo*
SWW:	yeye	alikuwa	uani	kwa sababu	hapo
EWW:	he	he was	on the property	because	there
MMG:	*ewe* 3SG.PRON	*a-aa-ri* 3SG-PST-COP	*mu=i-siku* NC18=NC5.SG-property	*kwa sababu* for reason	*ha-yo* NC16-RDEM

Zak6b:	***tukiigwa***	*ne*	*ezimburi*	***zirarira***
SWW:	tulisikia	na	mbuzi	zinalia
EWW:	we heard	with	goats	they are crying
MMG:	*tu-ka-igw-a* 1PL-NAR-hear-FV	*ne* CCONJ.AUG10	*e-zi-mburi* AUG10-NC10.PL-goat	*zi-ra-rir-a* SBJ10-PROG-cry-FV

Zak6c:	*munyuumba*	*muyo.*
SWW:	nyumbani	humo
EWW:	in house	in there
MMG:	*mu=Ø-nyuumba* NC18=NC9.SG-house	*mu-yo* NC18-RDEM
SFT:	‘Alienda shambani kwa sababu mbuzi walikuwa wanalia nyumbani humo.’	
EFT:	‘He went out to the farm because the goats were making noise inside the house.’	

Zak7a:	*Bhoono*	*hano*	*Kusaya*	***yaari***	***kutwa***
SWW:	sasa	hapa	Kusaya	alikuwa	kuchuma
EWW:	now	here	Kusaya	he was	to pick
MMG:		*ha-no* NC16-PDEM.REL		*a-aa-ri* 3SG-PST-COP	*ku-tw-a* INF-pick-FV

Zak7b:	*amachuungwa*	*gayo*,	***yaari***	*na*
SWW:	machungwa	hayo	alikuwa	na
EWW:	oranges	those	he was	with
MMG:	*a-ma-chuungwa* AUG6-NC6.PL-orange	*ga-yo* AGR6-RDEM	*a-aa-ri* 3SG-PST-COP	*na* CCONJ

Zak7c:	*ebhisehi*	*bhitatu.*
SWW:	vikapu	vitatu
EWW:	baskets	three
MMG:	*e-bhi-sehi* AUG8-NC8.PL-basket	*bhi-tatu* AGR8-three
SFT:	'Sasa, Kusaya alikuwa akichuma machungwa na alikuwa na vikapu vitatu.'	
EFT:	'Now, Kusaya had three baskets with him while he was picking oranges.'	

Zak8a:	*Na*	*kira*	*ekisehi*	***kyaari***
SWW:	na	kila	kikapu	kilikuwa
EWW:	and	each	basket	it was
MMG:	*na* CCONJ		*e-ki-sehi* AUG7-NC7.SG-basket	*ki-aa-ri* SBJ7-PST-COP

Zak8b:	***kiizwiirwemu***	*na*	*amachungwa*	*gayo.*
SWW:	kimejazwamo	na	machungwa	hayo
EWW:	it has been filled inside	with	oranges	those
MMG:	*ki-izur*[274]*-irwe=mu* SBJ7-be.full-PASS.ANT=LOC18	*na* CCONJ	*a-ma-chungwa* AUG6-NC6.PL-orange	*ga-yo* AGR6-RDEM
SFT:	'Na kila kikapu kilikuwa kimejazwa na machungwa.'			
EFT:	'Each basket had been filled with oranges.'			

Zak9a:	*Bhoono*	*hano*	*Kusaya*	***yaari***	***atiiriri***
SWW:	sasa	hapa	Kusaya	alikuwa	amepanda
EWW:	now	here	Kusaya	he was	he has climbed
MMG:		*ha-no* NC16-PDEM.REL		*a-aa-ri* 3SG-PST-COP	*a-tiir-iri* 3SG-climb-ANT

Zak9b:	*kumuti,*	***yaari***	***yiizwiirye***	*ebhisehi*
SWW:	kwenye mti	alikuwa	amejaza	vikapu
EWW:	on the tree	he was	it has been filled	baskets
MMG:	*ku=mu-ti* NC17=NC3.SG-tree	*a-aa-ri* 3SG-PST-COP	*ya-izur-irye* 3SG-be.full-CAUS.ANT	*e-bhi-sehi* AUG8-NC8.PL-basket

[274] The underlying form of the verb root *izur* surfaces as *izw* when followed by the *–iri* suffix (Gray & Smith 2010b: 17-18). This type of elision with *r*-final roots before the Anterior suffix is found other places in the text as well.

Zak9c:	*bhibhiri.*
SWW:	viwili
EWW:	two
MMG:	*bhi-bhiri* AGR8-two
SFT:	'Sasa, Kusaya, alipokuwa amepanda kwenye mti, alikuwa amejaza vikapu viwili.'
EFT:	'After Kusaya had climbed up the tree, he had filled two baskets (with oranges).'

Zak10a:	*Bhoono*	***yaareenda***	***kwiizuririrya***
SWW:	sasa	alikuwa anataka	kukijaza
EWW:	now	he wanted	to fill (it)
MMG:		*a-ara-end-a*[275] 3SG-PST.PROG-want-FV	*ku-izur-ir-iry-a* INF-be.full-APPL-CAUS-FV

Zak10b:	*kya*	*katatu.*
SWW:	cha	tatu
EWW:	of	third
MMG:	*ki-a* AGR7-ASC	*ka-tatu* AGR12-three
SFT:	'Sasa alitaka kukijaza cha tatu.'	
EFT:	'Now he was about to fill the third.'	

Zak11a:	*Bhoono*	*hano*	***yaari***	***kweenda***
SWW:	sasa	hapa	alikuwa	kutaka
EWW:	now	here	he was	to want
MMG:		*ha-no* NC16-PDEM.REL	*a-aa-ri* 3SG-PST-COP	*ku-end-a* INF-want-FV

Zak11b:	***kwiizuririrya***	*kya*	*katatu,*	*niho*
SWW:	kukijaza	cha	tatu	ndipo
EWW:	to fill (it)	of	third	then
MMG:	*ku-izur-ir-iry-a* INF-be.full-APPL-CAUS-FV	*ki-a* AGR7-ASC	*ka-tatu* AGR12-three	*ni-ho* FOC-NC16

[275] A long-vowel cannot surface word-initially so a *y* is inserted at the beginning of words like this.

Zak11c:	*hasuuhwiigo*	*Waandibha*	***akarweera***
SWW:	muda kidogo	Wandiba	akatokea
EWW:	little while	Wandiba	he came from
MMG:	*ha-suuhu=igo* NC16-little=like.that		*a-ka-rw-er-a* 3SG-NAR-come.from-APPL-FV

Zak11d:	*eyo*	*ne*	*emburi.*
SWW:	huko	na	mbuzi
EWW:	there	with	goat
MMG:	*e-yo* NC23-RDEM	*ne* CCONJ.AUG9	*e-m-buri* AUG9-NC9.SG-goat
SFT:	'Sasa, alipokuwa akitaka kujaza cha tatu, baada ya kidogo Wandiba akatokea na mbuzi.'		
EFT:	'While he was preparing to fill the third basket, after a little while, Wandiba came out with a goat.'		

Zak12:	*Waandibha*	*wuyo*	***yaaragya***	*wa*	*Kusaya.*
SWW:	Wandiba	huyo	alikuwa anakwenda	kwa	Kusaya
EWW:	Wandiba	that	he was going	to (place of)	Kusaya
MMG:		*wu-yo* AGR1-RDEM	*a-ara-gy-a* 3SG.PST.PROG-go-FV	*u-a* AGR23?[276]-ASC	
SFT:	'Wandiba alikuwa akienda kwa Kusaya.'				
EFT:	'Wandiba was headed toward Kusaya.'				

Zak13a:	*Hano*	***akagya,***	***akahika***	*wa*
SWW:	hapa	akaenda	akafika	kwa
EWW:	here	he went	he arrived	to (place of)
MMG:	*ha-no* NC16-PDEM.REL	*a-ka-gy-a* 3SG-NAR-go-FV	*a-ka-hik-a* 3SG-NAR-arrive-FV	*u-a* AGR23?[277]-ASC

Zak13b:	*Kusaya,*	***akabhona***	*Kusaya*	***akyaari***	*kumuti*
SWW:	Kusaya	akapata	Kusaya	bado	kwenye mti
EWW:	Kusaya	he found	Kusaya	still	at the tree
MMG:		*a-ka-bhon-a* 3SG-NAR-find-FV		*a-kyaa-ri* 3SG-PERS-COP	*ku=mu-ti* NC17=NC3.SG-tree

[276] In closely related Ikizu, a *wa* associative shows Class 23 agreement (Gray 2013: 34); however, more research is needed to establish this as Class 23 for Zanaki. In Zanaki, predicative adjectival agreement seems to be indicated by a Class 16 marker *ha-*, as the following sentence shows: *Wa Kusaya ni hazomu* 'Kusaya's place is nice'. This is also seen in Ikizu (Gray 2013: 38).

[277] See footnote in line Zak12.

Zak13c:	***aratwa***	*amachuungwa.*
SWW:	anachuma	machungwa
EWW:	he is picking	oranges
MMG:	*a-ra-tw-a* 3SG-PROG-pick-FV	*a-ma-chuungwa* AUG6-NC6.PL-orange
SFT:	'Alipoenda, akafika kwa Kusaya, akapata Kusaya bado anachuma machungwa kwenye mti.'	
EFT:	'When he reached Kusaya, he saw that Kusaya was still at the tree picking oranges.'	

Zak14a:	*Ewe*	***akahita***	*ne*	*emburi*
SWW:	yeye	akapita	na	mbuzi
EWW:	he	he passed by	with	goat
MMG:	*ewe* 3SG.PRON	*a-ka-hit-a* 3SG-NAR-pass.by-FV	*ne* CCONJ.AUG9	*e-m-buri* AUG9-NC9.SG-goat

Zak14b:	*yaye,*	***akagya***	*mumugye.*
SWW:	yake	akaenda	kwenye mji
EWW:	his	he went	to the compound
MMG:	*e-aye* AGR9-3SG.POSS	*a-ka-hit-a* 3SG-NAR-pass.by-FV	*mu=mu-gye* NC18=NC3.SG-compound
SFT:	'Wandiba akapita na mbuzi yake, akaenda kwenye mji.'		
EFT:	'Wandiba passed by with his goat and went into the compound.'		

Zak15a:	*Niho*	*hasuuhwiigo*	***akaaza***	*omwaana*
SWW:	ndipo	muda kidogo	akaja	mtoto
EWW:	then	little while	he came	child
MMG:	*ni-ho* FOC-LOC16	*ha-suuhu=igo* NC16-little=like.that	*a-ka-az-a* 3SG-NAR-come-FV	*o-mu-ana* AUG1-NC1.SG-child

Zak15b:	*owuundi,*	*no*	*omwaana*	*wuyo*
SWW:	mwingine	na	mtoto	huyo
EWW:	another	with	child	that
MMG:	*o-wu-ndi* AUG1-AGR1-other	*no* CCONJ.AUG1	*o-mu-ana* AUG1-NC1.SG-child	*wu-yo* NC1-RDEM

Zak15c:	***ndarora***	***kubha***	***yaari***	*ni omwiibhi.*
SWW:	ninaona	kuwa	alikuwa	ni mwizi
EWW:	I am seeing	that	he was	is a thief
MMG:	*n-ra-ror-a* 1SG-PROG-see-FV	*ku-bh-a* INF-be-FV	*a-aa-ri* 3SG-PST-COP	*ni=o-mu-ibhi* COP=AUG1-NC1.SG-thief
SFT:	'Halafu, baada ya kidogo, akaja mtoto fulani ambaye alionekana kuwa mwizi.'			
EFT:	'Then, after a little bit, I saw a kid coming who looked like a thief.'			

Zak16a:	*Omwaana*	*wuyo*	***yaari***	***yiibhohiri***
SWW:	mtoto	huyo	alikuwa	amevaa
EWW:	child	that	he was	he is wearing
MMG:	*o-mu-ana* AUG1-NC1.SG-child	*wu-yo* NC1-RDEM	*a-aa-ri* 3SG-PST-COP	*ya-ibhoh*[278]*-iri* 3SG-wear-ANT

Zak16b:	*engosira*	*ne*	*ebhasikeeri.*
SWW:	kofia	na	baiskeli
EWW:	hat	with	bicycle
MMG:	*e-n-gosira* AUG9-NC9.SG-hat	*n-e* CCONJ.AUG9	*e-Ø-bhasikeeri* AUG9-NC9.SG-bicycle
SFT:	'Mtoto huyo alikuwa amevaa kofia, akawa na baiskeli.'		
EFT:	'That boy was wearing a hat and was riding a bicycle.'		

Zak17a:	*Hano*	***akaaza,***	***akarora***	*Kusaya*
SWW:	hapa	akaja	akaona	Kusaya
EWW:	here	he came	he saw	Kusaya
MMG:	*ha-no* NC16-PDEM.REL	*a-ka-az-a* 3SG-NAR-come-FV	*a-ka-ror-a* 3SG-NAR-see-FV	

Zak17b:	***yaari***	***akyaari***	*kumuti*	*gurya*
SWW:	alikuwa	bado	kwenye mti	ule
EWW:	he was	still	on the tree	that
MMG:	*a-aa-ri* 3SG-PST-COP	*a-kyaa-ri* 3SG-PER-COP	*ku=mu-ti* NC17=NC3.SG-tree	*gu-rya* AGR3-DDEM

[278] The etymology of this verb root is a reflexive form (*i-*) of *boh* 'close' that has become lexicalized as *iboh* 'get dressed'. Semantic change from 'close oneself' to 'get dressed' is of interest typologically.

Zak17c:	***aratwa***	*amachuungwa.*
SWW:	anachuma	machungwa
EWW:	he is picking	oranges
MMG:	*a-ra-tw-a* 3SG-PROG-pick-FV	*a-ma-chuungwa* AUG6-NC6.PL-orange
SFT:	'Alipofika, akaona Kusaya aliyekuwa bado kwenye mti ule akichuma machungwa.'	
EFT:	'When he reached the tree, he saw Kusaya still picking oranges.'	

Zak18a:	*Hano*	***yaamuruuzi***	*Kusaya,*	*kyaamwe*	*Waandibha*
SWW:	hapa	alimwona	Kusaya	kisha	Wandiba
EWW:	here	he saw him	Kusaya	so	Wandiba
MMG:	*ha-no* NC16-PDEM.REL	*a-a-mu-ruuzi*[279] 3SG-PST-3SG.OBJ-see.ANT			

Zak18b:	***akarora***	*Kusaya*	***ari***	*iguru*	*bhukong'u.*
SWW:	akaona	Kusaya	yuko	juu	mno
EWW:	he saw	Kusaya	he is (there)	high	very
MMG:	*a-ka-ror-a* 3SG-NAR-see-FV		*a-Ø-ri* 3SG-PRS-COP	*i-guru* NC23-up	*bhu-kong'u* NC14-very
SFT:	'Alipomwona Kusaya, Wandiba akaona yuko juu sana.'				
EFT:	'When he saw Kusaya, Wandiba he observed that he was up very high.'				

Zak19a:	***Akagega***	*ekisehi*	*kimwe*	*kya*
SWW:	akabeba	kikapu	kimoja	cha
EWW:	he carried	basket	one	of
MMG:	*a-ka-geg-a* 3SG-NAR-carry-FV	*e-ki-sehi* AUG7-NC7.SG-basket	*ki-mwe* AGR7-one	*ki-a* AGR7-ASC

Zak19b:	*amachuungwa,*	***akatuura***	*ekisehi*	*kiyo*
SWW:	machungwa	akaweka	kikapu	hicho
EWW:	oranges	he put	basket	that
MMG:	*a-ma-chuungwa* AUG6-NC6.PL-orange	*a-ka-tuur-a* 3SG-NAR-put-FV	*e-ki-sehi* AUG7-NC7.SG-basket	*ki-yo* AGR7-RDEM

[279] This form deserves more attention from a comparative perspective, since it shows spirantization, which is very infrequently observed in the Mara languages. It comes from *ror* 'see' + *-iri* 'ANT'. A similarly spirantized form for this combination of morphemes occurs in Kabwa as well.

Zak19c:	*kubhasikeeri*	*yaye,*	***akahurura.***
SWW:	kwenye baiskeli	yake	akaanza safari
EWW:	on the bicycle	his	he started a trip
MMG:	*ku=Ø-bhasikeeri* NC17=NC9.SG-bicycle	*y-aye* AGR9-3SG.POSS	*a-ka-hurur-a* 3SG-NAR-start.trip-FV
SFT:	'Akachukua kikapu kimoja cha machungwa, akaweka kwenye baiskeli yake, akaanza safari.'		
EFT:	'He took one basket of oranges, put it on his bike and took off.'		

Zak20a:	*Hano*	***yaatanwiiri***	*no*
SWW:	hapa	aliondoka	na
EWW:	here	he left	with
MMG:	*ha-no* NC16-PDEM.REL	*a-a-tanur-iri* 3SG-PST-depart-ANT	*no* CCONJ.AUG11

Zak20b:	*orugeendo*	*rwaaye,*	***akagya.***
SWW:	safari	yake	akaenda
EWW:	journey	his	he went
MMG:	*o-ru-geendo* AUG11-NC11.SG-journey	*ru-aye* AGR11-3SG.POSS	*a-ka-gi-a* 3SG-NAR-go-FV
SFT:	'Alipoondoka na safari yake, akaenda.'		
EFT:	'When he left on his trip, he was gone.'		

Zak21a:	*Hano*	***yaagiiri,***	***akasikana***	*no*
SWW:	hapa	alienda	akakutana	na
EWW:	here	he left	he met	with
MMG:	*ha-no* NC16-PDEM.REL	*a-a-gi-iri* 3SG-PST-go-ANT	*a-ka-sik-an-a* 3SG-NAR-meet-RECP-FV	*no* CCONJ.AUG1

Zak21b:	*omuukya*	*wuundi,*	*woosi*	***ari***	*kubhasikeeri*
SWW:	binti	mwingine	naye	yuko	kwenye baiskeli
EWW:	girl	another	and she	she is (there)	on the bicycle
MMG:	*o-mu-ukya* AUG1-NC1.SG-girl	*wu-ndi* AGR1-other	*wa-osi* AGR1-all	*a-Ø-ri* 3SG-PRS-COP	*ku=Ø-bhasikeeri* NC17=NC9.SG-bicycle
SFT:	'Alipkuwa ameenda, akakutana na binti ambaye yuko kwenye baiskeli.'				
EFT:	'After he had gone on, he met a girl riding a bike.'				

Zak22a:	*Eno*	*ewe*	***yiibhiri***	*amachuungwa*
SWW:	huku	yeye	ameiba	machungwa
EWW:	here	he	he has stolen	oranges
MMG:	*e-no* NC23-PDEM	*ewe* 3SG.PRON	*a-ibh-iri* 3SG-steal-ANT	*a-ma-chuungwa* AUG6-NC6.PL-orange

Zak22b:	*kubhasikeeri*	*ne*	*ekikapu*	*kyaaye.*
SWW:	kwenye baiskeli	na	kikapu	chake
EWW:	on the bicycle	with	basket	his
MMG:	*ku=Ø-bhasikeeri* NC17=NC9.SG-bicycle	*ne* CCONJ.AUG7	*e-ki-kapu* AUG7-NC7.SG-basket	*ki-aye* AGR7-3SG.POSS
SFT:	‘Na huku ambaye ameiba yuko kwenye baiskeli na kikapu chake cha machungwa.’			
EFT:	‘And the one who stole the oranges was there with them in the basket on his bike.’			

Zak23a:	*Hano*	***yaasikeeni***	*no*	*omuukya*
SWW:	hapa	alikutana	na	binti
EWW:	here	he left	with	girl
MMG:	*ha-no* NC16-PDEM.REL	*a-a-sik-eeni* 3SG-PST-meet-RECP.ANT	*no* CCONJ.AUG1	*o-mu-ukya* AUG1-NC1.SG-girl

Zak23b:	*wuyo,*	***bhakiitemana***	*ne*	*ebhasikeeri.*
SWW:	huyo	wakagongana	na	baiskeli
EWW:	that	they ran into each other	with	bicycle
MMG:	*wu-yo* AGR1-RDEM	*bha-ka-i-tem-an-a* 3PL-NAR-REFL-hit-RECP-FV	*ne* CCONJ.AUG9	*e-Ø-bhasikeeri* AUG9-NC9.SG-bicycle
SFT:	‘Alipokutana na binti huyo, baiskeli wakagongana na baiskeli.’			
EFT:	‘When they reached the same spot (on the path), their bikes ran into each other.’			

Zak24a:	*Kyaamwe*	*omumura*	*wuyo,*	*amachuungwa*
SWW:	kisha	kijana	huyo	machungwa
EWW:	so	young man	that	oranges
MMG:		*o-mu-mura* AUG1-NC1.SG-young.man	*wu-yo* AGR1-RDEM	*a-ma-chuungwa* AUG6-NC6.PL-orange

Zak24b:	*gayo*	***gakiitika***	*haasi,*
SWW:	hayo	yakamwagika	chini
EWW:	those	they spilled	down
MMG:	*ga-yo* AGR6-RDEM	*ga-ka-itik-a* SBJ6-NAR-be.spilled-FV	*ha-esi*[280] NC16-NC9.SG.land

Zak24c:	*woosi*	***akagwa.***
SWW:	naye	akaanguka
EWW:	and he	he fell
MMG:	*wo-osi* AGR1-all	*a-ka-gw-a* 3SG-NAR-fall-FV
SFT:	'Kisha, kijana huyo akaanguka, machungwa yake yote yakamwagika.'	
EFT:	'So, the boy fell down and his oranges spilled everywhere.'	

Zak25a:	*Hano*	***yaagwiiri,***	*bhasi*	*omuukya*	*wurya*
SWW:	hapa	alianguka	basi	binti	yule
EWW:	here	he fell	then	girl	that
MMG:	*ha-no* NC16-PDEM.REL	*a-a-gw-iri* 3SG-PST-fall-ANT		*o-mu-ukya* AUG1-NC1.SG-girl	*wu-rya* AGR1-DDEM

Zak25b:	***akeenderera***	*no*
SWW:	akaendelea	na
EWW:	she continued	with
MMG:	*a-ka-enderer-a* 3SG-NAR-continue-FV	*no* CCONJ.AUG11

Zak25c:	*orugeendo*	*rwaaye.*
SWW:	safari	yake
EWW:	journey	her
MMG:	*o-ru-geendo* AUG11-NC11.SG-journey	*ru-aye* AGR11-3SG.POSS
SFT:	'Kijana alipoanguka, binti yule akaendelea na safari yake.'	
EFT:	'When the boy fell, that girl just continued on her journey.'	

[280] The fact that the vowel assimilates to *a* instead of *e* may have to do with the antiquity of this lexicalized form. A related form in Ikizu appears to have an unspecified vowel for this root *-Vsɨ* 'land' (Gray 2013: 15, 18, 21, 33). This is an area for further study.

Zak26a:	*Bhoono*	*hano*	*ewe*	***yaagwiiri,***	***akawaanza***
SWW:	sasa	hapa	yeye	alianguka	akaanza
EWW:	now	here	he	he fell	he started
MMG:		*ha-no* NC16-PDEM.REL	*ewe* 3SG.PRON	*a-a-gw-iri* 3SG-PST-fall-ANT	*a-ka-wanz-a* 3SG-NAR-start-FV

Zak26b:	***kukumania***	*amachuungwa*	*garya.*
SWW:	kukusanya	machungwa	yale
EWW:	to gather (together)	oranges	those
MMG:	*ku-kum-an-i-a* INF-gather-RECP-CAUS-FV	*a-ma-chuungwa* AUG6-NC6.PL-orange	*ga-rya* AGR6-DDEM
SFT:	'Kijana akaanza kukusanya machungwa yale yote.'		
EFT:	'Then the boy started to gather the oranges together.'		

Zak27a:	*Ne*	*engosira*	*yaaye*
SWW:	na	kofia	yake
EWW:	with	hat	his
MMG:	*ne* CCONJ.AUG9	*e-n-gosira* AUG9-NC9.SG-hat	*ya-aye* AGR9-3SG.POSS

Zak27b:	*yosi*	***ekagwa***	*haasi.*
SWW:	nayo	ikaanguka	chini
EWW:	it also	it fell	down
MMG:	*y-osi* AGR9-all	*e-ka-gw-a* SBJ9-NAR-fall-FV	*ha-esi* NC16-NC9.SG.land
SFT:	'Na kofia yake nayo ilikuwa imeanguka chini.'		
EFT:	'His hat had also fallen down.'		

Zak28a:	*Niho*	*hasuuhwiigo*	***bhakaaza***	*abhamura*
SWW:	ndipo	muda kidogo	wakaja	wavulana
EWW:	then	little while	they came	boys
MMG:	*ni-ho* FOC-LOC16	*ha-suuhu=igo* NC16-little=like.that	*bha-ka-az-a* 3PL-NAR-come-FV	*a-bha-mura* AUG2-NC2.PL-boy

Zak28b:	*abhaandi*	*bhatatu*	*neebho*	*bhosi*
SWW:	wengine	watatu	nao	pia
EWW:	other	three	with them	also
MMG:	*a-bha-ndi* AUG2-AGR2-other	*bha-tatu* AGR2-three	*na-ebho* CCONJ-3PL.PRON	*bha-osi* NC2.PL-all

Zak28c	*ni abhaana*	*abharikyaaye.*
SWW:	ni watoto	wenzake
EWW:	are children	his companions
MMG:	*ni=a-bha-ana* COP=AUG2-NC2.PL-child	*a-bha-riky-aye* AUG2-AGR2-companion-3SG.POSS
SFT:	'Baada ya kidogo, watoto watatu wengine wakaja ambao ni wenzake.'	
EFT:	'After a little while, three children who were the boy's friends came (along the path).'	

Zak29a:	*Hano*	*bhosi*	***bhakahika,***	***bhakabhona***
SWW:	hapa	pia	wakafika	wakapata
EWW:	here	also	they arrived	they found
MMG:	*ha-no* NC16-PDEM.REL	*bha-osi* NC2.PL-all	*bha-ka-hik-a* 3PL-NAR-arrive-FV	*bha-ka-bhon-a* 3PL-NAR-find-FV

Zak29b:	*wurya*	*kyeego*	*amachuungwa*	***giitikiri,***
SWW:	yule	jinsi	machungwa	yamemwagika
EWW:	that (boy)	how	oranges	they have spilled
MMG:	*wu-rya* NC1-DDEM		*a-ma-chuungwa* AUG6-NC6.PL-orange	*ga-itik-iri* SBJ6-be.spilled-ANT

Zak29c:	***bhakamusakirya***	***kugazoora,***	***bhakazoora***
SWW:	wakamsaidia	kuyakusanya	wakakusanya
EWW:	they helped him	to gather them	they gathered (them)
MMG:	*bha-ka-mu-sakiry-a* 3PL-NAR-3SG.OBJ-help-FV	*ku-ga-zoor-a* INF-OBJ6-gather-FV	*bha-ka-zoor-a* 3PL-NAR-gather-FV

Zak29d:	***bhakiizurya***	***mukisehi***	*murya.*
SWW:	wakajaza	katika kikapu	mle
EWW:	they filled	in basket	in there
MMG:	*bha-ka-izur-y-a* 3PL-NAR-be.full-CAUS-FV	*mu=ki-sehi* NC18=NC7.SG-basket	*mu-rya* AGR18-DDEM
SFT:	'Walipofika, wakapata mtoto akikusanya machungwa, wakamsaidia kuyakusanya, wakakijaza kikapu.'		
EFT:	'When the three boys reached him, they saw that he was gathering the oranges, so they stopped to help him and refilled the basket.'		

Zak30a:	***Bhakatuura***	*kubhasikeeri,*	*wurya*	***akahorura,***
SWW:	wakaweka	kwenye baiskeli	yule	akaongoza
EWW:	they put	on the bicycle	that	he led
MMG:	*bha-ka-tuur-a* 3PL-NAR-put-FV	*ku=Ø-bhasikeeri* NC17=NC9.SG-bicycle	*wu-rya* NC1.SG-DDEM	*a-ka-horur-a* 3SG-NAR-lead-FV

Zak30b:	***akeenderera***	*no*
SWW:	akaendelea	na
EWW:	he continued	with
MMG:	*a-ka-enderer-a* 3SG-NAR-continue-FV	*no* CCONJ.AUG11

Zak30c:	*orugeendo*	*rwaaye.*
SWW:	safari	yake
EWW:	journey	his
MMG:	*o-ru-geendo* AUG11-NC11.SG-journey	*ru-aye* AGR11-3SG.POSS
SFT:	‘Wakaweka kikapu kwenye baiskeli na mtoto huyo akaendelea na safari tena.’	
EFT:	‘The put the basket on his bike and then the boy started out again on his journey.’	

Zak31a:	*Hano*	***yaagiiri***	***okubhuuka,***	***akabhona***
SWW:	hapa	alienda	kuamka	akapata
EWW:	here	he left	to arise	he found
MMG:	*ha-no* NC16-PDEM.REL	*a-a-gi-iri* 3SG-PST-go-ANT	*o-ku-bhuuk-a* AUG15-INF-arise-FV	*a-ka-bhon-a* 3SG-NAR-find-FV

Zak31b:	***atigireho***	*engosira*	*yaye.*
SWW:	ameacha hapo	kofia	yake
EWW:	he left there	hat	his
MMG:	*a-tig-ire=ho* 3SG-leave-ANT=LOC16	*e-n-gosira* AUG9-NC9.SG-hat	*y-aye* AGR9-3SG.POSS
SFT:	‘Alipoenda kuamka, akapata ameachia kule kofia yake.		
EFT:	‘When he regained his senses, he realized that he had left his hat behind.’		

Zak32a:	*Bhoono*	*hano*	***yaakyaagya,***	***akagya,***	***akagya,***
SWW:	sasa	hapa	bado alienda	akaenda	akaenda
EWW:	now	here	he was still going	he went	he went
MMG:		*ha-no* NC16-PDEM	*a-a-kyaa-gi-a* 3SG-PST-PERS-go-FV	*a-ka-gi-a* 3SG-NAR-go-FV	*a-ka-gi-a* 3SG-NAR-go-FV

Zak32b:	***akagya,***	***akagya,***	***akasikana***	*na*
SWW:	akaenda	akaenda	akakutana	na
EWW:	he went	he went	he met	with
MMG:	*a-ka-gi-a* 3SG-NAR-go-FV	*a-ka-gi-a* 3SG-NAR-go-FV	*a-ka-sik-an-a* 3SG-NAR-meet-RECP-FV	*na* CCONJ

Zak32c:	*abhamura*	*abhaandi*	*bhatatu.*
SWW:	wavulana	wengine	watatu
EWW:	boys	other	three
MMG:	*a-bha-mura* AUG2-NC2.PL-boy	*a-bha-ndi* AUG2-AGR2-other	*bha-tatu* AGR2-three
SFT:	‘Sasa, bado alikuwa akienda, akaenda, akaenda, akaenda, akakutana na wavulana wengine watatu.’		
EFT:	‘He was still going down the path, on and on, and he met three more boys.’		

Zak33:	*Abhamura*	*bhayo*	***bhaari***	***bharabharaana.***
SWW:	wavulana	hao	walikuwa	wanacheza
EWW:	boys	those	they were	they are playing
MMG:	*a-bha-mura* AUG2-NC2.PL-boy	*bha-yo* AGR2-RDEM	*bha-aa-ri* 3PL-PST-COP	*bha-ra-bharaan*[281]*-a* 3PL-PROG-play-FV
SFT:	‘Wavulana hao walikuwa wanacheza.’			
EFT:	‘These boys were playing.’			

Zak34a:	*Hano*	***bhaamuruuzi,***	***bhakahitirana***
SWW:	hapa	walimwona	wakapishana
EWW:	here	they saw him	they passed by each other
MMG:	*ha-no* NC16-PDEM.REL	*bha-a-mu-ruuzi* 3PL-PST-3SG.OBJ-see.ANT	*bha-ka-hit-ir-an-a* 3PL-NAR-pass.by-APPL-RECP-FV

[281] This is superficially similar to *bhar* ‘count’ + *-an* ‘RECP’, but a semantic link between the two has not yet been established.

Zak34b:	*nawe.*
SWW:	naye
EWW:	with him
MMG:	*na-we* CCONJ-3SG.PRON
SFT:	'Walipomwona, wakaanza kupishana naye.'
EFT:	'After they saw him, they got to where he was (on the road).'

Zak35a:	*Bhoono*	*ewe*	***yaari***	***arahuna***
SWW:	sasa	yeye	alikuwa	anasukuma
EWW:	now	he	he was	he was pushing
MMG:		*ewe* 3SG.PRON	*a-aa-ri* 3SG-PST-COP	*a-ra-hun-a* 3SG-PROG-push-FV

Zak35b:	*ebhasikeeri*	*yaye*	*eno*	***arasugura.***	
SWW:	baiskeli	yake	huku	anachechemea	
EWW:	bicycle	his	here	he was limping	
MMG:	*e-Ø-bhasikeeri* AUG9-NC9.SG-bicycle	*y-aye* AGR9-3SG.POSS	*e-no* NC23-PDEM	*a-ra-sugur-a* 3SG-PROG-limp-FV	
SFT:	'Sasa, yeye alikuwa anasukuma baiskeli yake wakati alikuwa anachechemea.'				
EFT:	'Now, the boy was pushing his bicycle while limping.'				

Zak36a:	*Boono*	*kuumbe*	*hano*	***yaagwiiri***	*na*
SWW:	sasa	kumbe	hapa	alianguka	na
EWW:	now	amazingly	here	he fell	with
MMG:		*kumbe* EMPH	*ha-no* NC16-PDEM.REL	*a-a-gw-iri* 3SG-PST-fall-ANT	*na* CCONJ

Zak36b:	*amachuungwa*	***gakiitika,***	***akeebhaho***	
SWW:	machungwa	yakamwagika	akasahau pale	
EWW:	oranges	they spilled	he forgot there	
MMG:	*a-ma-chuungwa* AUG6-NC6.PL-orange	*ga-ka-itik-a* SBJ6-NAR-fall-FV	*a-ka-ebh-a=ho* 3SG-NAR-forget-FV=LOC16	

Zak36c:	*engosira*	*yaye.*
SWW:	kofia	yake
EWW:	hat	his
MMG:	*e-n-gosira* AUG9-NC9.SG-hat	*y-aye* AGR9-3SG.POSS
SFT:	'Sasa, alipoanguka na machungwa yalimwagika, kumbe akasahau kofia yake.'	
EFT:	'When the boy had fallen and the oranges had fallen, incredibly he had also forgotten his hat.'	

Zak37a:	*Abhamura*	*bhayo*	***bhakahitirania***
SWW:	wavulana	hao	wakapishana
EWW:	boys	those	they passed by each other
MMG:	*a-bha-mura* AUG2-NC2.PL-boy	*bha-yo* AGR2-RDEM	*bha-ka-hit-ir-an-i-a* 3PL-NAR-pass.by-APPL-RECP-CAUS-FV

Zak37b:	*nawe,*	***bhaari***	***bhana***	*enaamba*	*ya*
SWW:	naye	walikuwa	wana	hamu	ya
EWW:	with him	they were	they have	desire	of
MMG:	*na-we* CCONJ-3SG.PRON	*bha-aa-ri* 3PL-PST-COP	*bha-na* 3PL-have	*e-Ø-naamba* AUG9-NC9.SG-desire	*i-a* AGR9-ASC

Zak37c:	*amachuungwa,*	*nawe*	***bhataamusabhiri.***
SWW:	machungwa	lakini	hawakumwomba
EWW:	oranges	but	they did not ask him
MMG:	*a-ma-chuungwa* AUG6-NC6.PL-orange		*bha-ta-a-mu-sabh-iri* 3PL-NEG-PST-3SG.OBJ-ask-ANT
SFT:	'Wavulana hao wakapishana naye, wakawa na hamu ya machungwa, lakini hawakumwomba.'		
EFT:	'As the boys passed by him, they were hungry for the oranges, but they didn't ask him for any.'		

Zak38a:	***Bhakahitirania,***	***bhakageenda***	*tee*
SWW:	wakapishana	wakatembea	mpaka
EWW:	they passed by each other	they walked	until
MMG:	*bha-ka-hit-ir-an-i-a* 3PL-NAR-pass.by-APPL-RECP-CAUS-FV	*bha-ka-geend-a* 3SG-NAR-walk-FV	

Zak38b:	*kuungosira*	*harya.*
SWW:	kwenye kofia	pale
EWW:	to the hat	there
MMG:	*ku=n-gosira* NC17=NC9.SG-hat	*ha-rya* NC16-DDEM
SFT:	'Wakatembea mpaka kwenye kofia pale.'	
EFT:	'They walked to where the hat was.'	

Zak39a:	*Hano*	***bhakeenda***	***kuhitira***	*engosira,*
SWW:	hapo	wakataka	kupitia	kofia
EWW:	here	they wanted	to pass by	the hat
MMG:	*ha-no* NC16-PDEM.REL	*bha-ka-end-a* 3PL-NAR-want-FV	*ku-hit-ir-a* INF-pass.by-APPL-FV	*e-n-gosira* AUG9-NC9.SG-hat

Zak39b:	*kyaamwe*	*owumwe*	*waabhu*	*wuno*	***yaari***
SWW:	kisha	mmoja	wao	ambaye	alikuwa
EWW:	so	one	of them	who	he was
MMG:		*o-wu-mwe* AUG1-NC1.SG-one	*wa-abhu* AGR1-3PL.POSS	*wu-no* AGR1-PDEM.REL	*a-aa-ri* 3SG-PST-COP

Zak39c:	*omukuru,*	*omukuru*	*waabhu,*	***akatoora***
SWW:	mkubwa	mkubwa	wao	akaokota
EWW:	older person	older person	of them	he picked up
MMG:	*o-mu-kuru* AUG1-NC1.SG-big	*o-mu-kuru* AUG1-NC1.SG-big	*wa-abhu* AGR1-3PL.POSS	*a-ka-toor-a* 3SG-NAR-pick.up-FV

Zak39d:	*engosira*	*yiyo.*
SWW:	kofia	hiyo
EWW:	hat	that
MMG:	*e-n-gosira* AUG9-NC9.SG-hat	*yi-yo* AGR9-RDEM
SFT:	'Walipotaka kupitia kofia, mmojawapo aliyekuwa mkubwa akaokota kofia hiyo.'	
EFT:	'As they were about to pass by where the hat was, the big one picked it up.'	

Zak40a:	*Hano*	***akatoora***	*engosira,*	***akarurya,***
SWW:	hapo	akaokota	kofia	akapiga mluzi
EWW:	here	he picked up	the hat	he whistled
MMG:	*ha-no* NC16-PDEM.REL	*a-ka-toor-a* 3SG-NAR-pick.up-FV	*e-n-gosira* AUG9-NC9.SG-hat	*a-ka-rury-a* 3SG-NAR-whistle-FV

Zak40b:	***akabhirikira***	*omumura*	*wuyo*	*na*
SWW:	akamwita	kijana	huyo	na
EWW:	he called for	young man	that	with
MMG:	*a-ka-bhirikir*[282]*-a* 3SG-NAR-call-FV	*o-mu-mura* AUG1-NC1.SG-young.man	*wu-yo* AGR1-RDEM	*na* CCONJ

Zak40c:	*amachuungwa*	*gaaye,*	*eno*	***arasukuma***
SWW:	machungwa	yake	huku	anasukuma
EWW:	oranges	his	here	he is pushing
MMG:	*a-ma-chuungwa* AUG6-NC6.PL-orange	*ga-aye* AGR6-3SG.POSS	*e-no* NC23-PDEM.REL	*a-ra-sukum-a* 3SG-PROG-push-FV

Zak40d:	*ne*	*ebhasikeeri*	*yaye*	***arahuna.***
SWW:	na	baiskeli	yake	anasukuma
EWW:	and	bicycle	his	he is pushing
MMG:	*ne* CCONJ.AUG9	*e-Ø-bhasikeeri* AUG9-NC9.SG-bicycle	*y-aye* AGR9-3SG.POSS	*a-ra-hun-a* 3SG-PROG-push-FV
SFT:	‘Baada ya kuokota kofia, akapiga mluzi kumwita kijana huyo na machungwa yake ambaye anasukuma baiskeli yake.’			
EFT:	‘After picking up the hat, he whistled for the boy with the oranges who was pushing his bike.’			

Zak41a:	***Akatema***	*omururyo*	***arabhirikira,***
SWW:	akapiga	mluzi	anamwita
EWW:	he hit	whistle	he is calling
MMG:	*a-ka-tem-a* 3SG-NAR-hit-FV	*o-mu-ruryo* AUG3-NC3.SG-whistle	*a-ra-bhirikir-a* 3SG-PROG-call-FV

Zak41b:	***arabhuga***	*“swiii,”*	*niho*	*wurya*
SWW:	anasema	swiii	ndipo	yule
EWW:	he is saying	swiii	then	that
MMG:	*a-ra-bhug-a* 3SG-PROG-say-FV		*ni-ho* FOC-LOC16	*wu-rya* NC1.SG-DDEM

[282] The APPL *–ir* appears to be lexicalized in this form.

Zak41c:	***akiikibhuuka.***
SWW:	akageuka (kwa kichwa)
EWW:	he turned his head
MMG:	*a-ka-ikibhuuk-a* 3SG-NAR-turn.head.around-FV
SFT:	'Akampigia mluzi, "swiii," halafu yule akageuka (kwa kichwa).'
EFT:	'After whistling, "swiii," the boy turned his head.'

Zak42a:	*Hano*	***yaamuruuzi,***	*kyaamwe*	***akagega***
SWW:	hapo	alimwona	kisha	akabeba
EWW:	here	he saw him	so	he carried
MMG:	*ha-no* NC16-PDEM.REL	*a-a-mu-ruuzi* 3SG-PST-3SG.OBJ-see.ANT		*a-ka-geg-a* 3SG-NAR-carry-FV

Zak42b	*engosira,*	***akamutunirira.***
SWW:	kofia	akamfuata
EWW:	hat	he followed him
MMG:	*e-n-gosira* AUG9-NC9.SG-hat	*a-ka-mu-tunir*[283]*-ir-a* 3SG-NAR-3SG.OBJ-follow-APPL-FV
SFT:	'Alipomwona yule anageuka, akampelekea kofia yake.'	
EFT:	'When he saw the boy turn around, he started to take the hat to him.'	

Zak43:	*Hano*	***akagya,***	***akamuhikaku.***[284]
SWW:	hapo	akaenda	akamfikia huko
EWW:	here	he went	he reached him there
MMG:	*ha-no* NC16-PDEM.REL	*a-ka-gy-a* 3SG-NAR-go-FV	*a-ka-mu-hik-a=ku* 3SG-NAR-3SG.OBJ-arrive-FV=LOC17
SFT:	'Alipoenda, akamfikia kule.'		
EFT:	'He went and got to where the boy was.'		

Zak44a:	*Hano*	***yaamuhikiriku,***
SWW:	hapo	alimfikia huko
EWW:	here	he reached him there
MMG:	*ha-no* NC16-PDEM.REL	*a-a-mu-hik-iri=ku* 3SG-PST-3SG.OBJ-arrive-ANT=LOC17

[283] The verb *tunir* 'follow' is a lexicalized applicative from the verb *tun* 'search'. After the lexicalization was complete, *tunir* could take its own applicative.
[284] Gray (pc) notes the similarity of this to Ikizu, where a LOC appears to license an object and the Applicative verb form is not necessary.

Zak44b:	***akamuha***	*engosira*	*yirya.*
SWW:	akampa	kofia	ile
EWW:	he gave him	hat	that
MMG:	*a-ka-mu-h-a* 3SG-NAR-3SG.OBJ-give-FV	*e-n-gosira* AUG9-NC9.SG-hat	*yi-rya* AGR9-DDEM
SFT:	'Alipokuwa amefikia huko, akampa kofia yake.'		
EFT:	'After reaching that spot, he gave the boy the hat.'		

Zak45a:	*Bhoono,*	*hano*	*omwaana*	*wurya,*	*hano*
SWW:	sasa	hapo	mtoto	yule	hapo
EWW:	now	here	child	that	here
MMG:		*ha-no* NC16-PDEM	*o-mu-ana* AUG1-NC1.SG-child	*wu-rya* NC1.SG-DDEM	*ha-no* NC16-PDEM.REL

Zak45b:	***yaaruuzi***	*wurya*	***amureteeri***
SWW:	aliona	yule	amemletea
EWW:	he saw	that	he has brought him
MMG:	*a-a-ruuzi* 3SG-PST-see.ANT	*wu-rya* NC1.SG-DDEM	*a-mu-reet*[285]*-eeri* 3SG-3SG.OBJ-bring-APPL.ANT

Zak45c:	*engosira*	*yaye,*	***akagega***	*amachuungwa*
SWW:	kofia	yake	akabeba	machungwa
EWW:	hat	his	he carried	oranges
MMG:	*e-n-gosira* AUG9-NC9.SG-hat	*y-aye* AGR9-3SG.POSS	*a-ka-geg-a* 3SG-NAR-carry-FV	*a-ma-chuungwa* AUG6-NC6.PL-orange

Zak45d:	*atatu,*	*kwa*	*sababu*	***bhaari***	*bhaana*	*bhatatu,*
SWW:	matatu	kwa	sababu	walikuwa	watoto	watatu
EWW:	three	for	reason	they were	children	three
MMG:	*a-tatu* AGR6-three			*bha-aa-ri* 3PL-PST-COP	*bha-ana* NC2.PL-child	*bha-tatu* AGR2-three

[285] The shortening of the underlying root *reet* to *ret* in the surface form is not yet understood and bears further investigation.

Zak45e:	***akamuha***	*amachuungwa*	*atatu.*
SWW:	akampa	machungwa	matatu
EWW:	he gave him	oranges	three
MMG:	*a-ka-mu-h-a* 3SG-NAR-3SG.OBJ-give-FV	*a-ma-chuungwa* AUG6-NC6.PL-orange	*a-tatu* AGR6-three
SFT:	'Mtoto yule alipoona mwingine anamletea kofia, akachukua machungwa matatu, kwa walikuwa watato watatu, akampa.'		
EFT:	'When that boy saw the other one bringing his hat to him, he took three oranges, since there were three boys, and he gave them to him.'		

Zak46a:	*Na*	*wurya*	***akamuha***	*engosira*
SWW:	na	yule	akampa	kofia
EWW:	and	that	he gave him	hat
MMG:	*na* CCONJ	*wu-rya* NC1.SG-DDEM	*a-ka-mu-h-a* 3SG-NAR-3SG.OBJ-give-FV	*e-n-gosira* AUG9-NC9.SG-hat

Zak46b:	*yaye,*	***akagaruka.***
SWW:	yake	akarudi
EWW:	his	he returned
MMG:	*y-aye* AGR9-3SG.POSS	*a-ka-garuk-a* 3SG-NAR-return-FV
SFT:	'Na yule akampa kofia yake, akarudi.'	
EFT:	'And that boy gave him the hat and returned to his friends.'	

Zak47a:	*Hano*	***yaagarukiri***	*omwaana*	*wurya,*
SWW:	hapo	alikuwa amerudi	mtoto	yule
EWW:	here	he had returned	child	that
MMG:	*ha-no* NC16-PDEM.REL	*a-aa-garuk-iri* 3SG-PST.ANT-return-ANT	*o-mu-ana* AUG1-NC1.SG-child	*wu-rya* NC1.SG-DDEM

Zak47b:	***akeenderera***	*ne*	*ekisehi*	*kyaaye*
SWW:	akaendelea	na	kikapu	chake
EWW:	she continued	with	basket	his
MMG:	*a-ka-enderer-a* 3SG-NAR-continue-FV	*ne* CCONJ.AUG7	*e-ki-sehi* AUG7-NC7.SG-basket	*ki-aye* AGR7-3SG.POSS

Zak47c:	*na*	*amachuungwa*	*gaaye.*
SWW:	na	machungwa	yake
EWW:	with	oranges	his
MMG:	*na* CCONJ	*a-ma-chuungwa* AUG6-NC6.PL-orange	*ga-aye* AGR6-3SG.POSS
SFT:	'Mtoto yule aliporudi, akaendelea na kikapu ambacho kimejaa machungwa.'		
EFT:	'As the boy returned to his friends, the boy on the bike continued on his journey with a full basket of oranges.'		

Zak48:	*Na*	*bhano*	*bhoosi*	***bhakakyoora.***
SWW:	na	hao	wote	wakarudi
EWW:	and	these	all	they returned
MMG:	*na* CCONJ	*bha-no* NC2.PL-PDEM	*bha-osi* AGR2-all	*bha-ka-kyoor-a* 3PL-NAR-return-FV
SFT:	'Na watatu wote wakarudi.'			
EFT:	'All the boys returned.'			

Zak49a:	*Wuno*	*hano*	***yaahikiri***	*kubhaana*
SWW:	huyu	hapo	alikuwa amefika	kwa watoto
EWW:	this (person)	here	he had arrived	to the children
MMG:	*wu-no* AGR1-PDEM	*ha-no* NC16-PDEM.REL	*a-aa-hik-iri* 3SG-PST.ANT-arrive-ANT	*ku=bha-ana* NC17= NC2.PL-child

Zak49b:	*bharya*	*bhabhiri,*	***bhakasoonda***	*kira*	*omuutu*
SWW:	wale	wawili	wakagawana	kila	mtu
EWW:	those	two	they divided	each	person
MMG:	*bha-rya* NC2-DDEM	*bha-bhiri* AGR2-two	*bha-ka-soond-a* 3PL-NAR-divide-FV		*o-mu-utu* AUG1-NC1.SG-person

Zak49c:	*erichuungwa*	*rimwe*	*rimwe,*	***bhakabha***	***bhararya.***
SWW:	chungwa	moja	moja	wakawa	wanakula
EWW:	orange	one	one	they were	they are eating
MMG:	*e-ri-chuungwa* AUG5-NC5.SG-orange	*ri-mwe* AGR5-one	*ri-mwe* AGR5-one	*bha-ka-bh-a* 3PL-NAR-be-FV	*bha-ra-ry-a* 3PL-PROG-eat-FV
SFT:	'Mtoto na machungwa matatu alipokuwa amefika wengine wawili, akawagawia chungwa moja moja, wakawa wanakula.'				
EFT:	'When the boy who had been given the three oranges returned to his companions, he gave them each an orange, and they ate them.'				

Zak50a:	*Kuumbe*	*hano*	***bhakagya,***	***bhakagya***	*tee*
SWW:	kumbe	hapo	wakaenda	wakaenda	mpaka
EWW:	amazingly	here	they went	they went	until
MMG:	*kumbe* EMPH	*ha-no* NC16-PDEM.REL	*bha-ka-gy-a* 3PL-NAR-go-FV	*bha-ka-gy-a* 3PL-NAR-go-FV	

Zak50b:	***bhakahikira***	*kuunzira*	*harya*
SWW:	wakafikia	njiani	pale
EWW:	they reached	to the path	there
MMG:	*bha-ka-hik-ir-a* 3PL-NAR-arrive-APPL-FV	*ku=n-zira* NC17=NC9.SG-path	*ha-rya* NC16-DDEM

Zak50c:	*Kusaya*	***yaari***	*kumuchuungwa*	*gurya.*
SWW:	Kusaya	alikuwa	kwenye mchungwa	ule
EWW:	Kusaya	he was	in the orange tree	that
MMG:		*a-aa-ri* 3SG-PST-COP	*ku=mu-chuungwa* NC17=NC3.SG-orange	*gu-rya* AGR3-DDEM
SFT:	‘Kumbe, walipoenda, wakaenda mpaka kufikia njiani kwenda kwenye mchungwa wa Kusaya.’			
EFT:	‘When they left, they went until they reached the path that led to where Kusaya was in the tree picking oranges!’			

Zak51a:	*Hano*	*Kusaya*	***yiikiri***	*kumuchuungwa*
SWW:	hapo	Kusaya	aliteremka	kwenye mchungwa
EWW:	here	Kusaya	he climbed down	in the orange tree
MMG:	*ha-no* NC16-PDEM.REL		*a-a-ik-iri* 3SG-PST-go.down-ANT	*ku=mu-chuungwa* NC17=NC3.SG-orange

Zak51b:	*gurya,*	***akabhona***	*ekisehi*	*kimwe*
SWW:	ule	akakuta	kikapu	kimoja
EWW:	that	he found	basket	one
MMG:	*gu-rya* AGR3-DDEM	*a-ka-bhon-a* 3SG-NAR-find-FV	*e-ki-sehi* AUG7-NC7.SG-basket	*ki-mwe* AGR7-one

Zak51c:	***kitareho.***[286]
SWW:	hakipo
EWW:	it is not there
MMG:	*ki-ta-re=ho* SBJ7-NEG-COP=LOC16
SFT:	'Kusaya alipokuwa ameshuka, akakuta kwamba kikapu kimoja hakipo.'
EFT:	'When Kusaya had climbed down, he found that one of his baskets was missing.'

Zak52a:	*Bhoono*	*Kusaya*	***akiitakya,***
SWW:	sasa	Kusaya	akashtuka
EWW:	now	Kusaya	he was surprised
MMG:			*a-ka-i-taky*[287]*-a* 3SG-NAR-REFL-surprise-FV

Zak52b	***akamagamaga,***	***akawaanza***	***kubhara***
SWW:	akageukageuka	akaanza	kuhesabu
EWW:	he turned around and around	he started	to count
MMG:	*a-ka-maga~mag-a* 3SG-NAR-turn.around~RDPL-FV	*a-ka-waanz-a* 3SG-NAR-start-FV	*ku-bhar-a* NC15.INF-count-FV

Zak52c:	*ebhisehi*	*bhyaaye*	*bhiyo,*	***akarora***
SWW:	vikapu	vyake	hivyo	akaona
EWW:	baskets	his	those	he saw
MMG:	*e-bhi-sehi* AUG8-NC8.PL-basket	*bhi-aye* AGR8-3SG.POSS	*bhi-yo* AGR8-RDEM	*a-ka-ror-a* 3SG-NAR-see-FV

Zak52d:	***bhireho***	*ebhisehi*	*bhibhiri,*	*nawe*
SWW:	vipo	vikapu	viwili	lakini
EWW:	they are there	baskets	two	but
MMG:	*bhi-Ø-re=ho* SBJ8-PRS-COP=LOC16	*e-bhi-sehi* AUG8-NC8.PL-basket	*bhi-bhiri* AGR8-two	

[286] Gray (pc) mentions that this is similar to Ikizu where the vowels of Class 16 and 23 locative clitics affect the vowel height of the Copula.

[287] The basic verb root is *tak* 'get' + *y* 'CAUS'; the concept 'cause to get' has lexicalized as 'surprise (someone)'. When the reflexive is added to this new root it has the sense of 'be surprised'.

Zak52e:	*ekiindi*	*kino*	***kyaari***	***kiizwiiri***
SWW:	kingine	ambacho	kilikuwa	kimejaa
EWW:	other	which	it was	it has filled
MMG:	*e-ki-ndi* AUG7-NC7.SG-other	*ki-no* AGR7-PDEM.REL	*ki-aa-ri* SBJ7-PST-COP	*ki-izur-iri* SBJ7-be.full-ANT

Zak52f:	***kitareho.***
SWW:	hakipo
EWW:	it is not there
MMG:	*ki-ta-re=ho* SBJ7-NEG-COP=LOC16
SFT:	'Sasa, Kusaya akashtuka, akageukageuka, akaanza kuhesabu vikapu, akaona viwili vipo, lakini kingine kilichojaa machungwa hakipo.'
EFT:	'Now, Kusaya was shocked and looked all over. Then he started to count the baskets, saw that there were two and realized that he was missing the one that had been full of oranges.'

Zak53:	*Bhoono*	***akabha***	***akyaaruguura.***
SWW:	sasa	akawa	bado anashangaa
EWW:	now	he was	he is still surprised
MMG:		*a-ka-bh-a* 3SG-NAR-be-FV	*a-kyaa-ruguur-a* 3SG-PERS-be.surprised-FV
SFT:	'Sasa, bado alikuwa ameshangaa.'		
EFT:	'He was still in a state of shock.'		

Zak54a:	*Hano*	***yaakyaaruguura,***	*hasuuhwiigo,*
SWW:	hapo	alikuwa akishangaa bado	muda kidogo
EWW:	here	he was still surprised	little while
MMG:	*ha-no* NC16-PDEM.REL	*a-a-kya-ruguur-a* 3SG-PST-PERS-be.surprised-FV	*ha-suuhu=igo* NC16-little=like.that

Zak54b:	*abhaana*	*bharya*	***bhakagya,***
SWW:	watoto	wale	wakaenda
EWW:	children	those	they went
MMG:	*a-bha-ana* AUG2-NC2.PL-child	*bha-rya* AGR2-DDEM	*bha-ka-gy-a* 3PL-NAR-go-FV

Zak54c:	***bhakamuhikaku.***
SWW:	wakamfikia huko
EWW:	they reached him there
MMG:	*bha-ka-mu-hik-a=ku* 3PL-NAR-3SG.OBJ-arrive-FV=LOC17
SFT:	'Alipokuwa bado akishangaa, baada ya kidogo, watoto wale wakafikia huko.'
EFT:	'While he was still in shock, the three boys reached the tree where he was.'

Zak55a:	*Hano*	***bhakamuhikaku,***	***akarora***	*kira*
SWW:	hapo	wakamfikia huko	akaona	kila
EWW:	here	they reached him there	he saw	each
MMG:	*ha-no* NC16-PDEM.REL	*bha-ka-mu-hik-a=ku* 3PL-NAR-3SG.OBJ-arrive-FV=LOC17	a-ka-ror-a 3SG-NAR-see-FV	

Zak55b:	*owumwe*	***ana***	*erichuungwa*	*ryaaye*
SWW:	mmoja	ana	chungwa	lake
EWW:	one	he has	orange	his
MMG:	*o-wu-mwe* AUG1-NC1.SG-one	*a-na* 3SG-have	*e-ri-chuungwa* AUG5-NC5.SG-orange	*ri-aye* AGR5-3SG.POSS

Zak55c:	*mukubhoko.*
SWW:	mkononi
EWW:	in hand
MMG:	*mu=ku-bhoko* NC18=NC15.SG-hand
SFT:	'Walipomfikia, akaona kila mmoja ana chungwa mkononi.'
EFT:	'When they reached him, he saw that each one had an orange in his hand.'

Zak56a:	*Bhoono*	***akabha***	***ariibhuurya***	*kana:*
SWW:	sasa	akawa	anajiuliza	hivi
EWW:	now	he was	he is asking himself	this
MMG:		*a-ka-bh-a* 3SG-NAR-be-FV	a-ra-i-bhuury-a 3SG-PROG-REFL-ask-FV	

Zak56b:	*"Gano*	***bhagabhweeni***	*hayi*
SWW:	haya	wamepata	wapi
EWW:	these	they have got	where
MMG:	*ga-no* AGR6-PDEM	*bha-ga-bhweeni* 3PL-OBJ6-get.ANT	*ha-yi* NC16-which

Zak56c:	*eno*	*ekisehi*	***bhatanakyo?"***
SWW:	huku	kikapu	hawanacho
EWW:	here	basket	they do not have them
MMG:	*e-no* NC23-PDEM	*e-ki-sehi* AUG7-NC7.SG-basket	*bha-ta-na=kyo* 3PL-NEG-have=OBJ7
SFT:	'Sasa, alikuwa anajiuliza hivi: "Hayo wameyapata wapi bila kikapu chake?"'		
EFT:	'Now, he was asking himself: "Where did they get those since they don't have the basket?"'		

Zak57a:	*Na*	*abhaana*	*bhayo*	***bhakamuhitira,***
SWW:	na	watoto	hao	wakampitia
EWW:	and	children	those	they passed him
MMG:	*na* CCONJ	*a-bha-ana* AUG2-NC2.PL-child	*bha-yo* AGR2-RDEM	*bha-ka-mu-hit-ir-a* 3PL-NAR-3SG.OBJ-pass.by-APPL-FV

Zak57b:	***bhakagya,***	***bhakeenderera***	*no*	*orugeendo*
SWW:	wakaenda	wakaendelea	na	safari
EWW:	they went	they continued	and	journey
MMG:	*bha-ka-gy-a* 3PL-NAR-go-FV	*bha-ka-enderer-a* 3PL-NAR-continue-FV	*no* CCONJ.AUG11	*o-ru-geendo* AUG11-NC11.SG-journey

Zak57c:	*rwaabhu.*
SWW:	yao
EWW:	their
MMG:	*ru-abhu* AGR11-3PL.POSS
SFT:	'Na watoto hao wakampitia, wakaendelea na safari yao.'
EFT:	'And the kids just passed by him and continued on their way.'

Zak58a:	*Bhoono*	*Kusaya*	***akatanga***	*kweeki*	***kubhara***
SWW:	sasa	Kusaya	akaanza	tena	kuhesabu
EWW:	now	Kusaya	he began	again	to count
MMG:			*a-ka-tang-a* 3SG-NAR-begin-FV		*ku-bhar-a* INF-count-FV

Zak58b:	*ebhisehi*	*bhyaaye*	*na*	*amachuungwa*
SWW:	vikapu	vyake	na	machungwa
EWW:	baskets	his	and	oranges
MMG:	*e-bhi-sehi* AUG8-NC8.PL-basket	*bhi-aye* AGR8-3SG.POSS	*na* CCONJ	*a-ma-chuungwa* AUG6-NC6.PL-orange

Zak58c:	*gaaye*	***akarora***	*ni masuuhu.*
SWW:	yake	akaona	ni machache
EWW:	his	he saw	it is few
MMG:	*ga-aye* AGR6-3SG.POSS	a-ka-ror-a 3SG-NAR-see-FV	*ni ma-suuhu* COP AGR6-little
SFT:	'Sasa, Kusaya akaanza tena kuhesabu vikapu vyake na machungwa yake, akaona ni machache.'		
EFT:	'Now, Kusaya started to count his baskets and oranges again, but he saw that he didn't have enough.'		

Zak59a:	*Bhoono*	***akiiseega***	*"mmmm,"*	***ataabhweeni***
SWW:	sasa	akafikiri	mmmm	hakupata
EWW:	now	he thought to himself	hmmm	he did not get
MMG:		*a-ka-i-seeg-a* 3SG-NAR-REFL-begin-FV		*a-ta-a-bhweeni* 3SG-NEG-PST-get.ANT

Zak59b:	*erikyooro*	*ryooryoosi.*
SWW:	jibu	lolote
EWW:	answer	any
MMG:	*e-ri-kyooro* AUG5-NC5.SG-answer	*ri-o-ri-osi* AGR5-all-AGR5-all
SFT:	'Akafikiri, "mmmm," hakupata jibu lolote.'	
EFT:	'He thought to himself, "hmmm," but he didn't understand.'	

Zak60a	*Hano*	***ataabhweeni***	*erikyooro*
SWW:	hapo	hakupata	jibu
EWW:	here	he did not get	answer
MMG:	*ha-no* NC16-PDEM.REL	*a-ta-a-bhweeni* 3SG-NEG-PST-get.ANT	*e-ri-kyooro* AUG5-NC5.SG-answer

Zak60b:	*ryooryoosi*	***araruguura***	*omweene,*
SWW:	lolote	anashangaa	mwenyewe
EWW:	any	he is surprised	himself
MMG:	*ri-o-ri-osi* AGR5-all-AGR5-all	*a-ra-ruguur-a* 3SG-PROG-be.surprised-FV	*o-mu-ene* AUG1-NC1.SG-self

Zak60c:	***arasiingisya***	*omutwe,*
SWW:	anatikisa	kichwa
EWW:	he is scratching	head
MMG:	*a-ra-siingisy-a* 3SG-PROG-scratch-FV	*o-mu-twe* AUG3-NC3.SG-head

Zak60d:	*eno*	***aramagamaga.***
SWW:	huku	anageukageuka
EWW:	here	he is turning around and around
MMG:	*e-no* NC23-PDEM	*a-ra-maga~mag-a* 3SG-PROG-turn.around~RDPL-FV
SFT:	‘Aliposhindwa kupata jibu, akawa ameshangaa, anatikisa kichwa na kugeukageuka.’	
EFT:	‘After failing to figure it out, he was surprised and started scratching his head and turning around and around.’	

Zak61a:	*Nawe*	***kana:***	*“Ekisehi*	*kyaane*
SWW:	lakini	hivi	kikapu	changu
EWW:	but	this	answer	my
MMG:			*e-ki-sehi* AUG7-NC7.SG-basket	*ki-ane* AGR7-1SG.POSS

Zak61b:	*eno*	***kigiiri***	*ni hayi?”*
SWW:	huku	kimeenda	ni wapi
EWW:	here	it has gone	it is where
MMG:	*e-no* NC23-PDEM	*ki-gi-iri* SBJ7-go-ANT	*ni ha-yi* COP NC16-which
SFT:	‘Lakini alijiuliza: “Kikapu changu kimeenda wapi?”’		
EFT:	‘But he asked himself again: “Where did my basket go?”’		

Zak62a:	*Nawe*	***akamagamaga***	***ataaruuzi***
SWW:	lakini	akageukageuka	hakuona
EWW:	but	he turned around and around	he did not see
MMG:		*a-ka-maga~mag-a* 3SG-NAR-turn.around~RDPL-FV	*a-ta-a-ruuzi* 3SG-NEG-PST-see.ANT

Zak62b:	*eno*	*ekisehi*	***kyeerekeeri.***
SWW:	huku	kikapu	kilielekea
EWW:	here	basket	it headed
MMG:	*e-no* NC23-PDEM.REL	*e-ki-sehi* AUG7-NC7.SG-basket	*ki-erek-eeri* SBJ7-head-APPL.ANT
SFT:	‘Lakini akageukageuka, hakuona kikapu kilikoelekea.		
EFT:	‘But he looked all over and could not see where the basket had gone.’		

Zak63a:	*Na*	*abhamura*	*bharya*	*kira*	*omuutu*
SWW:	na	wavulana	wale	kila	mtu
EWW:	and	boys	those	each	person
MMG:	*na* CCONJ	*a-bha-mura* AUG2-NC2.PL-boy	*bha-rya* AGR2-DDEM		*o-mu-utu* AUG1-NC1.SG-person

Zak63b:	***ahitiri***	***ana***	*erichuungwa*	*mukubhoko.*
SWW:	amepita	ana	chungwa	mkononi
EWW:	he has passed	he has	orange	in hand
MMG:	*a-hit-iri* 3SG-pass.by-ANT	*a-na* 3SG-have	*e-ri-chuungwa* AUG5-NC5.SG-orange	*mu=ku-bhoko* NC18=NC15.SG-hand
SFT:	‘Na kila mvulana ambaye alikuwa amepita, akawa na chungwa mkononi.			
EFT:	‘And every boy that had passed by had an orange in his hand.’			

Zak64a:	*Bhoono*	***akasiingisya***	*omutwe,*	***ataabhweeni***
SWW:	sasa	akatikisa	kichwa	hakupata
EWW:	now	he scratched	head	he did not get
MMG:		*a-ka-siingisy-a* 3SG-NAR-scratch-FV	*o-mu-twe* AUG3-NC3.SG-head	*a-ta-a-bhweeni* 3SG-NEG-PST-get.ANT

Zak64b:	*erikyooro.*
SWW:	jibu
EWW:	answer
MMG:	*e-ri-kyooro* AUG5-NC5.SG-answer
SFT:	'Sasa akatikisa kichwa, hakupata jibu.'
EFT:	'Now, he scratched his head, but he couldn't figure it out.'

Zak65:	*Kyaamwe*	*erigano*	*rino*	***rikahweera***	*hayo.*
SWW:	kisha	hadithi	hii	ikaishia	hapo
EWW:	so	story	this	it ended	here
MMG:		*e-ri-gano* AUG5-NC5.SG-story	*ri-no* AGR5-PDEM	*ri-ka-hweer-a* SBJ5-NAR-end-FV	*ha-yo* NC16-RDEM
SFT:	'Kisha, hadithi hii inaishia hapo.'				
EFT:	'So, that's the end.'				

Appendix G: Swahili Elicitation Sentences with English Translations

The following list presents the 91 Swahili sentences that were used to elicit much of the TAM data in this thesis. They are presented in the order they were elicited with English translations for each sentence.

1) Twaenda Dar Es Salaam.
'We go to Dar Es Salaam.'

2) Tulienda Dar Es Salaam zamani.
'We went to Dar Es Salaam a long time ago.'

3) Mara kwa mara, sisi huenda Dar Es Salaam.
'We go to Dar Es Salaam frequently.'

4) Leo, tunaenda Dar Es Salaam.
'Today, we are going to Dar Es Salaam.'

5) Kwa kawaida, sisi huenda Dar Es Salaam, lakini mara huu tunaenda Nairobi.
'Normally, we go to Dar Es Salaam, but this time we are going to Nairobi.'

6) Tunachimba tena na tena.
'We are digging over and over.'

7) Tunachimba siku nzima.
'We are digging all day long.'

8) Huwa tunachimba siku nzima.
'We are usually digging all day long.'

9a) Tumeongea naye.
We have spoken with him.

9b) Tumechimba shimo.
'We have dug a hole.'

10) Tumeshaongea naye.
'We have already spoken with him.'

11) Tumeongea naye dakika kumi zilizopita.
'We spoke with him ten minutes ago.'

12) Tumeongea naye masaa mawili yaliyopita.
'We spoke with him two hours ago.'

13) Tuliongea naye asubuhi hii.
'We spoke with him this morning.'

14) Jana, tuliongea naye.
'We spoke with him yesterday.'

15) Juzi, tuliongea naye.
'We spoke with him the day before yesterday.'

16) Tuliongea naye wiki iliyopita.
'We spoke with him last week.'

17) Tuliongea naye mwezi uliyopita.
'We spoke with him last month.'

18) Tuliongea naye mwaka uliyopita.
'We spoke with him last year.'

19) Tuliongea naye zamani.
'We spoke with him a long time ago.'

20) Tutaongea naye hivi karibuni.
'We will speak with him soon.'

21) Tutaongea naye baada ya dakika kumi.
'We will speak with him in ten minutes.'

22) Tutaongea naye baada ya masaa mawili.
'We will speak with him in two hours.'

23) Tutaongea naye jioni hii.
'We will speak with him this evening.'

24) Tutaongea naye kesho.
'We will speak with him tomorrow.'

25) Tutaongea naye kesho kutwa.
'We will speak with him the day after tomorrow.'

26) Tutaongea naye wiki ijayo.
'We will speak with him next week.'

27) Tutaongea naye mwezi ujao.
'We will speak with him next month.'

28) Tutaongea naye mwaka ujao.
'We will speak with him next year.'

29) Siku za mbele, tutaongea naye.
'We will speak with him in the future.'

30) Siku za mbele, tuongee naye.
'We should speak with him in the future.'

31) Mwezi ujao, tuongee naye.
'We should speak with him next month.'

32) Kesho kutwa, tuongee naye.
'We should speak with him the day after tomorrow.'

33) Kesho, tuongee naye.
'We should speak with him tomorrow.'

34) Tuongee naye hivi karibuni.
'We should speak with him soon.'

35) Tuongee naye.
'We should speak with him.'

36) Tutakuwa tunacheza (ngoma). [Form 1]
'We will be dancing.'

37) Tutakuwa tukicheza (ngoma). [Form 2]
'We will be dancing.'

38) Watakapopiga ngoma, tutakuwa tukicheza.
'When they (will) play the drums, we will be dancing.'

39) Wanapopiga ngoma, tunacheza.
'When they play the drum, we are dancing.'

40) Kila wanapopiga ngoma, tunacheza.
'Every time they play the drums, we are dancing.'

41) Kila walipopiga ngoma, tulicheza.
'Every time they played the drums, we danced.'

42) Kila walipokuwa wakipiga ngoma, tulicheza.
'Whenever they were playing the drums, we danced.'

43) Walipopiga ngoma, tulikuwa tukicheza.
'When they played the drums, we were dancing.'

44) Tulikuwa tunacheza. [Form 1]
'We were dancing.'

45) Tulikuwa tukicheza. [Form 2]
'We were dancing.'

46) Tumekuwa tunacheza.
'We have been dancing.'

47) Tulikuwa tumecheza.
'We had danced.'

48) Tutakuwa tumecheza.
'We will have danced.'

49) Kama tunacheza (ngoma) sasa, tutachoka badaaye.
'If we are dancing now, we will be tired later.'

50) Kama tulikuwa tunacheza jana, tumechoka leo.
'If we were dancing yesterday, we are tired today.'

51) Kama tumekuwa tunacheza tangu asubuhi hii, tunachoka sasa.
'If we have been dancing since this morning, we are getting tired now.'

52) (Kama) tukiondoka sasa, tutachelewa kufika.
'If we depart now, we will get there late.'

53) (Kama) tukiondoka sasa, hatutafika mapema.
'If we depart now, we won't get there early.'

54) Tusipoondoka sasa, hatutafika mapema.
'If we don't depart now, we won't get there early.'

55) Tusipoondoka sasa, tutachelewa kufika.
'If we don't depart now, we will get there late.'

56) Tungeondoka sasa, tungechelewa kufika.
'If we would depart now, we would arrive late.'

57) Tungeondoka sasa, tusingefika mapema.
'If we would depart now, we wouldn't get there early.'

58) Tusingeondoka sasa, tusingefika mapema.
'If we wouldn't depart now, we wouldn't get there early.'

59) Tusingeondoka sasa, tungechelewa kufika.
'If we wouldn't depart now, we would get there late.'

60) Tungaliondoka mapema, tungalifika mapema.
'If we would have departed early, we would have arrived early.'

61) Tungaliondoka mapema, tungefika mapema.
'If we would have departed early, we would arrive early.'

62) Tungaliondoka mapema, tusingalichelewa kufika.
'If we would have departed early, we wouldn't have arrived late.'

63) Tungaliondoka mapema, tusingechelewa kufika.
'If we would have departed early, we wouldn't arrive late.'

64) Tusingaliondoka mapema, tusingalifika mapema.
'If we wouldn't have departed early, we wouldn't have arrived early.'

65) Tusingaliondoka mapema, tusingefika mapema.
'If we wouldn't have departed early, we wouldn't arrive early.'

66) Tusingaliondoka mapema, tungalichelewa kufika.
'If we wouldn't have departed early, we would have arrived late.'

67) Tusingaliondoka mapema, tungechelewa kufika.
'If we wouldn't have departed early, we would arrive late.'

68) Chimba viazi!
'Dig up the potatoes!'

69) Chimbeni viazi!
Dig up (pl.) the potatoes!

70) Tulikuwa tunachimba shimo (sasa tumeacha).
'We were digging a hole (we have stopped now).'

71) Tulikuwa tunachimba shimo (tunaendelea sasa).
'We were digging a hole (we are continuing now).'

72) Tulikuwa walimu kwa miaka mingi.
'We were teachers for many years.'

73) Miaka mingi iliyopita, tulikuwa walimu kwa siku moja.
'Many years ago, we were teachers for a day.'

74) Jana, tulikuwa walimu kwa siku moja.
'Yesterday, we were teachers for the day.'

75) Tutaongea naye (kuna uhakika kwamba tendo litafanyika).
'We will speak with him (it is certain that it will happen).'

76) Tutaongea naye (tendo linawezekana).
'We will speak with him (possibly).'

77) Tutaongea naye (kuna shaka kwamba tendo litafanyika).
'We will speak with him (it is doubtful that it will happen).'

78) Hatuendi Dar Es Salaam.
'We are not going to Dar Es Salaam.'

79) Bado tunaongea naye.
'We are still speaking with him.'

80) Bado hatuongei naye.
'We still don't speak with him./We still aren't speaking with him.'

81) Bado hatujaongea naye.
'We haven't spoken with him yet.'

82) Bado tutaongea naye.
'We will still speak with him.'

83) Bado hatutaongea naye.
'We still won't speak with him.'

84) (Anaenda wapi?) Anaenda Dar Es Salaam.
'(Where is he going?) He's going to Dar Es Salaam.'

85) (Kwa nini anaenda Dar Es Salaam?) Anaenda kumtembelea mama yake.
'(Why is he going to Dar Es Salaam?) He's going to visit his mother.'

86) (Anaenda Dar Es Salaam vipi?) Anaenda kwa ndege.
'(How is he going to Dar Es Salaam?) He's going by plane.'

87) (Nani anaenda Dar Es Salaam?) Huyu anaenda.
'(Who is going to Dar Es Salaam?) This (person) is going.'

88) (Anafanya nini?) Anaenda Dar Es Salaam.
'(What is he doing?) He is going to Dar Es Salaam.'

89) Tunataka kuchimba shimo.
'We are about to dig a hole. (lit. We want to dig a hole.)'

90) Tunaweza tukachimba shimo.
'We could/might dig a hole.'

91) Tunaweza kuchimba shimo.
'We can dig a hole.'

References

Aunio, Lotta. 2010. Ikoma nominal tone. *Africana Linguistica* 16. 3-30.

Bailey, Charles-James N. 1982. *On the Yin and Yang nature of language*. Ann Arbor, MI: Karoma Publishers, Inc.

Bastin, Yvonne. 1983. Essai de classification de quatre-vingts langues bantoues par la statistique grammaticale. *Africana Linguistica* 9. 11-108.

Bastin, Yvonne, André Coupez & B. de Halleux. 1983. Classification lexicostatistique des langues bantuoues (214 relevés). *Bulletin des Séances du Académie Royale du Sciences d'Outre-Mer* 27(2). 173-199.

Bearth, Thomas. 2003. Syntax. In Nurse, Derek & Gérard Philippson (eds.), *The Bantu languages*, 121-142. London & New York: Routledge.

Bleek, W.H.I. 1862. *A comparative grammar of South African languages, part 1: Phonology*. Cape Town & London: J.C. Juta and Trübner & Co.

Botne, Robert. 1986. The temporal role of eastern Bantu –ba and –li. *Studies in African Linguistics* 17(3). 303-318.

Botne, Robert. 1989/1990. The historical relation of Cigogo to zone J languages. *Sprache und Geschichte in Afrika* 10/11. 187-222.

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

Botne, Robert. 2013. Organizational properties of Bantu T/A systems. Paper presented as background reading prior to the 5th International Conference on Bantu Linguistics [Bantu 5], Paris, June 12-15, 2013. http://bantu5.sciencesconf.org/conference/bantu5/pages/Bantu5_WS_2_Bantu_TA_position_paper.pdf. (5 August, 2013.)

Botne, Robert & Tiffany L. Kershner. 2008. Tense and aspect in cognitive space: On the organization of tense/aspect systems in Bantu languages and beyond. *Cognitive Linguistics* 19(2). 145-218.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Cammenga, Jelle. 2002. *Phonology and morphology of Ekegusii: A Bantu language of Kenya*. Köln: Rüdiger Köppe Verlag.

Cammenga, Jelle. 2004. *Igikuria phonology and morphology: A Bantu language of south-west Kenya and north-west Tanzania*. Köln: Rüdiger Köppe Verlag.

Chafe, Wallace (ed.). 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Compton, Alison & John B. Walker. 2010. *Suba-Simbiti orthography sketch*. Nairobi, Dar Es Salaam & Kampala: SIL International, Uganda-Tanzania Branch.

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

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

Dixon, R.M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.

Futakamba, Willy, Shem Koren, Holly Robinson & Oliver Stegen. 2013. *The grammar basics of Zanaki*. Unpublished draft ms.

Gray, Hazel. 2011a. *Discourse features of Ikizu*. Unpublished draft ms.

Gray, Hazel. 2011b. *Discourse features of Zanaki*. Unpublished draft ms.

Gray, Hazel. 2013. *Locatives in Ikizu*. Leiden: Leiden University MA Thesis.

Gray, Hazel & Michelle Smith. 2010a. *Ikizu orthography sketch*. Nairobi, Dar Es Salaam & Kampala: SIL International, Uganda-Tanzania Branch.

Gray, Hazel & Michelle Smith. 2010b. *Zanaki orthography sketch*. Nairobi, Dar Es Salaam & Kampala: SIL International, Uganda-Tanzania Branch.

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.

Guthrie, Malcolm. 1967-71. *Comparative Bantu*. Farnborough, England: Gregg International Publishers.

Harjula, Lotta. 2004. *The Ha language of Tanzania: Grammar, texts and vocabulary*. Köln: Rüdiger Köppe Verlag.

Haspelmath, Martin. 2010. Discussion notes: Comparative concepts and descriptive categories in crosslinguistic categories. *Language* 86(3). 663-687.

Heine, Bernd & Tania Kuteva. 2002a. On the evolution of grammatical forms. In Wray, Alison (ed.), *The transition to language*, 376-397. Oxford: Oxford University Press.

Heine, Bernd & Tania Kuteva. 2002b. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.

Henrici, Alick. 1973. Numerical classification of Bantu languages. *African Language Studies* 14. 81-104.

Higgins, Holly A. 2010. *Ikoma orthography sketch*. Nairobi, Dar Es Salaam & Kampala: SIL International, Uganda-Tanzania Branch.

Higgins, Holly A. 2011. *Ikoma vowel harmony: Phonetics and phonology*. Langley, BC: Trinity Western University MA Thesis.

Hill, Dustin, Anna-Lena Lindfors, Louise Nagler, Mark Woodward & Richard Yalonde. 2007. *A sociolinguistic survey of the Bantu languages in Mara region, Tanzania.* Dodoma, Tanzania: SIL International – Tanzania.

Hinnebusch, Thomas J. 1989. Bantu. In John Bendor-Samuel (ed.), *The Niger-Congo languages*, 450-473. Lanham, MD, New York & London: University Press of America.

Holden, Clare J. & Russell D. Gray. 2006. Rapid radiation, borrowing and dialect continua in the Bantu languages. In Forster, Peter & Colin Renfrew (eds.), *Phylogenetic methods and the prehistory of languages,* 19-31. Cambridge: McDonald Institute for Archaeological Research.

Huson, Daniel H. & David Bryant. 2004. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2). 254-267.

Kiraka, Pius & John B. Walker. 2010. *Kabwa orthography sketch*. Nairobi, Dar Es Salaam & Kampala: SIL International, Uganda-Tanzania Branch.

LaPolla, Randy J. 2009. Causes and effects of substratum, superstratum and adstratum influence, with reference to Tibeto-Burman languages. In Nagano, Yasuhiko (ed.), *Issues in Tibeto-Burman historical linguistics* (Senri Ethnological Studies 75), 227-237.

Lewis, M. Paul, Gary F. Simons, & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the world, seventeenth edition*. Dallas, TX: SIL International. http://www.ethnologue.com. (13 August 2013).

Maho, Jouni Filip. 2009. *NUGL online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages*. http://goto.glocalnet.net/mahopapers/nuglonline.pdf. (4 January, 2013.)

Mitterhofer, Bernadette, Hazel Gray & Shona Blatch. 2009. *Dialect survey among the Zanaki people*. Unpublished draft ms.

Mitterhofer, Bernadette & Holly Robinson. 2012. *Short report of the Suba-Simbiti survey*. Unpublished draft ms.

Mitterhofer, Bernadette & Misha Sandeen. 2012. *Short report of the Ikizu-Sizaki survey*. Unpublished draft ms.

Mould, Martin Joel. 1976. *Comparative grammar reconstruction and language subclassification: The North Victorian Bantu languages*. Los Angeles: UCLA Ph.D. Dissertation.

Mous, Maarten. 2000. Counter-universal rise of infinitive-auxiliary order in Mbugwe (Tanzania, Bantu F34). In Vossen, Rainer, Angelika Mietzner & Antje Meissner (eds.), *"Mehr als nur Worte...": Afrikanistische Beiträge zum 65 Geburtstag von Franz Rottland*, 470-481. Köln: Rüdiger Köppe Verlag.

Muzale, Henry R.T. 1998. *A reconstruction of the Proto-Rutara tense/aspect system*. St. John's, NL: Memorial University of Newfoundland Ph.D. Thesis.

Mwita, Leonard Chacha. 2008. *Verbal tone in Kuria.* Los Angeles: UCLA Ph.D. Dissertation. http://www.linguistics.ucla.edu/general/Dissertations/MwitaChachaDissertationUCLA2008.pdf. (25 March 2013.)

Nichols, Johanna. 1996. The comparative method as heuristic. In Durie, Mark & Malcolm Ross (eds.), *The comparative method reviewed: Regularity and irregularity in language change*, 39-71. New York & Oxford: Oxford University Press.

Nurse, Derek. 1999. Towards a historical classification of East African Bantu languages. In Hombert, Jean-Marie & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and Empirical perspectives*, 1-42. Stanford, CA: CSLI Publications.

Nurse, Derek. 2007. *Bantu tense and aspect systems: Appendix 1.* http://www.ucs.mun.ca/~dnurse/pdf/tabantu-app1.pdf. (25 March 2013.)

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

Nurse, Derek & Henry R.T. Muzale. 1999. Tense and aspect in Great Lakes. In Hombert, Jean-Marie & Larry M. Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives*, 517-544. Stanford, CA: CSLI Publications.

Nurse, Derek & Gérard Philippson. 1980. The Bantu languages of East Africa. In Polomé, E. & C.P. Hill (eds.), *Language in Tanzania*, 26-67. Oxford: Oxford University Press.

Nurse, Derek & Gérard Philippson. 2003a. Introduction. In Nurse, Derek & Gérard Philippson (eds.), *The Bantu languages*, 1-12. London & New York: Routledge.

Nurse, Derek & Gérard Philippson. 2003b. Towards a historical classification of the Bantu languages. In Nurse, Derek & Gérard Philippson (eds.), *The Bantu languages*, 164-181. London & New York: Routledge.

Nurse, Derek & Gérard Philippson. 2006. Common tense-aspect markers in Bantu. *Journal of African Languages and Linguistics* 27(2). 155-196.

Nurse, Derek & Franz Rottland. 1991/1992. Sonjo: Description, classification, history. *Sprache und Geschichte in Afrika* 12/13. 171-289.

Robinson, Holly. 2013. *The grammar basics of Ikoma*. Unpublished draft ms.

Roth, Tim. 2011. *The genetic classification of Wungu: Implications for Bantu historical linguistics*. Langley, BC: Trinity Western University MA Thesis.

Rundell, Oliver. 2012. *A comparative discourse analysis of some Bantu languages in Tanzania*. London: Middlesex University MA Dissertation.

Schadeberg, Thilo C. 2003. Historical linguistics. In Nurse, Derek & Gérard Philippson (eds.), *The Bantu languages*, 143-163. London & New York: Routledge.

Schoenbrun, David Lee. 1990. *Early history in eastern Africa's Great Lakes region: Linguistic, ecological, and archaeological approaches, ca. 500 B.C. to ca. A.D. 1000*. Los Angeles: UCLA Ph.D. Dissertation.

Shetler, Jan Bender. 2003. *Telling our own stories: Local histories from south Mara, Tanzania* (African Sources for African History 4). Leiden & Boston: Brill.

Smith, Bryan, Rachel Smith, Holly Higgins & Bernadette Mitterhofer. 2008. *Dialect survey amongst the Ikoma-Nata-Isenye peoples*. Unpublished draft ms.

University of California at Berkeley. 1975. *The pear film.* Online video. http://www.youtube.com/watch?v=bRNSTxTpG7U. (30 April, 2013.)

Walker, John B. 2010. *Kabwa dialect survey report.* Unpublished draft ms.

Walker, John B. 2011. *Discourse features of Kabwa*. Unpublished draft ms.

Whiteley, W.H. 1960. *The tense system of Gusii*. Kampala: East African Institute of Social Research.