Comparative Tone Analysis of Several Bantu D30 Languages (DR Congo)

KENT RASMUSSEN
University of Texas at Arlington
Department of Linguistics and TESOL
kent.rasmussen@mavs.uta.edu
kent_rasmussen@sil.org

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON
May 19, 2018
Version: 4.0 (indexed)

A Comparative Tone Analysis of Several Bantu D30 Languages (DR Congo)
by Kent Rasmussen
2016-2018

Acknowledgements

This dissertation, including field trips to accomplish it, was supported financially by the University of Texas at Arlington (UTA) College of Liberal Arts, The Department of Linguistics and TESOL (UTA), The Jerold A. Edmondson Research Endowment in Linguistics (UTA), The Linguistics and TESOL Alumni Endowment (UTA) and SIL International (Eastern Congo Group - ECG). Fieldwork carried out in support of this dissertation is covered under UTA IRB #2014-0381.

Fieldwork necessary to accomplish this dissertation was also supported logistically by Université Shalom de Bunia (Democratic Republic of the Congo - DRC), SIL International (ECG), and Mission Aviation Fellowship (MAF), Diguna in Bunia (DRC), and three church denominations of the DRC: Communauté Evangélique au Centre de l'Afrique (CECA/20), Communauté Evangélique du Christ au Cœur de l'Afrique (CECCA/16), and Communauté Emmanuel (CE/39). This work was also made possible by the prayer, financial, and personal support of numerous individuals, families, and churches who have partnered with us in our Wycliffe ministry. Thanks!

This work stands on the shoulders of a number of technologies, two of which I would like to specifically thank here. Much of my data was collected with development versions of WeSay (www.wesay.org), which involved vigorous dialog with WeSay developers over a period of years. Thanks to you all for championing a product that helps people with little to no computer experience get involved in work on their own language; my ability to include so many people in my work is a direct result of your efforts. I have also greatly benefited from extensive Email dialog with Andy Black, who supports the linguistic writing XML specification which I used for this dissertation. His contribution to XLingPaper (software.sil.org/xlingpaper/) made this document much more principled and less painful than my earlier graduate work, as well as vastly superior in terms of its technical layout.

An initial analysis of consonants, vowels, and lexical tone in nouns for Nyali-Kilo, Vanuma, and Bira was conducted at a workshop run by the Eastern Congo Group (ECG) of SIL International and directed by Constance Kutsch Lojenga, to whom I owe a great debt of gratitude for her visionary and analytical insights toward helping people read well. I directed the initial work in Nyali-Kilo under her supervision, then extended the same basic methodology into Ndaka and Mbo.

Neither this dissertation, nor the analysis that underlies it, would be possible without the generous support and personal contribution of numerous members of the Ndaka, Mbo, Nyali-Kilo, Vanuma, Budu, and Bɨra communities. Each workshop was a collaboration between a large number and wide variety of people. Those who participated more or less full time in at least one of the workshops should be listed below.

Ndaka data come from Dengaino Banapia Simon, Gbandu Camile Christia, Angama Gbandu Camile, Mazandu Ikpati Ngaga Simon, Mazandu Bangamosigo Jean-Pierre, as well as from Banapia Koingo Samuel, Bangamosibo Mazandu Marc, Bibengama Koingo Jean-Pierre, and Natumani Abedi Ramazani, then from Asibemani Frederic, Bunangane Alphonse, Mambo Mazuoko Pascal, Bagwetani Jean, Bapiagbani Victor, Atanza Sara, Kalunga Ainga Kolo, and Banajonio Bakusi Ivone.

Mbo data come from Amae Bapeido Vedaste, as well as from Iduanza Asani Vicky and Mamboya Ambasani Baudouin, then from Kasongo Bafogbombi Josue and Idimi Banasungwa Mussa. The above were aided in the first week of the second workshop by Kayumba Isiya Bugumba, Bausane Jean-Pierre, Bokumo Fundi Leonard, Sabeti Adengana Tabo Pierre, Kokyaneno Valili, and Ikiango Baboda Jada Sermali.

Nyali-Kilo data come from Koko Walu David, Basugodite Ezambi Martin, and Akilisende Bandenonga John.

Vanuma data come from Bakabukingi Bolanao William, Babote Rwabongo Gustave, Kosianja Kosianja Michel, and Beli Bokyou.

Budu data come from Bhambéníkombó Gilbert and Ingyáyékombó Emmanuel (each from Wamba, bafaKoya).

Bɨra data come from Malali Zanabo Dieu-donné, Katanabo Hemukabo Jean, Ngamakali Sumbabo Dieu-donné and Jean Katanabo Hemukabo.

A number of Ndaka community members also made significant personal contributions to the running of the workshops in Nya-Nya, including Kopikama Ruth Joséphine, Bakumbu Banapia Esthère, and Anjandongo Banapia Régine.

I would also like to acknowledge numerous people who have been instrumental in my academic development as a linguist. Doris Payne gave me my first introduction to a phoneme, as well as to fieldwork methodology and African languages. Each of these continue to shape most of what I do today. Scott Delancey gave me my first taste of historical and comparative linguistics, which again has born fruit I did not imagine at the time.

While beating my head against Maasai [mas] verb tone, I ran into Keith Snider, who has since become a very valued mentor and friend. His dedicated and detailed efforts over the last decade and a half have helped me grow as a linguist, particularly in the study of tone.

This dissertation would not exist (at least not in its present state) without the work of Mary Bradshaw, who —I am grateful to say— once had the opportunity to set me straight face to face. Her doctoral advisor, David Odden, has also made himself available on a number of occasions, leading to very fruitful conversations on a range of issues. I am grateful for his comments on my work, especially in pushing me beyond limits I had set for myself. But perhaps even more helpful was the chance to think through phonological issues with him on a broader level, making connections between my work and that of others, including theoretical work still in development. I would like also to thank Joey Sabbagh, who has given me the opportunity to rethink my positions on syntax and semantics, but also the free reign to chase my phonological data where they lead me. And finally, I have received much help from Cindy Kilpatrick, who directed this dissertation. Her numerous and copious comments on my work, as well as her careful and personal attention to all questions linguistic and political, have been invaluable to my progress in this program.

Words cannot express the debt that I owe to my wife Kimberly. Her love, devotion, and support have enabled almost everything listed above. Thank you for standing beside me in the hard times and places, and for sticking with me for better and for worse. I love you and look forward to becoming more indebted to you in the decades to come.

May 4, 2018

A Comparative Tone Analysis of
Several Bantu D30 Languages
(DR Congo)

Kent Rasmussen, Ph.D.

The University of Texas at Arlington, 2018

Committee members:

____
Cynthia Kilpatrick, Ph.D. (chair)

____
Keith Snider, D.Litt.

____
Joseph Sabbagh, Ph.D.

Abstract

A Comparative Tone Analysis of
Several Bantu D30 Languages
(DR Congo)

Kent Rasmussen, Ph. D.

The University of Texas at Arlington, 2018

This dissertation presents an analysis of the tonal systems of the Bantu D30 languages. There are 34 Bantu D languages in the Democratic Republic of the Congo (DRC), few of which have developed orthographies (Lewis, Simons, & Fennig 2013). These languages make use of contrastive pitch (tone), which is a source of difficulty in their analysis. I begin, therefore, with a review of literature arguing for a number of positions regarding the collection, analysis, and representation of tone data. Additionally, I describe the fundamentals of consonant-tone interaction, and some background on the Bantu D30 languages, including their genetic and geographic context.

My research describes infinitive and finite verb forms, mostly in Ndaka [ndk] and Mbo [zmw], though also including data from other Bantu D33 languages Budu [buu], Nyali-Kilo [nlj], Vanuma [vau] and Bantu D32 Bɨra [brf]. I further compare the Bantu D30 languages, in order to provide insight into the analysis of consonant-tone interaction (CTI). Specifically, these descriptions and comparisons form the basis for my argument that a synchronic application of Bradshaw (1999)'s [L/Voice] cannot account for Bantu D30 data. Rather, a historical and comparative perspective is needed to account for data not predicted by Bradshaw (1999), if her fundamental claims are to remain intact.

This research has two primary aims. First, I describe the three lexical verb root classes by tone in Ndaka and Mbo, along with data from other Bantu D30 languages. Second, I provide comparative data in the Bantu D30 languages. These comparative data allow one a better view of certain consonants of the Bantu D30 languages. A diachronic view of these consonants is further helpful in understanding current tone patterns as the result of historical consonant-tone interaction (CTI). Ultimately, this work contributes to the understanding of the genetic

relationships between the Bantu D30 languages, which is somewhat debatable at present.

Perhaps the greatest theoretical contribution of this dissertation is a new model of tonogenesis (the creation of new tone melodies). This new model provides an account for the large majority of available Bantu D30 data fitting a distribution as expected by Bradshaw (1999)'s [L/Voice], while at the same time accounting for a non-trivial number of exceptions to that distribution. This model accounts for the Bantu D30 data without rejecting Bradshaw (1999), but also provides a model for how tonogenesis may have occurred, which can then also be applied to the origin of tone melodies in other languages.

Table of Contents

To the Glory of God

List of Figures

List of Tables

List of Abbreviations

ISO 639-3 codes used in this document:

bas	Basaá (A43)
bhy	Bhele (D31)
bip	Bila (D31)
brf	Bɨra, Bera, Plains Bira (D32)
buu	Budu (D33)
kkq	Kaiku (D31)
kmw	Komo (D20)
lik	Lika, Liko (D20)
ndk	Ndaka (D33)
nlj	Nyali-Kilo (D33)
swc	Congo Swahili (G40)
sqm	Suma (Adamawa-Oubangi)
tvu	Nen, Tunen (A44)
vau	Vanuma (D33)
zmw	Mbo (D33)

Other abbreviations used in this document:

DRC	The Democratic Republic of the Congo
BLR3	Bantu Lexical Reconstructions 3 (Bastin, et al 2002)
ECG	the Eastern Congo Group of SIL International (www.sil.org)
MAF	Mission Aviation Fellowship (www.maf.org)
GIMP	GNU Image Manipulation Program (www.gimp.org)
Bɨra	Bantu D32 language of the DRC; ISO 639-3: [brf]
D30	Bantu D30 languages, including D31, D32, D33 and other non-vigorous languages
D31	Bantu D31 languages: Bila [bip], Bhele [bhy], and Kaiku [kkq]
D32	Bantu D32 language: Bɨra [brf]
D33	Bantu D33 languages: [ndk], [zmw], [buu], [nlj], and [vau] and other non-vigorous languages
IPA	The International Phonetic Alphabet

TBU	Tone Bearing Unit (mora)
C1	First root consonant
C2	Second root consonant
px	prefix
FV	final vowel (nonlexical; required by CV structure)
H	high tone feature
L	low tone feature
F_0	Fundamental Frequency
$\mathcal{h}$	higher register feature
ℓ	lower register feature
[L/Voice]	Privative feature combining low tone and voice (Bradshaw 1999)
[ℓ/Voice]	Privative feature combining lower register and voice
RTT	Register Tier Theory (Snider 1999)
ATR	Advanced Tongue Root (vowel harmony feature)
TAM	Tense, aspect, and mood
past	Formal TAM category with at least some past functions.
future	Formal TAM category with at least some future functions.
CTI	Consonant-tone interaction
TRN	Tonal Root Node (Snider 1999)
1s	first person singular
1p	first person plural
2s	second person singular
2p	second person plural
3s	third person singular
3p	third person plural

Chapter 1
Introduction

This dissertation provides an introduction to the tone systems found in some of the Bantu D30 languages, a group of closely related languages spoken in the northeastern corner of the Democratic Republic of the Congo (DRC). These languages are as interesting as they are unstudied. There are few descriptions of any kind for these languages, yet what is already known shows data that speak to interactions between consonants and tone, as well as to tonogenesis (the origin and development of new tone melodies) in Africa.

It is my hope that the comparative method employed here will serve as a model for others to follow when studying related languages, but also that this study will provide help to the communities that speak the Bantu D30 languages, as these communities move forward with their own decisions on how to write their languages.

1.1 Problem Statement

The languages of northeastern Democratic Republic of the Congo (DRC) pose a particular conundrum for linguists. Our understanding of language generally is based on cross-linguistic surveys, but the Bantu D30 languages are not included in those studies because of the lack of data generally available on them. Yet Congolese languages, and in particular the Bantu D30 languages in northeast DRC, are rich and unique in their position as Bantu languages on the edge of the "Macro-Sudan Belt" (Güldemann 2008, c.f. Dalby 1970), the convergence of Nilo-Saharan, Niger-Congo (Bantu and non-Bantu), and Afro-Asiatic languages. For instance, while proto-Bantu has been reconstructed with only seven vowels (Hyman 2003a, Odden 2014), at least six Bantu D30 languages have nine contrastive vowels (Kutsch Lojenga 2006, 1994, Morgan & Van Otterloo 2009, Rasmussen 2012, 2014, 2015a).[1] Furthermore, several of the Bantu D30 languages show at least some variety of consonant-tone interaction (CTI). At least some of the languages with CTI show evidence of lexicalization, leading to more complex tone systems (Rasmussen 2012, 2015a) than the two tone system reconstructed for proto-Bantu (Odden 2014).

[1] This claim is laid out in more detail in Section 2.3.1.

The Bantu D30 languages, therefore, have the potential to provide data which are very interesting for the study of new vowel and tone contrasts. Yet there remains very little relevant data collected in these languages. The lack of data is in part due to the difficulty associated with doing tone work, compounded by the difficulty typical of doing fieldwork in central Africa. One can only expect this paucity of data to remain, until such a time as we take the initiative to collect and analyze tone data in eastern DRC. This dissertation seeks to address the lack of data in the languages of eastern DRC by increasing both the breadth and depth of what has been published on these languages.

1.2 Purpose

The purpose of this study is threefold: empirical, theoretical, and developmental. Empirically, I aim to reduce the number of African languages in which no tonal analysis has been done, both by analyzing a limited number of languages and by providing a model for others to follow. Theoretically, at least some of the Bantu D30 languages share a number of properties which may be considered novel within the Bantu languages, most notably a nine vowel ATR system (Kutsch Lojenga 1994, Rasmussen 2012, 2015a, 2015b) and depressor consonants (Rasmussen 2012, 2015a, 2015b, c.f. Bradshaw 1999). A comparative analysis of the Bantu D30 languages sheds more light on these innovations, helping us better understand their genealogy.

With regards to development, data collection and analysis has been done in a manner that encourages the involvement of those who speak the languages under study, so far as possible (i.e., participatory methodology; Kutsch Lojenga 1996). Because of this participatory focus, my research has benefited from a richer perspective on the languages in question, and the community benefits from a grassroots involvement in the analysis, as well as a better grasp of that analysis and its implications (e.g. for literacy).

1.3 Background

The eighteenth edition of the Ethnologue lists 2,138 languages spoken in Africa, 210 of which are spoken in the Democratic Republic of the Congo (DRC; Lewis, Simons, & Fennig 2015). Conservative estimates, based on my own experience as well as conversations with others who have worked in the area, indicate that at least half of those spoken in DRC have contrastive pitch (tone), yet there are relatively few published tonal analyses for DRC languages. Bantu languages generally provide fertile ground for comparative work, because of the number of comparatively closely related languages. But the data to be compared must be first collected and analyzed. For instance, there are 39 languages in the Bantu D grouping (summarized in appendix A and discussed in more detail in Section 2.3.1). Perhaps 14 Bantu D languages have some level of development, typically including a draft orthography, or some basic literacy materials. But I am aware of no tone analysis for any of these languages, except for Lika [lik] (de Wit 2010)[2].

The reason for the lack of general language development in Bantu D is at least twofold. First, there is a lack of human development in the DRC generally, even more so than elsewhere in Africa. The DRC ranked last in human development in the world in both 2011 and 2013 (UNDP 2011, 2013) and in the welfare of women and children (Geoghegan 2013). The lack of human development coincides naturally with rampant illiteracy (Henderson, Rohloff & Henderson 2014, Alege 2012), aggravating the lack of language development already present.

In addition to the difficulty of accessing languages in areas with such a low state of human development, the languages of northeast DRC have a level of complexity beyond what is normally seen in Bantu. These languages are situated on the edge of what has been called the "Sub-Saharan Fragmentation Belt", or the “Macro-Sudan Belt” (Güldemann 2008), the convergence of Nilo-Saharan, Niger-Congo (Bantu and non-Bantu), and Afro-Asiatic languages (Good 2013, Dalby 1970). These genetically distinct languages meeting in northeast DRC provides a level of complexity that makes language data collection and analysis more difficult than elsewhere, since expectations based on related languages often do not hold. Furthermore, linguistic and material poverty interact, as minimal language development fuels and is fueled by the economic and military turmoil of the region.

The difficulties mentioned above are especially concerning in the collection and analysis of tone data. Problems particularly associated with tone data are in part due to the difficulty of analyzing tone systems generally, but also due to the lack of linguists trained to collect and analyze data which are useful in a principled tone analysis.[3] Whatever the cause, few published tone analyses means we lack broad typological data on which to make valid claims about the nature and origin of tone in Africa.

But the difficulty of doing linguistic fieldwork in DRC does not negate the fact that the languages of the DRC are much in need of analysis, both for the sake of

[2] I have reason to believe that some tone work has been done in Budu, but I have been unable so far to find a copy of it, or even to clarify who did the work or when. Asangama (1983) contains what appear to be tone marks on much of its data, but no particular tonal analysis, per se.

[3] See Section 2.1 for some of the issues involved.

the communities that speak them, and for the sake of a more truly cross-linguistic basis for our understanding of tone and other facets of language.

1.4 Questions

Given the little data from eastern DRC incorporated into cross-linguistic surveys, one must ask the question as to whether that data, once provided, will match well with the conclusions already drawn. Specifically, I have collected and analyzed tone data from Bantu D30 languages to understand their tone systems, but also to see how that data hold up to our current understanding of how tone works.

This study looks at consonant-tone interaction (CTI), a phonological treatment of phonetic correlations observed since at least Beach (1923), where voiced consonants systematically lower the pitch on a following vowel. Bradshaw (1999) provides a cross-linguistic survey of consonant-tone interaction, including a proposal of a feature [L/Voice].[4] The question remains, however, whether Bradshaw (1999)'s [L/Voice] can account for the data coming from other languages, such as the Bantu D30 languages.[5]

The ability of Bradshaw (1999) to account for the Bantu D30 languages is relevant because three of these languages (Nyali-Kilo [nlj], Ndaka [ndk] and Mbo [zmw]) have already been shown to have consonants that appear with tone patterns otherwise associated with depressor consonants, where the consonants themselves are not typical depressor consonants (Rasmussen 2012, 2015a, and 2016, but see also Chapters 3–5 of this dissertation). Furthermore, at least some of the consonants in question do not consistently depress; that is, they appear both with tone patterns typical of depressor consonants, and with tone patterns typical of non-depressor consonants (Rasmussen 2012, 2015a). Thus, while Bradshaw (1999)'s [L/Voice] covers over 80% of the data I present in this dissertation, there is some data which [L/Voice] does not adequately predict.

This difficulty leads to a secondary question: Where the analysis of Bradshaw (1999) doesn't seem to work for the Bantu D30 languages, can a historical and comparative approach enable us to understand the Bantu D30 languages, without needing to modify the essential claims of Bradshaw (1999)? It has already been suggested that at least some of the difficulty accounting for the Bantu D30 languages with Bradshaw (1999) can be resolved by a historical and comparative historical analysis (Kutsch Lojenga 2006, Rasmussen 2012), but resolving this

[4] Details are given in Section 2.2.

[5] This subgroup of the Bantu D languages is discussed in more detail in Section 2.3.1. See also appendix A for a chart of these languages in their genetic context.

difficulty requires more comparative data for the subgroup. The amount of data already documented as exceptional to a strict [L/Voice] analysis (Rasmussen 2012, 2015a, and 2016, but see also Chapters 3–5 of this dissertation) is enough to pose serious problems for Bradshaw (1999), but in this dissertation I will resolve them through a historical and comparative analysis.

1.5 Significance

Given the lack of tone analyses in African languages generally, the primary value of this dissertation is broadening the base of descriptive data from which we can draw typological generalizations about tone. I do this by providing a comparative analysis of the Bantu D30 languages, thereby making data and some analysis for these languages available. Methodologically, I show that comparative tone analysis can provide data for multiple languages more advantageously than the analysis of one language at a time. As such, I hope to provide a model for those seeking to analyze tone languages in Africa, helping fill the gap of languages that remain unanalyzed today. In terms of advancement of our theory of language, more and better language data available means more and better data on which to build our understanding of language more generally. And specifically, this dissertation provides a model by which we can understand tonogenesis, the creation of new tone melodies.

1.6 Scope

This dissertation continues in Chapter 2 with a review of relevant literature on tone, consonant-tone interaction, and language classification, followed by additional background relevant to the segmental phonology of the Bantu D30 languages, including their vowels, consonants, and criteria to identify infinitive verbs.

Because my work has been intentionally comparative, it is limited in the number of questions addressed. This is done to maintain a manageable scope, but also to focus on useful points of comparison. The first and perhaps most basic question in any tone analysis is how many tone melodies there are for a given grammatical category of morphemes, and how they can be characterized. The next question is how the observed tone melodies compare between contexts, within a language. These two questions will be treated in Chapters 3, 4, and 5, for each of Ndaka [ndk], Mbo [zmw], and Nyali-Kilo [nlj], respectively. In each of these chapters I compare two basic contexts: infinitive forms in isolation[6], and simple conjugated forms preceded by a pronoun and followed by a time adverb. The surface melodies in each of these contexts should correlate in some manner; what do observed correlations say about underlying verb root classification by tone?

Because the Bantu D30 languages also show interesting consonant-tone interaction (CTI), I also ask how consonant types correlate to one or more of the observed verb root classes by tone. Because Bradshaw (1999) predicts a clear correlation between segments specified for voicing and CTI effects, I end each of these chapters with a summary of the extent to which this prediction is borne out in the data, with the actual observed distribution of consonant types across verb root classes by tone.

Beginning in Chapter 6, I ask how the inventory of tone melodies compares between languages, both generally, and in terms of specific sets of cognates. This chapter also sets up a number of regular sound correspondences across the Bantu D30 languages, which will form a basis for comparison in Chapter 7. And finally, I examine the correlation between voicing specification and tone melodies between each of the languages.

Chapter 7 then seeks to provide a model by which we can understand the consonant-tone interactions unexpected by Bradshaw (1999) as part of a natural historical process which includes the central thesis of Bradshaw (1999), but which ultimately goes beyond Bradshaw (1999) to create new lexical tone melodies.

The body of the dissertation concludes with Chapter 8, bringing together the data presented in Chapters 3–6 with the model and data presented in Chapter 7, to summarize the overall contributions of this dissertation.

Due to the number of lines of research that have come out of this study, but which don't directly impact the analysis of the verb root classes by tone in Chapters 3–6, I include a number of appendixes. The first (A) is a family tree of the Bantu D30 languages for the convenience of the reader, compiled from Lewis, Simons, & Fennig (2015).

Appendix B describes the methodology I used to create the figures in this dissertation which consist of multiple pitch traces overlaid on top of each other.

Appendix C treats the relative functional load of tone in the Bantu D30 languages. This appendix addresses the relative need for tone to be expressed in

[6] It is generally advisable to start a tone analysis with the part of the system that looks to be the simplest and most straightforward (Snider 2018). Because of the potential confounding variable of tones attached to noun classes (c.f., Snider 2014), and because a cursory examination shows many more tone melodies on nouns in isolation than on verbs, I begin with verbs. Working with verbs also allows for a more straightforward cross-linguistic comparison.

orthography, a very real and important question for the people who speak these languages.

Appendix D treats briefly (V)C roots, which are excluded from the body of this dissertation as non-canonical. Their study does provide, however, a number of important and interesting observations, including the permanence of an ATR feature in absence of a vowel in the underlying form.

Appendix E discusses possibilities for integrating the feature geometries contained in Snider (1999) and Bradshaw (1999), two theoretical works of practical importance to tone and consonant-tone interaction, respectively.

Appendix F treats a possible extension of the model presented in the body of the dissertation to other cases, including proto-Bantu and tonogenesis *ex nihilo*.

Also for the convenience of the reader, I include two indexes of language data: one of segmental word forms, the other of glosses. This should help the reader find particular data of interest in this dissertation; these appear after the references.

Chapter 2
Literature Review

Having briefly introduced the subject of this dissertation in Chapter 1, this chapter provides more detailed background on selected topics helpful to understanding and answering my research questions. Section 2.1 is addressed to those without a particular training in tone analysis. The principles of good phonological analysis are sometimes easily transferable to tone studies, but the methods and expectations available to those with a more general linguistic training may not be the most helpful in the analysis of tone. Topics covered include the relationship between fundamental frequency (F_0), pitch, and tone (Section 2.1.1), the relative nature of tone data (Section 2.1.2), the nature and proof of tone contrasts (Section 2.1.3), the relationship between tone and register (Section 2.1.4), and certain problems inherent in using diacritics to present phonetic data for tone analysis (Section 2.1.5).

Section 2.2 treats a topic less relevant to some tone analyses, but central to this dissertation: consonant-tone interaction (CTI). It gives a brief overview of our understanding of consonant-tone interaction, focusing on Bradshaw (1999)'s thesis that CTI can be accounted for by a single, privative feature: [L/Voice]. I assume Bradshaw (1999); as mentioned in Section 1.4, I also evaluate its ability to account for the tone systems found in Bantu D30 languages.

Section 2.3 addresses the Bantu D30 languages, in three dimensions. The first covers their larger contexts of genetics, geography, and two points of larger phonological interest, in Section 2.3.1. The next two sections each describe methods of categorizing languages. In Section 2.3.2, the comparative method is introduced as a means to compare languages one to another in terms of their genetic relationships. Areal linguistics in introduced in Section 2.3.3. This field, which could also be called geographical typology, organizes languages into units of geographically close languages with similarities which are accounted for by geographic contact, rather than by genetics or universal principles. Together, Section 2.3 gives the reader a background on the Bantu D30 languages, in terms of some context for the languages themselves, their genetic relationships (between each other and with other languages), and their similarities to neighboring, unrelated languages.

This literature review finishes with Section 2.4, which provides more specific background relevant to the Bantu D30 languages, including their vowels and consonants, but also the functional and syntactic criteria I use to identify infinitive verbs, and the categories and frames I use to analyze them.

Section 2.5 gives a brief overview and criticism of a small portion of Kutsch Lojenga (2006), that portion of the book chapter which describes the addition of new tone through loss of voicing in Bila [bip], a Bantu D30 language not otherwise treated in this dissertation.

2.1 Tone Data and Analysis

This section covers essential background for linguists who are not specifically trained in the study of tone. The following subsections deal with the relationship between fundamental frequency (F_0), pitch, and tone (Section 2.1.1), the relative nature of tone data (Section 2.1.2), the nature and proof of tone contrasts (Section 2.1.3), the relationship between tone and register (Section 2.1.4), and particular problems inherent in using diacritics to present phonetic data for tone analysis (Section 2.1.5).

2.1.1 Fundamental Frequency (F_0), Pitch, and Tone

Yip (2002:5) defines three categories relevant to the phonetic and phonological study of tone: fundamental frequency (F_0), pitch, and tone. F_0 tracks the number of vibrations per second in the vocal folds; it is objectively measurable in hertz (hz). On the contrary, pitch is what is perceived by human physiology. It includes what the auditory apparatus is capable of hearing, less the details filtered out by the human cognitive function. Pitch may be used to describe linguistic and non-linguistic events, and may arise as a result of lexical and/or prosodic functions. Finally, tone is a linguistic category, which applies to relative pitch differences that impact lexical or grammatical meaning.

From the above definitions of F_0, pitch, and tone, a number of important principles can be gleaned. First of all, the terms F_0, pitch, and tone are not even roughly equivalent. They describe different kinds of things, in addition to whatever level of detail might distinguish them.

Secondly, F_0 is the only objectively measurable value. However, it contains detail that is unimportant to pitch or tone, including that which is imperceptible to the human ear. Understanding the capacity of F_0 to carry information beyond what is meaningful in language is critical to not overvaluing the details in a pitch trace (calculated F_0 over time), which carries information which varies from one production of a word to another, without necessarily making any meaningful (or perhaps even noticeable) differences in production.

Third, there are two layers of abstraction between F_0, pitch, and tone. The human auditory apparatus takes in F_0 and abstracts or interprets pitch, including only the information relevant to human physiology and cognition. This abstraction is similar to someone taking in the sound waves resulting from production of a consonant, and hearing [b], [p˺] or [pʰ]. Alternatively, one may hear some non-linguistic sound, such as a raspberry (interlabial trill). But whether the sound is judged to be language or not, our ears and brain take in and process a subset of what is present in the sound wave. Tone is a further abstraction, including only information relevant to human language. The linguistic category of tone underlies and categorizes pitch in the same way a segmental phoneme (e.g., /p/), may underlie and categorize one or more allophones (e.g., [b], [p˺] or [pʰ]).

Practically, this distinction between F_0, pitch, and tone means that we need some way to get from what is objectively observable (i.e., F_0) to what is subjectively perceived as phonetics (i.e., pitch), in order to use that as a basis for analyzing phonological categories (i.e., tone). I do that by overlaying multiple F_0 traces of utterances judged to have the same tone, such as in Figure 1 and Figure 2, previewed here from Chapter Section 3. For each of Figure 1 and Figure 2, some 50 utterance F_0 traces are superimposed on each other. This allows one to observe the general trend of pitch; I include my judgement of that trend through the thicker bands[1].

[1] See appendix B for more information on how I make these images more generally, including the trendlines.

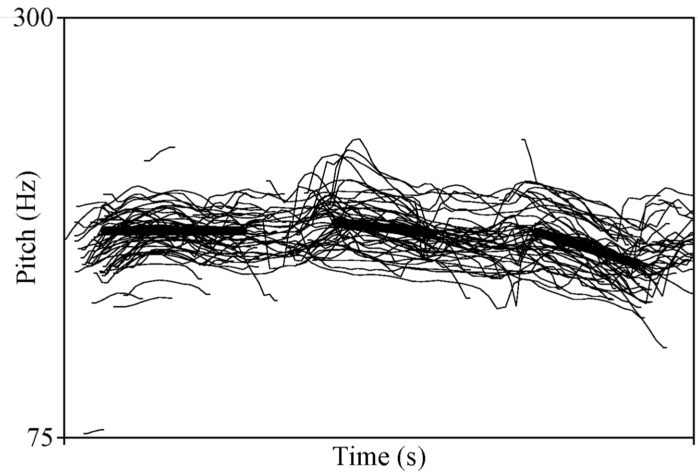


Figure 1. Overlay of infinitive pitch traces for Ndaka [ndk] Low verbs

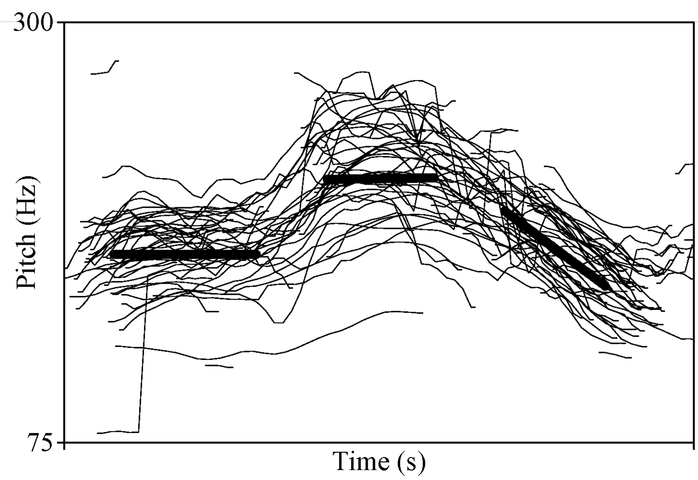


Figure 2. Overlay of infinitive pitch traces for Ndaka [ndk] High verbs

The reader is of course free to disagree with my judgements of the trends in these figures, but this possibility of disagreement is to some extent the point. I seek to justify my phonetic transcriptions in a manner than allows real, constructive criticism. Because we are transcribing something which is by nature to some degree subjective, I must of course accept some level of disagreement. Even if one may disagree with either transcription on some small point, I trust there is agreement that [– – –] is an essentially faithful transcription of Figure 1, and that [– ¯ ⟍] is an essentially faithful transcription of Figure 2. To push the point just a bit further, it is my goal to move from objective to subjective observations in a principled and defensible manner, rather than to justify distinctions between objective observations with statistical tests, which would in this case ultimately analyze differences which are not important to language[2]. Any defense of these subjective judgements, however principled, must acknowledge that phonetic pitch transcription is always to some degree subjective.

To maintain the distinction between high or low pitch on the one hand, and high or low tone on the other, I will use “H” and “L” for phonemic units of tone, and “high” and “low” as more descriptive or phonetic terms (though including the description of tone when explicitly indicated, as I have done in this sentence). I also use the capitalized “Low” and “High” as names of verb root classes by tone, e.g., in Section 3.1 and Section 3.2, respectively.

Despite what may be a near universal gut feeling that tone is fundamentally different from other sounds we study, the number of relevant categories, and the relationships between them, are not significantly different from those in the study of consonants and vowels. In each case, we perceive with our ears less than all the information carried in the sound wave, and we interpret as relevant to linguistic meaning less than all that we are capable of hearing.

2.1.2 The Relative Nature of Tone Data

Though the number of relevant categories, and the relationships between them, are not significantly different from those in the study of consonants and vowels, we must still come to terms with the nature of the abstractions mentioned in Section 2.1.1, i.e., that pitch and tone are relative categories (Beach 1923; Snider & van der Hulst 1993). As both pitch and tone are impacted, it follows that relevant data are relative both phonetically and phonologically.

[2] I say this not to disparage statistics, which certainly have their place in science. Rather, I say this because I have frequently heard the comment that I should base these trendlines on a statistical analysis of the underlying F_0. But in addition to the fundamental difficulty of making meaningful statistical tests for the amount of data in these overlays, I question the point of digging further into the details of these F_0 traces, given that we do not hear most of those details. The result of such a statistical analysis would be to analyze differences which are not important to language, the opposite of my point in making these images. I seek to capture what is relevant to human language, which must ultimately be a subjective process. If the reader essentially agrees with my transcriptions, then these images have accomplished their purpose. If the reader essentially disagrees with my transcriptions, then let's discuss that. But in neither case do I think that statistics will help us decide and move on to the phonological analysis.

Phonetically, pitch is relative both within and between tones. Within a given tone bearing unit (TBU), one can describe its pitch as "level", "rising", or "falling", only by comparing the pitch at various points across the TBU to each other. Between TBU's, one can describe a tone as phonetically "high" or "low" only by comparing it with the tones that surround it. Tone is especially relative when dealing with a variety of speakers, as the F_0 of one person's high tone may be below that of another person's low tone, even when both tones are in the same context (e.g., utterance initial). Interpersonal relativity is particularly relevant when considering pitch differences between men and women; F_0 values are essentially meaningless until compared with what comes before and after them. Because of these multiple types of relativity, tone cannot be interpreted outside of a larger pitch context. Because of this need for context, without surrounding context for comparison the phonetic transcriptions [¯], [–], and [_] are indistinguishable, as each represents a single level pitch.

Phonologically, the value of a tone category is relative to what else could occur in that same position of an utterance. For instance, a high tone is generally considered high because it is higher in opposition to something else that could occur at the same position —depending on the analysis, that opposition might be to a low tone[3]. But neither high nor low tones have absolute values; they are each defined in terms of the other.

The implications of high and low tones being defined in terms of one another may not be immediately obvious, so I propose a comparison to segmental data. The phonetic implementation of what is heard as [p] may vary from person to person, and from utterance to utterance. But there is a limit to the variation allowed in what could be transcribed [p]; in the extreme case, someone who pronounces [p] without using the lips would certainly have communication issues. In fact, one would generally say such a production isn't [p] at all, since IPA [p] denotes, by convention, a voiceless **bilabial** plosive. But there is no similar articulatory requirement placed on what is or isn't a high or low tone; at no point can one look at the F_0 of a vowel (or the position of the glottis), and say that it cannot be a particular tone.

Similarly, the phonological category /p/ may include a number of allophones (e.g., [pʰ], [p˺] and [p]), yet we would be surprised to have /p/ include [k], for instance, or for it to include [b] or [ɓ] without also including [p]. High tone, on the other hand, makes no claim whatsoever about the F_0 pitch at which it is pronounced, except that it is higher than it would be if it were a low tone.

[3] Alternatively, a high tone may oppose the lack of underlying tone specification, as in a privative H system (c.f., Hyman 2000)

So the phonetic implementation of pitch in a tone language shows a variation and lack of constraint that we don't expect in segments. As a result, the F_0 associated with a low tone early in an utterance may easily be higher in value than the F_0 associated with a high tone later in the same utterance. Or, the highest F_0 in the entire utterance of a male speaker may be lower than the lowest F_0 of a female making the same utterance. The implication of these facts for tone analysis is that it is impossible to draw a categorical correspondence between underlying linguistic categories and F_0. We may collect F_0 data objectively, or else transcribe pitch data subjectively, but it is impossible to correctly interpret either outside of at least some minimal context.[4]

The relative nature of tone is a source of much of the difficulty presented to researchers collecting, describing, and analyzing tone data. Anyone who approaches tone data armed only with expectations built by experience with segmental data should expect trouble with tone. One should expect difficulty faithfully transcribing and analyzing what is important in tone data, if unequipped to deal with the critically important relativity of tone.

2.1.3 Establishing Tone Contrast

Since the appearance of Chomsky & Halle (1968), the analysis of words into segmental phonemes has proceeded by a fairly straightforward process. We understand segments to be composed of distinctive features, which are distinctive precisely because each may be the only difference between the value of one segment and that of another, and hence between one meaningful word and another. These segmental features have been understood as operating within a particular geometry, but tone features did not initially figure greatly in any geometry, even when describing tone languages (c.f., Halle 1995 and Yip 1989).

The insights of Leben (1973) and Goldsmith (1979)'s Autosegmental Theory provided an elegant way of capturing the behavior of phonological information which is not strictly attached to segments, but which is nonetheless organized and in some manner associated to segments. The classic example of autosegmental information is tone, which is clearly associated to vocalic length/weight (mora), but which also spreads and moves from one mora to another.

Still, the analysis of tone has not always proceeded in a principled manner, for a number of reasons. Perhaps foremost is the fact that researchers expect tone to be

[4] I have heard that some native speakers of some tone languages can give tone values for monosyllables out of context. While some may be able to do so, the fact remains that they are interpreting language data from a language internal perspective. As a result, it would be difficult to separate the perception of tone via pitch from the perception of other redundant cues that would also identify a given word with its tone value.

different, and so treat it differently. Yet while tone data must be treated differently than segmental data (c.f. Section 2.1.2), and while underlying tone categories are less strictly bound to the segmental tier than segments are, the principles of establishing contrast should remain the same for tone.

As a result, it is possible to treat tone as too different, and too much the same, at the same time. If tone is treated as having special principles of establishing contrast (against what is said here), yet being essentially the same kind of data as segmental data (against what is said in Section 2.1.2), then we will both misunderstand our data, and treat them badly. That kind of analysis is not the most helpful to develop a comprehensive and consistent tone analysis for a language. A principled tone analysis acknowledges the uniquely relative nature of tone (as described in Section 2.1.2), while maintaining consistent reasoning for establishing phonological contrast.

Snider (2018) discusses a methodology to establish underlying tone contrasts in a principled manner, extending Snider (2014). His basic claim is that underlying contrast must be established through the comparison of tone data which are truly comparable. Comparisons must control for every factor which may affect the surface tone melody (the collection of pitches across an utterance). This includes phonological factors, such as the syllable structure, including the presence and sonority of syllable codas, as well as syntactic factors, such as grammatical categorization and surrounding morphemes.

It is particularly important to control for syntactic factors because one does not know where floating (unassociated) tones exist, especially in the initial analysis. If a particular floating tone is categorically attached to all nouns, for instance, then comparing a noun and a verb shows a difference between nouns and verbs, but not the difference between the values of tone (e.g., high v low v mid) on a set of roots. Similar problems occur when comparing words with and without extra morphemes, which almost certainly bear their own surface tones, but may also bear floating tones. Because of the problems for tone analysis inherent in not controlling syntactic context, I compare verbs with verbs, in one morphosyntactic context at a time.

One final point is important to establishing tone contrast: the melodic nature of tone. It is often clear early in a tone analysis, that there are a certain number of part of speech subcategorizations by tone, or groups of morphemes according to tone behavior. But what is the nature of the lexical categorization by which that tone behavior is determined? Leben (1978) provides compelling evidence, which is used by Snider (1999, 2014 and 2018) to argue persuasively that tone languages distinguish morphemes using tone melodies (i.e., set of lexical tone contrasts lexically assigned to a morpheme), rather than some smaller unit used to build those melodies, such as a L or H “toneme”. One implication of this fact is that a system which opposes H and L melodies on verbal roots, for instance, is simpler than one that opposes H, L, HL, and LH —despite the fact that they both may be described as “two tone” systems. As a result, a system that has gone from {H v L} to {H v L v HL v HL} has added new phonemic categories (the HL and LH tone melodies), and the mechanism by which that could be done must be interesting to one interested in the development of new tone contrasts.

2.1.4 Tone and Register

Snider & van der Hulst (1993) contains a survey of the development of tone theory, including the relative nature of tone and the need for multiple features, the latter needed to account for more than two tone heights as well as for register effects. The relative nature of tone and the presence of multiple tone heights can be best understood by seeing pitch as resulting from the interaction of independent features of tone and register. One specific theoretical analysis of tone and register interaction is given in Snider (1999) under the name Register Tier Theory (RTT).

Descriptively, register is the domain of relative pitches within which tones are pronounced, i.e., the range of lowest to highest possible pitches at a given moment. Because each tone is pronounced relative to that domain, when the register raises (or lowers), the pitch of that tone is higher (or lower) as a result. It is perhaps helpful to think of music, with notes in a melody, which is played in a certain key. The same melody may be played in one key, then another, without changing the melody – though it would be played on different notes. In the same way, a given tone melody may be pronounced in another register with essentially the same meaning, but on different pitches. The difference in key/register is a different kind of difference than a difference in tone/melody, and they may mean different things in the system.

We can then derive the phonetic implementation of pitch from tone and register by seeing pitch as the contribution of register plus/minus that of tone, as described in more detail in Snider (1999), and as illustrated in Figure 3:

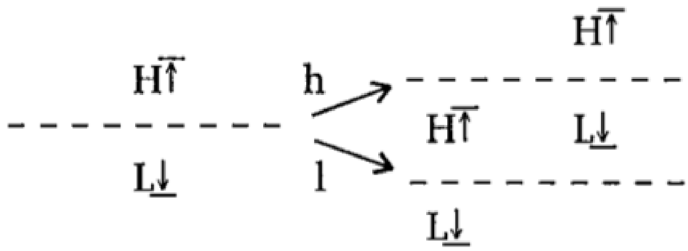


Figure 3. Phonetic implementation of tone and register, from Snider (1999:25)

On the left of Figure 3, H (high tone) creates a pitch higher than the center of the register (represented by the dotted line), while L (low tone) creates a pitch lower than the center of the register. To the right of those tones, a higher register feature (lower case "h") may produce a higher register, in relation to which H and L are now pronounced (far right). Alternatively, a lower register feature (lower case "l") may produce a lower register, in relation to which H and L are pronounced (near right). So given the implementation of pitch and register as in Figure 3, H on a lower register (near right) as slightly higher than L on a higher register (far right). In this way, the simple diagram presented in Figure 3 shows two tone values on three registers producing six different pitch levels (three different H pitch levels and three different L pitch levels) –and this is done with only two distinct tone units: H and L.

The relationship between tone and register may be easily confused by poorly distinguishing between multiple kinds of pitch lowering, along with multiple terms used for them (e.g., "downdrift", "systematic downdrift", "intonational downdrift", "downstep", "phonemic downstep", "downtrend", "downglide", "key lowering", and "declination"). But this multiplicity of terms covers only three distinguishable phenomena, which I will follow Stewart (1964, 1968, 1983) in calling Automatic Downstep, Non-automatic Downstep, and Declination.

Automatic Downstep is a register lowering phenomena which is also referred to as "systematic", "intonational", and "non-distinctive"; it is a language-wide lowering of register with each low tone. That is, a HLH sequence is pronounced [¯ _ –] rather than [¯ _ ¯]. But because it is the register that is lowered, it is not just the pitch of the next tone which is lowered; all tones after this second high tone are pronounced with lower pitches as well, such that a HLHH sequence would be pronounced [¯ _ — —], rather than [¯ _ — ¯]. Furthermore, because of the relative nature of the change, it applies successively, such that a HLHLHLH sequence would be pronounced [¯ – ¯ – – _ –]. The interaction of tone and register in automatic downstep is exemplified with tone and register features in Figure 4:

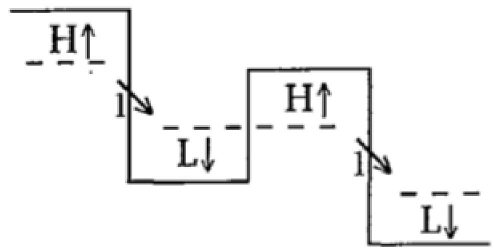


Figure 4. Tone and register in automatic downstep, from Snider (1999:28)

In a language with Automatic Downstep, as in Figure 4, each low tone bears a lower register feature, dropping the register with each low tone. As a result, the two high tones in Figure 4 are pronounced on different registers, as are the two low tones. Automatic Downstep is thus a pitch lowering that is predictable on the basis of the surface representation of an underlying phonological category (i.e., Low tone). This phonological lowering of register is therefore redundant with a low tone which is realized in the surface form, so the register lowering is not the only surface indicator of that underlying phonological category. This redundancy is why this type of downstep is called automatic —it correlates absolutely with surface low tones.

Non-automatic Downstep is a similar lowering of the register, but without a surface low tone. That is, what appears to be a sequence of high tones is pronounced [¯ –], with the second lowered. Non-automatic Downstep then contrasts with no lowering (i.e., [¯ ¯]), but also with a HL sequence, which would have a greater drop in pitch (e.g., [¯ _]). The operation of tone and register in Non-automatic Downstep is represented in features in Figure 5:

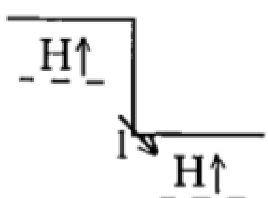


Figure 5. Tone and register in Non-automatic Downstep, from Snider (1999:81)

But following tones are again lowered, such that a HHLHLH sequence would be pronounced [¯ ¯ – – _ –], assuming both Automatic and Non-automatic Downstep were operating. Because there is no surface low tone triggering Non-automatic Downstep, this register drop is also described as "unpredictable" or "phonemic".

Non-automatic Downstep is thus a register lowering that is not predictable on the basis of the surface representation of any underlying phonological category. Not only does Non-automatic Downstep make reference to an underlying phonological category (i.e., low tone), the register drop is the only surface indication of it.

Automatic and Non-automatic Downstep, then, are essentially the same phenomenon (a lowering of register through a low tone), despite the fact that they have been taken for two distinct phenomena (Stewart 1964, 1968, 1983). The difference, Stewart (1964, 1968, 1983) argues, is whether the low tone itself is visible in the surface pronunciation ("Automatic Downstep"), or if it is only visible through its effect on the register ("Non-automatic Downstep").

The third phenomena to introduce here is declination, a pitch lowering phenomenon which could be described as physiological rather than phonological. That is, as a speaker progresses through a breath group, there is less air pressure behind what is spoken. Declination leads to gradually lower pitch as a matter of course, as opposed to the result of a speaker's making linguistic distinctions. Declination can be observed in a pitch trace as a more gradual lowering (generally 1-5hz/syllable), but also as a more gradient lowering (the slope of pitch descent remains relatively constant within and between syllables, with the possible exception of the end of a breath group). Declination has alternatively been called "downdrift", "downtrend", "declination" and "automatic downstep", among others, though it is descriptively and functionally distinct from the Automatic Downstep of Stewart (1964, 1968, 1983). In any case, the implementation makes no reference to phonological categories (e.g., high and low tones, or higher or lower register); it does not interact with the phonological system at all. As a result, declination will be largely ignored in this dissertation.

Given the above account, it should be clear that the implementation of tone and register has been a major terminological issue, as well as a theoretical and descriptive one. The multitude and confusion of terms must necessarily disturb attempts at language description, along with any theory development that is based on those descriptions. While Stewart (1964, 1968, 1983) made a clear step forward in showing the relationship between automatic and Non-automatic Downstep, the broader terminological question is more directly addressed by Connell (2001).

Connell (2001) proposes specifically that we use Stewart (1983)'s "downstep" for the successive lowering of register due to low tones, with the subcategories of "automatic downstep" (when those low tones are realized on the surface, making the downstep predictable) and "Non-automatic Downstep" (when those low tones are floating, making the downstep not predictable). "Declination", on the other hand, should be reserved for the universal, gradual phonetic lowering of pitch across an utterance. Declination occurs across all languages, is not categorical in nature, and is not tied to the presence of low tones (Connell 2001).

Because the term "downdrift" has been used synonymously with either downstep or declination, without referring to a clearly distinct phenomenon (Connell 2001) I will avoid using "downdrift", while using "downstep" and "declination", following Connell (2001).

The question of whether tone has features at all has been raised (Clements, Michaud & Patin 2011, Hyman 2011, Odden 2011), though that position is still a matter of debate (McPherson 2015). My orientation presumes at least some interaction between tone and register, as described in this section. Specifically, I assume a relationship between tone and register features where each is associated to a tonal root node, which is associated to a mora, along the lines of that presented under the name Register Tier Theory in Snider (1999), and schematized in Figure 6:

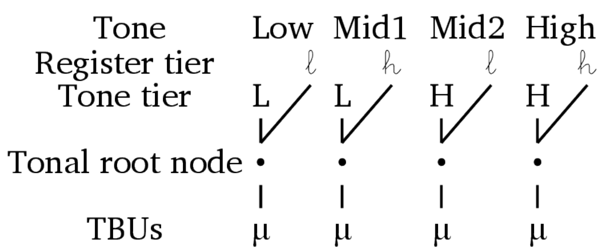


Figure 6. Review of tonal feature geometry in Register Tier Theory (Snider 1999)

Register and tone combine to provide a variety of possible pitch changes, even occasionally cancelling each other out. That is, if the phonetic contribution of tone and register in a given language are of similar value (i.e., same change in pitch for each), then a low tone on a higher register will sound the same as a high tone on a lower register. As a result, it may be difficult to know when one is dealing with a tone or register change, as in "total downstep" languages (Meeussen 1970 and Stewart 1993).

To see the interaction of tone and register working in a specific example, consider the following Mbo [zmw] data:

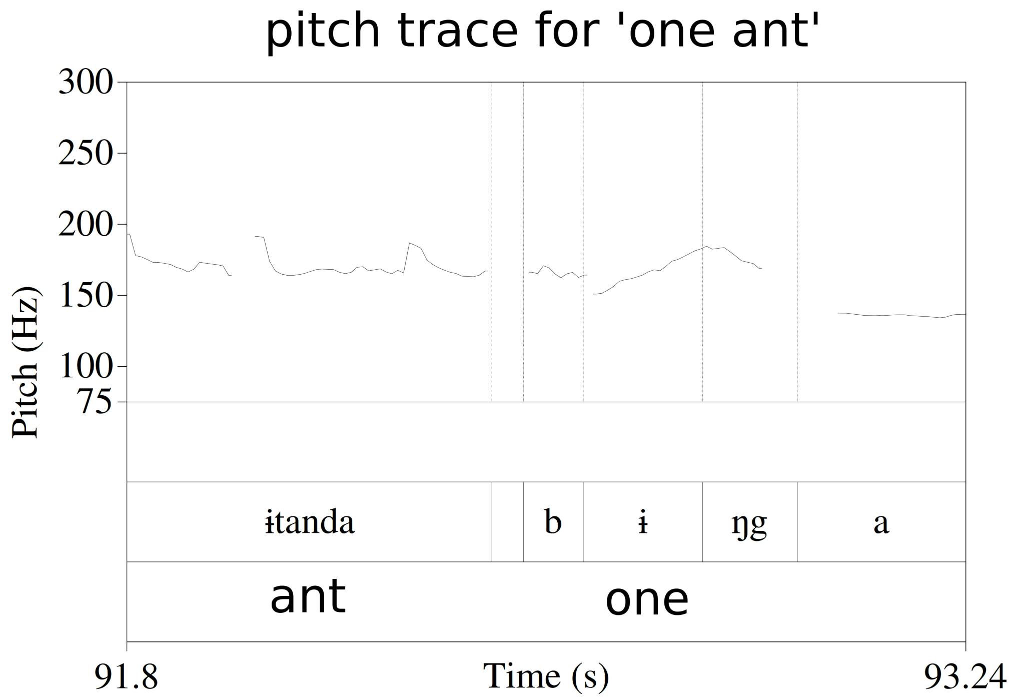


Figure 7. Pitch trace of one ant in Mbo [zmw], showing phonetic implementation of tone and register

The F_0 targets for each syllable in Figure 7 are given in Table 1, along with differences in F_0 between syllables which are attributable to declination, tone, and downstep.

TBUs	ɨ	(t)a	(nd)a	#	(b)ɨ	(ŋg)a
Tone	L	L	L		L-H	L
Pitch targets	169hz	167hz	163hz		151-182hz	137hz
declination (L-L)		-2hz	-4hz			
Tone (L-H/H-L)					+31hz (L-H)	-45hz (H-L)
downstep (L#L, LHL)					-12hz (L#L)	-14hz (LHL)

Table 1. Pitch targets from Figure 7, with declination, tone, and downstep

Table 1 shows three kinds of changes in F_0. Declination can be observed within the first word, where *a drop of less than 5hz occurs between each of the three low toned syllables*. I call this <5hz/TBU change declination because it is a minimal change of essentially one slope across adjacent TBU's of the same tone value. Because the change between the first three syllables in Table 1 is relatively constant, of minimal value, and not apparently triggered by phonological categories, it is reasonable to call it declination, a lowering in pitch due to physiology, rather than phonology.

The second and much more obvious change in F_0 comes in the second word, where there is *a transition from low to high tone* within the first syllable, followed by a transition back to low tone for the second syllable. These low-high and high-low transitions are each more than 30hz, significantly greater than the $<$5hz change seen for declination in the first word.

The third factor impacting F_0 in Figure 7, downstep, occurs at the beginning of each syllable of the second word. The first occurrence happens because there is a low tone on each side of the word break (i.e., L#L), such that the last syllable of the first word, and the first syllable of the second word, are not associated to the same low tone. At the point of the adjacency of the two low tones there is a 12hz drop, over twice that observed for declination.

A similar drop (i.e., downstep) is observed on the second syllable of that word, where another low tone appears 14hz below the prior low tone, though this time with an intervening high tone. This occurrence is an elegant example, however, in that one can see the H-L (45hz) transition is approximately the value of the LH (31hz) transition, plus one downstep (~13hz), due to a new low tone in HL. This shows that one can have a ~30hz transition due to tone (either LH or HL), plus a ~13 hz transition due to downstep (for each new L), or both[5], as exemplified in Figure 7 by LH, L#L, and HL, respectively.

So there is *a drop of about 13hz at each introduction of a new low tone*, whether across wordbreaks or after a High tone. I again follow Connell (2001) in calling this lowering “downstep”.

To summarize the above, Figure 7 shows a $<$5hz non-phonemic F_0 drop across syllables associated to the same tone, a $>$30hz change between High and Low tones, and a ~10hz drop in register with each new low tone. The last (register) drop I follow Connell (2001) and Stewart (1964, 1968, 1983) in calling **downstep**.

In summary, a phonetic change in pitch can result from a change in register, tone, or both. As a result, a researcher must pay attention to changes in pitch which result from a change in register, as well as to those that arise from changes in tone, for both the collection and analysis of tone data.

[5] While tone and register independently impact the phonetic implementation of pitch, in a language with automatic downstep, a HL transition will always include a drop in pitch due to downstep, since downstep is triggered by a low tone.

2.1.5 Diacritics and the Presentation of Phonetic Tone Data

Given the multiple factors impacting pitch discussed in Section 2.1.4, one might ask if it matters how we transcribe pitch, in either analysis, or in the presentation of phonetic data and justification of that analysis to others. In this section, I hope to argue not only that it matters very much, but also that the *de facto* method of tone transcription with diacritics causes a number of problems for tone analysis.

Based on the material presented in Sections 2.1.1–2.1.4, an analysis able to handle the particular nature of tone must maintain a number of distinctions: between the kinds of information collected (Sections 2.1.1 and 2.1.2), between the contexts in which that information is collected (Section 2.1.3), and between the analytical categories used to understand that information (Section 2.1.4). I would further assert that our means of representing data must be appropriate for the purpose of that representation.

There are many different ways to represent tone data, each of which may be more appropriate at one stage of the analytical process than at another. As with any phonological analysis, initial data collection should represent data in their most raw, unadulterated, and unanalyzed form. At some point we do the most basic and simple analysis of those data, and ultimately present and justify it to others. That presentation and justification should establish the analysis up to a certain point, after which that analysis may inform the collection of data without circularity.

This cycle of data collection and analysis continues as the analysis becomes less basic and more interesting, and ultimately more general theoretical claims are made based on that data and analysis. But while data collection is almost always eventually informed by analysis, it should not be initially. Before an analysis can be used to inform data collection, it should be defended on the basis of data that is collected and presented without presuming it, to avoid circularity. Or, in terms of phonological categories, we should not collect phonemic transcriptions until we have justified a phonemic analysis on the basis of phonetic data. Once we know what the underlying categories are (on the basis of a phonological study grounded in phonetics), then it is more appropriate to collect data on the basis of those categories, if one should desire to do so. But if our transcription methodology is better suited to phonological categories than phonetics, then it doesn't serve us well for the initial analysis.

Assuming it is our intention to justify our analyses with phonetic data (as other phonologists do), then our method of transcription must be appropriate to representing phonetic data. Furthermore, a phonetic transcription is useless to the reader if it does not allow at least some reasonable approximation of a faithful reconstruction of the pitch of the utterance (recall from Section 2.1.2 that "high" and "low" do not constrain possible productions in the same way that [p] and [b] do). In this section I hope to show that the *de facto* method of tone data transcription does not allow the reader to faithfully reproduce pitch contours. Furthermore, transcriptions using this method both imply and require some kind of analysis to decode them. For these two reasons, this method is unsuitable for transcription of phonetic pitch data as the basis for a tone analysis[6].

Put another way, our phonological analyses must be based on verifiable phonetic data. It is an empirical problem, then, when a researcher assumes an analysis of tone (whether knowingly or not) in the collection and representation of the initial tone data themselves. At best, the researcher's presentation of data is biased by the assumed analysis; at worst, the analysis itself is unfounded. And far too often, we readers are simply asked to trust that the researcher has faithfully analyzed what is presented.

It is therefore particularly important, for an analysis to be empirically valid, to have a clear distinction between what forms are presented as primary data, and what forms are presented as analysis of those data. Furthermore, the forms presented as data should be objectively interpretable by others, as the basis for and defense of the analysis presented. But the presentation of objectively interpretable tone data is made difficult by the *de facto* standard for presenting them, which uses diacritics to represent tone and/or pitch, and which must necessarily assume some level of analysis.

This *de facto* standard[7] for representing tone data uses an acute accent (e.g., *á*) for high tone, and a grave accent (*à*, or else unmarked) for low tone; if both high and low are marked, unmarked vowels may be used for a mid tone. This system is ostensibly used to represent either pitch or tone, as the need arises. In the case where only two or three[8] pitch levels need to be represented, this system may seem to work to represent pitch. But as shown in Section 2.1.4, there can easily be more

[6] This is especially the case when that phonetic pitch data are presented as grounds for the very phonological analysis required to interpret those phonetic data. One can hope that this kind of circularity doesn't happen much, but I will show in this section that reputed scholars have come much closer to this than we ever should.

[7] That this is the broadly accepted manner of transcribing tone data can be seen almost anywhere tone descriptions are published. The force of this broad acceptance is especially notable in that despite the difficulties outlined here, I have been criticized when attempting to publish a description of a tone language, on the basis that "pitch overlays are not a standard phonological representation for tone".

[8] As typically used, this system allows for three distinct levels. But with the addition of *a̋* for extra high and *ȁ* for extra low, the system can describe five levels. But as I hope to show in this section, the problems with this system go beyond the simple level of how many levels the system can represent.

than three pitch levels in an utterance, so transcription with diacritics is almost certain to lose phonetic information.

In addition to the loss of information, assigning a pitch melody to a series of diacritics is a non-trivial matter. For instance, the tone melody [– ¯] may be transcribed with diacritics as “low mid” *à-a*, “mid high” *a-á*, or “low high” *à-á*. But the choice between these three representations is one of phonological analysis, not phonetic transcription, since they categorize the utterance with reference to the system as a whole.[9] The transcription [– ¯], on the other hand, makes no claims about how the two pitch levels relate to other pitch levels[10]; such a melody after [¯ –] combination may be pronounced [¯ _ _ ¯], [¯ _ _ –], [¯ _ – –], [¯ _ – ¯], [¯ – _ –], or otherwise. While the ambiguity inherent in these symbolic pitch traces might seem undesirable, it matches well the relativity of pitch data as described in Section 2.1.2. That is, a bisyllabic utterance cannot be interpreted absolutely outside of some broader context, so a transcription method that corresponds with that fact is superior to one that forces the researcher (and reader) to presume a particular relationship to the pronunciation of other utterances[11].

With more than two syllables, the situation becomes more complex. For example, there are four pitch levels in the four syllables represented in Figure 4 from page 19, repeated for the reader's convenience in Figure 8:

[9] While it may seem desirable to categorize a pitch transcription in relation to the entire register of possible pitches, any attempt to do so must either result in data dependent on F_0 (which tells us nothing about how it relates to the system at work in the language), or else subjective data based on the analysis to date of the language (since we don't immediately know the full range of possible pitches in a given language). As a result, the only reliable presentation of pitch data is one that compares pitch levels with other pitches present in the same utterance.

[10] Such a claim may be implied with the height of the phonetic brackets as a reference (if it is presented in phonetic brackets), but such an implication may or may not be helpful, as a reference point outside of the utterance must depend on the researcher's subjective evaluation.

[11] That a surface tone pattern cannot be rightly interpreted in isolation is particularly true given the fact that two different underlying tone melodies may have the same surface form in a particular context, with the difference only showing up in some other context. The fact that the surface pitch of one utterance in a particular context has a certain relationship with to the surface pitch of another utterance in that context doesn't absolutely determine the relationship between their underlying tone melodies.

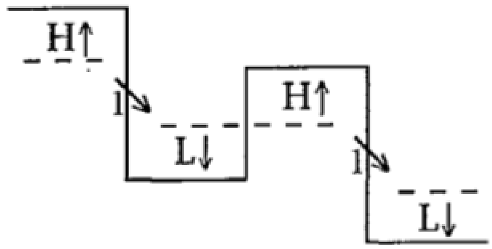


Figure 8. Multiple pitch levels, few features (=Figure 4; Snider 1999:28)

Changes in register are often handled by other symbols, superscript ! or ↓ for downstep (e.g., ꜝá or ꜜá), and ↑ for upstep (e.g., ꜛá). But these register change symbols make an even greater analytical claim than acute and grave accents do. That is, calling a lower pitch a downstep makes claims about what one expects to find after that point (i.e., a lower pitch register), but it also makes specific analytical claims about the underlying form of that data, e.g., the presence of a floating low tone (c.f., Stewart 1968 and 1983).

So if one of the goals of phonologists working in tone languages is to defend a tone analysis on the basis of objectively interpretable data (as I hope it is), I assert that diacritic marking is problematic to say the least. So far this assertion is explained and justified on the basis of hypothetical examples; the remainder of this section discusses real examples from two chapters in The Bantu Languages (Nurse & Philippson 2006), a leading authority on the (mostly tonal) Bantu languages.

The first example comes from (Hyman 2003b:262), where tone/pitch data are presented on two different tiers:

(1) Basaá [bas] (Hyman 2003b:262)

```
a   bí     ꜝtí   ɓɔ́     ꜝkɔ́nc
|   |      |     |      |
L   H L    H     H L    H
```

he gave them beans

I assume it is the author's intention to communicate phonetic data through the diacritics placed on the segmental tier, and phonological data/analysis through the series of H's and L's on the tonal tier[12]. The problem is that each seem to

[12] If this assumption were not correct, not only would the two tiers would be redundant, but there would also be no phonetic data at all, and thus no ground for the phonological analysis. My assumption therefore grants the greater benefit of the doubt, in terms of the value of the information present to justify phonological claims made. I believe this greater benefit of the doubt is warranted, given the author's established credibility.

be essentially phonological forms. That is, there is a one to one correspondence between the symbols on the two tiers: between H on the tonal tier and acute marked vowels on the segmental tier, between floating (unassociated) L on the tonal tier and downsteps (superscript ! in *ꜝtí* and *ꜝkónc*) on the segmental tier, and between associated L on the tonal tier and unmarked vowels on the segmental tier. Because each element representing tone or pitch corresponds exactly between the two tiers, there is no information contained on one tier which is not also on the other. The two tiers then either represent a system where phonetic implementation is completely transparent given phonological forms, or else they both represent the same underlying phonological categories. In other words, the relationship between the two representations is either uninteresting or tautological.

But continuing for a moment with the assumption that the segmental tier is intended to be a phonetic representation, it lacks important phonetic detail, on the one hand, and forces the reader into a particular analysis, on the other. Because downsteps are used to describe the phonetic data in (1), the reader is left to trust the author that the downstep analysis is correct, and to reconstruct the phonetics based on that analysis, in order to evaluate the analysis based on the data used to support it. But unfortunately Hyman (2003b) contains no prose description detailing the phonetic implementation of downstep in Basaá [bas]. As a result, one could reconstruct the phonetics of (1) in at least the following three ways:

(2) Phonetic implementation of LHꜝHHꜝH from (1)
  a. [_ ¯ ‾ ‾ —] - assuming the L→H transition is more than 2 downsteps
  b. [_ ¯ – – _] - assuming the L→H transition is about 2 downsteps
  c. [— ¯ – – _] - assuming the L→H transition is about 1 downstep

The ambiguity presented in (2) exists because the diacritics given in (1) do not tell us if the final (doubly downstepped) acute marked (i.e., high) vowel is pronounced at, above, or below the pitch of the initial unmarked (i.e., low) vowel. This ambiguity regarding the relative pitches in (1) has a number of consequences regarding the reader's interpretation of the data.

If the lack of a clearly reproducible pronunciation were not enough, in whatever manner the reader reconstructs the pronunciation, the transcription in (1) is not the only option for transcribing that pronunciation in diacritics. If the reader reconstructs the phonetics of (1) as in (2a), then the data might just as well have been transcribed *à bí ꜝtí ɓɔ́ konc* (with a downstep then a mid tone) or *à bí ti ɓɔ ꜝkonc* (with three mid tones, the last of which is downstepped).

But if the reader reconstructs the phonetics of (1) as in (2b), it would have been equally well transcribed with diacritics as *à bí ti ɓɔ kònc* (with mid tones) or as *à* *bí tꜜí ɓꜜɔ́ kònc* (with downstepped high tones), since the first and last syllables are pronounced at the same pitch, and there is only one pitch level between that and the high pitch level, which may be interpreted as either mid or downstepped high tone.

And if the reader reconstructs the phonetics of (1) as in (2c), the data could have been equally well transcribed with diacritics as *a bí ti ɓɔ kònc* (with mid tones) or *à bí tì ɓɔ̀ ꜜkònc* (with low tones, one of which is downstepped), since the initial unmarked vowel and the first downstepped acute marked vowels would be pronounced at the same level, with only one higher and one lower pitch in the utterance.

So however the reader reconstructs[13] the pitch of the data in (1), the transcription of that pitch melody with diacritics is non-trivial; in no case is a single transcription with diacritics obviously correct, until it is informed by some level of analysis. Yet only one alternative is presented to the reader, who is obligated to accept or reject it without argument. Likely one might say that the transcription with downsteps is superior, because it captures a generalization that would be lost using the alternatives. But the presence of this kind of generalization is precisely the problem. Data should not be presented in a way that forces the reader to a particular generalization of them, because such a generalization is analysis. One cannot evaluate an analysis based on presented data, when those data are presented in a manner which presupposes the analysis to be evaluated.

Considering tone data in this way, it becomes clear that the phonetic values of data presented in diacritics are almost never unambiguous. Yet the data presented in (1) are not outliers; rather, they exemplify the way tone data are almost always presented (whether for traditional reasons or because of the historical difficulty of transcribing pitch in papers[14]). Another variation on the diacritic data presentation scheme is shown in (3):

[13] I write of speaker reconstruction because we do not have access to the original pronunciation, outside of the author's transcription. Were it the case that any of the plausible reconstructions allowed only one possible transcription, one might argue that the author used that transcription because it was the only one available, therefore that reconstruction must match the original pronunciation, but this is not the case. Rather, we are left with multiple possible reconstructions, each of which, if it were the original pronunciation, could have been transcribed in multiple alternate ways with diacritics. As a result, at no point can we can conclude that the author chose a particular transcription because it was the only one available.

[14] However true it may be that tradition pushes us to continue the use of diacritics to transcribe pitch, if that tradition does not serve our field, it should be changed. And regarding historical difficulties, with the advent of unicode and freely available quality fonts (as used in this dissertation), the historical difficulties of representing pitch without diacritics is obsolete.

(3) Nen (A44; Mous 2003:287; cited as DL 166)
bá-ná nùíyì sàk-ak-´ → báná ꜜ núíyí sàkák
They made a bridge of lianas across the river

Mous (2003) is fairly clearly providing a derivation, likely between underlying and surface forms. Yet both forms are provided in the same diacritical notation. The presentation of phonetic (pitch) and phonological (tone) data in the same notation system creates two fundamental problems for the data in (3). First, the potential ambiguity between a downstepped high and mid tone is avoided, instructing the reader to interpret the pitch on *nuiyi* as coming from a downstepped high tone (and thereby removing a set of possible analyses). Secondly, it differentiates between two kinds of downstep, which are likely phonetically very similar, if not indistinguishable. It is almost certainly the author's intent to communicate that the initial *bá* and the final floating ´ on the left side have the same tone value (i.e., they are both underlying high tones), but it is almost certainly not the case that the initial *bá* and the final *kák* on the right side have the same phonetic implementation/pitch, even though they are marked with the same diacritic. These syllables should not be pronounced at the same pitch because (in addition to a universal declination, or lowering of pitch across an utterance), as the author describes in prose, "Highs after Lows are realized lower (automatic downstep)" Mous (2003:286). That is, the right side of the data in (3) shows Non-automatic Downstep (via ꜜ), but an accurate phonetic account (based on the prose) should also include automatic downstep, each of which would likely have an identical phonetic effect on following pitches[15]. So the phonetic implementation should be something like (4b), rather than as in (4c):

[15] One may say that distinguishing automatic and Non-automatic Downstep here is justified precisely because the one is predictable/automatic, and the other is not. Or, given a certain knowledge of the system, the phonetics can be clearly reconstructed from the information given, so there is no ambiguity in the phonetic data. But dependence on a particular understanding of the system is precisely what is inappropriate in the presentation of data. In the presentation of phonetic data, portions of the data that are the same phonetically ought to look the same. To leave something out on the basis that it is predictable is to deprive our readers of the data that show that predictability. More critically for the peer review process, leaving out such data deprive our readers the chance to objectively evaluate the claims made, including that of predictability.

(4) a. bánáꜜnúíyí sàkák ‘They made a bridge of lianas across the river’

b. [¯ ¯ ¯ ¯ _ —] (implementation based on prose)

c. *[¯ ¯ ¯ ¯ _ —] (wrong implementation, based on diacritics)

So in (3) a given diacritic does not correspond to a given phonetic value, but rather to a relative value within a phonological context[16]. Thus the use of diacritics in (3), as is generally the case, presupposes an analysis of the tone, rather than giving objectively interpretable phonetic data. On the basis of the above, which could be placed alongside many similar examples, I assert that the diacritic transcription methodology is not appropriate for the presentation of phonetic data on which a tone analysis is to be defended.

If one were inclined to be cantankerous, one might go so far as to say that the data in (3) actually argues against the prose description quoted above, since they contain a high after low which is **not** marked as downstepped in the output. While (4c) is almost certainly not what the author intended, it remains the case that the prose is critical to correctly reconstructing the pronunciation of the data in (3), so the data in (3) are not reliably interpretable as phonetic data on their own.

To summarize the difficulties posed by transcription of pitch with diacritics, there are a number of principles at stake. First, we must base our phonological analyses on clearly presented phonetic data. Starting with phonological categories should not be acceptable in any phonological analysis, unless it is based on another analysis which has established the validity of those categories on the basis of phonetics. Neither should we present phonetic data in a manner that makes it difficult for the reader to reconstruct the pronunciation of that data. Presuming familiarity with the language is not sufficient, if our work is to be understood by those not familiar with the language of study. And finally, we should acknowledge that the above cannot be done with a system that represents only three pitch levels, except perhaps in very extreme cases. In almost any tone language utterance with more than three syllables, there is a good chance of at least three pitch levels. As a result, I believe that a commitment to good tone scholarship requires abandoning

[16] One could argue that the data in (4a) could be resolved by modifying the phonetic transcription to read *bánáꜜnúíyí sàkꜜák*, such that both phonetic downsteps would be represented, but I imagine that for some such a resolution (with a second explicit downstep marked) would now be ambiguous with a possible doubly downstepped interpretation, where the final syllable should be both automatically and non-automatically downstepped. In any case, ꜜ*a* makes a different analytical claim than that of either *á* or *a*, and the diacritic representational scheme forces the author to make such an analytical distinction in the presentation of phonetic data, where it does not belong.

the representation of tone data with diacritics, except as an orthographic representation of underlying categories, and only where those categories have previously been worked out and justified on the basis of phonetic data.

Given my desire to present phonetic data in a way that is reliably interpretable, and in order to justify phonological claims with data that is more clearly objective and phonetic in nature, I present data with symbolic pitch traces, as in (4b) and (4c) above. I further illustrate and justify those symbolic pitch traces with F_0 traces made with Praat (Boersma & Weenink 2014), as will be seen in Chapter 3, and discussed in more detail in appendix B. This method of presenting data provides more phonetic detail, allowing the reader to see the correspondence between my symbolic pitch traces and more narrow phonetic data, and to better evaluate the analysis based on those data. It is important that theoretical claims are built on a solid empirical foundation, as empirical claims justify more theoretical claims, such as those of Bradshaw (1999) in Section 2.2.

2.2 Consonant-Tone Interaction and Bradshaw (1999)'s [L/Voice]

The theoretical and analytical contributions of this dissertation depend on an understanding of consonant-tone interaction, as in Bradshaw (1999), and exemplified in (5):

(5) Tone and voicing in Suma [sqm] (Bradshaw 1995:263)

a.	ɓórɑ́	shrunken	bɔ̀rɑ́	long
b.	fókɑ́	flowing	zìkɑ́	turning
c.	kírí	look for	gɑ̀lɑ́	soaking
d.	fɔ́ɗí	stir briskly	vɑ̀tɑ́	stirring up

While there is certainly more that could be said about the surface forms of the data in (5), the forms on the left have initial consonants which are not specified for voice, and they begin with high tones[17]. On the other hand, the forms on the right of (5) begin with consonants that are specified for voice, and they begin with low tones.

The correlation between voicing and lower F_0 has been documented at length (Meyer 1937; House & Fairbanks 1953; Lehiste & Peterson 1961; Mohr 1971; Lea 1973; Hombert 1978; Hombert, Ohala & Ewan 1979; Mohr 1968; Wright & Shryock 1993). This correlation is discussed in phonological terms as early as Beach (1923), but see also Welmers (1962); Hobley (1964); Greenberg (1970); Hyman (1973); Hyman & Schuh (1974); Hombert (1975). This correlation is discussed in terms of "depressor consonants" or "depressors" as early as Cope (1970); Louw (1971), and is also called consonant-tone interaction (CTI). The volume *Tone: A Linguistic Survey* (Fromkin 1978) includes three chapters referring to ongoing research into the interaction between segments and tone, and which contains several observations which remain useful today. One Chapter (Schuh 1978) claims that consonants (as opposed to vowels) interact with tone. Schuh (1978) further claims that consonants affect tone, and not the other way around (though the direction of the effect is disputed and defended in works contemporary to Schuh 1978, e.g., Maddieson 1974, Hyman & Schuh 1974, and Hyman 1976; this point may remain open to debate today).

Hombert (1978), a second chapter in Fromkin (1978), addresses consonant-tone interaction from a different angle, reviewing tonogenesis as contrastive tone

[17] I presume that these representations are phonemic, rather than surface pitches (c.f., Section 2.1.5). In any case, what remains clear from this data is that voicing contrast correlates absolutely with grave marked vowels (presumably bearing low tone), and the lack of voicing specification (with or without surface voicing) correlates with acute marked vowels (presumably bearing high tone).

which develops on vowels due to a loss of a voicing distinction on prevocalic consonants. In this scenario, previously voiced consonants precede new lower pitches, and previously voiceless consonants precede new higher pitches. The number of contrastive tone levels increases twofold through this process, which is well attested, though only in Asia (see abundant references covering multiple Asian languages, in Hombert 1978:78-79). This dissertation provides evidence of similar operation in Africa.

The third chapter in Fromkin (1978) that deals with tones and consonants includes a proposal to "construct a unified account of voicing, aspiration, glottalization, and other laryngeal phenomena, as well as tone, in terms of a simple system of four features" (Anderson 1978:161), but is forced to conclude "Unfortunately, this review of proposals that have been made for the features of tone does not allow us to come to a definite resolution of the question" (Anderson 1978:172).

While the three Fromkin (1978) chapters described above address consonant-tone interaction, they do not provide a theoretically unified means to interpret the phenomenon. Many researchers continue to work under such a deficit. For instance, one book chapter on harmony systems includes a treatment of laryngeal harmony, including voicing harmony (Rose & Walker 2011:244). Yet two of those cases are acknowledged to allow for alternate interpretations, such that the harmony would be conditioned not by consonant types, but by tone.

But progress in understanding consonant-tone interaction has been made; a good description of the phonetic basis for the interaction of voicing and pitch is given in Bradshaw (1997), including the summary of previous research that "voiced and voiceless obstruents are both characterized by consistent pitch effects, realized on the following vowel, which are cues to the voicing of the obstruents" (Bradshaw 1997:20). While the pitch effects described in Bradshaw (1997) can be taken as an indication of voicing, they might also (or alternatively) be interpreted as a difference in tone, however much it would correspond to voicing. The potential ambiguity between voicing and pitch as causing the other allows for the reanalysis of the one system as the other system. This could be referred to as the phonologization of the phonetic effect of voicing on pitch. The new phonological element of the tone system would continue to correspond to voicing, at least until something changed that correspondence. I will take up an example of this in Chapter 7.

The perspective just outlined, of consonant-tone interaction as reanalysis of physiology, is based on a feature geometry proposed in Bradshaw (1999), as will be discussed in Section 2.2.2. Downing & Gick (2001) and Downing (2009) criticise this proposal as inadequate to account for each of the range of consonants interacting with tone and the range of consonant-tone interactions observed. Rather, they promote Optimal Domains Theory (Cassimjee 1998 and Cassimjee & Kisseberth 1998) as a replacement for autosegmental representations (Leben 1973 and Goldsmith 1979). Unfortunately, the differences between Bradshaw (1999) and Downing (2009) include at least a number of fundamental theoretical assumptions, such as the value of autosegmental representations. Because of this, and because Downing (2009) appears to differ with much of the material I presented in Section 2.1[18], I will set aside this criticism for the remainder of this work, perhaps to be taken up again in later work.

The following subsections address various aspects of consonant-tone interaction (CTI). They begin with the phonetic impact of voicing on pitch, and the physiological correlation between the two (Section 2.2.1) before introducing a phonological account through the feature [L/Voice], from Bradshaw (1999) (Section 2.2.2). The operation of [L/Voice] in phonological space is covered in Section 2.2.3, followed by a discussion of consonant types, and how Bradshaw (1999)'s [L/Voice] would expect each to behave in a CTI system (Section 2.2.4). For the interested reader, I have included an appendix (E) which discusses possibilities for integrating Bradshaw (1999)'s [L/Voice] with (Snider 1999)'s Register Tier Theory (RTT).

2.2.1 Correlating pitch and voicing

This section discusses the connection between phonetics and phonology for consonant tone interaction (CTI). There is an observed correlation between phonetic voicing and lower pitch, and in many languages there is also a correlation between voicing and a phonological low tone. Given that the second may well arise from a reanalysis of the first, we should understand these correlations in terms of their phonetic and phonological implications, before trying to establish a unified account of them (in Section 2.2.2).

The phonetic correlation between phonetic voicing and lowered pitch has been described elsewhere, but it can also be seen in Bantu D30 languages, as in the Nyali-Kilo [nlj] data in Figure 9:

[18] There are two points of difference which are particularly important. First, Downing (2009) appears to attempt to ground consonant-tone interaction in the phonetics of pitch lowering, but all data in Downing (2009) are presented in a phonemic notation system, which I consider problematic as discussed in Section 2.1.5. It is therefore non-trivial to reconstruct her understanding of how pitch and tone interact in the data presented; it is in fact difficult to reconstruct the pitch on any of that data. The second point is that her analytical solution depends on a distinction between tone and register, though it is unclear what is meant by "register" in this context; it is described as "local", and as such does not appear to align with what was presented in Section 2.1.4.

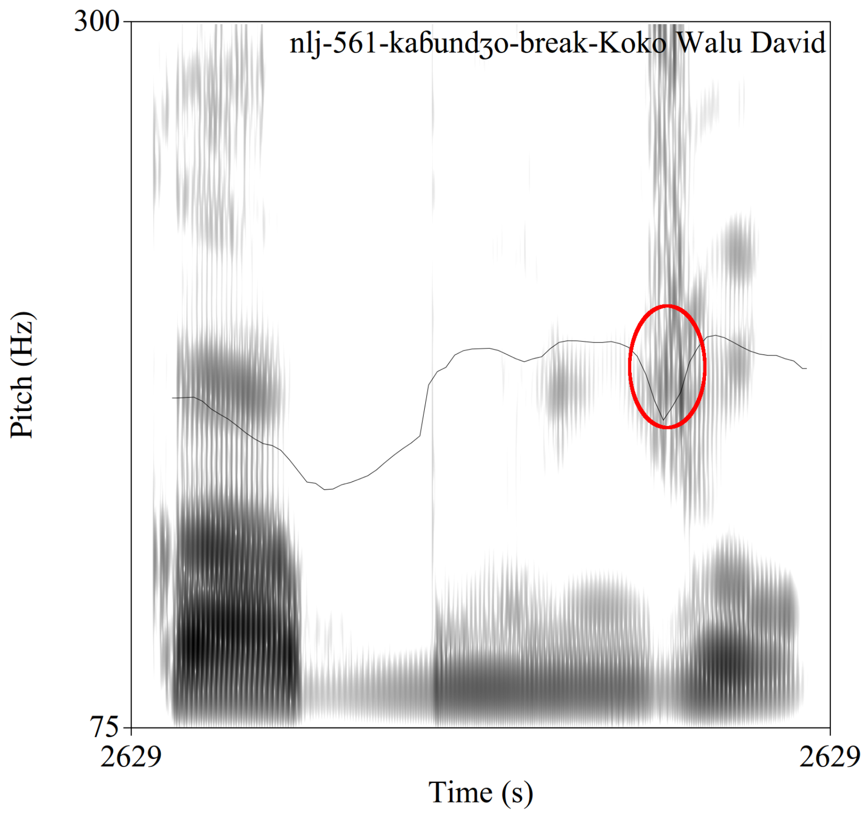


Figure 9. Localized depression on *dʒ* in Nyali-Kilo [nlj] [kaɓundʒo] [– ¯ ¯] 'break'

There is about a 15hz drop in pitch over the consonant itself (circled in Figure 9), the pitch over the vowels on each side are at about the same level. The lack of impact on adjacent segments is why I describe the depression over [dʒ] as localized; the pitch on surrounding vowels is not immediately impacted.

The phonetic correlation between lowered pitch and voicing can be seen again in Figure 10, which shows a additional contrast between voiced and voiceless stops:

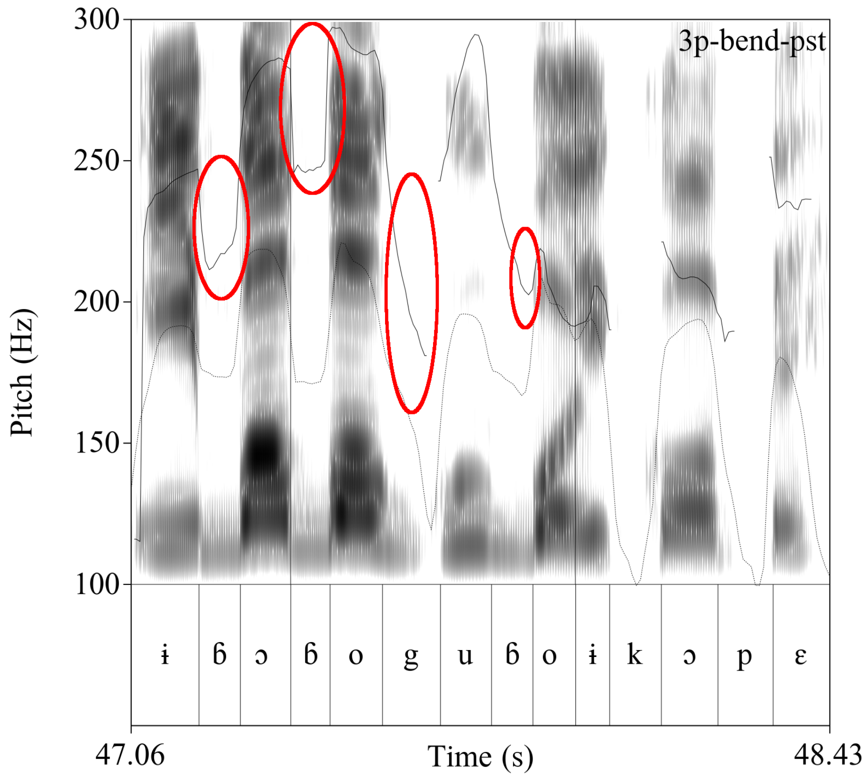


Figure 10. Analyzing Ndaka [ndk] consonants and their effect on pitch

The circled portions of Figure 10 show lowered pitch, localized over each voiced consonant. The pitch drops during each of these consonants, but the pitch of surrounding segments is unaffected: there is essentially one level on the vowels before and after each of these consonants except where the tone on the vowel changes (e.g., the L-H transition for [ɨɓɔ] [– ¯], and the H-L transition in [uɓo] [¯ –]).

The pitch depressions in Figure 10 correlate with voicing, which can be seen in the spectrogram F_0 (<120hz on the pitch trace axis). The consonants on the left are voiced, with noise in that voice bar area, while the voiceless consonants (the last two, on the right) are essentially quiet in the spectrogram F_0.

In languages with consonant-tone interaction (CTI), the phonetic correlation between pitch and voicing has been reanalyzed as a phonologically significant part of the tone system.

Similar data, but with a clearer change in tone, are shown in Figure 11:

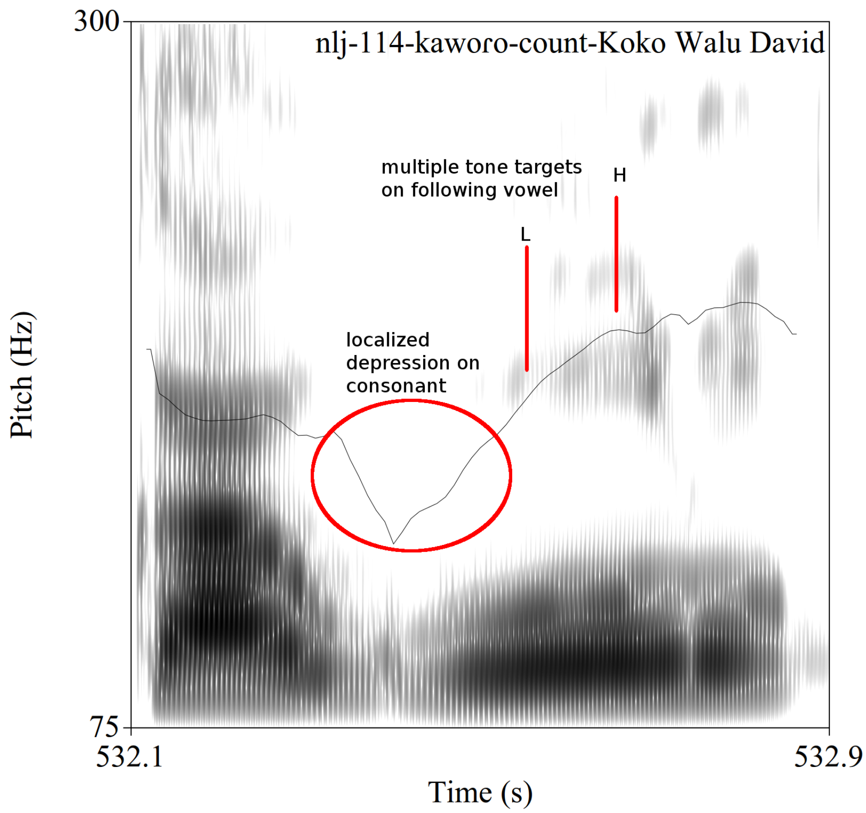


Figure 11. Localized depression on *w* in [nlj] [kawoɾo] [– ⟋ ¯] ‘count’, followed by additional tone target on vowel

Given that the phonetic correlation between pitch and voicing (as in Figure 9 and Figure 10) can be reanalyzed and used by speakers of a language as active parts of a phonological system (as in Figure 11), it would be helpful to know which phonological categories are used, whether they are consistent across languages with CTI, or not. Bradshaw (1997) ultimately shows that the phonetic effects on pitch of [voice] are phonologized in tone, while those of [-voice] are not, arguing that [voice] is a privative feature, rather than existing as a binary feature [±voice].

The most important implication of the privative nature of [voice], for this dissertation, is that there is something of a mismatch between the segments which produce phonetic depression of pitch, and those that depress tone phonologically (Bradshaw 1997). In Figure 10, pitch is depressed phonetically by egressive and implosive obstruents alike. But if [voice] is a privative phonological feature, then it would not be specified on implosives, which do not typically exist in voiced and voiceless pairs. As a result, when the effect of phonetic voicing is reanalyzed as a rule triggered by phonological voicing, implosives (and any other phonetically voiced segments which do not exist in voiced and voiceless pairs) should not trigger such a rule, since they would not bear the privative phonological feature [voice]. And this is just what Bradshaw (1999) has found, as will be discussed further in Section 2.2.4.

So we see a phonetic correlation between surface voicing and lowered pitch, which in some languages is phonologized into a rule corresponding a privative feature [voice] with low tone. Given that these correlations have been observed in a number of different languages (Bradshaw 1999), we ought to be able to account for them in some kind of unified way. The account I use will be addressed in the next section.

2.2.2 Accounting for consonant-tone interaction with [L/Voice]

Bradshaw (1999) proposed that consonant-tone interaction (CTI) be accounted for by a single, privative feature [L/Voice]. Perhaps the point most central to this thesis is that [L/Voice] is low tone and voicing in one feature, as opposed to a separate feature for each of low tone and voicing (as in most systems). The geometry and activity of that feature will be taken up in Section 2.2.3.

Based on the single feature hypothesis of Bradshaw (1999)'s [L/Voice], then, one would expect that CTI should involve the presence of a low tone with **all** segments which **are** phonologically specified for voicing, but **not** with segments which **are not** phonologically specified for voicing.

Before moving on to important specifics of this feature in a system (in Section 2.2.3 and Section 2.2.4), there are two things the above is not saying. First, it is not saying that correlation between voicing and pitch is only present with contrastive voice. That correlation is based on physics, and is present regardless of the phonologization of it, as seen in Figure 9. The [L/Voice] claim impacts the featural specification in a phonological system, not the phonetic implementation of pitch.

And second, Bradshaw (1999)'s [L/Voice] speaks to the geometric configuration of these features; it does not specify any particular derivation for either the underlying or surface forms in consonant-tone interaction. That is, there may be a straightforward association of the [L/Voice] feature to a following mora, or the feature may remain floating and cause downstep, or it may simply block the association of another tone, or something else. But whatever CTI there is, it should be based on the unity of low tone and contrastive voice, as the two are phonologically instantiated in a single feature [L/Voice].

2.2.3 [L/Voice] and geometry

Bradshaw (1999) claims that a single feature [L/Voice] (including both voicing and low tone) may attach to either or both of a laryngeal node or a mora, as in Figure 12, which is reproduced from example (5) of Bradshaw (1999:51):

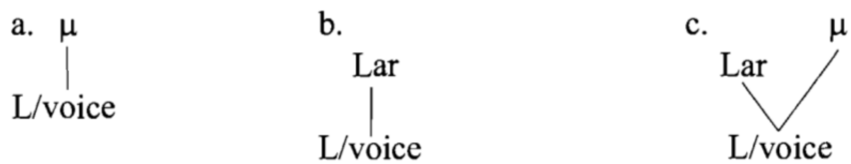


Figure 12. [L/Voice] geometries; example (5) from Bradshaw (1999:51)

That is, [L/Voice] can attach to a mora (as in Figure 12a), to the laryngeal node of a consonant (as in Figure 12b), or to both (as in Figure 12c). Though the feature specifies both voicing and low tone, each aspect of the one feature would be realized based on where the feature is associated. The low tone aspect of [L/Voice], for instance, would be expressed when [L/Voice] is associated to a tone bearing unit (TBU), or mora (as in Figure 12a), but not otherwise[19]. In the same way, voicing would be expressed when [L/Voice] is associated to a laryngeal node (as in Figure 12b); on a TBU it would likely be redundant, as TBU's are typically voiced on the surface. Where [L/Voice] is associated to both a laryngeal node and a TBU, both tone and voicing would normally be expressed contrastively (as in Figure 12c).

Figure 12 is clearly simplified[20], and does not include any processes that might be required to achieve the geometries presented. Adding in the consonants and vowels (but not the TRN or T node to keep the figure fairly simple), I interpret the processes present in a depressor consonant system under Bradshaw (1999) as in Figure 13, including the addition of a consonant-tone interaction (CTI) association rule, making [L/Voice] spread to a following mora, from where it could be expressed and otherwise interact with the tone system.

[19] This is typically the case with any tone not associated to a TBU, commonly referred to as a floating tone.

[20] For instance, it lacks a timing tier, typically acknowledged as needed where H and L are on separate tiers. The point of this image is not to address all questions of the geometry of [L/Voice], but rather to show how it interacts with both the consonantal system and the tone system.

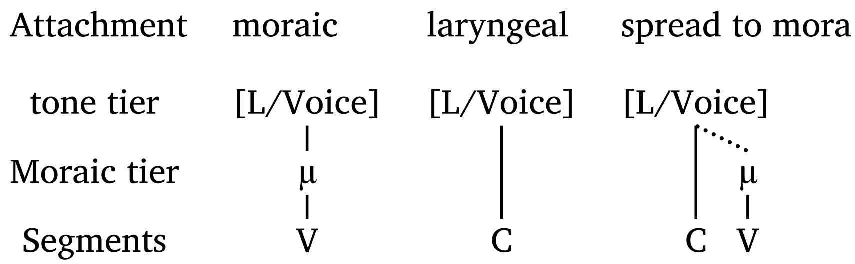


Figure 13. Options for associating [L/Voice]; my interpretation of Figure 12

That is, vowels associated to the [L/Voice] feature, as in the left column of Figure 13, would exhibit a low tone. The voicing specification would be superfluous[21], but the Low tone would be heard. Alternatively, [L/Voice] could be associated to a consonant without being associated to a TBU, as in the center column of Figure 13. Here the low tone would not be realized, as it is not associated to a TBU[22]. I interpret the geometry in 5c of Figure 12 to arise from the fact that in languages with consonant-tone interaction (CTI), the [L/Voice] feature lexically associated to a consonant **systematically spreads** to a mora, as in the right column of Figure 13[23]. A CTI rule as in Figure 13 allows for expression both of the contrastive voicing on the consonant, and the Low tone in the tone system, as well as providing the intrinsic connection between the two.

2.2.4 [L/Voice] and consonant types

Bradshaw (1999)'s [L/Voice] leads to certain predictions about which consonant types will participate in consonant-tone interaction (CTI). Given the fact that CTI is the phonologization of a physiological relationship between voicing and low tone (as in Section 2.2.1), one might expect that all consonants with surface voicing would be implicated. But the premise that [L/Voice] is a single, privative feature

[21] Voicing specification would typically be superfluous on most TBU's, unless there was also a voicing contrast for vowels (or other TBU's) in the language.

[22] A tone not associated to a TBU is typically not realized, though it may impact the system as would other floating tones. It could of course also spread to a TBU.

[23] I will assume the CTI rule formulation as in Figure 13 in this dissertation, but there could of course be a number of variations on this theme. For instance, the CTI rule operational in Bantu D30 rules is somehow sensitive to syllable structure, or perhaps lexical tones, in that C1 determines CTI in most cases, rather than C2. In some cases (especially in Nyali-Kilo) C2 appears to depress, but never following a C1 depressor in the same verb root. Nouns do not seem to have this constraint, either because they have two lexical TBU's, or else because they have two lexical tones, or for some other reason. In any case, I assume that any given language would be free to formulate one or more CTI rules to meet its needs, though for simplicity's sake I will continue to use the rule as given in Figure 13.

(as in Section 2.2.2), implies that CTI should only operate where segments are underlyingly specified for voicing. Bradshaw (1999)'s [L/Voice], then, accounts for the fact that consonant tone interaction (CTI) involves those consonants which are **specified for voice**. Furthermore, the [L/Voice] analysis predicts which segments will not participate in CTI, i.e., those segments which are voiced phonetically, but do not have a voiceless counterpart which contrasts with it phonologically, such as nasals, liquids, and implosive stops in most languages. As given in Bradshaw (1999), the behavior of each consonant type can be summarized as in (6):

(6)
a. Voiced obstruents (e.g., *b*, *d*, *g*, *z*, and *v*) are the primary consonants that depress tone.
b. Voiced double articulated stops (e.g., *gb*) pattern with the other voiced obstruents.
c. In some languages voiced sonorants (e.g., *m*, *n*, *w*, *l* and *j*) depress tone as well.
d. Nasal-obstruent sequences will occasionally lower tone, depending in part on their phonological status in the language.
e. Implosive stops (e.g., *ɓ* and *ɗ*) do not pattern with the other voiced obstruents.
f. In no case would we expect voiceless consonants to lower tone.

The generalizations in (6) can be further broken down by phonological relevance as follows:

(7)
a. Contrastively voiced: voiced obstruents (single and double articulated) are most likely to depress tone (6a-b).
b. Contrastively voiced in some languages: Voiced sonorants and Nasal-obstruent sequences depress tone in some languages (6c-d).
c. Not contrastively voiced: Implosive stops and voiceless consonants are unlikely to ever depress tone (6e-f).

These generalizations lead to the following predictions:

(8) Predictions derived from Bradshaw (1999)
a. CTI should occur with **all** segments which **are** phonologically specified for voicing (e.g., *b*, *d*, and *g* in most languages), and

b. CTI should occur with **no** segments which **are not** phonologically specified for voicing (e.g., voiceless *p, t, k*, as well as *ɓ, ɗ, ɠ, m, l*, and *r* in most languages, where they are best analyzed as underlyingly unspecified for voice)

To preview the outcome of these predictions, the data shown in Chapters 3–6 will show (8a) to be upheld, while the data that argues against (8b) will be addressed in Chapter 7.

The predictions in (8) flow naturally from Bradshaw (1999)'s [L/Voice], because there is no phonological specification for voice for sonorants and implosives, there is no [L/Voice] feature, and therefore no impact on the tone system, even though they are phonetically voiced, and therefore no impact on pitch as in Figure 10 on page 19.[24]

Note that in Figure 9 (on page 36) [ndʒ] is not contrastively voiced in Nyali-Kilo; there is no [ntʃ][25]. As a result, there is no expectation of consonant-tone interaction here, though there remains a phonetic depression of pitch, which is localized over the consonant.

In summary, for languages with CTI, Bradshaw (1999) predicts that it will operate where the consonant in question is phonologically specified for voice. Because there is an intrinsic relationship between low tone and voice (i.e., they are one feature [L/Voice]), there should be consonant-tone interaction everywhere voicing is specified, and nowhere where it is not.

I assume Bradshaw (1999)'s [L/Voice] as the basis for my analysis, though I also evaluate its ability to account for the tone systems of Bantu D30 languages. Wherever data are found which do not match the predictions generated by Bradshaw (1999), we must be prepared to understand either some unexpected subtlety in the data, or else a need to rethink this understanding of CTI. This will become particularly relevant as we consider what appears to be a lexicalization of a CTI system, as described in Rasmussen (2012 and 2015a), but also presented in Chapters 3–6, and analyzed in Chapter 7 of this dissertation.

[24] Why *w* in Figure 11 does seem to interact with the tone system (i.e., despite the lack of a /w̥/ phoneme) is a more complicated question, which will be addressed in Chapter 7.

[25] See Table 4 and Table 5 in Section 2.4 for an overview of the Bantu D30 consonant systems, including where to expect consonants to be specified for voice (i.e., depressors).

2.3 Language Classification

This section addresses the Bantu D30 languages, in three dimensions. Section 2.3.1 provides some context for the Bantu D30 languages in terms of genetic affiliation, geography and phonology, and is followed by two sections on comparative methodologies. Section 2.3.2 surveys the comparative method, which will be used to compare the languages one to another in terms of their genetic relationships. Section 2.3.3 surveys areal linguistics, or geographical typology, which organizes languages into units of geographically close languages, rather than by genetics or universal principles. Sections 2.3.2 and 2.3.3 are included because a study of a language or group of languages in isolation is not as productive as studying it in its historical and geographic context.

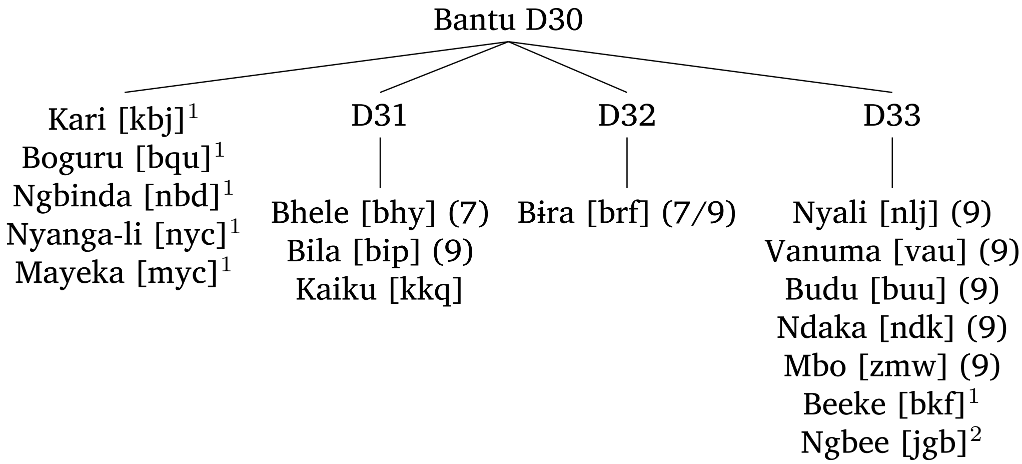
Together Sections 2.3.1, 2.3.2 and 2.3.3 give the reader a background on the Bantu D30 languages, in terms of some facts about the languages themselves, their genetic relationships (to each other and with other languages), and their relationships to neighboring languages, apart from genetic relationship. And ultimately we find that the Bantu D30 languages contain an interesting combination of Bantu and non-Bantu features (apparently from both genetic and non-genetic sources). So in order to see the whole picture we need to look at both the genetic and non-genetic factors which have brought the Bantu D30 languages to where they are today.

2.3.1 The Bantu D30 Languages

This section gives a brief overview of the Bantu D30 languages, including mention of two phonological properties a number of them hold in common: nine vowel systems and consonant-tone interaction.

The Bantu D30 languages are one subgroup of languages within the Bantu D languages, which are a further subgroup of the Narrow Bantu languages in the Niger-Congo language family (Simons, & Fennig 2018). The Bantu D30 languages are shown in Figure 14, with "dying" (EGIDS[26] 8-9) and unsubgrouped languages (left), as well as three subgroupings (right) for languages which are "developing", "vigorous" or "in trouble" (EGIDS 7 or greater) as given in Simons, & Fennig (2018), and with the number of contrastive vowels where that information is available:

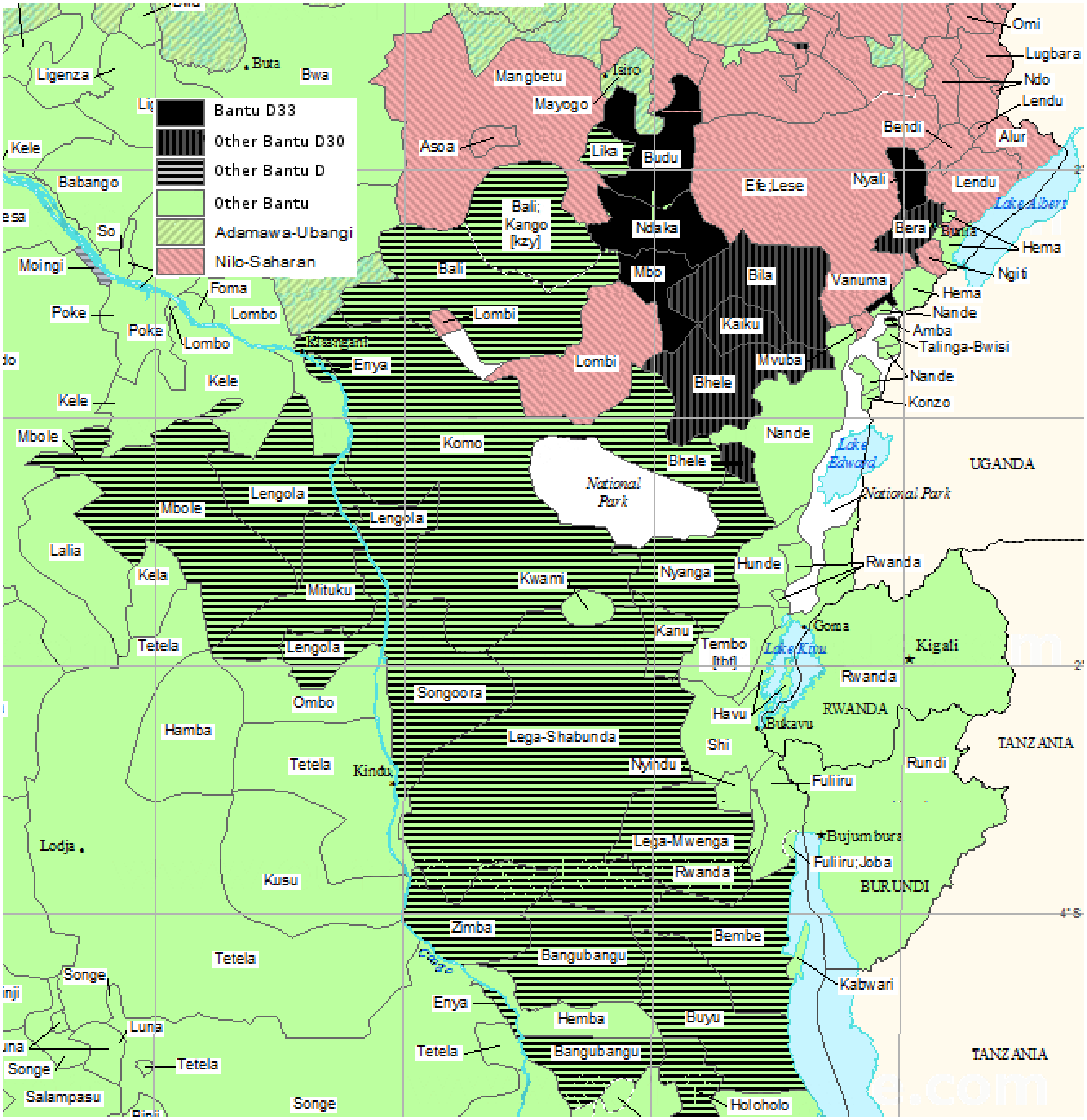
[26] The Expanded Fishman's Graded Intergenerational Disruption Scale (EGIDS), an evaluative framework of language endangerment, is documented in Lewis & Simons (2010), and is used as the indication of language vitality and endangerment in the Ethnologue (Simons, & Fennig 2018).



1. Dying; 2. Extinct; source: Simons & Fennig (2018)

Figure 14. Genetic Affiliations of the Bantu D30 Languages

The Bantu D languages are located in the northeastern corner of the Democratic Republic of the Congo. The map in Figure 15 shows this region of the DRC, along with bordering countries Uganda, Rwanda, Burundi, and Tanzania, to the west of the DRC on this map. While not on this map, this area is also south of South Sudan and Central African Republic. The Bantu D33 languages are shaded in black, essentially surrounded by Nilo-Saharan languages. Other Bantu D languages are shaded horizontally or vertically:



Adapted, with permission, from Ethnologue: Languages of the World, 18 th online edition.

Figure 15. Map of the Bantu D Languages, with highlighting D30 and D33 subgroups

While there is very little description of the Bantu D30 languages, it is already possible to do a little comparative work showing novel vowel contrasts. Proto-Bantu has been reconstructed with only seven vowels (Hyman 2003a and Odden 2014), but a number of Bantu D30 languages show more. For instance, while Bhele [bhy] (D31) has only seven vowels (Hartell 1993), Bila [bip] (also D31) has nine (Kutsch Lojenga 2006). ATR Harmony (Casali 2003 and 2008, c.f., Starwalt 2008, Stirtz 2009, and Gafos & Dye 2011) is found in a number of Bantu D30 languages. Bɨra [brf] (D32) has seven contrastive vowels, though there are nine phonetic vowels through ATR harmony (Morgan & Walker 2009). Nyali-Kilo [nlj] has nine contrastive vowels (Rasmussen 2012), as do other D33 languages Vanuma [vau] (Morgan & Van Otterloo 2009), Budu [buu] (Kutsch Lojenga 1994), Ndaka [ndk] (Rasmussen 2015b), and Mbo [zmw] (Rasmussen 2015a). These nine vowel languages therefore include all the vigorous Bantu D33 languages. Their genetic relationship can be confirmed by the shared innovation (c.f., Section 2.3.2) of extra vowels.

Another point of interest regarding the Bantu D30 languages is that depressor consonants are reported to be historically involved in Bila [bip] (D31; Kutsch Lojenga 2006), and to some extent in each D33 language for which information is available (c.f. sources cited for each language above). Given that consonant-tone interaction (CTI) is not reconstructed for Proto-Bantu, the presence of CTI is an innovation in the tone system, which may help us better understand how new tone contrasts develop (tonogenesis, c.f., Chapter 7).

A majority of the data in this dissertation comes from Ndaka [ndk] and Mbo [zmw], though I also include data from Nyali-Kilo [nlj], Vanuma [vau], Budu [buu], and Bɨra [brf] where I have them. That is, it includes the vigorous Bantu D33 languages and the one Bantu D32 language. As a matter of pure convenience, I will refer to these as the Bantu D30 languages throughout this work, though I make no claim to represent the Bantu D31 languages Bhele [bhy], Bila [bip], and Kaiku [kkq]. Generalizations also do not always represent Bɨra, but that will hopefully be evident at each point by the degree to which it conforms to the Bantu D33 data, when Bɨra data is provided.

Given this brief overview of the larger context of the Bantu D30 languages, the following sections will provide background on two different methods of comparing them, with and without presumed genetic affiliation.

2.3.2 Comparative (Genetic) Classification

For more than half a century, linguistics has used at least two main means of classifying languages, which may at times be at odds with one another. COMPARATIVE linguistics makes claims about genetic/historical origins (Greenberg 1960, Newman 2000), while TYPOLOGICAL linguistics (c.f. Section 2.3.3) does not –in fact, typological classifications generally appear stronger when they include genetically unrelated languages. This interaction between classification strategies has remained active, e.g. Greenberg (1960), Newman (2000), Heine & Nurse (2000, 2008). I will take each in turn, to provide the historical context of the Bantu D30 languages (this section) followed by their geographical context in Section 2.3.3. Understanding both genetic and geographical factors should provide a more insightful analysis than would be obtained by studying the Bantu D30 languages without respect to their genetic and geographical contexts.

The aim of historical and comparative linguistics is to learn of the history of a language or group of languages, where there there is no access to a written history (Newman 2000). There are two essential steps in the process; the first establishes the presence of relationship, and the second specifies the type of relationship. Establishing relationship between Bantu languages is not particularly relevant today, as they have been generally understood as a genetic group for some time (Greenberg 1955). Nevertheless, an overview of the process can be found in Newman (2000:261-3).

Given the understanding that Bantu (and thus the Bantu D30) languages are related, the question then becomes what kind of relationships exist between the individual languages. Subcategorization within a family is established on the basis of shared innovations. That is, two languages that share a novel feature (of any kind) may be taken to have been one language until after developing that feature (Newman 2000). Establishing shared innovations is done through comparing cognate word forms in related languages, to see which forms differ in which regular ways. With enough regular sound correspondences, one can deduce at least some proto word forms, or reconstructions. With enough reconstructions, one can describe the system of the proto-language, including which features are retained or lost in each of the daughter languages (Newman 2000).

In any case, sound changes are assumed to be regular, not isolated or idiosyncratic (Newman 2000). But this assumption (along with the status of a given word as a cognate) must be held with care. Language change must begin somewhere before generalizing, both in terms of the system of a language, and in terms of its dialects (as throughout The Handbook of Historical Sociolinguistics, Hernández-Campoy & Conde-Silvestre 2012).

Despite the difficulties, a great deal of proto-forms have been reconstructed for Proto-Bantu (the ancestor of the Bantu languages). The Bantu Lexical Reconstructions 3 (BLR3; Bastin, et al 2002) is perhaps the most authoritative current compilation of reconstructions for proto-Bantu, and as such is the source of all proto-Bantu reconstructed forms in this dissertation. But while many proto-Bantu forms have been reconstructed, the reflexes in proto-daughters may not also have been reconstructed. One product of this dissertation is a beginning of the reconstruction of Proto-Bantu D, a descendant of proto-Bantu and ancestor of the Bantu D languages. It will at times be convenient that I should propose a plausible D30 reconstruction for a given word as part of a derivation, but I will not spend time defending my proposed proto-Bantu D30 forms in this dissertation.

2.3.3 Typological (non-Genetic) Classification

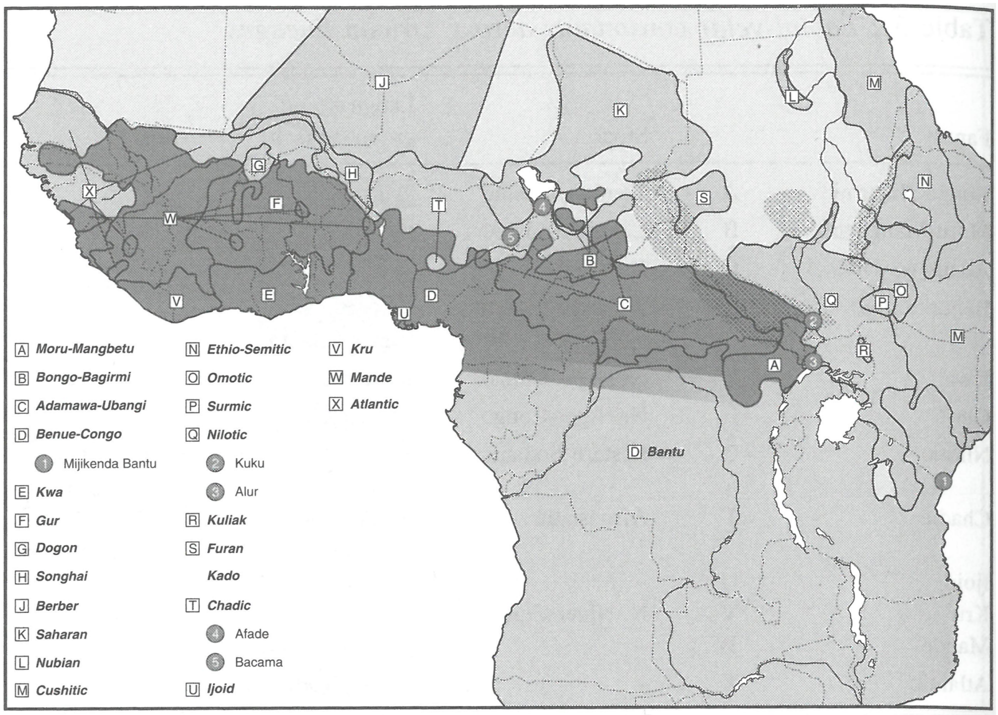
While the comparative method answers many questions about the history of specific languages, it would be a mistake to believe that all similarities between any two given languages must come from a common genealogy. This issue is addressed throughout Heine & Nurse (2008), *A linguistic geography of Africa*. One pair of authors describe language areas as "…defined exclusively in terms of linguistic parameters without reference to the historical forces that gave rise to them" (Heine & Leyew 2008:16). While there may be some disagreement over the extent to which language contact is necessary to the definition of a language area, all agree that language areas should not be based on similarities which originate from genetics or universal principles (Heine & Leyew 2008). That is, areal linguistics is a typological classification within a geographical rather than a universal or genetic domain.

With this basic understanding of the interaction of genetic and geographic influences in mind, there are a number of specific language similarities where areal linguistics can shed some light. For instance, the multitude of nine vowel languages with ATR harmony in Bantu D30 languages seems very strange from a genealogical perspective, since nine vowel ATR systems are generally not found in Bantu languages. They are very common, however, in Central Sudanic and other Nilo-Saharn languages, which happen to be geographically very close to the North and East of the Bantu D30 languages.

Furthermore, one survey shows 80 of 99 African languages having tone (Heine & Leyew 2008:29). So while we may understand tone to be inherited from a proto-language, there is also something particularly African about having tone. But not every issue is as clear as we might like; Heine & Nurse (2000:5) specifically mentions the origin of features defining the Khoisan language family as an open debate, for instance. That is to say, what features exist in a set of languages may by all appearances be genetic in origin, but the reality may be more debatable, when one considers both genetic and other factors.

Clements & Rialland (2008) looks at possible zones within Africa. The “Sudanic belt” has distinctive consonants in labial flaps, labial-velar stops, and implosives, each of which are infrequent or absent in the other African areas and outside of Africa (based on a survey of 495 languages, 100 of which are from the Sudanic belt; Clements & Rialland 2008:40). The Bantu D30 languages have two of the three criteria mentioned (labial-velar stops and implosives).

Güldemann (2008) describes essentially the same area, but in greater detail. He gives six features that may be used to define the area: Logophoricity, Labial-velar consonants, ATR vowel harmony, S-(AUX)-O-V-X word order, V-O-NEG word order, and labial flap consonants. This area is further described as north of the Congo basin (Güldemann 2008:152), which along with the six maps showing distribution of the six features, indicates that the area ends north of Bantu. The Moru-Mangbetu is the southernmost “frequent” marked grouping in each, and “most of Narrow Bantu” is labelled as not frequent for each feature (Güldemann 2008). The Bantu D languages are surrounded by Moru-Mangbetu (“A” in Figure 16; north, east, and west of Budu in Figure 15), as illustrated in Figure 16:



Map 5.2 Labiovelar consonants across African lineages

Figure 16. Labiovelars in Africa, from map 5.2 in Güldemann (2008:158) [used by permission]

Though the Bantu D33 languages are not counted as part of the zone indicated by shading in Figure 16, they meet two of the three featural criteria for which I have data; labialvelar stops and ATR vowel harmony are present in every vigorous Bantu D33 language. The presence of these features in Bantu D30 indicates that the zone may go a bit further south than indicated in Güldemann (2008), to include at least some Bantu languages. As such, the Bantu D30 languages contain an interesting combination of Bantu and non-Bantu features, likely from both genetic and non-genetic sources.

One of the main difficulties in determining the origin of the similarities between languages is the lack of good data from a broad number of African languages. Heine & Nurse (2000) indicates the presence of a “reasonably accurate and comprehensive reference grammar available” in less than 100 African languages (Heine & Nurse 2000:5). The prevailing lack of relevant data is a good reason to work in the Bantu D30 languages, as mentioned in Section 1.5. There is not only a need for descriptions of African languages in general, but the Bantu D30 languages speak directly to the question of inheritance vs geography. That is, they are clearly a subgroup of Bantu, yet they also likely belong to what Güldemann (2008) describes as the “Macro-Sudan Belt”, though they are not included as such in that work. This dissertation, then, seeks to show that the “Macro-Sudan Belt” of Güldemann (2008) needs to be extended South at least a bit, on the one hand, but also that at least some of the features of the Bantu D30 languages may come from their neighbors, rather than from their parents.

2.4 Bantu D30 Segments and Verbal Structure

This section provides details on segmental phonology and verbal structure which are common to Bantu D30 languages, which should provide a helpful framework within which to understand the data in the chapters to follow. Section 2.4.1 covers the basic morphological structure of Bantu D30 verbs. Section 2.4.2 describes the four paradigmatic[27] verb root classes by tone observed in conjugated tone data (by subject and TAM), as well as the frames used to control for them. Section 2.4.3 lays out briefly the consonants and vowels in the Bantu D30 languages.

2.4.1 Bantu D30 verbal structure

As mentioned in Section 1.6, this dissertation is limited to a small number of verbal contexts. Infinitive verb forms for this study were elicited based on formal and functional criteria. Functionally, Bantu D30 infinitive verb forms fit into an auxiliary verb construction, of the form glossed ‘I want to X’. Formally, they match verb forms as generally the case in Bantu, in other Bantu D33 languages, and in other verbs of the same language. As with many Bantu Languages, the Bantu D30 languages have canonical word forms for verbs[28], as shown in Figure 17:

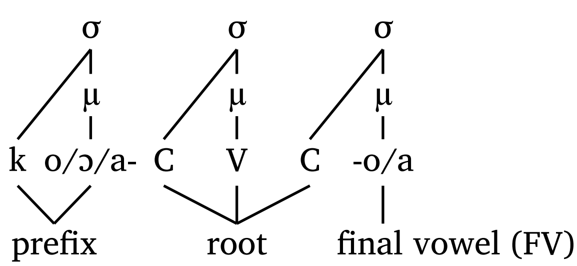


Figure 17. Syllabification of lexical items in canonical (CVC) Bantu D30 verbs.

[27] Root classes by tone are covered in Chapters 3–5.

[28] Nouns also have canonical forms, the vast majority having the form px-CVCV, where the prefix (px) is either V- or CV-. This dissertation, however, is mostly limited to verbal data.

That is, infinitive verbs are generally of form px-CVC-FV[29], where the prefix (px) is *kɔ-/ko-* for Mbo and Ndaka and *kʉ-/ku-* for Budu (each according to ATR harmony), and *ka-* for Nyali-Kilo and Vanuma[30]. The final vowel (FV) is *-o/-a* (also according to ATR harmony) for Nyali-Kilo, Vanuma, Budu, Mbo and Ndaka, though just *-a* (without harmony) for D32 Bɨra [brf]. As can be seen in Figure 17, the second consonant of the root is syllabified with the FV. As a result, canonical infinitive verb forms begin the root on the *second* syllable in a three syllable word, and the final mora is not part of the lexical root. This structure is therefore a typical Bantu verb structure, though I am not for the moment considering extensions, which may occur between the root and FV (Givón 1971, Hyman 1993, and Bresnan & Moshi 1990). The part of this syllable structure that is the most critical to this dissertation is the first root consonant (C1), which seems to be the trigger for consonant-tone interaction (CTI) in Bantu D30. The Bantu D30 canonical infinitive verb pattern is exemplified in the Mbo data in (9):

(9)	a.	ko-sis-o	[_ ¯ ¯]	move forward
	b.	kɔ-kɨj-ɑ	[— — —]	act
	c.	ko-ɓund-o	[_ ¯ ¯]	break
	d.	kɔ-ɓʉt-ɑ	[_ ¯ —]	become long
	e.	ko-ɓeɲ-o	[_ ¯ ¯]	wink
	f.	kɔ-kɛk-ɑ	[_ ¯ ¯]	decorate
	g.	ko-sok-o	[_ ¯ ¯]	cackle
	h.	kɔ-mvɔɗ-ɑ	[_ ¯ ¯]	suck
	i.	kɔ-bɑb-ɑ	[¯ _ ⟋]	carry

See appendix D for data on verbs lacking a full CVC root structure, including a discussion of what those data say on the permanence of ATR.

2.4.2 Tone frames and verb root classes by paradigmatic tone

Conjugated verb forms are presented in Chapters 3 and 4 in a tone frame consisting of a free pronoun preceding the verb, and an adverbial time word following the verb, for a whole utterance of <pronoun|verb|adverb>. This frame fulfills two functions in the tone analysis. First, it provides a tone frame for the words, so that we are not just looking at them in isolation. And secondly, it allows us to clearly identify and control for the four conjugated forms for each verb, which are

[29] See appendix D for a brief discussion of (V)C roots.

[30] D32 Bɨra [brf] does not have an infinitive verb prefix.

described in the remainder of this section. Verb root classes by tone for Ndaka, Mbo, and Nyali-Kilo will be addressed in Chapters 3, 4 and 5, respectively.

For at least Ndaka, Mbo and Budu (though somewhat differently for Bɨra), free subject pronouns fall into two classes by tone, with 1s and 2s being low, and 3s and all plural subjects being low-high (see appendix C for more details). Bound subject pronouns fall into the same two classes by tone, though their forms differ from those of the free subject pronouns. The bound subject pronouns are of form (C)V-, and replace the infinitive prefix in Figure 17 to form conjugated forms of canonical CV structure (C)V-CVC-FV, which is essentially the same as for infinitive forms in Figure 17, apart from the prefix.

The 2s and 3s bound subject pronouns were chosen for this study because they represent a minimal difference between the two observed tonal classes into which the bound subject pronouns split conjugated verb forms. They also contain minimal segmental information (a single vowel[31]), so the impact of other segments is minimalized.

Because 2s and 3s conjugated forms are only distinguishable by tone, the presence of a preceding free subject pronoun (i.e., 2s [ʉwɛ] [‒ ‒] or 3s [ɨjɛ] [– ¯]) makes the subject of the verb clear, in addition to providing a tone frame to better interpret the verbal tone. Further information on the full range of pronouns in the Bantu D30 languages, and how they split into classes by tone, can be found in appendix C.

Conjugated Bantu D30 verb forms also split into at least two other classes by tone, indicating tense-aspect-modality (TAM) categories I am for this document calling 'past' and 'future'. These names are not the result of any kind of rigorous semantic study; rather they reflect the glossing of these forms in French and Swahili. These TAM categories may well indicate primarily aspect, and not tense at all. While there are certainly other TAM categories in these languages, these two formal categories in my data serve a similar purpose to that of the bound subject pronouns, in that their lexical tone differs, but their segmental content does not. The purpose of this investigation was not to establish all the forms and functions used to represent TAM categories in these languages, but rather to establish the identity and number of lexical verb root classes by tone through conjugated data, which I have done in Chapters 3 and 4. Because I have not done any serious semantic study on these categories, I resist the temptation to give these categories

[31] These bound subject pronouns are more likely best analyzed as only tone. This same vowel is present across all bound subject pronouns for at least Ndaka, Mbo, and Budu. In any case, the 2s and 3s pronouns contain the least segmental information and represent the two classes by tone, so they seem ideal for comparison. More details on the information contained in consonants, vowels, and tone in Ndaka, Mbo, Budu and Bɨra can be found in appendix C.

more meaningful labels than past (for the TAM category that includes what is glossed as 'past tense' in French and Swahili) and future (for the TAM category that includes what is glossed as "future tense" in French and Swahili). Nevertheless, these categories differ one from the other formally only in tone, and they can be controlled by the addition of an adverb after the conjugated verb, either [ɨkɔpɛ́] [¯ – −] 'yesterday (n)' or [kubá] [¯ –] 'tomorrow (n)' (or [ɨpáɓɔ] [– ¯ –] 'today (n)') for Ndaka and [ɨkɔɸɛ́] [¯ – −] 'yesterday (n)' or [naɨbá] [¯ – −] 'tomorrow (n)' (or [íweí] [¯ – −] 'today (n)') for Mbo. So these following adverbs not only provide consistent tone frames to interpret the verbal tone, they also make the TAM category clear.

The binary tone split according to (2s/3s) subject, combined with the binary split according to (past/future) TAM category gives four minimally different underlying tone patterns on any given verb root (i.e., there are different tone and meaning for four forms, each of which have the same segments, for each root), at least for Ndaka, Mbo and Budu, as in Table 2:

subj/TAM	past	future
2s	2s-past	2s-future
3s	3s-past	3s-future

Table 2. The combinations of subject and TAM classes by tone in this work

For at least Ndaka and Mbo, then, any given root has four possible conjugated tone melodies. These sets of four melodies will be used to characterize and distinguish verb root classes by tone in Chapters 3 and 4.

The presence of bound subject pronouns before the conjugated verb, and the TAM-marking adverb following the conjugated verb, allows the non-native speaker of these languages to see which subject-TAM combination has been selected. This can be important because several of these forms are segmentally minimal, for at least Ndaka and Mbo. Note however, that these free pronouns and TAM adverbs have been added for the convenience of this study; the conjugated verb words are pronounceable in isolation as complete sentences. One implication of this fact is that each verb root can form four different full sentences (i.e., with different meanings), but which are distinct only by tone, in terms of their pronunciations. Other implications of these tonally minimal distinctions, including the relative importance of tone for the orthography of these languages, are addressed in appendix C.

2.4.3 Bantu D30 segments

As mentioned in Section 2.3.1, all vigorous Bantu D33 languages have nine (9) vowels. They are transcribed in this dissertation in the following manner, according to ATR harmony:

+ATR Vowels		-ATR Vowels	
i	u	ɨ	ʉ
e	o	ɛ	ɔ
		a	

Table 3. The nine vowels found in Bantu D33 languages

Tables 4 and 5 show the consonants of Bantu D30, from two different perspectives. Table 4 shows one chart for the subfamily, with notes for which consonants appear in which language(s). Table 5 shows one column for each language, though with parallel columns for voiced and voiceless segments for each language, and with shading for contrastively voiced segments. In either case, the data come from the same sources for each language: Bɨra [brf] (D32; Morgan & Walker 2009), Nyali-Kilo [nlj] (Rasmussen 2012), Vanuma [vau] (Morgan & Van Otterloo 2009), Budu [buu] (Kutsch Lojenga 1994), Ndaka [ndk] (Rasmussen 2015b), and Mbo [zmw] (Rasmussen 2015a). Evidence of Ndaka [ndk] and Mbo [zmw] [k^w] contrasting with [g^w] has not been published before, but exists, as in (10) and (11):

(10) Ndaka [ndk] k^w and g^w
  a. [$kog^w ejo$] [– ⁄ –] gather
  b. [$kok^w ejo$] [− − −] wrap package

(11) Mbo [zmw] k^w and g^w
  a. [$kog^w ajo$] [¯ _ ¯] gather
  b. [$kok^w ejo$] [− − −] wrap package

		Bilab.	**Lab-dent**	**Alv.**	**Postalv.**	**Lab-Vel**	**Velar**	**Glottal**
Plos.	**egr.**	p b		t d	t͡ʃ d͡ʒ	k͡p g͡b	k g	
	prenas.	^{m}b		^{n}d	^{n}d͡ʒ	^{n}g͡b	$^{\text{ŋ}}$g	
	impl.	ɓ		ɗ				
	palatal			d$^{j}_{4}$			k$^{j}_{8}$ g$^{j}_{4}$	
							$^{\text{ŋ}}$g$^{j}_{8}$	
	labial						k$^{w}_{9}$ g$^{w}_{9}$	
Fric.			f v$_{3}$	s z$_{5}$				
		ɸ$_{1}$ β$_{2}$	$^{\text{ɱ}}$v	^{n}z$_{6}$	ʃ$_{6}$			h$_{10}$
Nasal		m		n	ɲ			
Approx.				l ɾ$_{7}$	j	w		

Table 4. Summary of Consonants found in Bantu D32 and D33 languages[32]

[32] Consonants with subscripts are not found in all languages. 1: [zmw] only; 2: [vau] and [brf] only; 3: except [zmw]; 4: [buu] only; 5: except [nlj] and [vau]; 6: [brf] only; 7: [brf] only, but allophone of /l/ in [vau] and [nlj]; 8: [vau] and [buu] only; 9: [ndk] and [zmw] only; 10: [brf], [buu] and [vau]

	Ndaka [ndk]		Mbo [zmw]		Nyali-Kilo [nlj]		Vanuma [vau]		Budu [buu]		Bɨra [brf]	
Plosives	p	**b**	p	**b**	p	**b**	p	**b**	p	**b**	p	**b**
	t	**d**	t	**d**	t	**d**	t	**d**	t	**d**	t	**d**
	t͡ʃ	**d͡ʒ**	t͡ʃ	**d͡ʒ**	t͡ʃ	**d͡ʒ**	t͡ʃ	**d͡ʒ**	t͡ʃ	**d͡ʒ**	t͡ʃ	**d͡ʒ**
	k͡p	**g͡b**	k͡p	**g͡b**	k͡p	**g͡b**	k͡p	**g͡b**	k͡p	**g͡b**	k͡p	**g͡b**
	k	**g**	k	**g**	k	**g**	k	**g**	k	**g**	k	**g**
prenasalized		^{m}b		^{m}b		^{m}b		^{m}b		^{m}b		^{m}b
		^{n}d		^{n}d		^{n}d		^{n}d		^{n}d		^{n}d
		^{n}d͡ʒ		^{n}d͡ʒ		^{n}d͡ʒ		^{n}d͡ʒ		^{n}d͡ʒ		^{n}d͡ʒ
		^{n}g͡b		^{n}g͡b		^{n}g͡b		^{n}g͡b		^{n}g͡b		^{n}g͡b
		^{ŋ}g		^{ŋ}g		^{ŋ}g		^{ŋ}g		^{ŋ}g		^{ŋ}g
implosive		ɓ		ɓ		ɓ		ɓ		ɓ		ɓ
		ɗ		ɗ		ɗ		ɗ		ɗ		ɗ
palatalized										d^{j}		
							k^{j}		k^{w}	**g** ^{w}		
								^{ŋ}g^{j}		^{ŋ}g^{j}		
labialized	k^{w}	**g** ^{w}	k^{w}	**g** ^{w}								
Fricatives			ɸ					β				β
	f	**v**	f		f	**v**	f	**v**	f	**v**	f	**v**
	s	**z**	s	**z**	s		s		s	**z**	s	**z**
											ʃ	
							h		h		h	
		^{ɱ}v		^{ɱ}v		^{ɱ}v		^{ɱ}v		^{ɱ}v		^{ɱ}v
												^{n}z
Nasals		m		m		m		m		m		m
		n		n		n		n		n		n
		ɲ		ɲ		ɲ		ɲ		ɲ		ɲ
Approximants		l		l		l~ɾ		l~ɾ		l		l ɾ[33]
		j		j		j		j		j		j
		w		w		w		w		w		w

Table 5. Consonants of Bantu D30 by language[34]

[33] Morgan & Walker (2009) claim complementary distribution between [l] and [ɾ], though both seem to be included in their alphabet chart. My data contradict their proposed environmental distribution (before [a]).

[34] sources: Morgan & Walker 2009, Rasmussen 2012, Morgan & Van Otterloo 2009, Kutsch Lojenga 1994, Rasmussen 2015b, Rasmussen 2015a, *k*^{w} and *g*^{w} as above; consonants specified for voicing shaded in bold

Recall that the presence of phonologically specified voicing is the defining feature for depressor consonants. As a result, the consonantal systems in Table 5 allow us to see where to expect depressor effects. For instance, recall from (7b) on page 42 that nasal-obstruent sequences are contrastively voiced in some languages but not in others. In the languages in Table 5, they are not. The same is true for the implosive stops, as well as for all the sonorants. The existence of both [ɸ] and [β] in the subgroup might lead one to expect [β] to be a depressor, but it does not exist in opposition to [ɸ] in any language in Table 5 (i.e., any given language has [ɸ] or [β], or neither, but not both), so bilabial fricatives in Bantu D30 are either voiceless or voiced without a voiceless counterpart, and neither should be expected to interact with tone, as neither would be phonologically specified for voicing.

2.5 Kutsch Lojenga (2006) and Tonogenesis

There is little published work addressing the development of tone in the Bantu D30 languages, though there is a brief foray in Kutsch Lojenga (2006) worth mentioning, from the D31 Bila [bip] chapter of *The Bantu Languages* (Nurse & Philippson 2006). The basic outlines of that story align with what I will say in this dissertation, but there are some major differences in scope and objectives between the two. In this section, I outline a number of crucial gaps in the data and analysis presented in support of claims of tonogenesis in Kutsch Lojenga (2006). In this way, I set the stage for how I will approach the question of tonogenesis in the Bantu D30 languages.

Any discussion of the argument for tonogenesis in Kutsch Lojenga (2006) must acknowledge its intended scope. The section on tone covers less than two pages of the 25 page chapter, which covers ethnography, phonology, nominal morphology and verbal morphology. The last two of these cover more than half the chapter. Given that the entire phonology section covers only about five and a half pages, it would be unfair to imagine that a few paragraphs on tonogenesis were central to the chapter on Bila.

Yet there are two claims made about the tone system of Bila in Kutsch Lojenga (2006) that I will discuss here. The first deals with the analysis of tone in Bila generally, and the second deals with tonogenesis.

The first claim begins “Minimal pairs and sets like the following rule out an analysis into two tones underlyingly” (Kutsch Lojenga 2006:457). This is followed by three triplets and four pairs, which is followed by “My conclusion is therefore that Bila has three underlying tones synchronically, High, Mid, and Low, which

I call High (H, acute accent), Low (L, zero marking), and extra-Low (xL, grave accent) respectively, for the sake of comparison with the neighboring languages Komo and Bhele." (Kutsch Lojenga 2006:457)

However, this claim and the data presented with it does not indicate a principled methodology to establish contrast, such as that outlined in Snider (2014), and as discussed in Section 2.1.3. In terms of the data chosen to support this claim, only one of the sets of data could be taken as three tonemes contrasting in the same environment. The other triplet has two words with initial H tone, and the third with initial xL tone, so the word final contrast is not compared to a consistent initial tone. The remaining data have only two words each, so none show the contrast between three different pitch levels in the same environment. As a result, there is not clear evidence of xL as a distinct tonal unit. It would be nice to see if her "xL" contrasts in a comparable environment with each of L and H, as a true toneme would (this kind of test for tone contrast is discussed in Snider 2014).

There are also problems with the representation of the data mentioned above. As described in the above quote, the data in Kutsch Lojenga (2006) are presented in a phonemic transcription system, making the phonetics difficult to recover (c.f. Section 2.1.5). There is no justification for this phonemic marking system, neither in terms of phonetic data, nor in terms of other work cited, which might potentially justify those underlying categories on the basis of phonetics.

Kutsch Lojenga (2006) does provide some basis for reconstructing some of the phonetics. According to her prose, downstep "…lowers the register in such a way that a following H tone is lowered one step, namely to the level of L, as exemplified with the words *míkí* 'child (n)' *kíma* 'thing (n)', and *ɓá*ꜜ*wá* 'who(m)?'" (Kutsch Lojenga 2006:458). According to this prose, the pitches on 'thing' and 'who(m)?' should be indistinguishable, though the author does not explicitly say so. Based on the placement of the data, it may also be that the author intended them to contrast (otherwise, why put them together?[35]), but that is also unclear. Furthermore, with this implementation it should be impossible to distinguish a H/M/L contrast from a H/L/xL contrast phonetically, nor from a system which was merely H/L, with downstep (assuming Low could be downstepped). This makes the need to establish underlying categories all the more relevant.

This lack of clarity is due in part to the fact that there are only four sentences presented as evidence of downstep. This is a problem mostly because all are in

[35] It is of course possible that the author thought the prose already cited sufficiently clear to establish that 'thing' and 'who(m)?' had the same surface pitches. But in this case, placing them side by side would serve to point out that they have the same pitch, yet different underlying forms (hence different orthographies). I personally would consider this scenario interesting enough to merit a more explicit mention, so I remain unclear as to the author's intent.

the phonemic marking system as quoted above, so it is difficult to reconstruct the phonetics of any of them. Given the potential connection between a third underlying tone unit and register features (e.g., downstep), the lack of phonetic data on downstep makes it hard to see whether her three toneme analysis covers anything beyond words in isolation. This is particularly the case given the fact that the sentential data provided to support downstep do not include any instances of xL tone, making its complete distribution as a toneme suspect.

The result of the above is that I see no convincing evidence in Kutsch Lojenga (2006) of a third phonemic tone unit (toneme), beyond the one triplet already mentioned. This represents a crucial gap in the data presented to justify the use of a third (xL) toneme, which is in turn crucial to the argument of tonogenesis. In this dissertation, I will not be positing such an additional tone unit, but rather a third lexical tone melody. This lexically distinct tone melody is justified for Ndaka and Mbo on the basis of five different phonetic forms (i.e., infinitive forms and four conjugated forms) for each of the three lexical verb root classes by tone, in Chapters 3 and 4, respectively.

The tonogenesis claim from Kutsch Lojenga (2006) is as follows: "I therefore believe that the xL tone is new in the system, and has emerged as a result of a process in which depressor consonants once caused tonal depression, presumably lowered L tones, and apparently also H tones, to an xL level. Subsequently, a sound shift must have taken place, in which voiced egressive stops, the depressor consonants, became voiceless, and thus caused the xL level to become contrastive, which has resulted in the present-day three-tone system." (Kutsch Lojenga 2006:458)

That is, the proposed xL tone arose as a result of devoicing in a system which previously had depressor consonants. This story hangs on at least three pieces: the historical process of devoicing in Bila, the historical presence of depressors in Bila, and some mechanism by which that devoicing makes a depressor tone pattern contrastive. While this story is not implausible, there are significant gaps in the evidence provided in Kutsch Lojenga (2006) for the second and third pieces of this story. I will discuss each of these three pieces in turn.

The first piece of the claim of tonogenesis in Kutsch Lojenga (2006) is well established by the data in support of it, namely that there has been categorical loss of contrastive voicing in Bila. This can be seen synchronically in the consonant chart, which lacks voiced stops (in fact, it lacks any voiced segment which exists in opposition to a voiceless counterpart). The loss of contrastive voicing is also established historically through comparative data, which I reproduce in full in Figure 18:

Komo	Bhele	Bila	
b	b	p	
tɛ́bá	*tɛ́bá*	*tɛ́pà*	'to laugh'
ntábe	*ntábi*	*tápì*	'branch'
d	d	t	
demá	*demá*	*tìmá*	'to deceive'
didá	*didá*	*tìtò*	'to descend'
j	j	s	
jɔ́	*jɔ́*	*sɔ̀*	**'basket'**
jóngá	*jóngá*	*sùngá*	**'to boil (intr.)'**
g	g	k	
ágɔ́	*ágɔ́*	*ákɔ̀*	**'bat'**
gabá	*gabá*	*kàpà*	**'to divide'**
gb	gb	kp	
gbómá	*gbómá*	*kpùmá*	**'to bark'**
gbútu	*gbútu*	*kpǔtu*	**'stick, stool'**

Figure 18. Loss of voicing in Bila from Kutsch Lojenga (2006:456-7)

In each case, Bila shows voiceless consonants corresponding to voiced consonants in Komo [kmw] and Bhele [bhy][36]. Given that Bhele and Bila are both Bantu D31 languages, while Komo is Bantu D20 (Lewis, Simons, & Fennig 2015), these data argue that an ancestor of Bila, Komo, and Bhele (i.e., at least proto-Bantu D, if not also Proto-Bantu) had voiced consonants, from which Bila innovated a loss of voicing. There is also voiceless *t* across all three languages in Figure 18, in 'laugh' and 'branch (n)' and 'stick, stool (n)', showing Bila voiceless stops which were originally voiceless. There is therefore evidence in Figure 18 of both **d > t* and **t > t* for Bila, such that voiced and voiceless sources have merged into voiceless stops in Bila.

While devoicing has been well established, the impact this merger has had on the tone system of Bila remains to be seen, especially in relation to those of Komo and Bhele. The second piece of the tonogenesis claim in Kutsch Lojenga (2006) is the historical presence of depressors in Bila, which is not established in Kutsch Lojenga (2006). It seems evident that consonant-tone interaction (CTI) has been

[36] One of these, 'bark', has cognates in the Bantu D33 languages, in Table 34 of Section 6.3.

active in other languages of the area, as described in the body of this dissertation. But Komo and Bhele, which are presented as evidence of devoicing, do not have CTI (Kutsch Lojenga 2006:458). Evidence for the presence of CTI is limited to the mention of three languages with consonant-tone interaction (Mayogo, Adamawa-Ubangi; Budu, D33; Lika, D20), with a brief description of each system. However, no data or citations are presented to support the claim of CTI in those languages. It is difficult, therefore, to have any theoretical or analytical understanding of the mechanics by which those systems operate.

The third piece of the claim of tonogenesis in Kutsch Lojenga (2006) is some mechanism by which that devoicing makes a depressor tone pattern contrastive. The data presented in Figure 18 may be taken to show correspondences indicating reanalysis of voicing as tone, but such a reanalysis would be evidenced tenuously at best.

In terms of their tone, the examples for Bila in Figure 18 are presented as xL after voiceless segments which correspond with voiced segments in the other languages, which themselves precede either H tone (e.g., 'branch') or L tone (e.g., the first syllables of 'descend' and 'divide'). One exception to this generalization[37] is 'stick, stool' which has a xL-H tone in Bila corresponding with a H tone in the other languages. Granting for the sake of argument the validity of this data and the underlying categories they represent, the data in Figure 18 do not show a simple correspondence of voiceless to voiced obstruents. On the one hand, for each syllable with voiceless obstruents in Bila corresponding to voiced obstruents in Komo and Bhele, there is also xL tone in Bila corresponding with either H or L tone in Komo and Bhele. On the other hand, the original tone contrast (assuming this is inherited by Komo and Bhele) is neutralized in this context, given the presence of both *H > xL and *L > xL. The result of this is that the correspondence in Figure 18 is not simple, but rather complex, involving a change in voicing, as well as multiple changes in tone, which have not been otherwise accounted for.

The data in Figure 18, then, show Komo and Bhele voiced obstruents followed by H or L tones corresponding with Bila voiceless obstruents followed by xL tones. This indicates that Komo and Bhele represent systems **without** depressor consonant effects, while the Bila system has lost contrastive voicing, but is also **without**

[37] Consider also adding 'deceive', given at least a small amount of typographical ambiguity regarding the mark above the (italic) -ATR high front vowel. The word glossed 'deceive' and at least two other examples in the chapter ('leg' on p454, 'four' on p463) have similar marks which are smaller than other grave accents used throughout the chapter, including that for the (italic) -ATR high front vowel in the second word glossed 'squirrel, sp. (n)' on p455. But given that the (italic) small caps letter occurs without a dot, e.g., in 'dew (n)' on p454 and elsewhere in the chapter, it would seem likely that each of these marks is intended to be a grave accent.

depressor consonant effects. None of the data presented in Figure 18 represent, therefore, a depressor consonant system, either before or after devoicing.

What is missing in Figure 18 is data showing an intermediary state, such that the reader does not have to accept both devoicing and tone reanalysis in a single leap across a single set of forms in a cognate set. Such forms could show pitch which has remained stable despite the change in consonants, or else the change in tone without the devoicing. Perhaps this is intended to be shown in Figure 18, though as it explicitly shows phonemic categories, rather than pitch, it is hard to know how phonetic pitch (like systematic pitch depression due to voiced obstruents) would appear. As a result, there is no theoretical nor empirical mechanism in Kutsch Lojenga (2006) by which a syllable which had contrastively voiced onset with H (or else L) tone becomes a syllable with a voiceless onset and an xL tone. This represents another crucial gap in an argument for tonogenesis.

In this dissertation, I will provide evidence of such intermediary forms, to justify the mechanism proposed in my model. Chapters 3–6 show data which largely accord with a depressor consonant analysis, though they also include exceptional data which make such an analysis impossible. That these exceptional data are the result of voicing loss is established through cognate sets in Chapter 7, which includes words which have lost voicing both with and without depressor consonant effects.

Chapter 3
Ndaka [ndk]

This chapter presents Ndaka data collected during the course of my fieldwork. It is organized according to root class by tone, with the aim of characterizing those classes. This characterization will include the contrast between them, as well as evidence for the distribution of consonant types in each verb root class by tone.

The data in this chapter come from my own fieldwork notes and recordings, which were made in Nyanya, DRC, during a series of workshops to help the Ndaka and Mbo communities make advances in the development of their respective writing systems. The workshops were held December 1–15, 2006, May 20–June 6, 2014, and August 2–26, 2016. The workshops were held in a participatory manner (Kutsch Lojenga 1996), to engage the fullest participation of the community in the analysis of their language and development of their writing system. The participatory manner is described at length in (Kutsch Lojenga 1996), which includes citations of articles by Ursula Wiesemann, but it could be described briefly as involving as many speakers as possible in as much of the data collection, analysis, and discovery as possible. Decisions about how to go about the process are guided by what will help more people get involved, participate in, and own the analysis for themselves. Individual participants are listed in the acknowledgements on page vi. There are, unfortunately, a number of words which were sorted by infinitive form, but which became detached from their glosses. I have marked these glosses "no data" ("n.d.") until I can return to the language area to confirm them.

This chapter describes Ndaka verbs according to three observed root classes by tone, following the two questions laid out in Section 1.6 (i.e., the number and character of tone melodies, and the correlation between them). The first question (the number and character of tone patterns) will serve as a basis to organize the chapter, with one section for each tone pattern. Each section will provide a characterization of one verb root class by tone, beginning with the tone melody of the infinitive form, followed by the tone melodies of conjugated forms. In addition, each section after the first contains a subsection comparing its verb root class by tone with previously described verb root classes by tone. In this way, I hope to show both a clear unity for each lexical verb root class by tone across infinitive and conjugated forms, as well as a clear contrast between them.

The three Ndaka verb root classes by tone can each be identified by a set of tone melodies, as in Table 6. Recall from Section 2.4.2 that these data are presented in tone frames of the form <pronoun|verb|adverb>, both to provide a frame of reference for relative pitch, and to control for the four paradigmatic verb root class by tone combinations observed in these data.

<table>
<tr><th></th><th>inf</th><th>2s-past</th><th>2s-future</th><th>3s-past</th><th>3s-future</th></tr>
<tr><td>Low</td><td>[− − ¬]</td><td>[−|− ˏ ˎ|∨]</td><td>[−|_ _ ¯|⟍]</td><td>[ˊ|¯ ¯ −|˅]</td><td>[ˊ|¯ − ¯|⟍]</td></tr>
<tr><td>High</td><td>[− ¯ ⟍]</td><td>[−|− ˏ ˎ|∨]</td><td>[−|_ ¯ ¯|⟍]</td><td>[ˊ|¯ ¯ ¯|˅]</td><td>[ˊ|¯ ¯ ¯|⟍]</td></tr>
<tr><td>Rising</td><td>[− ⟋ ⟍]</td><td>[−|− ˏ ˎ|∨]</td><td>[−|ˎ _ ˏ|⟍]</td><td>[ˊ|¯ ¯ −|⌄]</td><td>[ˊ|¯ − −|⟍]</td></tr>
</table>

Table 6. Ndaka [ndk] infinitive and conjugated tone melodies according to verb root class by tone

The Low, High, and Rising verb root classes are named based on the second syllable of the infinitive form, rather than on any particular analysis or hypothesis of their underlying forms. These Low, High, and Rising verb root classes will be described in further detail in Sections 3.1, 3.2, and 3.3, respectively. These three Ndaka verb root classes by tone can be seen in the tone melodies across the columns of Table 6, which will be justified by row, one in each root class section. While the contrast between the three verb root classes by tone by root is neutralized for 2s-past forms, the distinction is maintained for 3s-past forms, as well as in the infinitive and both future forms.

In addition to the identity of each tonal root class and the contrast between them, each section will summarize the consonants that are found in verbs of that verb root class by tone, and how they correspond to expectations based on Bradshaw (1999). The implications of consonant distributions across verb root classes by tone will be addressed in Chapter 7.

These comparisons culminate in Section 3.4, which contains a summary comparison of all Ndaka verb tone patterns, showing visual comparisons in composite figures.

After laying out the basic distinctions between the three Ndaka verb root classes by tone in Sections 3.1–3.4, Section 3.5 provides an initial analysis that accounts for generalizations that have been made up to that point, including a proposal for the lexical specification of morphemes and rules and assumptions needed to derive surface forms.

Additionally, the basic analysis in this chapter provides a model and starting point for comparison with other Bantu D30 languages, some of which will be addressed in following chapters.

3.1 Ndaka [ndk] Low Verbs

The infinitive melody for Low verbs is low and level, as exemplified by the word *kɔkpata* [– – –] 'follow' in Figure 19:

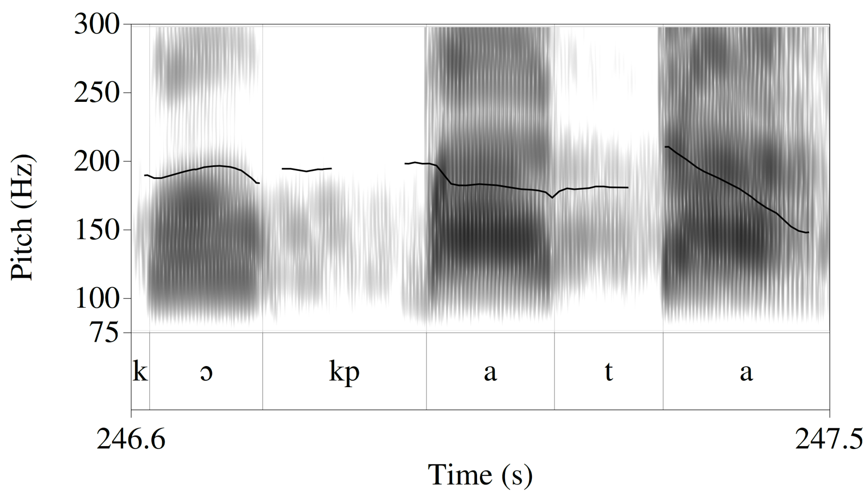


Figure 19. Ndaka Low [kɔkpata] [– – –] 'follow' spectrogram segmented with pitch trace superimposed showing low and level pitches

As mentioned in Section 2.1.1, there is a lot of information in a spectrogram and pitch trace, but two points are most relevant to my analysis. The first is that the first root syllable (the first [a]) is pronounced at essentially the same pitch as the prefix syllable ([ɔ]). The difference may not even be perceptible, and in any case should be understood as declination, the lowering of pitch over time for purely physiological reasons (c.f. Section 2.1.4).

The second point is the sharper drop in pitch on the final syllable. This final drop is notable for two reasons. First, it drops a further (∼70hz) than the (∼10hz) drop between the first two syllables. And secondly, the drop is spread over the entire length of the vowel, indicating that there is a tonal pitch target at each end of the vowel, and what is realized is a transition from the first to the second. It is possible that there are two lexical tones attached to this final syllable, though it has been observed that phrase final low tones generally have a falling phonetic implementation (Snider 2014), so it would be abnormal to see this final syllable pitch completely level. Some variety of this final fall will be seen on all Ndaka infinitive tone patterns.

Other details in Figure 19 are not as important for the analysis of tone. For instance, the drop in pitch in the transition from [kp] to [a] is typical of voiceless onsets transitioning to a following vowel. This pitch drop is essentially imperceptible, despite the fact that it it is visually rather distinctive in the pitch trace. There is also a pitch trace line within the area segmented as [kp]. This pitch trace line is left over from the transition from [ɔ] to [kp]; voicing analyzed by Praat as continuing after a vowel is not uncommon in my data set. But what is very clear when comparing the pitch trace to the spectrogram, is that the section segmented as [kp] has only trace amounts of periodic activity (and nothing in the spectrogram F_0), justifying marking the area as belonging to the voiceless consonant.

So the pitch trace in Figure 19 contains a lot of information, some of which is more relevant than others. One may isolate the important information by studying the detail more precisely, as is done above, but there are other means to highlight important generalizations from pitch trace data. Another important method for highlighting relevant information, while backgrounding what is not, is to overlay multiple pitch traces from words judged by native speakers to have the same tone pattern (c.f. Section 2.1.1 and appendix B for discussion of and technical information on these figures). An example of such a pitch trace overlay, with 52 tokens of 40 different Low verbs, is given in Figure 20:

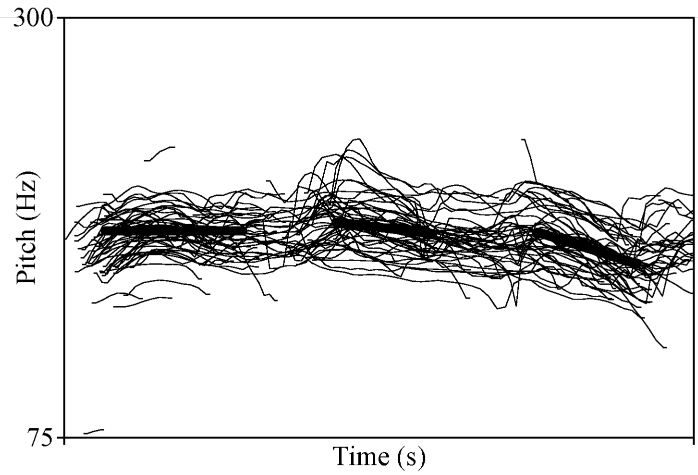


Figure 20. Overlay of infinitive pitch traces for Ndaka [ndk] Low verbs

In the overlay in Figure 20, one may see that despite lots of individual variation from word to word, there is a clear pattern of pitch moving at one level across the word. As indicated in the thicker trend bars, the second syllable pitch tends lower, and the third syllable pitch tends lower even more so. That is, the effects mentioned above which are due to the phonetic implementation relevant to only one or a few words is obscured, while the F_0 trends across the majority of words (i.e., declination and the final falling tone) come clearer into view. As a result, the pitch trace overlay in Figure 20 can be transcribed in a symbolic pitch trace as [⁻ ‒ ⌝].

I applied this methodology of overlaying individual pitch traces to show generalizations for conjugated forms (c.f. tone frames described in Section 2.4.2) to Low roots, resulting in Figure 21 and Figure 22:

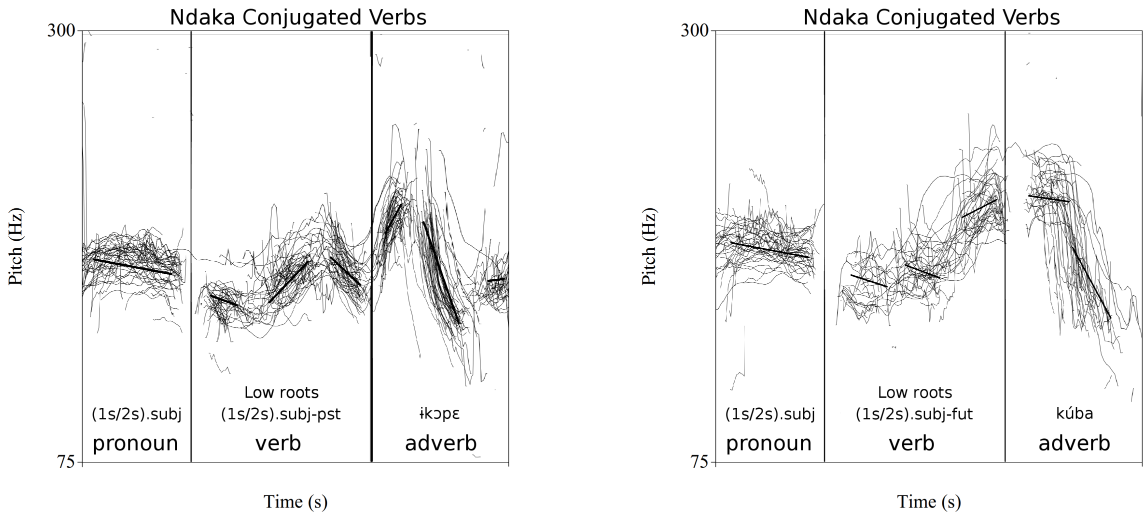


Figure 21. 2s past and future pitch traces for Low verbs in Ndaka [ndk]

The forms in Figure 21 are preceded by a (2s) low pronoun, and the verbs themselves also begin low. The past form (on the left) is the basic low and level tone pattern, which rises at the end of the penultimate TBU of the verb, falling again on the final verb TBU. The initial TBU of the following adverb is higher than any of the preceding TBU's, showing it to be high in comparison to preceding lows. But it is not as if those lows were all the same; there appears to be a register effect between the pronoun and the verb form, such that the lows on the verb are pronounced on a lower register than the preceding free pronoun low tones.

The future form (on the right) ends with a high TBU, which is at the same pitch level as the initial TBU of the following adverb. As a result, the pitch trace overlays in Figure 21 can be transcribed in symbolic pitch traces as [−|– ⸝ ⸜|V] and [−|_ _ ¯|⸌].

Unlike the forms in Figure 21, the forms in Figure 22 are preceded by a free pronoun with a (3s) low-high tone pattern, and the verb forms also begin with a high TBU:

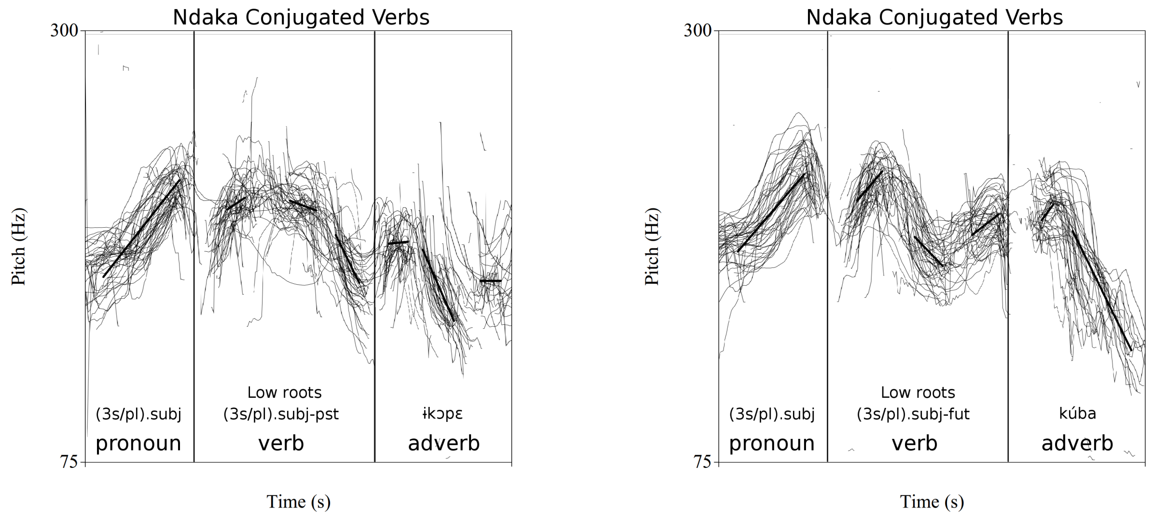


Figure 22. 3s past and future pitch traces for Low verbs in Ndaka [ndk]

In the past form (on the left) the high pitch on the initial TBU spreads to the second TBU. The initial high TBU does not spread in the future form (on the right). Like the 2s forms, the future form has a final high TBU at the same level as the following adverb initial high pitch (each of which is lower than the verb initial high pitch, c.f., Section 2.1.4). In each case, each low tone lowers the register (automatic downstep), such that high tones are lower after a low than before it. As a result, the pitch trace overlays in Figure 22 can be transcribed in symbolic pitch traces as [′|¯ ¯ –|⌄] and [′|¯ − ¯|∖].

The Low verb class tone melodies can be summarized as in Table 7, justifying the first row of Table 6 on page 66:

<table>
<tr><th>inf</th><th>2s-past</th><th>2s-future</th><th>3s-past</th><th>3s-future</th></tr>
<tr><td>[− − ¬]</td><td>[−|– ⸝ ⸜|∨]</td><td>[−|_ _ ¯|∖]</td><td>[′|¯ ¯ –|⌄]</td><td>[′|¯ − ¯|∖]</td></tr>
</table>

Table 7. Ndaka [ndk] Low verb infinitive and conjugated tone melodies

The following verbs are confirmed Low tone on the basis of their conjugations:

(12) Ndaka [ndk] Low verbs

a.	koɗijo	[– – ˥] oil (self)	
b.	kofuo	[– – ˥] fart	
c.	kokpeɲo	[– – ˥] chase	
d.	kɔkpɑkpɑ	[– – ˥] be blocked	
e.	kɔkpɑtɑ	[– – ˥] follow	
f.	kɔkɛkɑ	[– – ˥] crow (rooster)	
g.	kokuto	[– – ˥] close	
h.	kɔkʉɑ	[– – ˥] buy	
i.	kɔpɔwɑ	[– – ˥] spoil	
j.	kosiɓo	[– – ˥] wait	
k.	kɔsɨkɑ	[– – ˥] take	
l.	kɔsʉwɑ	[– – ˥] sing	
m.	komijo	[– – ˥] swallow	
n.	kɔmɔmɑ	[– – ˥] gather	
o.	kɔnɔtɑ	[– – ˥] give birth	
p.	kɔɲɛkɑ	[– – ˥] smile	
q.	kokwejo	[– – ˥] wrap (package)	
r.	kɔkwɑːnɑ	[– – ˥] cohabitate	
s.	kɔkwɑkɑ	[– – ˥] cut	chop
t.	kɔkwɛtɑ	[– – ˥] bite	
u.	kyɛndɑ	[– – ˥] walk	

The following other verbs are taken to be Low on the basis of their infinitive tone melodies, though they do not appear in conjugated tone data (glosses are from database where unambiguous):

(13) Ndaka [ndk] Low verbs by infinitive form only

a.	koɗeto	[– – ˥] ascend	
b.	kɔkɑŋgɑ	[– – ˥] knead	stare
c.	kɔkɨjɑ	[– – ˥] act	
d.	kɔkɔmɑ	[– – ˥] bewitch	
e.	kokuɗo	[– – ˥] crawl (lizard)	
f.	kokumo	[– – ˥] flee	
g.	kɔkʉwɑ	[– – ˥] buy	
h.	kɔkpɑŋgɑ	[– – ˥] n.d.	
i.	kɔpɛtɑ	[– – ˥] deceive	
j.	kopingo	[– – ˥] wring out	

k.	kopipo	[− − ¬] finish	
l.	kɔpɨŋgɑ	[− − ¬] peel	
m.	kopuno	[− − ¬] stir up	
n.	kɔsɑŋgɑ	[− − ¬] construct	
o.	kɔsɨjɑ	[− − ¬] hire	
p.	kɔmɑnɑ	[− − ¬] n.d.	
q.	kɔmɑtɑ	[− − ¬] increase	
r.	kɔmɛjɑ	[− − ¬] n.d.	
s.	kɔmɔtɑ	[− − ¬] n.d.	
t.	konejo	[− − ¬] rain	
u.	kɔnʉwɑ	[− − ¬] laugh	
v.	kowejo	[− − ¬] vomit	
w.	kɔwɛjɑ	[− − ¬] thatch	
x.	kowewɑ	[− − ¬] blow away	
y.	kɔkwɛtɑ	[− − ¬] bite	sting
z.	kɔkwɨjɑ	[− − ¬] n.d.	
aa.	kokwejo	[− − ¬] pack	
ab.	kombombo	[− − ¬] choose	
ac.	komvɨtɑ	[− − ¬] n.d.	

There are no C1 consonants specified for voicing in the Low verbs. The consonants found in C1 position of the verbs in (12) and (13) are all implosives (12a; 13a), voiceless obstruents (12b-l; 13b-o), nasals (12m-p; 13p-u), other sonorants (13v-x), or obstruents modified by sonorants (12q-u; 13y-ac). As none of the consonants in C1 position are specified for voice (c.f. Section 2.2.4) in any of the verbs in the Low verb root class, the Low verb root class fits well as a verb root class by tone not associated with depressor consonants under Bradshaw (1999).

3.2 Ndaka [ndk] High Verbs

In the High verb root class, the infinitive form is appreciably higher on the root mora, both in comparison to the root mora of Low infinitive forms, and in comparison to the surrounding morae in the High infinitive forms. The final syllable falls from a level higher than the initial low prefix syllable, as seen in Section 3.1 for Low roots. The High verb root class is exemplified by [kotito] [– ¯ ⟍] 'push' in Figure 23:

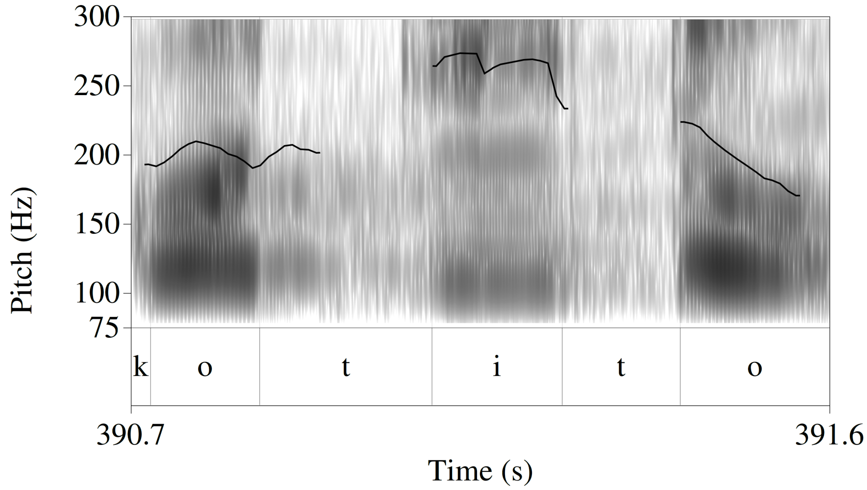


Figure 23. Ndaka High [kotito] [– ¯ ↘] ‘push’ spectrogram segmented with pitch trace superimposed showing high tone on the root mora

As in Section 3.1, one must sort through a number of details in Figure 23 to see what is relevant. For instance, the prefix tone is around 200hz again, but here the root syllable pitch is over 250hz, though still basically at one level. The pitch over the final vowel is again falling, from much the same place as seen with Low verbs in Figure 19, starting between the low of the prefix syllable and the high of the first root syllable.

Other details are phonetically imperceptible, or phonologically unimportant. For instance, we see the continuation of the pitch trace into the first root consonant [t], as was observed in Figure 19. Furthermore, the first root syllable shows two peaks, with a trough at about the midpoint of the vowel [i]. Despite those local maximums and minimum, however, the pitch trace is clearly more level there than on the final syllable, and clearly more high than either of the other syllables.

The High verb root class infinitive data can be again generalized through the overlay of multiple pitch traces from words judged to have the same tone. This generalization through superimposition is shown in Figure 24, an overlay of [kotito] [– ¯ ↘] ‘push’ in a total of 42 tokens of 32 different high verbs:

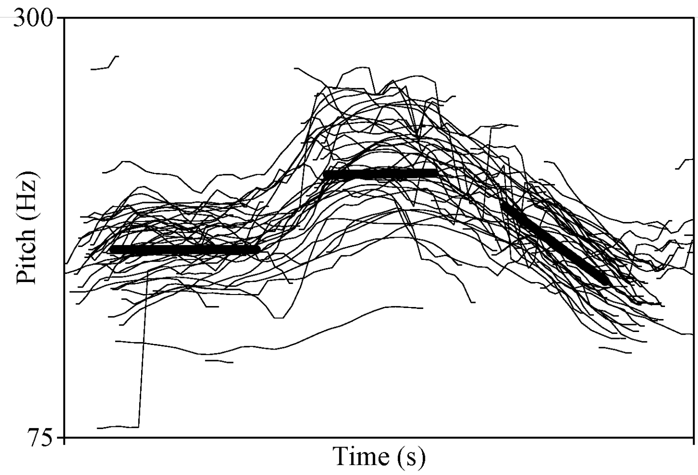


Figure 24. Overlay of infinitive pitch traces for Ndaka [ndk] High verbs

Here again there is a lot of variation, but there remain consistent trends, as expressed in the thicker lines. The root syllable is basically level, but significantly higher than the prefix syllable. The final syllable starts between the high and low levels, and moves down to a level lower than the prefix low, in a straight line. As a result, the pitch trace overlay in Figure 20 can be transcribed in a symbolic pitch trace as [– ¯ ⍀].

Conjugated forms for High verbs begin again with those with a low (2s) subject pronoun in Figure 25:

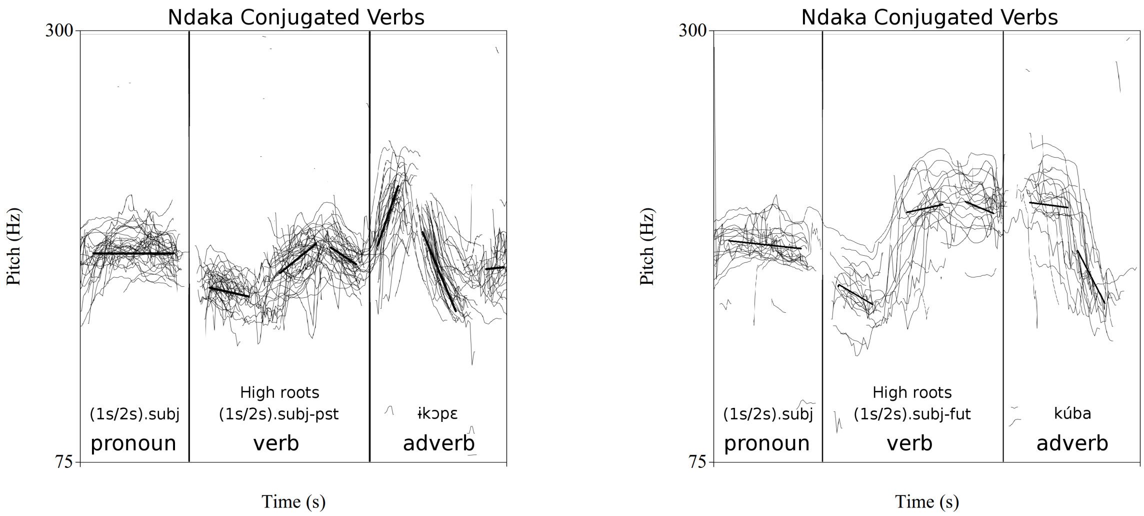


Figure 25. 2s past and future pitch traces for High verbs in Ndaka [ndk]

The past form (on the left) is the same basic low and level tone pattern seen for Low verbs in Figure 21, rising at the end of the penultimate TBU of the verb, falling again on the final verb TBU. Again, the initial TBU of the following adverb is higher than any of the preceding TBU's, and again the lows on the verb are pronounced on a lower register than the free pronoun low tones.

The future form (on the right) again ends with a high TBU, which is at the same pitch level as the initial TBU of the following adverb, as was the case for Low 2s-past verbs. But for High verbs, the final high TBU follows a high root TBU; this fact can be used to distinguish them from Low verbs, where the final high TBU follows a low TBU in 2s-future forms, in Figure 21.

As a result, the pitch trace overlays in Figure 25 can be transcribed in symbolic pitch traces as [−|_ ⸝ ⸜|\/] and [−|_ ¯ ¯|⸌].

Moving from low (2s) free pronoun forms in Figure 25 to the low-high (3s) free pronoun forms in Figure 26, we see again an initial high TBU on the verb:

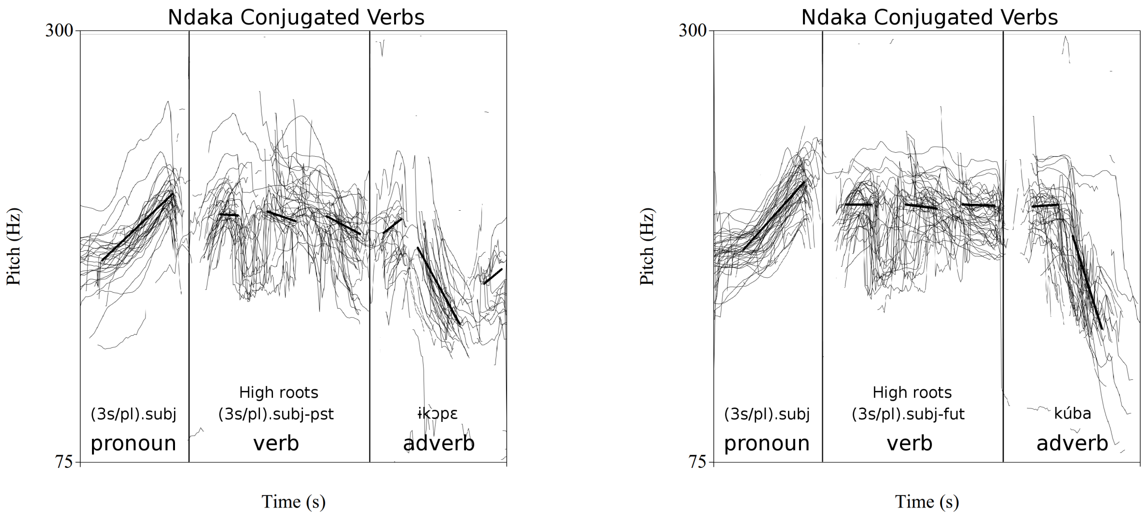


Figure 26. 3s past and future pitch traces for High verbs in Ndaka [ndk]

The past form (on the left) is high and level across the verb, at a level lower than the previous free pronoun ended, and at the level of the initial high of the following adverb. One detail which is potentially unimportant, is that the high tone of the verb drops somewhat on the final syllable. The future form (on the right) has the same form, except without that drop. However, native speakers indicated that these forms were homophonous, so that drop may be inconsequential.

The pitch trace overlays in Figure 26 can thus be transcribed in symbolic pitch traces as [ˊ|¯ ¯ ¯|⌄] and [ˊ|¯ ¯ ¯|⟍]. This results in the High verb class tone melodies in Table 8, justifying the second row of Table 6 on page 66:

inf	2s-past	2s-future	3s-past	3s-future
[– ¯ ⟍]	[¯|– ⸝ ⸜|\/]	[¯|_ ¯ ¯|⟍]	[ˊ|¯ ¯ ¯|⌄]	[ˊ|¯ ¯ ¯|⟍]

Table 8. Ndaka [ndk] High verb infinitive and conjugated tone melodies

The following verbs are confirmed High tone on the basis of their conjugations:

(14) Ndaka [ndk] High verbs
- a. kɔɓɛndɑ [– ¯ ⟍] hit
- b. kɔɓɛtɑ [– ¯ ⟍] shell (peanuts)
- c. koɓejo [– ¯ ⟍] protect
- d. koɓongo [– ¯ ⟍] wrap around

e. koɓoo [– ¯ ⟍] push
f. koɓundo [– ¯ ⟍] break
g. koɗijo [– ¯ ⟍] lick
h. kotito [– ¯ ⟍] push
i. kɔtɨnɑ [– ¯ ⟍] cut   harvest
j. kotoko [– ¯ ⟍] spit
k. kotoo [– ¯ ⟍] urinate
l. kotuo [– ¯ ⟍] forge
m. kɔtʉwɑ [– ¯ ⟍] return
n. kɔtɑmbɑ[1] [– ¯ ⟍] play
o. kɔkʉndɑ [– ¯ ⟍] love
p. komijo [– ¯ ⟍] make pottery   press oil[2]
q. kolindo [– ¯ ⟍] avoid
r. kɔjɑnɑ [– ¯ ⟍] be tired
s. komvono [– ¯ ⟍] smell
t. kɔkwɑnɑ [– ¯ ⟍] exchange

The following verbs are also taken to be in the High class, based on infinitive forms:

(15) Ndaka [ndk] High verbs by infinitive form only
a. koɗuko [– ¯ ⟍] be confused
b. kokuɓo [– ¯ ⟍] wither   be stunted
c. kopoɓo [– ¯ ⟍] read
d. kotiso [– ¯ ⟍] return
e. kɔtʉkɑ [– ¯ ⟍] insult
f. kosigo [– ¯ ⟍] talk with someone
g. kosoko [– ¯ ⟍] call
h. kɔfʉɓɑ [– ¯ ⟍] swell
i. kofoto [– ¯ ⟍] hold
j. komiso [– ¯ ⟍] dry out
k. kɛnɨdʲɑ [– ¯ ⟍] n.d.
l. kɛnɨɲɑ [– ¯ ⟍] lay (eggs)
m. kɛnɨkwɑ [– ¯ ⟍] fall

[1] I also have this form with irregular (or [nlj]/[vau]) *ka-*

[2] These two glosses were initially treated as distinct words, yet I have not found a formal distinction so far, so they may turn out to be either homophones or multiple senses of the same word (which is not too far fetched, given that oil may be pressed from palm nuts with the hands, not unlike the pressing of clay).

n. kɔwɑkɑ [– ¯ ↘] n.d.
o. kɔwɛtɑ [– ¯ ↘] sharpen
p. kowiso [– ¯ ↘] fill
q. kɔwɨjɑ [– ¯ ↘] speak
r. koliŋgo [– ¯ ↘] n.d.
s. kojoko [– ¯ ↘] be silent
t. kɔmvɔmvɑ [– ¯ ↘] n.d.
u. kondʒono [– ¯ ↘] n.d.

As with the Low verbs presented in Section 3.1, there are no C1 consonants specified for voicing in the High verbs. The consonants found in C1 position of the verbs in (14) and (15) are all implosives (14a-g; 15a), voiceless obstruents (14h-o; 15b-i), nasals (14p; 15j-m), other sonorants (14q-r; 15n-s), or obstruents modified by sonorants (14s-t; 15t-u). As none of the C1 consonants in (14) and (15) are specified for voice (c.f. Section 2.2.4), the High verb root class also fits well as a verb root class by tone not associated with depressor consonants under Bradshaw (1999).

3.2.1 Comparing High and Low Verbs

Having given tone melodies characteristic of the High and Low verb root classes, I now turn to comparing them, to show how they are distinct one from the other. The trendlines from Low verbs in Figure 21 and from High verbs in Figure 25 are extracted and superimposed in Figure 27:

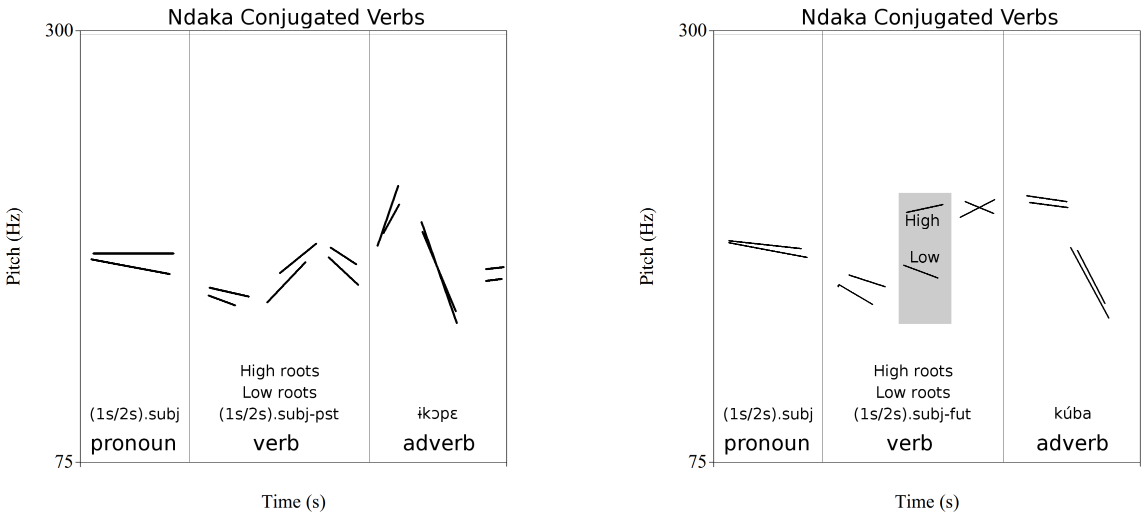


Figure 27. 2s past and future trendlines for High and Low verbs in Ndaka [ndk]

The past forms (on the left) are essentially the same for Low and High verbs, but the future forms (on the right) differ on the root mora (the second syllable of the verb), where High verbs are higher than Low verbs.

The trendlines from Low verbs in Figure 22 and from High verbs in Figure 26 are compared in Figure 28:

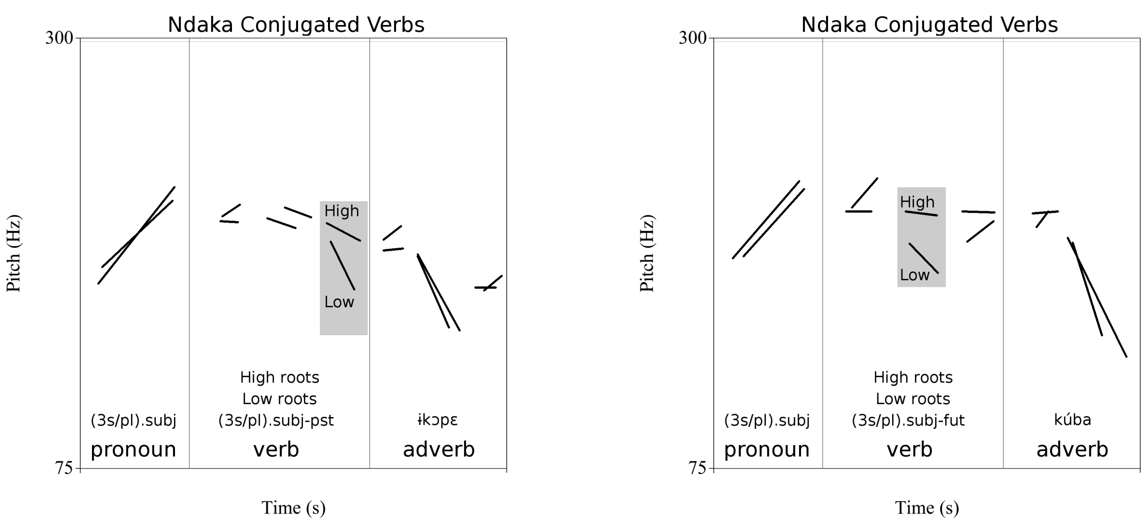


Figure 28. 3s past and future trendlines for High and Low verbs in Ndaka [ndk]

The past forms (on the left) are the same except on the third, final vowel (FV) syllable, where High roots remain high, and Low roots fall to a lower level. The future forms (on the right) differ on the root syllable, with Low roots lower than High roots. The final syllable appears different, but that difference may just amount to a return to the same pitch target from the lower second syllable (for Low roots), as opposed to maintaining a higher level (for high roots).

One generalization that can already be drawn from the data in Figure 27 and Figure 28 is that the contrast between High and Low verbs is expressed on only one TBU, wherever there is contrast. Further, that contrast shows High verbs high on that TBU, where Low verbs are low. Another generalization which begins to emerge here (and which will be confirmed with data from the other verb root classes by tone) is that the TBU of contrast between High and Low verbs on past forms is on the FV (where there is contrast at all), whereas on future forms that TBU is on the root mora (or second verb syllable).

3.3 Ndaka [ndk] Rising Verbs

The tone patterns in Sections 3.1 and 3.2 are observed on most verbs where the root consonants are voiceless obstruents, implosives, sonorants, or obstruents modified by sonorants. That is, the consonants that Bradshaw (1999) predicts would not be depressor consonants. But where voiced obstruents are present in the first root syllable, the following Rising infinitive tone pattern is observed:

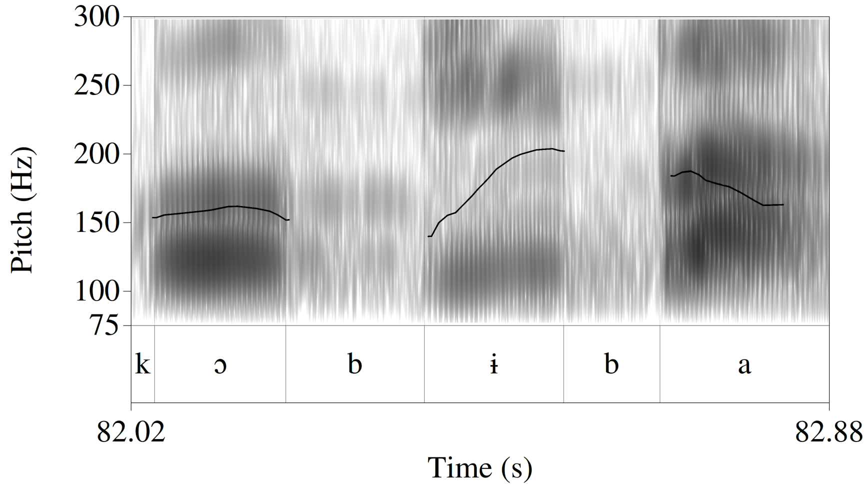


Figure 29. Ndaka Rising [kɔbɨba] [– ⟋ ⟍] ‘respect, admire’ spectrogram segmented with pitch trace superimposed showing rising the root mora

The root mora here has a rising pitch, a low to high transition. The final syllable has final falling pitch again, from an initial level between high and low, as defined by the two previous syllables. The Rising verb root class can be generalized in Figure 30 with [kɔbɨba] [– ⟋ ⟍] ‘respect, admire’ among 37 tokens of 26 different Rising verbs.

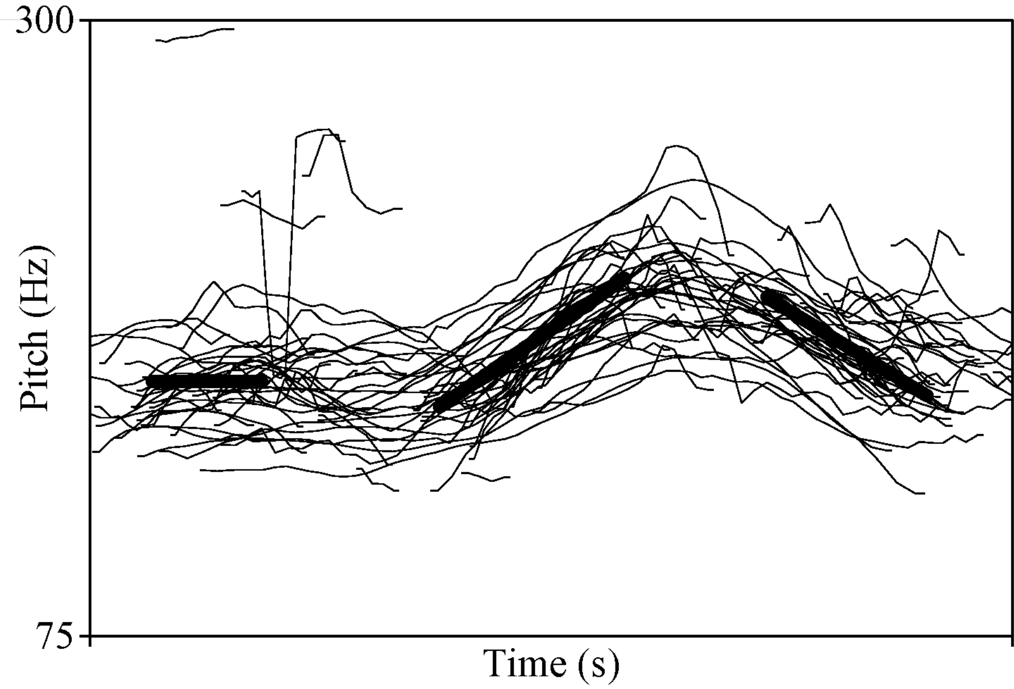


Figure 30. Overlay of infinitive pitch traces for Ndaka [ndk] Rising verbs

As a result, the pitch trace overlay in Figure 20 can be transcribed in a symbolic pitch trace as [– ⟋ ⟍].

Figures 31 and 32 show conjugated forms for Rising verbs, first with those with a low (2s) subject pronoun:

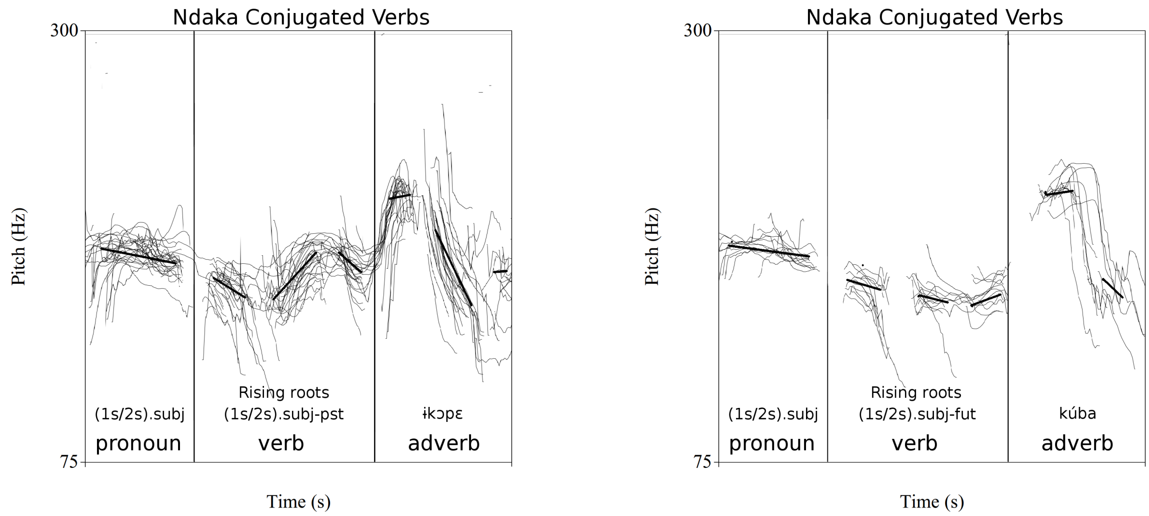


Figure 31. 2s past and future pitch traces for Rising verbs in Ndaka [ndk]

The past form (on the left) is the same basic low and level tone pattern seen for Low verbs in Figure 21 and for High verbs in Figure 25, rising at the end of the penultimate TBU of the verb, falling again on the final verb TBU. Again, the initial TBU of the following adverb is higher than any of the preceding TBU's, and again the lows on the verb are pronounced on a lower register than the free pronoun low tones.

The future form for rising verbs (on the right) does not end with a high TBU, but remains low before the initial high TBU of the following adverb. The Rising verb class, then, can be distinguished from both Low and High verbs by its final low TBU in 2s-future forms.

As a result, the pitch trace overlays in Figure 31 can be transcribed in idealized pitch traces as [−|– ⸝ ⸜|⋁] and [−|– – –|⌝\].

Moving from forms with low (2s) free pronouns in Figure 31 to those with low-high (3s) free pronouns in Figure 32, we see again an initial high TBU on the verb:

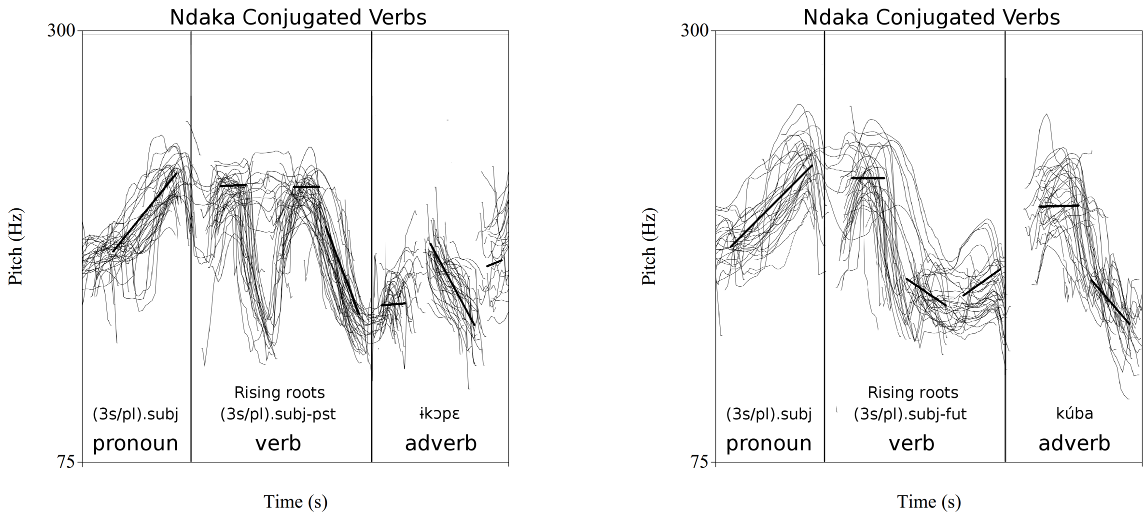


Figure 32. 3s past and future pitch traces for Rising verbs in Ndaka [ndk]

Here the past form (on the left) is high, then low, as for 3s High verb forms, but the following adverb begins low here, whereas it began high for High verbs. The future form (on the right) again ends low, unlike 3s forms for Low or High verbs.

As a result, the pitch trace overlays in Figure 32 can be transcribed in idealized pitch traces as [/|¯ ¯ –|◡] and [/|¯ – –|∖].

The Rising verb class tone melodies can be summarized as in Table 9, justifying the third row of data in Table 6 on page 66:

inf	2s-past	2s-future	3s-past	3s-future
[– ⁄ –]	[–|– ⁄ ∖|∨]	[–|∖ _ ⁄|∖]	[/|¯ ¯ –|◡]	[/|¯ – –|∖]

Table 9. Ndaka [ndk] Rising verb infinitive and conjugated tone melodies

The following verbs are confirmed Rising tone on the basis of their conjugations:

(16) Ndaka [ndk] Rising verbs (C1 specified for voice)
  a. kogiɲo [– ⁄ –] refuse
  b. kogiso [– ⁄ –] throw away
  c. koguɓo [– ⁄ –] bend
  d. kogwejo [– ⁄ –] gather
  e. kɔgbɔkɑ [– ⁄ –] find

f. kɔgbɔmɑ [– ⁄ –] bark
g. kɔgɔwɑ [– ⁄ –] snore
h. kɔbɛtɑ [– ⁄ –] open (eye)
i. kɔbɨjɑ [– ⁄ –] wipe
j. kɔdʉwɑ [– ⁄ –] pound (pestle)

The following verbs are also taken to be in the Rising class, based on infinitive forms:

(17) Ndaka [ndk] Rising verbs by infinitive form only (C1 specified for voice)
a. kɔbɨbɑ [– ⁄ –] respect admire
b. kɔdɨkpɑ [– ⁄ –] close
c. kɔdɔɓɑ [– ⁄ –] harvest honey be impotent
d. kodʒɑŋgɑ [– ⁄ –] fart fail lose impede
e. kodʒiso [– ⁄ –] n.d.
f. kɔgɑmɑ [– ⁄ –] borrow
g. kɔgbɑnɑ [– ⁄ –] be timid

Unlike the Low forms in (12-13) and the High forms in (14-15), the Rising forms in (16-17) all contain C1 consonants which are obstruents specified for voice. In fact, all Ndaka verbs with C1 consonants which are specified for voice are in the Rising verb root class. This fact is relevant because the Rising verb root class therefore is where Bradshaw (1999) would expect to have consonant-tone interaction. Together, then, the data so far presented in Sections 3.2, 3.1, and 3.3 align nicely with Bradshaw (1999)'s prediction that consonant-tone interaction should be found with segments which are specified for voice, assuming that the Rising tone group contrasts with the High and Low groups on the basis of consonant-tone interaction. Under such an analysis, Ndaka has a lexical high versus low contrast, with consonant-tone interaction creating a third surface contrast: Rising verbs.

Unfortunately for the somewhat simplistic analysis just mentioned, Ndaka's Rising tone group also contains other verbs, which do not have contrastively voiced consonants in the C1 position:

(18) Ndaka [ndk] Rising verbs (C1 NOT specified for voice)
a. koɗuwo [– ⁄ –] pull (up)
b. kɔkɔwɑ [– ⁄ –] cough
c. kɔkpɑgɑ [– ⁄ –] ferment
d. kɔpɑnɑ [– ⁄ –] show
e. kɔpɔpɑ [– ⁄ –] winnow flutter

f. kɔnɑgbɑ [– ⸝ –] have headache
g. kɔjɑnɑ [– ⸝ –] cook
h. koŋguso [– ⸝ –] remove draw out leave take away
i. kouo [– ⸝ –] blow (air)
j. kɔʉnɑ [– ⸝ –] sow (seed)

The following verbs are also taken to be in the Rising class, based on infinitive forms:

(19) Ndaka [ndk] Rising verbs by infinitive form only (also with C1 NOT specified for voice)
a. kɔkɛɗɑ [– ⸝ –] pass overtake
b. kɔkɔfɑ [– ⸝ –] cough
c. kotigbɑ [– ⸝ –] shake
d. kɔmɨjɑ [– ⸝ –] n.d.
e. keniso [– ⸝ –] n.d.
f. kowɑso [– ⸝ –] look for[3]
g. kowiso [– ⸝ –] n.d.
h. kowoko [– ⸝ –] migrate
i. kɔwʉmɑ [– ⸝ –] n.d.
j. kojono [– ⸝ –] take away
k. kɔŋgwɑnɑ [– ⸝ –] leave exit

The consonants found in C1 position of the verbs in (18-19) are all implosives (18a), voiceless obstruents (18b-e; 19a-c), nasals (18f; 19d-e), other sonorants (18g; 19f-j), or obstruents modified by sonorants without voicing opposition (18h; 19k). There are even two words missing a C1 (18i-j) altogether. Put together with the data in (16-17), Rising verbs have both C1 consonants which are specified for voicing and C1 consonants which are not specified for voicing.

In summary, Rising verbs (alone) include C1 segments which are specified for voice, arguing in favor of the Rising verb root class being a result of consonant-tone interaction (Bradshaw 1999). But Ndaka Rising verbs also include a number of verbs with C1 segments which are not contrastive for voice (c.f. Section 2.2.4), causing something of a problem for a simple [L/Voice] analysis based on Bradshaw (1999). The impact of the data in (18) and (19) will be discussed more fully in Chapter 7.

[3] This is apparently another example of +ATR harmony for [a], unless this root is underlyingly /uas/, with the +ATR harmony coming from the /u/.

3.3.1 Comparison of High and Rising Verbs

Having given tone melodies characteristic of the High and Rising verb root classes, I now turn to comparing them, to show how they are distinct one from the other. The trendlines from High verbs in Figure 25 and from Rising verbs in Figure 31 are extracted and superimposed in Figure 33:

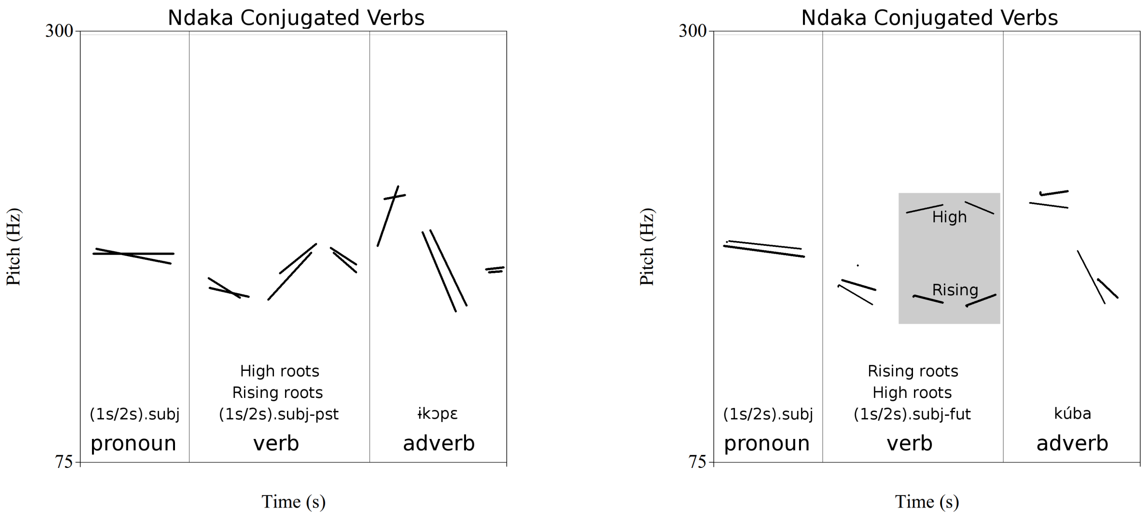


Figure 33. 2s past and future trendlines for High and Rising verbs in Ndaka [ndk]

The past forms (on the left) are the same for High and Rising roots. Together with the fact that High and Low roots are the same in this form, this points to a neutralization across the three root classes, for this conjugated form. The future forms (on the right) are high on the last two syllables for High verbs, whereas they are low on the last two syllables for Rising verbs. This is consistent with either spreading of a single root tone, or the presence of a second root tone. But in any case, the contrast between High and Rising roots here is expressed over two syllables, unlike in infinitive forms (where High, Low, and Rising contrast on the root mora only), and unlike the 2s-future contrast for High v Low, which is also expressed on only one mora (on the root).

The trendlines from High verbs in Figure 26 and from Rising verbs in Figure 32 are compared in Figure 34:

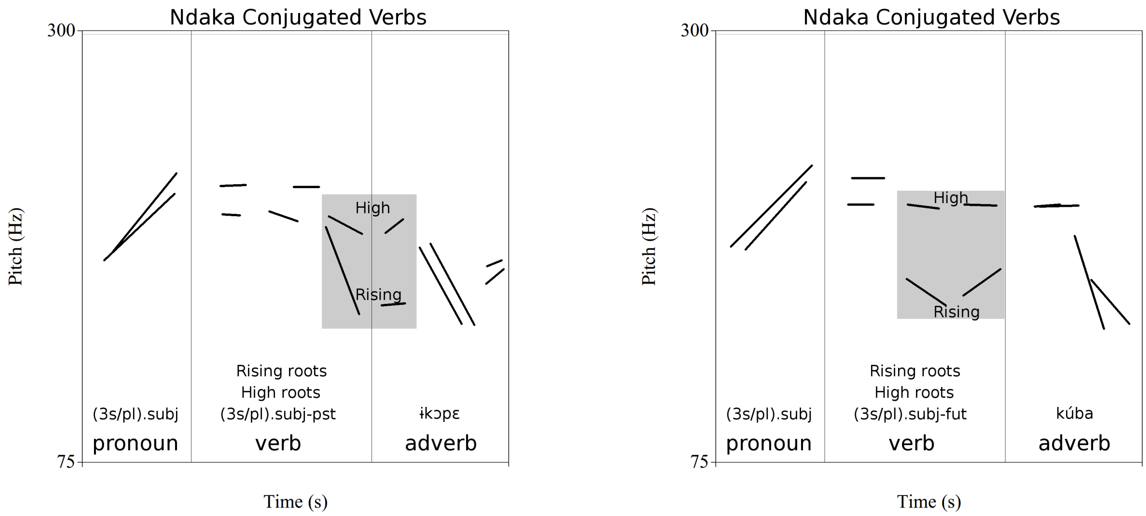


Figure 34. 3s past and future trendlines for High and Rising verbs in Ndaka [ndk]

The past forms (on the left) mostly differ over two syllables again, beginning with the syllable of the final vowel (FV). That is, they differ on the third syllable of the root and the first syllable of the following adverb. High verbs are high (though perhaps lower than previous syllables) on those two syllables, where Rising verbs are low. The future forms (on the right) bear a similar difference, except that the two syllable contrast begins on the root mora, rather than on the FV.

The contrasts in Figure 33 and Figure 34 show a two-TBU contrast between High and Rising verbs, where any contrast exists at all, with High verbs high over those two TBU's, and Rising verbs low over those two TBU's. These data also confirm the generalization mentioned in Section 3.2.1, that contrasts between roots on past forms begin on the FV, whereas contrasts between roots on future forms begin on the root. These two generalizations are in addition to the generalization that root tone contrasts are neutralized for 2s-past forms, as mentioned above.

3.3.2 Comparison of Low and Rising Verbs

This section compares the Low and Rising verb root classes, to show how they are distinct one from the other. The trendlines from Low verbs in Figure 21 and from Rising verbs in Figure 31 are extracted and superimposed in Figure 35:

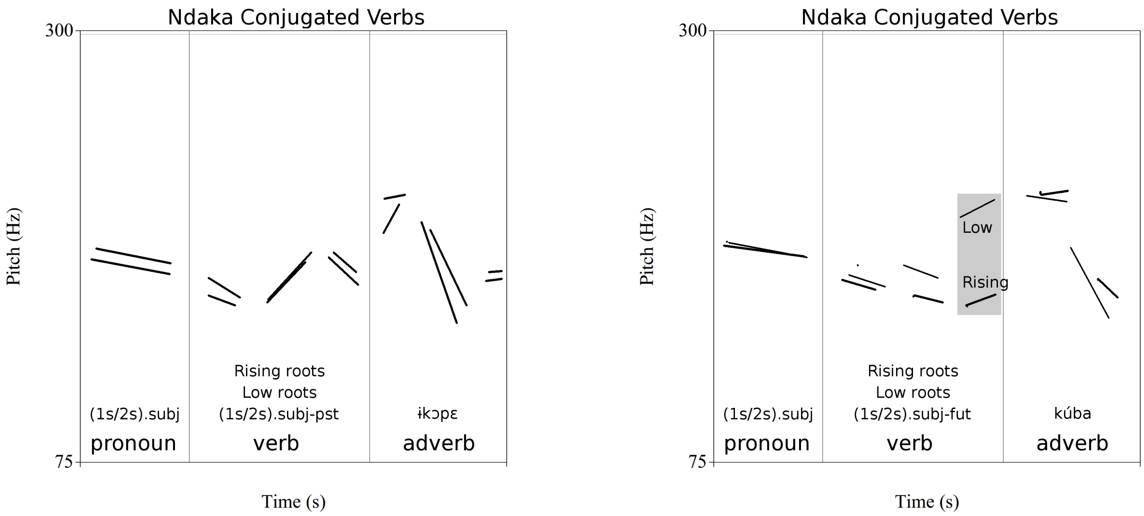


Figure 35. 2s past and future trendlines for Low and Rising verbs in Ndaka [ndk]

The past forms (on the left) are again essentially the same. The future forms (on the right) are different on the FV only. Given the generalization above, that future forms contrast roots beginning on the root syllable, these forms would appear to have the same initial lexical tone specification (i.e., low tone).

The trendlines from Low verbs in Figure 22 and from Rising verbs in Figure 32 are extracted and superimposed in Figure 36:

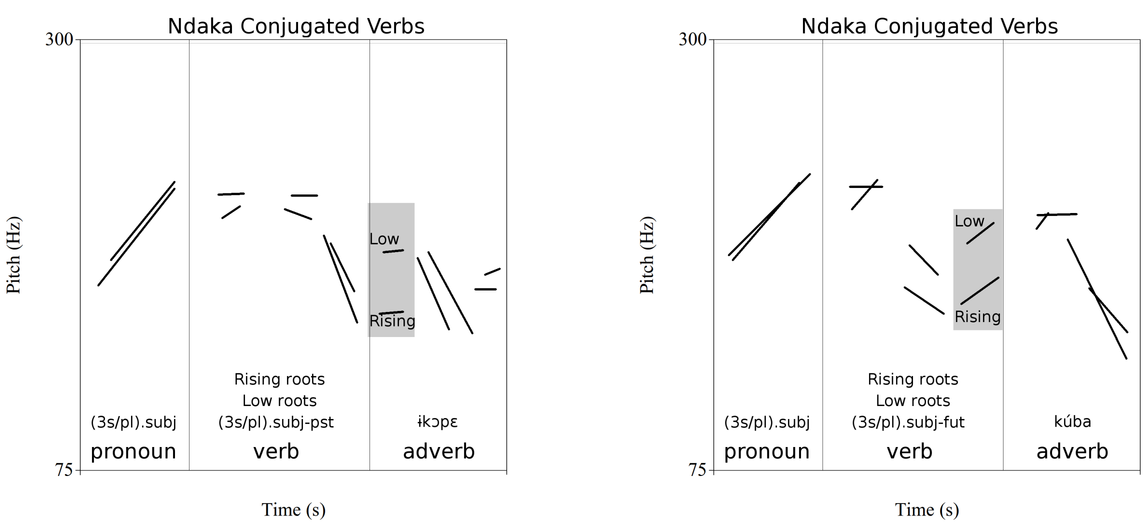


Figure 36. 3s past and future trendlines for Low and Rising verbs in Ndaka [ndk]

The past forms (on the left of Figure 36) show a contrast on the following adverb first syllable only, perhaps again because Low and Rising roots bear the same tone at the beginning of their lexical tone melody (the contrast of which begins on the FV).

The future forms (on the right of Figure 36) may appear to differ on two syllables, with the root syllable higher for Low verbs than for Rising verbs. While there is an observable difference, there are several reasons to analyze Low and Rising roots as bearing the same tone on the root syllable of their 3s-future forms. First, the difference may be understood as conditioned by their differing environments (namely, through anticipatory assimilation). That is, a low TBU before a high TBU (for Low verbs), would not be pronounced as low as a low TBU before another low TBU (for Rising verbs). One might describe the Low 3s-future forms (on the right in Figure 36) as high-low-high, while the Rising 3s-future forms (also on the right in Figure 36) could be described high-low-low. In such an analysis, the low root TBU on Low verbs raises somewhat in pitch in anticipation of the following high TBU (which does not happen for Rising 3s-future forms, where the low root TBU precedes a low TBU on the FV).

A second reason to analyze Low and Rising roots as bearing the same tone on the root syllable of their 3s-future forms has to do with evidence of downstep. The adverbs following the verbs begin at the same height after both Low and Rising 3s-future verbs. This fact provides evidence that their root TBU's are identical, in two ways. First, given that they end up with a high TBU at the same pitch level, one should not expect Low and Rising to be distinguished by a register shift (downstep), unless that was then subsequently reset. And second, because the following adverb (and final High TBU) are pronounced lower in pitch than the verb initial high TBU, we see evidence of downstep in each of Low and Rising 3s-future verbs. Together, there is evidence of neither more nor less than one downstep in each of Low and Rising 3s-future forms, presumably because each has a low tone on the root TBU.

Because there is evidence of neither more nor less than one downstep in each of Low and Rising 3s-future forms, and because the difference in their root TBU can be accounted for by anticipatory assimilation, I will continue with the basis that Low and Rising 3s-future forms are not underlyingly distinct on their root TBU, being both best described as low. As a result, the difference between Low and Rising 3s-future forms amounts to a high/low contrast on a single TBU, in line with the difference between Low and Rising roots for 2s-future and 3s-past forms. Furthermore, that single TBU of contrast is on the FV, in line with 2s-future forms.

This comparison of Low and Rising verbs has established that there is a difference between the two verb root classes in every form compared, with the sole exception of 2s-past forms. This is strong evidence of the lexical distinction between verbs of the two root classes. In particular, we find the difference on the final vowel (FV) of the future forms, and on the first TBU of the following word for 3s-past forms.

3.3.3 A Subset of Rising Tone Pattern Verbs

A small subset of the verbs with rising infinitive tone patterns as in Figure 30 seem to differ in one of their conjugated forms, i.e., that with a high bound subject pronoun (3s) in the past TAM group. Unfortunately, this apparent behavior is as unexpected as it is unconfirmed. Several of these forms have suspicious syllable structure, so it is entirely possible that they are morphologically complex. But without the data to decide this question at this point, and rather than attempting to ignore this apparent difficulty, I present here the verbs in question for future inspection:

The following verbs are in this subset of Rising verbs:

(20)
a. koɗuwo [– ⸍ –] pull (up)
b. kɔpɑnɑ [– ⸍ –] indicate show
c. kɔpɔpɑ [– ⸍ –] winnow
d. kɔnɑgbɑ [– ⸍ –] have headache
e. kɔjɑnɑ [– ⸍ –] cook
f. kɔʉnɑ [– ⸍ –] sow (seed)
g. kouo [– ⸍ –] blow air

The C1 consonants in this group cover the broad range of consonants not specified for voicing, and in fact constitute the lion's share of the Rising verb data without C1 specified for voice in (18). In no case, however, does the presence of this data diminish the importance of the potentially problematic distribution of consonants in the Rising verb root class. First, not all of that problematic data are in this subset. That is, there are Rising verbs with C1 not specified for voicing in (18), which do not behave as the verbs in this subset. And second, these data agree with Low roots in only one of the conjugated forms where Low and Rising roots normally contrast. That is, with regard to infinitive and future forms, these data agree with other Rising forms, so they must constitute a class distinct from Low roots in any case.

The following presents the conjugated patterns with trend lines for the verbs in (20), for the sake of completeness:

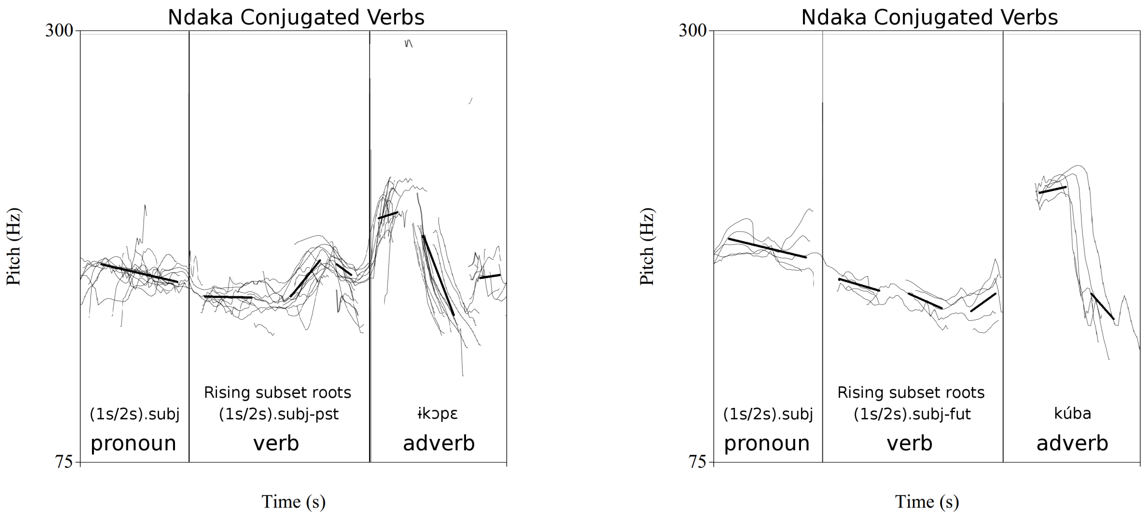


Figure 37. 2s past and future pitch traces for a subset of Rising verbs in Ndaka [ndk]

The past form (on the left) shows the same form as for all 2s-past conjugated verbs in Ndaka. The future form (on the right) is the same as for other Rising verbs, low and level across the verb, below the level of the initial low free pronoun.

The tone patterns in Figure 37 are compared to those for other Rising verbs (as presented in Figure 31, above) in Figure 38:

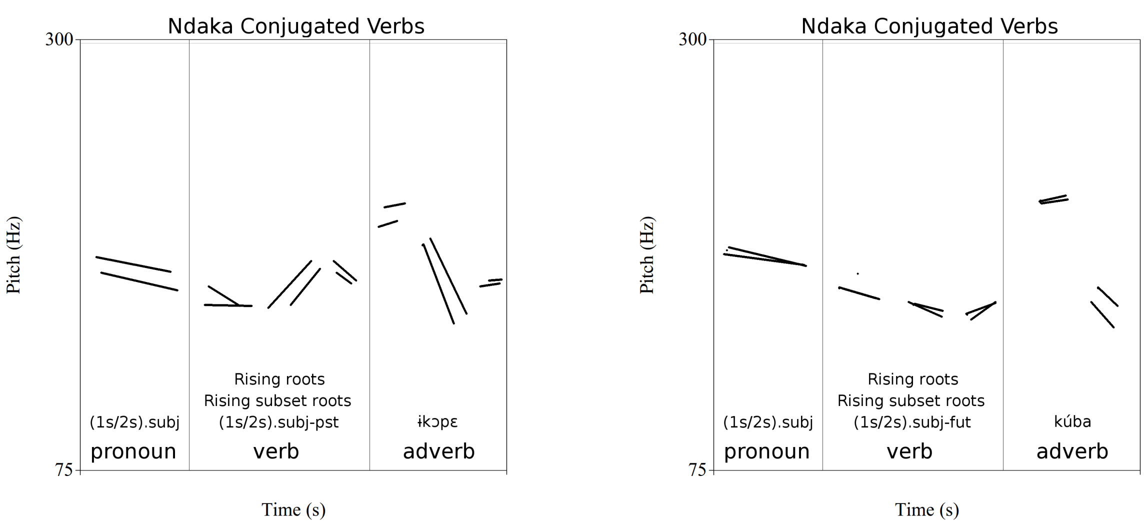


Figure 38. 2s past and future trendlines for (all) Rising verbs in Ndaka [ndk]

The past forms (on the left), along with the future forms (on the right), show no substantial difference for this subset of Rising verbs, for 2s subjects.

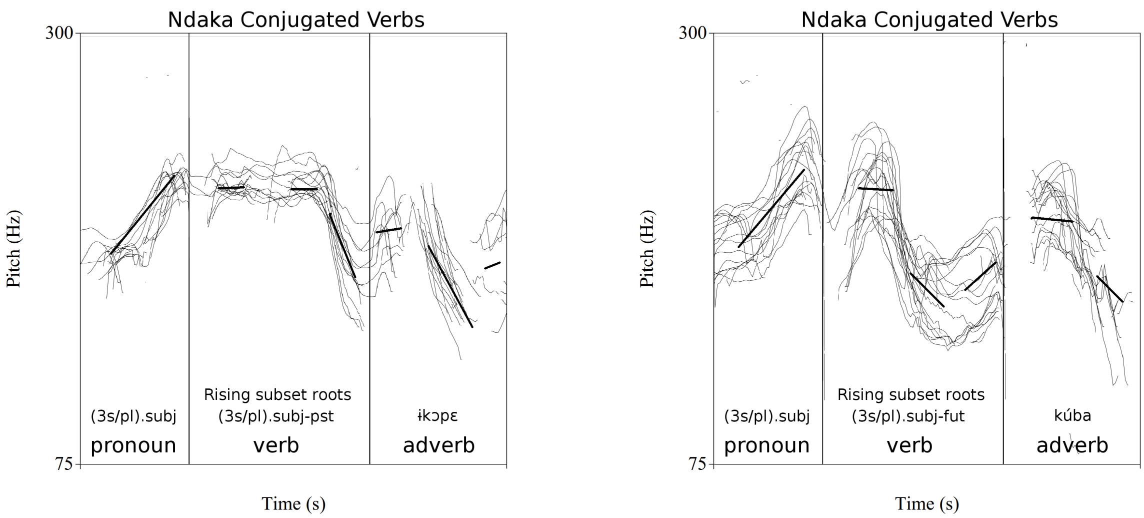


Figure 39. 3s past and future pitch traces for a subset of Rising verbs in Ndaka [ndk]

The past form (on the left) shows high and level tones on the first two syllables of the verb, followed by a low tone which downsteps the following adverb's initial high tone. The future form (on the right) shows high only on the first verb syllable, followed by two low syllables, the second of which rises slightly.

The tone patterns in Figure 39 are compared to those for other Rising verbs (as presented in Figure 32, above) in Figure 40:

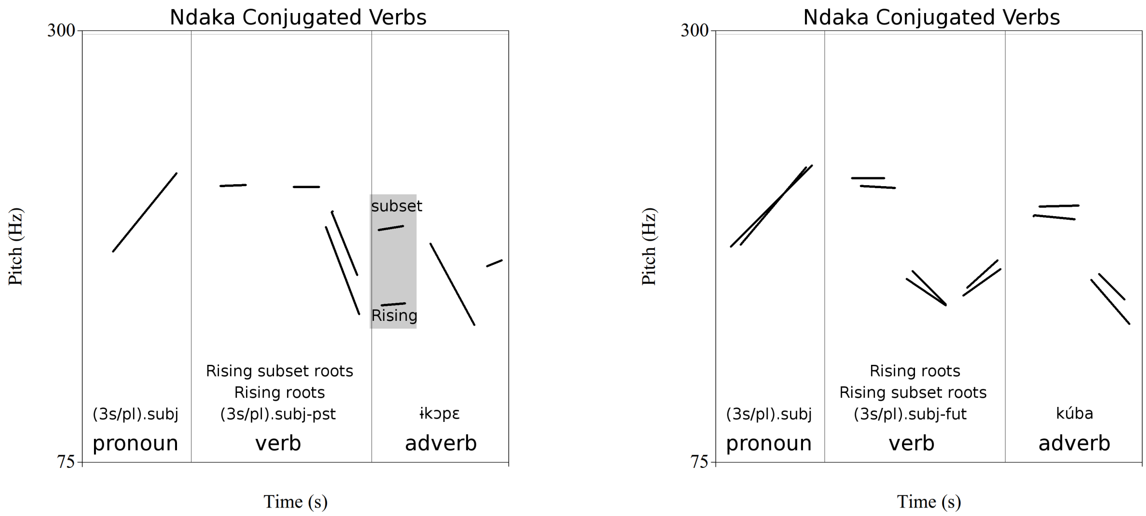


Figure 40. 3s past and future trendlines for (all) Rising verbs in Ndaka [ndk]

The past forms (on the left) show the one position in the conjugation where this subset of Rising verbs differs from other Rising verbs. The contrast is on the first syllable of the following adverb, and looks essentially the same as the 3s-past contrast between Low and Rising verbs in Figure 36. The future forms (on the right) show no substantial difference for this subset of Rising verbs for 3s subjects.

To summarize the comparisons in Figure 38 and Figure 40, the subset of Rising verbs in (20) show identical tone patterns, except for the 3s-past forms. In fact, for the 3s-past conjugation, this subset of Rising verbs stands with Low and High verbs on the adverb-initial syllable, in opposition to other Rising verbs, as shown in Figure 41:

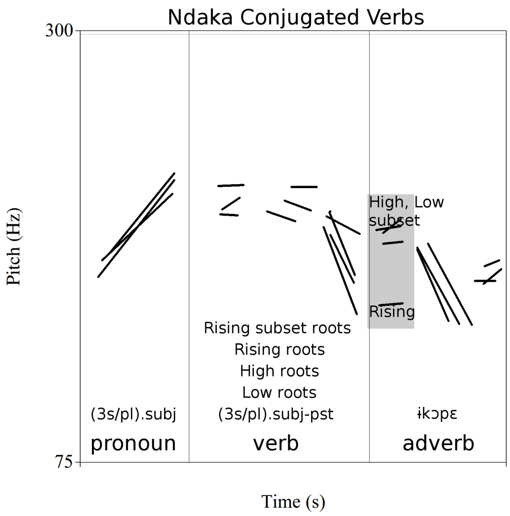


Figure 41. 3s past trendlines for Low, High, and (all) Rising verbs in Ndaka [ndk]

Figure 38, Figure 40, and Figure 41 show that this subset of Rising verbs behaves like Low verbs in 3s-past forms, but like Rising verbs in future forms. Specifically, they have a higher pitch on the following word than other Rising verbs do in the 3s-past form.

The operation of downstep in Ndaka is not completely understood, so the difference between Low and Rising tone patterns here may come down to the optional application of downstep across word breaks. Alternatively, the difference may result from a low spreading onto the following word in some cases, but not in others, for some reason currently not understood. In any case, given the lack of C1 specified for voice in any of the subset of Rising verbs (20), the difference may come down to differences in underlying forms due to historical processes (c.f., Chapter 7). In addition to whatever other factors may impact the tone on these verbs, at least two are plausibly morphologically complex (i.e., ‘cook’ and ‘indicate, show’, which may each have *-na* reflexive morpheme obligatorily attached).

3.4 Summary Comparison of Ndaka [ndk] Verb Root Classes by Tone

Sections 3.1, 3.2, and 3.3 have characterized the Low, High and Rising verb root classes for Ndaka, including for each the infinitive forms, four conjugated forms, and the C1 consonants of their verbs. They have furthermore shown the differences between these verb root classes by tone through pairwise comparisons of the four conjugated forms for each. This section provides figures summarizing those pairwise comparisons in a single three-way comparison for each of the four conjugated forms. The trendlines from 2s past and future forms with Low roots (from Figure 21), High roots (from Figure 25), and Rising roots (from Figure 31) are collated in Figure 42:

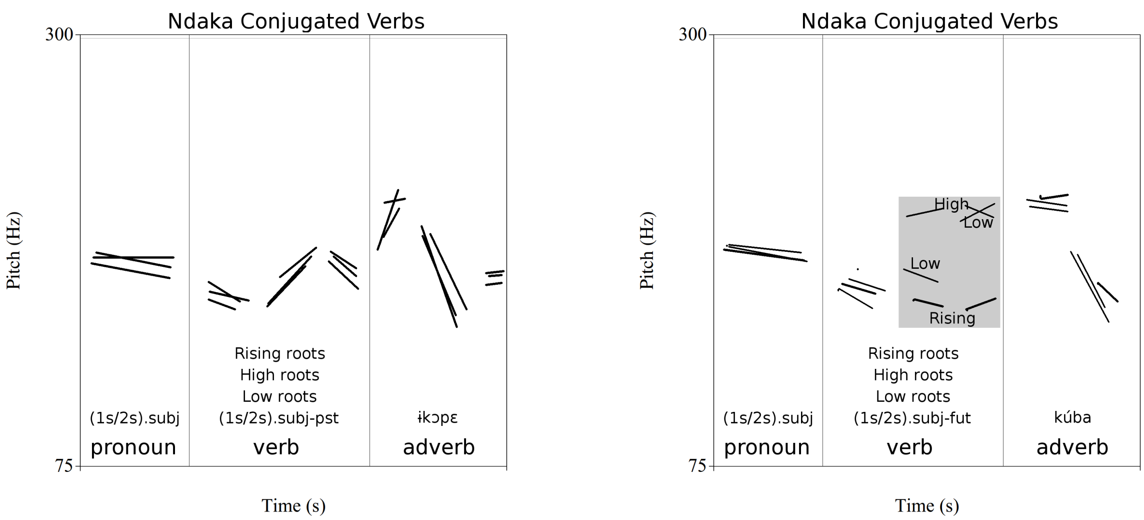
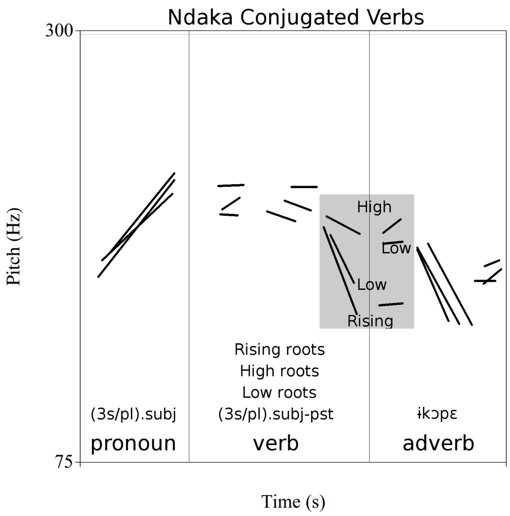


Figure 42. 2s past and future trendlines for High, Low, and Rising verbs in Ndaka [ndk]

The past forms (on the left) show essentially one tone pattern across all verb root classes by tone. The future forms (on the right) show High roots differing from Low and Rising roots on the root (or second) mora, and High and Low roots differing from Rising roots on the FV. As indicated in Section 3.2.1 and Section 3.3.1, future forms show verb root class by tone contrasts beginning on the root mora.

The trendlines from 3s past and future forms with Low roots (from Figure 22), High roots (from Figure 26), and Rising roots (from Figure 32) are collated in Figure 43:



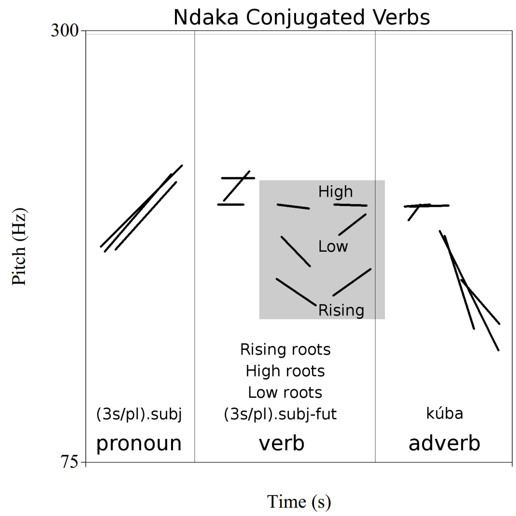


Figure 43. 3s past and future trendlines for High, Low, and Rising verbs in Ndaka [ndk]

The past form (on the left) shows a similar pattern of contrast to that in the future form in Figure 42, except that the contrast starts on the FV. High roots are opposed to both Low and Rising roots on the FV, and High and Low roots are opposed to Rising roots on the first mora of the following adverb.

The future form (on the right) shows contrast over two syllables, again starting on the root syllable, as for the future forms in Figure 42. The root syllable has three different pitch levels, but if Low roots are interpreted as high-low-high (as mentioned following Figure 36 in Section 3.3.2), the root mora in the future form of Figure 43 is high for High roots, and low for Low or Rising roots. Similarly, the FV would have high tone for High and Low verbs, as opposed to low tone for Rising verbs.

Rising verbs distinguish themselves as the only verb root class by tone which lowers the initial tone on the following adverb, in 3s-past forms. They are also the only verbs which have final low future forms, for both 2s-future and 3s-future forms. Together this indicates that Rising verbs may have more tonal material specified in the lexicon.

To summarize the forms in Figures 42 and 43 which show contrast across verb root class by tone (i.e., 2s-future, 3s-past and 3s-future, but not 2s-past), there is contrast across two TBU's in each case, as indicated by shading. Those two TBU's begin on the root mora in future forms, and on the final vowel (FV) on past forms.

And finally, those two TBU's are both high for High roots, both low for Rising roots, and low-high for Low roots.

But beyond whatever generalizations may be drawn from the forms presented in Figures 42 and 42, and however they impact the analysis of underlying forms (c.f., Section 3.5), one point should be absolutely clear at this point: there are three distinct Ndaka verb root classes by tone, as observed in infinitive and conjugated tone patterns. Those three verb root classes by tone are distinct everywhere they have been compared, with the notable exception of the 2s-past forms, which are neutralized across all root tone patterns. Further discussion of these three verb root classes by tone will be reserved for Section 3.6.

3.5 Initial Analysis of Ndaka [ndk] Tone System

Section 3.4 ended with a brief summary of generalizations based on the Ndaka verb root class by tone data provided in Sections 3.1–3.3, and summarized in Figures 42 and 43 on page 97 and page 98, respectively. These figures justify the idealized pitch traces used to describe the conjugated surface tone melodies provided in this chapter, first summarized in Table 6, and repeated here in Table 10, arranged with High class first:

<table>
<tr><th></th><th>2s-past</th><th>2s-future</th><th>3s-past</th><th>3s-future</th></tr>
<tr><td>High</td><td>[−|_ ⸝ ⸜|∨]</td><td>[−|_ ¯ ¯|⸌]</td><td>[ˊ|¯ ¯ ¯|⸜⸝]</td><td>[ˊ|¯ ¯ ¯|⸌]</td></tr>
<tr><td>Low</td><td>[−|_ ⸝ ⸜|∨]</td><td>[−|_ _ ¯|⸌]</td><td>[ˊ|¯ ¯ _|⸜⸝]</td><td>[ˊ|¯ _ ¯|⸌]</td></tr>
<tr><td>Rising</td><td>[−|_ ⸝ ⸜|∨]</td><td>[−|⸜ _ ⸝|⸌]</td><td>[ˊ|¯ ¯ _|⌣]</td><td>[ˊ|¯ _ _|⸌]</td></tr>
</table>

Table 10. Ndaka [ndk] conjugated pitch melodies by verb root class (∼Table 6)

The idealized pitch traces for the conjugated melodies in Table 10 can be converted into surface tones of binary value[4] as in Table 11. Here and at other times in this section I abbreviate [L/Voice] as L, for the sake of brevity in representations; the use of L should not be understood as different from [L/Voice].

[4] This conversion is admittedly nontrivial, particularly when comparing the 2s-past forms with the Rising future forms. Nontrivial aspects of this binary surface tone assignment will be argued for in the course of this section.

	2s-past |H	2s-future |H	3s-past |H	3s-future |H
High	[LLL|H]	[LHH|H]	[HHH|H]	[HHH|H]
Low	[LLL|H]	[LLH|H]	[HHL|H]	[HLH|H]
Rising	[LLL|H]	[LLL^{H}|H]	[HHL|L]	[HLL^{H}|H]

Table 11. Ndaka [ndk] conjugated surface tone melodies

The goal of this section is to account for these melodies on the basis of a consistent set of underlying forms and autosegmental rules and assumptions. I will take up briefly the question of the tone on bound subject pronouns, before moving to a fuller discussion of the tone on roots, in Section 3.5.1. I will address past and then future forms in Section 3.5.2. In Section 3.5.3 I will connect underlying tones with the assumptions needed to derive the surface representations in Table 11. And finally, I will briefly discuss the merits of this analysis in Section 3.5.4, including potential objections to and theoretical and practical advantages of the analysis.

3.5.1 Bound subject pronouns and roots

I begin with the most straightforward and superficially obvious part of this analysis, which is the subject bound pronouns. Table 12 arranges the tone patterns in Table 11 to compare forms which differ only by subject bound pronoun. For each pair of 2s/3s melodies, the part that differs between them is underlined:

	past |H	future |H
High	[LLL|H] (2s)	[LHH|H] (2s)
	[HHH|H] (3s)	[HHH|H] (3s)
Low	[LLL|H] (2s)	[LLH|H] (2s)
	[HHL|H] (3s)	[HLH|H] (3s)
Rising	[LLL|H] (2s)	[LLL^{H}|H] (2s)
	[HHL|L] (3s)	[HLL^{H}|H] (3s)

Table 12. Ndaka [ndk] conjugated surface tone melodies (by pronoun)

Table 12 provides for three observations about the tone of the bound subject pronouns, regarding the placement (21a), tonal value (21b), and length (21c) of the contrast:

(21) Observations about bound pronouns, based on Table 12
a. Conjugated 2s and 3s forms differ at the beginning of the word (at least).
b. Conjugated 2s forms consistently begin low (in Figure 42), whereas conjugated 3s forms consistently begin high (in Figure 43).
c. Conjugated future forms show contrast between 2s and 3s forms on one TBU only, while past forms show this contrast on either two (for Low or Rising verbs) or on all three TBU's (for High verbs).

Observation (21a) matches the segmental placement of these morphemes, which are prefixes. Observation (21b) speaks to the value of the contrast, i.e., a simple High/Low binary contrast between bound subject pronouns which are phonetically high (3s), versus those that are phonetically low (2s)[5]. This indicates a binary underlying contrast, with three logically possible hypotheses for their underlying representations:

(22) Possible underlying forms for bound subject pronouns 2s [L] and 3s [H]
a. {H,L}: Full specification
b. {H,∅}: Privative H / underspecified L
c. {∅,L}: Privative L / underspecified H

Without compelling reason to assume an unspecified tone for either prefix, I will assume the lexical forms are fully specified as L (2s) and H (3s), according to (22a) and as in Figure 44. This assumption is justified in light of the fact that there are no exceptions to observation (21b). That is, there is no case where a 2s form begins high, nor where a 3s form begins low. Because of this, positing underspecification in either case (as in 22b and 22c) would require other mechanisms to maintain the strong generalization of (21b).

[5] Recall that I have been using 3s and 2s as representatives of the binary tone contrast in Ndaka and Mbo bound subject pronouns. The full completement of pronouns is 1s and 2s (initially low) and 3s, 1p, 2p, and 3p (initially high). More information on their forms can be found in appendix C.

pronoun:	2s	3s
	L	H
	|	|
	μ	μ
	|	|
Lexical form:	ɔ-	ɔ-

Figure 44. Ndaka [ndk] 2s and 3s bound pronoun lexical forms

Including all the pronouns (c.f., appendix C) leads to a full complement of bound subject pronouns as in Figure 45:

tone class:	1s, 2s	1p, 2p, 3p, 3s
	L	H
	|	|
	μ	μ
	|	|
Lexical form:	(n)ɔ-	(k,n,ɓ)ɔ-

Figure 45. All Ndaka [ndk] bound pronoun lexical forms

The lexicon of Ndaka bound subject pronouns could be summarized as follows:

morpheme	**tone**
2s (and 1s)[6]	L-
3s (and 1p, 2p and 3p)	H-

Table 13. Draft lexicon of Ndaka [ndk] tones (bound subject pronouns)

So far I have addressed the first two points of (21): the placement (21a) and value (21b) of the bound subject pronoun tonal contrast. Observation (21c) speaks to the length of that contrast, and in doing so provides information about the tone of Ndaka roots. Specifically, the length of pronominal contrast indicates the amount of tonal information contained in each root class.

[6] Recall that there is a binary split for tone on bound subject pronouns, with 1s and 2s low, and the rest high.

Tonologists do not often write about the amount of tonal information contained in a morpheme, but the concept is nonetheless present in any analysis with morphemes which are not fully specified, or where one morpheme has more or fewer underlying tones than another. And furthermore, when language data indicates specification contrary to expectations (as for Ndaka), it is advisable to address this question up front.

For instance, in this analysis I will propose a lack of specification for High roots and a /LL/ specification for Rising roots, each of which will likely give most readers pause. The first goes against what is expected for Bantu languages generally (Hyman 2000), and the second contains what many would consider an OCP (or Twin Sister Convention) violation in the lexical form. So each of these two points is *prima facia* dubious. But I hope to show that they are each required to make sense of the data, beginning with the fact that Ndaka roots are not equally specified for tone.

To this end, I break (21c) down further in (23). The future data (in the right column) of Table 12 shows a single TBU contrast across all verb root classes, so these don't indicate any helpful differences. But looking at the past data (in the left column) of Table 12, one sees differences between the three verb root classes by tone as in (23). Note that these differences are specifically looking at the amount of tonal information, not the value (H/L) of that tonal information. The value of the tones in roots will be taken up in following pages.

(23) Observations about root class tones, based on past the data in Table 12
   a. High verbs have a single tone contrast across the whole word.
   b. Low verbs have a contrast over the first two TBU's, then a consistent surface low tone on the third TBU.
   c. Rising verbs appear the same as Low verbs, but with a low on the following word for 3s forms.

Comparing the first two verb root classes first, one could say that past High roots exhibit a single contrast only, presumably due to the contrast in bound pronouns, whereas the past Low verbs have both the pronominal tone contrast and another contrast, presumably due to the root. This difference might be summarized in saying that High verbs have less tonal information than do Low verbs.

Considering the Rising verbs, we see even more tonal information than for Low verbs. This is seen in the fact that past Rising verbs have a low tone pushed onto the following word, at least for 3s forms; neither High nor Low verbs ever push tonal information onto a following word. This could be summarized as in (24):

(24) Amount of tonal content: High < Low < Rising

The implications of (24) are nontrivial, and in some need of defense, because it reverses what is typically expected of tone in Bantu languages. That is, Hyman (2000) and others have proposed that most Bantu languages can be analyzed as privative H systems, with a phonologically active High tone or its absence, but not a phonologically active Low tone. Given that in the paradigm high tones would be phonologically specified, and low tones would not be, one would expect a different hierarchy of amount of tonal content, as in (25):

(25) Amount of tonal content: High > Low (implied by privative H)

While the privative H hypothesis does not deal directly with the Ndaka Rising verb root class, the relative amount of underlying content in surface high and low tones is reversed. Rather than a phonologically active high tone with a surface low which is phonologically underspecified (as in a privative H analysis), the hierarchy in (24) in a privative system would imply a phonologically active low tone, with a phonologically underspecified high tone —though (24) by itself does not require a privative system of any kind. Note that I am not arguing for a privative low tone generally in Ndaka, but rather that Ndaka roots are not equally specified for tone, they do not have equal amounts of tonal content. High verbs seem to carry the least tonal information, Low verbs more, and Rising verbs the most, regardless how high and low tones operate in the language generally. This observation, summarized in (24), will help inform the choice of underlying forms for Ndaka roots, but should not be taken to imply a particular kind of system in the language generally.

Rearranging Table 12 to see the difference by verb root class results in Table 14. For each triplet of High/Low/Rising melodies, the tones that differ between them is underlined (shaded cells show no contrast):

<table>
<tr><th></th><th>past |H</th><th>future |H</th></tr>
<tr><td>2s</td><td>[LLL|H] (High)
[LLL|H] (Low)
[LLL|H] (Rising)</td><td>[LHH|H] (High)
[LLH|H] (Low)
[LLL^{H}|H] (Rising)</td></tr>
<tr><td>3s</td><td>[HHH|H] (High)
[HHL|H] (Low)
[HHL|L] (Rising)</td><td>[HHH|H] (High)
[HLH|H] (Low)
[HLL^{H}|H] (Rising)</td></tr>
</table>

Table 14. Ndaka [ndk] conjugated surface tone melodies (by root class)

Table 14 provides for five observations about the contrasts between the root classes, again considering the length (26ab), values (26c), and placement (26de) of the contrast:

(26) Observations about root class tones, based on Table 14

a. Conjugated 2s-past data does not show different tone melodies by verb root class.
b. The contrast between the three verb root classes (where it exists) occurs over two TBU's (underlined in Table 14).
c. Where there is contrast, Low roots group with Rising verbs (and oppose High verbs) on the first TBU; Low verbs group with High verbs (and oppose Rising verbs) on the second TBU. In other words, the two TBU's of contrast are [HH] for High roots, [LL] for Rising roots, and [LH] for Low roots.
d. Conjugated future forms show root contrast beginning on the root vowel (second TBU) of the verb, whereas
e. Conjugated past forms show root contrast (where it exists) beginning on the final vowel (FV; third TBU) of the verb.

Observations (26a-c) provide the framework for the remainder of this section; observations (26de) will be addressed in Section 3.5.2.

Observations (26ab) speak to the length of the contrast; there is no contrast for 2s-past data, and a contrast expressed over two TBU's elsewhere. These two observations speak to the amount of tone specification on roots again: if all the root classes are fully specified for tone, one should not expect neutralization of that contrast anywhere in Table 14. While it is possible that the root tone contrast is neutralized by a grammatical tone paradigm (a replacive grammatical tone), the fact that it would apply for just one combination of subject-TAM makes this argument somewhat suspect. A more plausible hypothesis is that one of the tone classes is underspecified, allowing for neutralization in the context that provides the same tonal information as in the specified class(es).

Considering the value of the two TBU's expressing root contrast, the summary in (26c) provides for surface forms of [HH] for High, [LH] for Low, and [LL] for Rising verb roots. This provides for four basic possibilities for their underlying representations:

(27) Possible underlying forms for High [HH], Low [LH], and Rising [LL] roots
a. {HH,LH,LL}: Full specification (1 tone per TBU)
b. {H,LH,L}: Full specification (1 tone per surface tone height)
c. {HH,H,∅}: Privative H (underspecified L / Rising = ∅)
d. {∅,L,LL}: Privative L (underspecified H / High = ∅)

I will ultimately argue that (27d) is the best set of underlying forms for Ndaka verb roots, but first I want to consider the fuller implications of each of these hypotheses. Each of the four hypotheses in (27) makes a different claim about the amount of lexical material in each verb root class: (27a) posits two tones on each of the three verb root classes; (27b) posits more tonal content on Low verbs than on either High or Rising verbs; (27c) posits High verbs as having the most tonal content and Rising the least; (27d) posits High verbs as unspecified, with Rising verbs having the most tonal content.

Considering again the observations made so far about the relative amount of tonal information in each of the three verb root tone classes, observation (27b) provides for a High verb tone specification which has less tonal information than for Low verbs, but only (27d) also has both more tone in the Rising verb tone specification (satisfying all of 24) and an underspecified tone (to allow for verb root tone neutralization in 2s-past forms). As a result, I will take (27d) as the correct form, resulting in the lexical entries of /∅/ (High), /L/ (Low), and /LL/ (Rising), as in Figure 46:

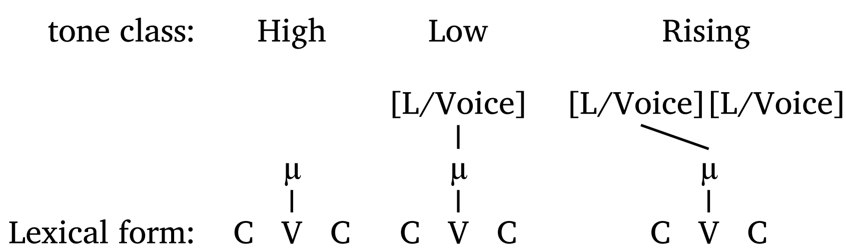


Figure 46. Ndaka [ndk] verb root lexical forms by tone class

I have already mentioned the potential objection to this analysis coming from Hyman (2000), that tone in Bantu should typically understood to be privative H, whereas this analysis has an unspecified High verb root class. Another likely objection to this analysis comes from the /LL/ specification of Rising verb roots. Having two of the same feature adjacent would be considered an OCP violation by many, and while some would allow it in derived environments, it is more difficult to accept in the lexical form. Similarly, the Twin Sister Convention could be invoked to claim that a second [L/Voice] attached to the same mora is meaningless at best —though note that the second [L/Voice] on the right of Figure 46 is floating, rather than attached to the same mora.

But there are three reasons why this /LL/ specification for Rising roots is justified for Ndaka. The first has already been mentioned: Ndaka Rising roots have more lexical tone content than other roots, based on the fact that they alone push tonal content onto following words.

A second justification for /LL/ specification for Rising roots comes from the fact that verbs with these roots have a consistent [LL] surface form. If the underlying specification were /L/ associated to two TBU's, one would expect to see one or other of those associations broken in favor of another tone at some point —but this never happens.

The third reason to accept /LL/ specification for Rising roots comes from the fact that a second underlying [L/Voice] feature in Rising roots accords very nicely with the analysis proposed in Chapter 7 for the development of the Rising verb root class. If the Rising verb root class arose through the addition of an extra [L/Voice] feature to the lexical melody (as I claim in Chapter 7), then one should expect to find an extra [L/Voice] feature in the underlying form of verb roots of that class. So despite the fact that some may say that /LL/ should never exist as an underlying representation for a single morpheme, the /LL/ specification for Rising roots in Ndaka matches the available data in terms of surface form length and value, as well as the proposed history for this particular verb root class.

To summarize the analysis so far, the lexicon of Ndaka bound subject pronouns and verb roots is as follows:

morpheme	**tone**
2s (and 1s)	L-
3s (and 1p, 2p and 3p)	H-
High roots	∅
Low roots	L
Rising roots	LL

Table 15. Draft lexicon of Ndaka [ndk] tones (bound subject pronouns and roots)

3.5.2 TAM categories (past and future)

Having addressed the observations from (26) on the length and value of the tone contrast by verb root class, we now turn to the position of that contrast, as expressed in observations (26d) and (26e). Taken together, these observations show that the root contrast on Ndaka past forms (where it exists) are one TBU to the right of the comparable future forms. This can be accomplished in at least three different ways: a prefixed past morpheme with some specified tone, a rightward shift in the past only, or reduplication of other prefixed material. I will ultimately argue for the reduplication of prefixed material, but first I will address the more inherently plausible possibilities.

A simple H or L prefixed tone to indicate past is ruled out by the data in the left (past) column of Table 14. Neither a H nor a L past prefix could generate these forms, as the prefix would have to be low in the case of 2s forms, and high in the case of 3s forms. This is because the 2s forms have two low TBU's before the root tone contrast, whereas the 3s forms have two high TBU's before the root tone contrast. But if the past content were H, one would expect high-high versus low-high (or else high-low, if the past tone preceded the pronominal tone). And if the past content were L, one would expect low-low versus high-low (or else low-high, if the past tone preceded the pronominal tone). Because none of those expectations are met, the past form cannot be reduced to an independent tone value.

A categorical rightward shift of root tones for all past forms is likewise excluded by the data in the left (past) column of Table 14, because 2s and 3s forms do not behave in the same manner. While there is a rightward shift of Rising tones onto the following word for 3s-past Rising forms (i.e., [HHL|L], where the following word normally begins [H]), there is no impact on the following word for 2s-past Rising forms (i.e., [LLL|H]). Because of this (and assuming Rising forms have only one underlying tone melody), a single shift for all past forms cannot account for the data in Table 14. A shift of tone one TBU to the right accounts for 3s-past Rising forms, but would predict ungrammatical *[LLL|L] for 2s-past Rising forms. Conversely, any formulation of a categorical shift that would **not** put 2s-past Rising tones onto the following word would predict ungrammatical *[HHL|H] for 3s-past Rising forms, which do in fact shift L onto the following word.

What is needed then, to account for the data in the left (past) column of Table 14, is a process that moves the root tones one TBU to the right for 3s (H) forms, but not for 2s (L) forms. This can be accomplished by reduplication of prefixed H tones. The reduplication of bound pronoun lexical entries could be formalized as in (28), and would produce the forms in the left (past) column of Table 14 when operating on lexical forms for the bound subject prefixes as in Figure 45.

(28) RED{H_{px}}
$\mathrm{H} \rightarrow \mathrm{HH} \,/\, [\_[\ldots]_{root}]_{past}$

The stipulation of this rule will be addressed in Section 3.5.4. At this point, it should suffice to say that it accounts for the observed data, unlike the other possible hypotheses mentioned above.

Rearranging Table 14 to see the difference by TAM category results in Table 16. For each pair of TAM melodies, the tones that differ between them is underlined (shaded cells show no contrast):

<table>
<tr><th></th><th>2s |H</th><th>3s |H</th></tr>
<tr><td>High</td><td>[LLL|H] (past)
[LHH|H] (future)</td><td>[HHH|H] (past)
[HHH|H] (future)</td></tr>
<tr><td>Low</td><td>[LLL|H] (past)
[LLH|H] (future)</td><td>[HHL|H] (past)
[HLH|H] (future)</td></tr>
<tr><td>Rising</td><td>[LLL|H] (past)
[LLL^{H}|H] (future)</td><td>[HHL|L] (past)
[HLL^{H}|H] (future)</td></tr>
</table>

Table 16. Ndaka [ndk] conjugated surface tone melodies (by TAM)

Again we see an unexpected neutralization of surface tone melodies (in shaded cells), and again it involves High roots. And again we see the additional material on Rising roots. These facts confirm the summary hierarchy of information in verb roots as given in (24), repeated here:

(29) Amount of tonal content: High < Low < Rising (=24)

Having discussed the past forms at some length, it remains to address the morphological marking on Ndaka future verbs. As may be observed in Table 16, conjugated Ndaka future forms end with high surface tone for all High and Low verbs. The only Ndaka future verb forms that end with low surface tone, then, are Rising forms.

The above allows for a simple -H suffix for future forms. This is in part because Rising forms have already been shown to have more tonal material than High and Low forms. And this extra material on Rising forms has also already been shown to push tonal material off the verb word, in 3s-past forms ([HHL|L]). A simple Ndaka future -H suffix is also possible because the Rising root tone melody /LL/ ends with an underlying low tone. Together, these facts make it entirely plausible that the high surface tone present on the end of High and Low forms is pushed off the end of Rising forms, in favor of the final low root tone on Rising verbs.

A simple -H suffix for Ndaka future forms also accounts for the presence of floating high tones in future Rising surface forms in Table 16. A future -H suffix would surface on High and Low forms, because their underlying tones are the same or fewer than their three TBU's. But the extra tone on Rising roots means a future -H suffix would be the fourth tone on those forms. As a result, the future -H suffix would remain floating (unassociated) at the end of Rising forms, as they have more underlying tones than TBU's (e.g., [LLL^{H}|H] for 2s-future Rising forms). This -H future suffix thus accounts for the floating H on Rising future forms. This floating H in turn accounts for the final flat low on Rising future forms (e.g., [−|⸌ _ ⸝|˥\]), as opposed to the final falling low pitch observed elsewhere (e.g., [−|_ ⸝ ⸜|\/]).

The lexicon of Ndaka morphemes discussed up to this point is as follows:

morpheme	**tone**
2s (and 1s)	L-
3s (and 1p, 2p and 3p)	H-
past	RED{H_{px}}
future	-H
High roots	∅
Low roots	L
Rising roots	LL

Table 17. Draft lexicon of Ndaka [ndk] tones

3.5.3 Putting it all together

At this point, I have provided basic justification for forms for each of the bound pronouns, roots classes, and TAM categories, as in Table 17, repeated in Table 18. The following section puts these morpheme lexical entries together into the underlying representation of words, and considers assumptions necessary to derive surface forms.

morpheme	tone
2s (and 1s)	L-
3s (and 1p, 2p and 3p)	H-
past	RED{H_{px}}
future	-H
High roots	∅
Low roots	L
Rising roots	LL

Table 18. Draft lexicon of Ndaka [ndk] tones (=Table 17)

Combining underlying forms for bound subject pronouns and verb root classes from Table 18 produces the matrix of underlying forms (less TAM) in Table 19:

pron \ roots	High /∅/	Low /L/	Rising /LL/
2s /L-/	/L[∅]/	/L[L]/	/L[LL]/
3s /H-/	/H[∅]/	/H[L]/	/H[LL]/

Table 19. Matrix of pronominal v root underlying forms without TAM (root tones in brackets)

Returning briefly to the question of the amount of tonal information in each form, the left column of Table 19 shows the High forms as having only one tone from subject bound pronoun and root, while the right column of Table 19 shows a total of three tones from the Rising and prefix morphemes. I will apply this matrix to past forms, then to future forms.

When the past reduplication rule in (28) operates on the forms in Table 19, the set of whole word underlying forms are as in Table 20:

past RED{H_{px}}	High (∅)	Low (L)	Rising (LL)
2s /L-/	/L[∅]/	/L[L]/	/L[LL]/
3s /H-/	/HH[∅]/	/HH[L]/	/HH[LL]/

Table 20. Matrix of Ndaka past underlying forms

The surface forms for Ndaka past data fall out rather naturally from the underlying forms in Table 20, given four basic assumptions:

(30) a. Tones are associated one per TBU from left to right.
b. Unassociated word-final [L/Voice] may associate to the first TBU on a following word.
c. H spreads to a following unassociated TBU.[7]
d. Unassociated TBU's are pronounced low.

The only intrinsic ordering in the assumptions in (30) is that (30a) occurs first, and (30d) at the end of whatever other derivation exists, as might be expected. An unassociated word-final [L/Voice], the context for (30b), is only found on 3s-Rising forms. Similarly, an empty TBU following a high tone, the context for (30c), is only found on verbs with High roots.

It should also be noted that (30cd) are configured for a system related to a privative H system, as found elsewhere in Bantu (Hyman 2000). This analysis would also work with these two inverted, with L spread and a default H. But without any particular reason to decide between the two on the basis of Ndaka data, I remain with the precedent established for Bantu analysis on this point.

The neutralization of all 2s-past forms is accounted for by the underlying forms in Table 20 and only (30d). That is, given the underlying forms /L/, /LL/, and /LLL/, with a default low pronunciation for empty TBU's, all roots would surface the same regardless of how tones were associated to TBU's, and without regard for any spreading or other association rules, as in Figure 47:

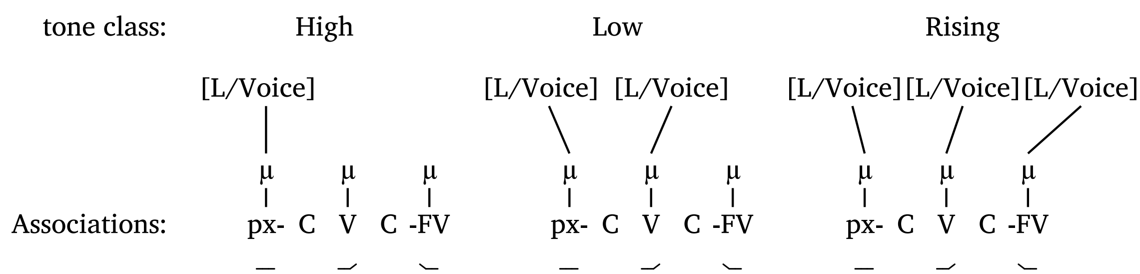


Figure 47. Ndaka [ndk] 2s-past forms

The 3s forms are similarly accounted for by simple left to right association (30ab) and H spread (30c), as in Figure 48:

[7] I have no data at this point to distinguish between bounded and unbounded spread, as I have no data where I would expect a high tone followed by two unassociated TBU's.

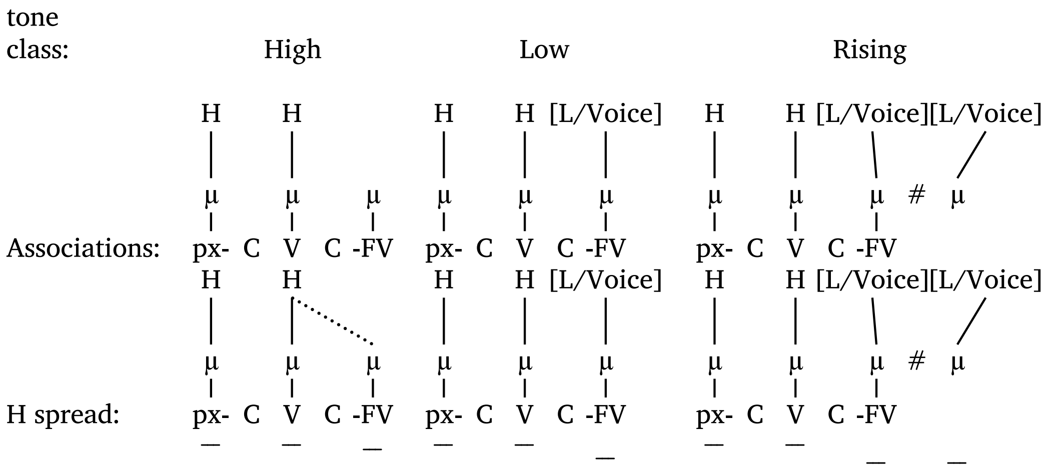


Figure 48. Ndaka [ndk] 3s-past forms, with H spread (30c)

With the future -H suffix applied to the matrix in Table 19, the set of whole word underlying forms are as in Table 21:

future -H	High (∅)	Low (L)	Rising (LL)
2s L-	/L[∅]H/	/L[L]H/	/L[LL]H/
3s H-	/H[∅]H/	/H[L]H/	/H[LL]H/

Table 21. Matrix of Ndaka future underlying forms

These underlying forms also derive surface forms fairly straightforwardly, with the association and spreading assumptions in (30a-c) above. The 2s forms are given in Figure 49:

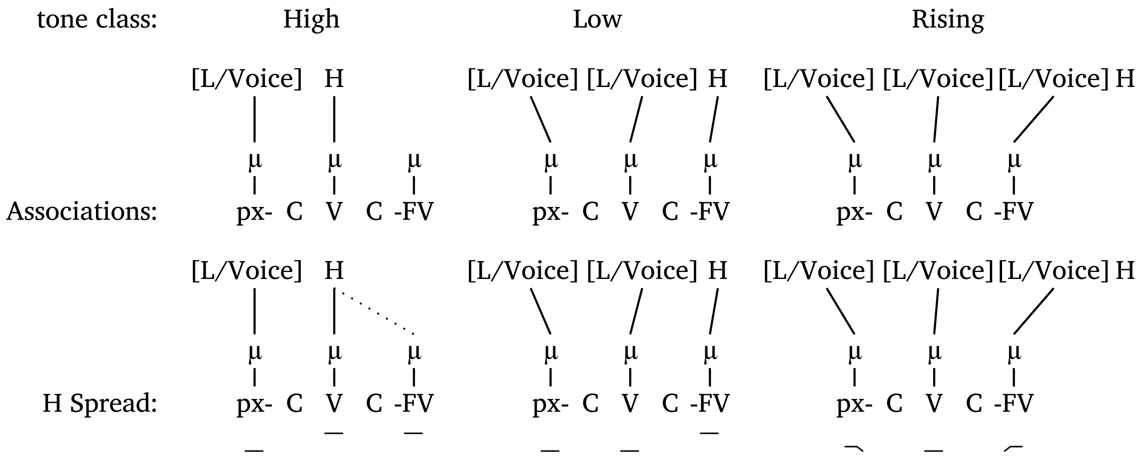


Figure 49. Ndaka [ndk] 2s-future forms, with H spread (30c)

The 3s forms are given in Figure 49:

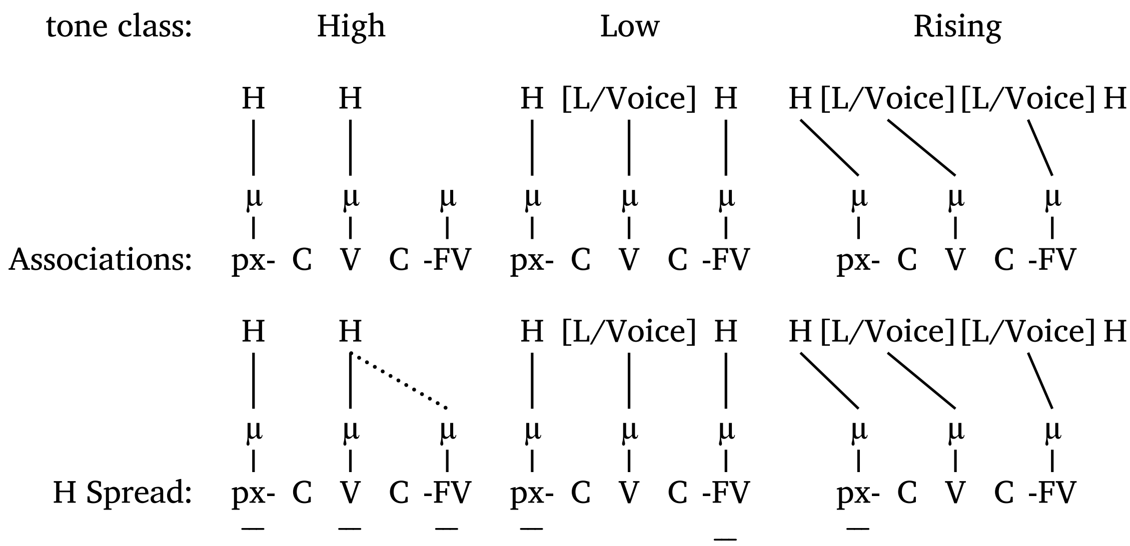


Figure 50. Ndaka [ndk] 3s-future forms

3.5.4 Discussion of initial tonal analysis of Ndaka verbs

Sections 3.5.1–3.5.3 show a fairly simple set of lexical items (in Table 18), and a fairly natural (and short) set of assumptions in (30), accounting for all the available conjugated data. This fact argues strongly that the analysis presented here is correct. Here I will address three potential criticisms of this analysis: the stipulation of reduplication for high prefix tones only, the differential treatment of high and low floating (unassociated) tones, and the interpretation of surface tones. Each will be addressed in turn.

Perhaps the weakest part of this analysis is the stipulation of reduplication of prefix high tones only, to mark past forms. But given that high tones are taken to be generally more phonologically active in Bantu (Hyman 2000)[8], reduplicating H but not L does not seem entirely implausible. This phonological activity on H only is perhaps a remnant from a time when the pronominal system was in fact a privative H system. And the data clearly show a shift of root tones for 3s-past forms only, so this must be accounted for in the analysis.

I have entertained numerous other possibilities to account for shift in this one set of forms only, but each proved significantly more difficult to pull together, and involving much more extensive stipulation. Particularly, if all prefix tones are reduplicated, then the Rising 2s-past form should have an underlying tone of /LL[LL]/, which should put a low onto the following word, which does not happen. To avoid this, one might posit an unspecified 2s prefix, which would give the Rising 2s-past an underlying tone of /∅[LL]/, which would correctly not put a low onto the following word. But this creates other problems, as some other mechanism is required to ensure the prefix tone is always low, in accordance with the observed surface tones (e.g., High 2s-future forms with /∅[∅]H/ > [LHH], not *[HHH] or *[HHL]). One could assume lexical specification of root tones, and thereby block tones from spreading to the prefix, but this would not help for High roots, where the root itself is also unspecified. In this case, one would need another mechanism to keep suffix High tones from spreading across an empty root to the prefix. One might call for cyclical association, such that the prefixes (and their tones) are added to the word only after the rest of it is built. While such an analysis may account for the available data, I believe it is simpler (and therefore preferred) to make a single stipulation of prefix high tones reduplicated to indicate past TAM.

One might similarly find the future -H analysis not entirely satisfactory. One might object that there is so far no evidence of a high tone added to following words (which might be expected in Rising forms, from either Figure 49 or Figure 50). This objection amounts to the fact that high and low tones are treated differently when

[8] Recall that I have not argued against Hyman (2000) in this analysis, only that Ndaka root tone specification does not follow a pattern expected in a privative H system. I think there is reason to believe that both H and L are phonologically present and active in Ndaka, beyond the specification of roots (e.g., in the bound subject pronouns), but that doesn't mean that Ndaka has never had a privative H system, nor that privative H systems are not found in many other Bantu languages.

unassociated at the end of a word. This difference exists in that low tones shift past the last TBU of the word and are associated to (and therefore pronounced on) the following word (as in Rising 3s-past forms, with /HH[LL]/ > [HHL|L]), while high tones shift past the last TBU of the word but remain floating (as in Rising future forms, e.g., 3s /H[LL]H/ > [HLL^{H}]). This analysis already depends on treating high and low tones differently, however, so this objection is not to be taken too seriously. And considering the impact of floating high tones (see the next point), one might argue that floating high and low tones are each expressed, but in different ways.

The final potential objection has to do with the interpretation of the surface tones. As mentioned in Section 2.1.1, my methodology depends on a somewhat subjective interpretation of pitch melodies into tone melodies, that subjective interpretation is necessarily (and rightly) subject to criticism. One might, for instance criticise my characterization of 2s-past forms, with (framed) surface pitches of [−|– ⸝ ⸜|∨] as composed of the surface tones [LLL]. This might appear particularly concerning when compared with 2s-future forms, with (framed) surface pitches of [−|⸌ _ ⸍|\], which I characterize as composed of the surface tones [LLL^{H}]. This characterization is justified, however, in light of the fact that final low tones fall almost universally, whereas low tones followed by floating tones typically do not. Additionally, I take the rise in the root mora of 2s-past forms to be due more to penultimate word stress than lexical tone. If there were a rule that insisted on some kind of higher tone (as may be the source of the penultimate/accented syllable rise in the 2s-past forms), then this rule might not apply where there existed already a high tone, as in the case of the 2s-future Rising forms. Thus, the presence of a floating high tone in the 2s-future Rising forms accounts for its flatness in two ways, by keeping the penultimate syllable from raising, and by keeping the final syllable from falling. Similar comments could be made about the 3s-future Rising forms.

This analysis provides two final benefits worth mentioning. First, it accounts for the difference between 2s-future Rising forms ([−|⸌ _ ⸍|\]) and 2s-past forms of any root class ([−|– ⸝ ⸜|∨]). If the 2s-past forms are (as I have stated) best understood as surface [LLL], and if the future is marked by a -H prefix, as in the analysis above, then the 2s-future Rising forms should surface [LLL^{H}]. This is because the 2s-future Rising forms should surface the same as the 2s-past forms, but with a following floating high tone (which marks future). This extra floating tone follows naturally from a single L- prefix tone, two root L tones, and a single -H suffix tone, associated left to right over three TBU's. In this way, the phonetic difference between the Rising 2s past and future forms as described in the previous paragraph correlates with the analytical difference required by marking the future with a H suffix.

The second benefit of this analysis is that it provides a way of understanding the data presented in Section 3.3.3. Recall that a subset of Rising verbs has the same tone patterns as other Rising verbs for every conjugated form but for 3s-past forms. As mentioned in Section 3.3.3, there may be a number of other reasons why this group of words behaves differently than expected (e.g., because they are morphologically complex). But the question remains as to why this group of verbs only differs in the one tone pattern. What sets that part of the paradigm apart?

This analysis provides an answer to the above question: the 3s-past Rising tone pattern is the only tone pattern which results in a floating low tone which may associate to a following word. As a result, any rule which targets (allowing or disallowing) floating low tones to associate to the first TBU of a following word would target this tone pattern only. If such a rule were to be optional in any way, this would explain why the Rising data are able to differ on only that one tone pattern, as it is the only tone pattern that results in floating low tones. The optionality of such a rule would have no impact on any other tone patterns, as they don't result in floating low tones.

I have one final comment on this analysis, before returning to a more general discussion of the Ndaka data and their implications for the development of new tone melodies in Section 3.6. Given the amount of data to test the analysis presented in this section, I have intentionally called this analysis “initial”. That is, the analysis in this section works as is for the data I have, but I assume that further data will either prove it correct, or else require modification to the proposed analysis. But neither outcome would impact the larger message of this chapter, which is that Ndaka has three clearly distinct verb root classes by tone, and there is a distribution of consonant types across those verb root classes by tone which doesn't easily submit to a simple [L/Voice] analysis. This will be discussed more fully in Section 3.6 and confirmed for Bantu D30 more generally by the Mbo data in Chapter 4, as well as by the Nyali-Kilo data in Chapter 5. The implications of these distributions will then be taken up in Chapter 7.

3.6 Discussion of Ndaka [ndk] Data

This section covers a number of generalizations that come out of the Ndaka data presented in Sections 3.1–3.3. Perhaps the most significant is that there are three verb root classes by tone, two of which (High and Low) have verb roots without any C1's which are specified for voicing. The third (Rising) verb root class has all the Ndaka roots with C1 specified for voicing, which indicates a complementary distribution between one verb root class by tone which derives from [L/Voice] specification and two that lack it (Bradshaw 1999). Such a generalization covers all Ndaka High and Low verbs, and 10 of 20 Rising verbs, for a total of 61 of the 71 Ndaka verbs for which I have conjugation data (i.e., 86%). But my data set also includes the other 10 of 20 Rising verb root class verbs (i.e., 14% of all Ndaka verbs in my data set) with C1's that are not specified for voicing, indicating that however strong the indication of complementary distribution is, the distribution of C1 types across tone patterns is not in fact complementary, but rather contrastive, as indicated in Table 22:

verb root class	C1 without [L/Voice]	C1 with [L/Voice]
Low	28	-
High	23	-
Rising (CTI)	10	10

Table 22. Distribution of consonants across melodies in Ndaka (numbers for conjugated data only)

The presence of the data in the (darker shaded) lower left corner prevents an analysis of complementary distribution, which would be necessary to allow a straightforward and synchronic [L/Voice] analysis to work. This is not the case only when examining larger categories as in Table 22 (i.e., with or without voicing specification); one is similarly forced to analyze the Ndaka data as contrastive on the basis of particular segments. There are examples of Low, High and Rising verbs with a variety of C1 segments not specified for voicing, including [k] (12f-h, 13b-g; 14o, 15b; 18b, 19a-b), [p] (12i, 13i-m; 15c; 18d-e), [ɗ] (12a, 13a; 14g, 15a; 18a), [m] (12m-n, 13p-s; 14p, 15j; 19d), [n] (12o, 13t-u; 15k-m; 18f, 19e) and [w] (13v-x; 15n-q; 19f-i). There are also examples of High and Rising verbs with C1 [t] (14h-n, 15d-e; 19c) and [j] (14r, 15s; 18g, 19j), and examples of Low and Rising verbs with C1 [kp] (12c-e, 13h; 18c). So there is no way to segment the data into smaller chunks than done in Table 22, to arrive at a complementary distribution of verb root class based on consonant type.

Yet a generalization expected by a simple [L/Voice] analysis covers a large portion of the Ndaka data, so it would seem inappropriate to reject [L/Voice] altogether. In terms of the predictions made by Bradshaw (1999)'s [L/Voice] in (8), the prediction in (8a), that **all** C1's specified for voice are engaged in CTI (the

Rising verb root class), is upheld. On the other hand, the prediction in (8b), that **only** C1's specified for voice are engaged in CTI (the Rising verb root class), is not. This last is because some verbs with C1 **not** specified for voice are found in the Rising verb root class. This tension between the fact that [L/Voice] covers a large part of the Ndaka data without covering all of them (or that one [L/Voice] prediction is upheld while the other is not), will be addressed in Chapter 7.

One can also draw a number of observations regarding the Ndaka **infinitive** verb forms in isolation. First, the contrast seems limited to the root syllable, even in the surface form. That is, the first and third syllables remain constant across the three infinitive tone patterns: the prefix syllable is always low, and the final syllable is always falling from a mid/low level. The stability of pitch on the first and third syllables of infinitive forms may follow from the lack of contrast in Ndaka prefixes and final vowels. Still, the observed stability of first and third syllables on infinitive forms indicates that there are not (so far observed) processes in Ndaka that bring changes to the surface tone of the first or third syllables on these words.

One gap that remains to be explained is the lack of a second depressor consonant verb root class by tone. That is, the infinitive form of the Rising verb root class (in Figure 30) could easily be taken for the result of a C1 depressor on a High root, but then we lack the result of a C1 depressor on a low root. Conversely, the analysis of conjugated forms in Section 3.5 indicates that the Rising verb root class has an underlying specification of two low tones. Assuming this arose as the result of a depressor on a low root, then we lack the result of a depressor on a high root[9].

It is possible that the lack of a second depressor consonant verb root class by tone is accidental, that low verbs with depressor consonants exist, though not in my database. Or there may be a number of reasons for this to be systematic. Voicing has been lost in other languages of the area (c.f., Kutsch Lojenga 2006), and it is possible that a sound shift has resulted in a loss of contrastive voicing in C1 position on verbs with low tone, in a manner than precluded CTI (see Chapter 7 for details). Alternatively, it is possible that verbs with depressor consonants have at some point been reanalyzed as high tone verbs, to highlight the contrast. Either of these hypotheses would need to deal with the fact that depressors do not seem to impact Low tones in Ndaka.

[9] Given that I have proposed tonally unspecified High roots in Section 3.5, one could assume that the result of a depressor on a high root would be identical to a low root without a depressor, i.e., an additional low tone added to nothing. However, this would lead one to expect to find Low forms with depressor consonants, which would then correspond with High forms in other languages. But there are no depressor consonants on Ndaka Low verbs, as reported in Section 3.1. One could further imagine that all of these depressors devoiced, though there is no evidence of this, either. As a result, this question will remain unanswered at least until more relevant data can be collected.

One final comment regards the nature of depressor consonants as a phonemic entity. The argument may be made that the lowering of pitch is purely phonetic, as pitch obligatorily lowers with voiced obstruents. But such a hypothesis would predict multiple depressor effects where there are multiple depressor consonants. Such a prediction turns out to be false, as shown in [kɔbɨba] [– ⟋ ⌐] (spectrogram and pitch trace in Figure 29), which has a second depressor consonant without any impact on the final vowel pitch. So Ndaka infinitive verbs are **limited to one depressor effect per word**. They are also limited to one lexical tone per word, and these may be related. I have not considered multiple hypotheses regarding the manner of association of the depressor consonant tone feature, but it is clear that only the C1 associates to a TBU on infinitive verbs.[10]

In conclusion, the most salient point coming from this chapter for this dissertation is that the three verb root classes by tone observed in Ndaka largely fit an [L/Voice] analysis as in Bradshaw (1999), but not strictly so. The implications of this near complementary distribution (as in Table 22), will be taken up again in Chapter 7.

[10] Nouns, which have two lexical morae on CVCV roots, exhibit two depressor effects per root, at least for Nyali-Kilo [nlj]. This indicates strongly that the limitation has to do with the number of underlying and lexical tone categories there are for the depressor consonants to interact with.

Chapter 4
Mbo [zmw]

This chapter presents Mbo data collected during the course of my fieldwork. It is organized according to verb root class by tone, with the aim of characterizing those verb root classes. This characterization will include the contrast between them, as well as evidence for the distribution of consonant types in each verb root class by tone.

The data in this chapter come from my own fieldwork notes and recordings, each of which were made in Nyanya, DRC, during a series of workshops to help the Ndaka and Mbo communities make advances in the development of their respective writing systems. The workshops were held December 1–15, 2006, May 20–June 6, 2014, and August 2–26, 2016. The workshops were held in a participatory manner (Kutsch Lojenga 1996), to engage the fullest participation of the community in the analysis of their language and development of their writing system. Individual participants are listed in the acknowledgements on page vi. There are, unfortunately, a number of words which were sorted by infinitive form, but which became detached from their glosses. I have marked these glosses "no data" ("n.d.") until I can return to the language area to confirm them.

Available Mbo data vary somewhat from the Ndaka data presented in Chapter 3, but mostly in detail; we see the same basic story of three verb root classes by tone, each with a set of infinitive and conjugated tone melodies as in Table 23; recall that conjugated data (here and in Sections 4.1, 4.2, and 4.3) are presented in the tone frames of the form <pronoun|verb|adverb>, both to provide a frame of reference for relative pitch, and to control for the four paradigmatic verb root class by tone combinations observed in this data.

	inf	2s-past	2s-future	3s-past	3s-future
Low	[− − ˎ]	[¯│_ ¯ −│∨]	[−│− ¯ −│∨]	[ˊ│− − −│∨]	[ˊ│¯ ¯ −│⌄]
High	[_ ¯ ¯]	[¯│− ¯ ¯│∨]	[⁄│_ − −│∨]	[ˊ│¯ ¯ ¯│⌄]	[ˊ│¯ ¯ ¯│∨]
Rising	[¯ _ ¯]	[¯│_ ¯ _│∨]	[⁄│_ ¯ −│⁄]	[ˊ│¯ _ _│∨]	[ˊ│¯ _ ⁄│∨]

Table 23. Mbo [zmw] infinitive and conjugated tone melodies according to verb root class by tone

The Low, High, and Rising verb root classes will be described in further detail in Sections 4.1, 4.2, and 4.3, respectively. These three Mbo verb root classes by tone can be seen in the infinitive and conjugated tone melodies across the columns of Table 23, which will be justified by row, one in each section. The names for the verb root class by tone (in the left column of Table 23) are established in correspondence with the Ndaka verb root classes by tone, rather than as any indication of an analysis of underlying form. Correspondences between Ndaka and Mbo verb root classes by tone will be discussed in more detail in Section 4.3.

Each section (after the first) contains a subsection comparing its tone patterns with those of previous sections, such that each section will show a clear unity for its lexical verb root class by tone across infinitive and conjugated forms, as well as contrast with other verb root classes by tone. These comparisons culminate in Section 4.4, which contains a summary comparison of all tone patterns, with visual comparisons in composite figures.

Each section will also summarize the consonants that are found in that verb root class by tone, and how they correspond to expectations based on Bradshaw (1999). The distribution of consonant types across tone patterns will be discussed in Section 4.5, but the implications of consonant distributions across verb root classes by tone will not be addressed until Chapter 7.

4.1 Mbo [zmw] Low Verbs

As described in Section 3.1 for Ndaka, and more generally and in more detail in appendix B, one can show the trends of a pitch melody by superimposing multiple pitch traces into a single image. For the Mbo Low verb infinitive melody, this results in what could be described as low and level across the verb:

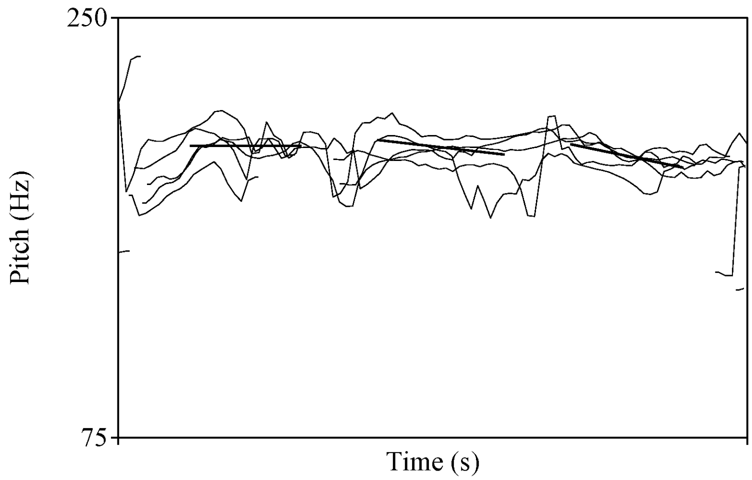


Figure 51. Overlay of infinitive pitch traces for Mbo [zmw] Low verbs [kɔɸɨnda] [‒ ‒ ‒] 'pay', [kotijo] [‒ ‒ ‒] 'be twisted', [kolijo] [‒ ‒ ‒] 'strain', [kosiɓo] [‒ ‒ ‒] 'wait', and [koɗijo] [‒ ‒ ‒] 'oil self'

The pitch trace overlay in Figure 51 can be transcribed in an idealized pitch trace as [‒ ‒ ¬].

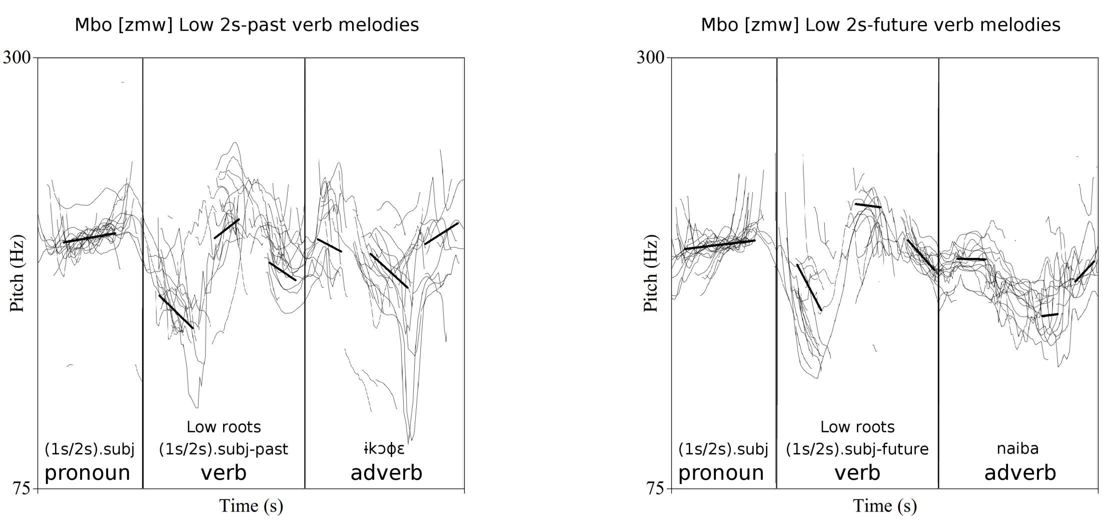


Figure 52. 2s past and future pitch traces for Low verbs in Mbo [zmw]

Distinguishing the two melodies in Figure 52, which might both be described as LHL, is aided by the use of the frames. The second syllable high is at about the level of the free pronoun in the past form (on the left), but it is well above that level in the future form (on the right). Similarly, the low third syllable is at the level of the initial high tone of the following adverb in the future form on the right, but it is below that level in the past form on the left. There is therefore something like an extra downstep (downward register shift) in the future form. As a result, the pitch trace overlays in Figure 52 can be transcribed in idealized pitch traces as [¯|_ ¯ –|∨] and [–|– ¯ –|∨].

Unlike the forms in Figure 52, the forms in Figure 53 are preceded by a pronoun with a (3s) low-high tone melody, and begin with a high TBU:

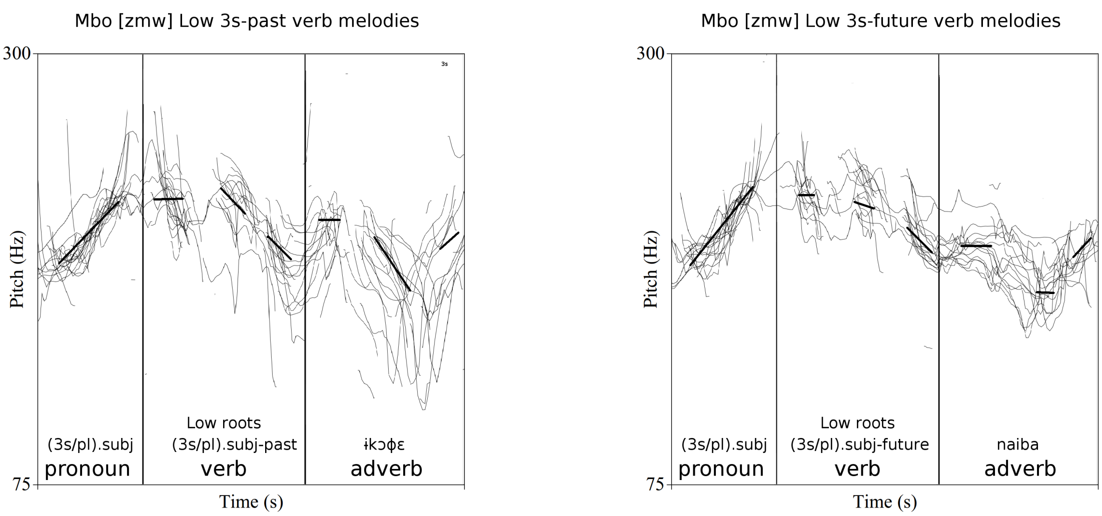


Figure 53. 3s past and future pitch traces for Low verbs in Mbo [zmw]

The melodies in Figure 53 are similarly difficult to distinguish, as they might both be described as HHL. The key difference is the initial high level of the following adverb, which is at the level of the verb final low in the future form (on the right), but above that final low level of the verb in the past form (on the left). There is thus again something like an extra downstep (downward register shift) in the future form. The pitch trace overlays in Figure 53 can thus be transcribed in idealized pitch traces as [′|¯ ¯ –|∨] and [′|¯ ¯ –|⌄].

As a result, the Low verb class tone melodies can be summarized as in Table 24, justifying the first row of Table 23 on page 122:

<table>
<tr><th>inf</th><th>2s-past</th><th>2s-future</th><th>3s-past</th><th>3s-future</th></tr>
<tr><td>[− − ¬]</td><td>[¯|_ ¯ −|∨]</td><td>[−|– ¯ −|∨]</td><td>[ʹ|¯ ¯ –|∨]</td><td>[ʹ|¯ ¯ –|ˇ]</td></tr>
</table>

Table 24. Mbo [zmw] Low verb infinitive and conjugated tone melodies

Low tone verbs in Mbo [zmw] include the following:

(31) Mbo [zmw] Low verbs (no segments specified for voice)
- a. koɗijo [− − ¬] oil (self)
- b. kopingo[− − ¬] wrap
- c. kɔtɛndɑ [− − ¬] walk
- d. kokuto [− − ¬] close
- e. kɔkpɑkpɑ [− − ¬] be blocked
- f. kɔkpɑtɑ [− − ¬] follow
- g. kokpeɲo [− − ¬] chase
- h. kɔɸʉwɑ [− − ¬] spoil
- i. kɔsɨkɑ [− − ¬] take
- j. kosiɓo [− − ¬] wait
- k. komiɲo [− − ¬] make pottery
- l. kɔmɔmɑ [− − ¬] gather
- m. kɔnɔtɑ [− − ¬] give birth
- n. kɔɲɛkɑ [− − ¬] smile
- o. kolindo [− − ¬] avoid escape
- p. kɔkwɑkɑ [− − ¬] cut chop
- q. kɔkwɑːnɑ [− − ¬] cohabitate
- r. kɔkwɛtɑ [− − ¬] bite
- s. kokwejo [− − ¬] wrap (package)

The following verbs are also taken to be Low on the basis of their infinitive forms, though not in conjugation data:

(32)
- a. kɔɸɨndɑ [− − ¬] pay
- b. kotijo [− − ¬] be twisted
- c. kolijo [− − ¬] strain

The consonants found in C1 position of the verbs in (31) and (32) are all implosives (31a), voiceless obstruents (31b-j; 32a-b), nasals (31k-n), other sonorants (31o;32c), or obstruents modified by sonorants without a voicing contrast (31p-s).

There are no C1 consonants specified for voicing in the Low verbs (c.f. Section 2.2.4), so the Low verb root class fits well as a tone pattern without depressor consonants under Bradshaw (1999).

4.2 Mbo [zmw] High Verbs

Mbo also has a High verb root class, with infinitive melody as exemplified in Figure 54. Comparing this with Low Mbo verbs in Figure 51, there are two syllables which are high in Figure 54, but low in Figure 51. So unlike the contrast on a single root syllable in Ndaka, the High/Low contrast for Mbo is spread over the final two of the three syllables in Figure 54:

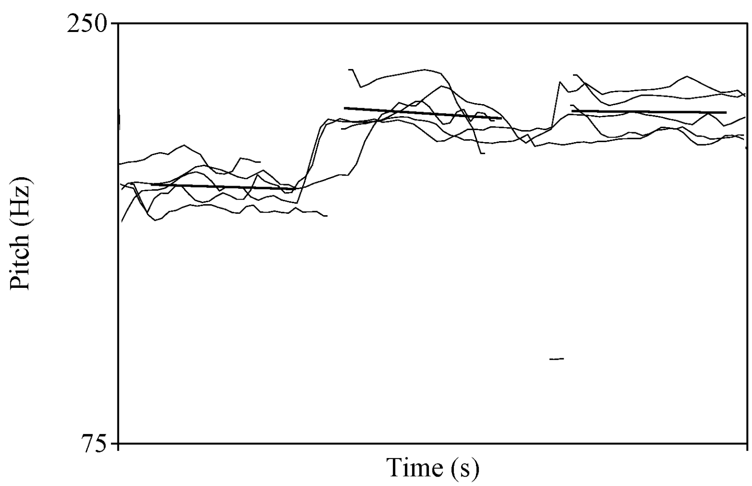


Figure 54. Overlay of infinitive pitch traces for Mbo [zmw] High verbs [kosiso] [– ¯ ¯] 'move forward', [koɓeɲo] [– ¯ ¯] 'wink', [koɓundo] [– ¯ ¯] 'break', [koweso] [– ¯ ¯] 'knock down', and [kɔtɨna] [– ¯ ¯] 'harvest'

The pitch trace overlay in Figure 54 can be transcribed in an idealized pitch trace as [– ¯ ¯].

The 2s past and future tone patterns for High verbs in Mbo are given in Figure 55:

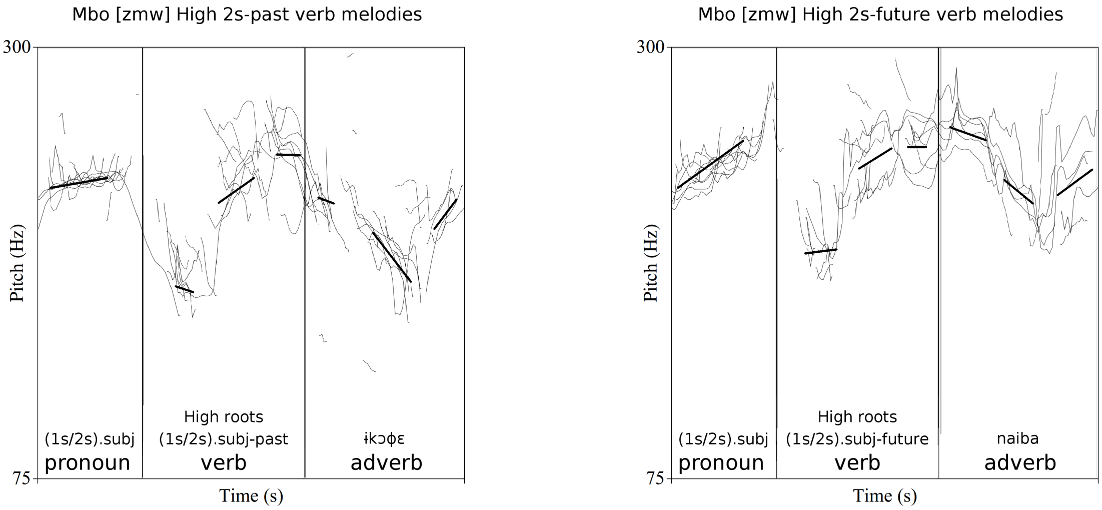


Figure 55. 2s past and future pitch traces for High verbs in Mbo [zmw]

While the tone patterns in Figure 55 might each be described as LHH, the past form (on the left) ends above the initial high level of the following adverb, whereas the future form (on the right) ends below it. The free pronoun before the verb is also rising in the future form (on the right). As a result, the pitch trace overlays in Figure 55 can be transcribed in idealized pitch traces as [˺|˗ ˉ ˉ|˅] and [ˊ|_ ˉ ˉ|˅].

Unlike the forms in Figure 55, the forms in Figure 56 are preceded by a pronoun with a (3s) low-high tone pattern, and begin with a high TBU:

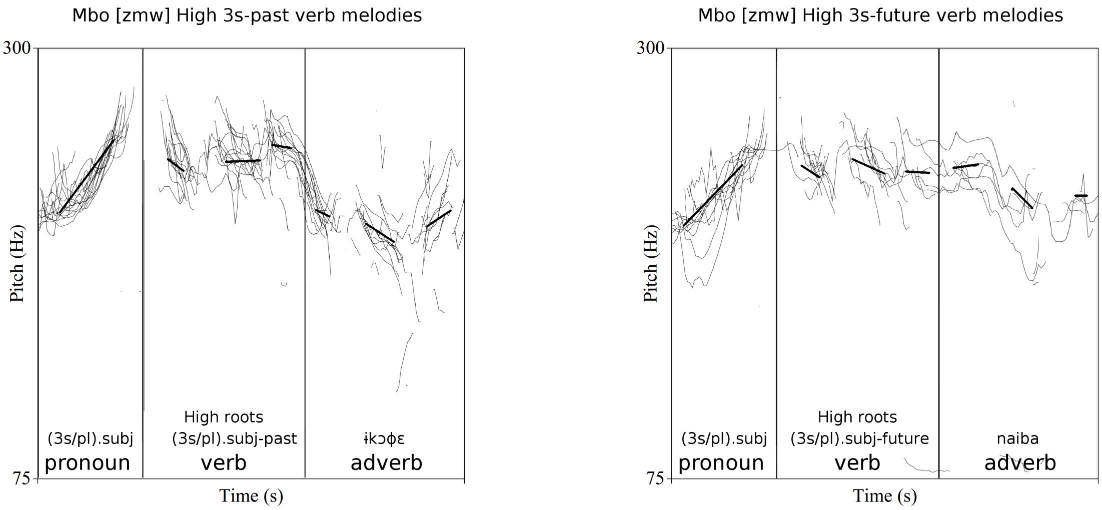


Figure 56. 3s past and future pitch traces for High verbs in Mbo [zmw]

Not only do the forms in Figure 56 begin high, but they end low. And again, while they may both be described as HHH, the past form (on the left) ends above the initial high level of the following adverb, whereas the future form (on the right) ends at the same level. This indicates a difference between past and future forms being a register shift, e.g., a final floating low tone in past forms.

As a result, the pitch trace overlays in Figure 56 can be transcribed in idealized pitch traces as [ˊ|¯ ¯ ¯|˅] and [ˊ|¯ ¯ ¯|ˇ], leading to the High verb class tone melodies as summarized in Table 25, justifying the second row of Table 23 on page 122:

inf	2s-past	2s-future	3s-past	3s-future
[– ¯ ¯]	[¯|– ¯ ¯|ˇ]	[ˊ|_ ¯ ¯|ˇ]	[ˊ|¯ ¯ ¯|˅]	[ˊ|¯ ¯ ¯|ˇ]

Table 25. Mbo [zmw] High verb infinitive and conjugated tone melodies

High tone verbs in Mbo [zmw] include the following:

(33) Mbo [zmw] High verbs (no segments specified for voice)
- a. koɓoo [– ¯ ¯] push open
- b. koɓundo [– ¯ ¯] break
- c. kɔɓɛndɑ [– ¯ ¯] hit
- d. kɔɓɛtɑ [– ¯ ¯] shell (peanuts)

e. koɗejo [– ¯ ¯] lick
f. kɔkɛkɑ [– ¯ ¯] crow (rooster)
g. kɔkʉndɑ [– ¯ ¯] love
h. kɔtɑmbɑ [– ¯ ¯] play
i. kɔtɨnɑ [– ¯ ¯] cut harvest
j. kɔtʉwɑ [– ¯ ¯] return
k. kotito [– ¯ ¯] push
l. kotoko [– ¯ ¯] spit
m. kotoo [– ¯ ¯] urinate
n. kotuo [– ¯ ¯] forge
o. komvono [– ¯ ¯] smell

The following verbs are also taken to be High on the basis of their infinitive forms, though they are not confirmed by conjugation data:

(34) a. koɓeɲo [– ¯ ¯] wink
b. koɓundo [– ¯ ¯] break
c. kosiso [– ¯ ¯] move forward
d. koweso [– ¯ ¯] knock down

The consonants found in C1 position of the verbs in (33) and (34) are all implosives (33a-e; 34a-b), voiceless obstruents (33f-n; 34c), other sonorants (34d), or obstruents modified by sonorants without a voicing contrast (33o). There are no C1 consonants specified for voicing in the High verbs (c.f. Section 2.2.4), so the High verb root class also fits well as a tone pattern without depressor consonants under Bradshaw (1999).

4.2.1 Comparing High and Low Verbs

Having given tone melodies characteristic of the High and Low verb root classes, I now turn to comparing them, to show how they are distinct one from the other. The trendlines from 2s Low verbs in Figure 52 and from 2s High verbs in Figure 55 are extracted and superimposed in Figure 57:

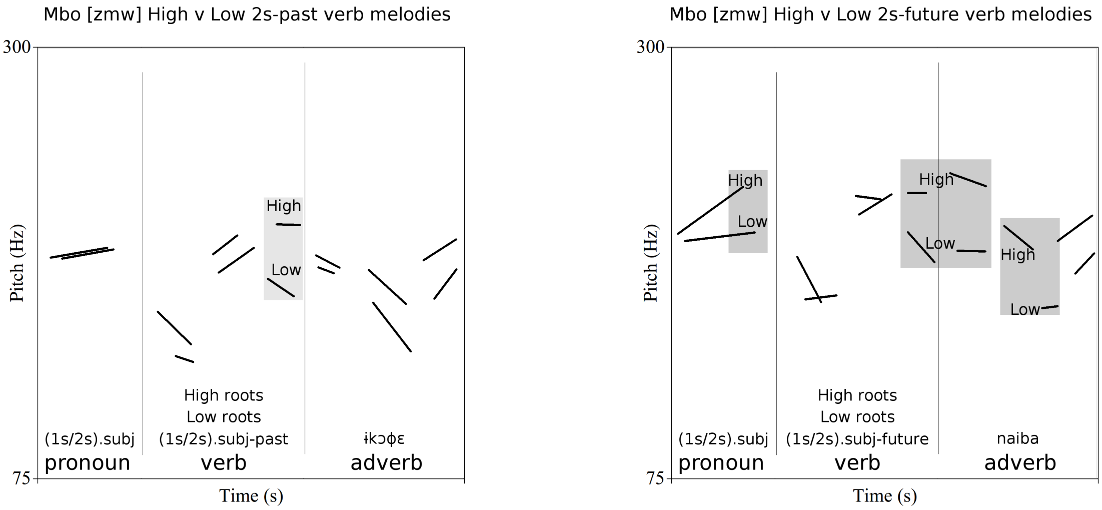


Figure 57. 2s past and future trendlines for Low and High verbs in Mbo [zmw]

The contrast between High and Low in the 2s-past forms (on the left) is essentially only on the final syllable of the verb, with High verbs higher than Low verbs. The contrast between High and Low in the 2s-future forms (on the right) spreads into the first two syllables of the following adverb, what may be a register shift over those syllables, with High verbs higher than Low verbs on each of these TBU's. One other contrast to note is on the pronoun before the verb. This is notable because the tone on free pronouns is very consistent for Ndaka (c.f., Chapter 3), so the comparative rise in the free pronoun for High verbs on the right in Figure 57 represents something of an anomaly, in terms of the data seen up to this point. This distinction may be related to what will be seen in Section 4.3 for the Rising verb infinitive form (c.f. Figure 59), where a prefix is high before a depressor consonant.

Whatever the source of the various contrasting TBU's in Figure 57, it is clear that the Low and High tone patterns in Mbo are not the same. This is also seen in Figure 58, where the trendlines from 3s Low verbs in Figure 53 and 3s High verbs in Figure 56 are superimposed and compared. They are again preceded by a pronoun with a (3s) low-high tone pattern, and begin with a high TBU:

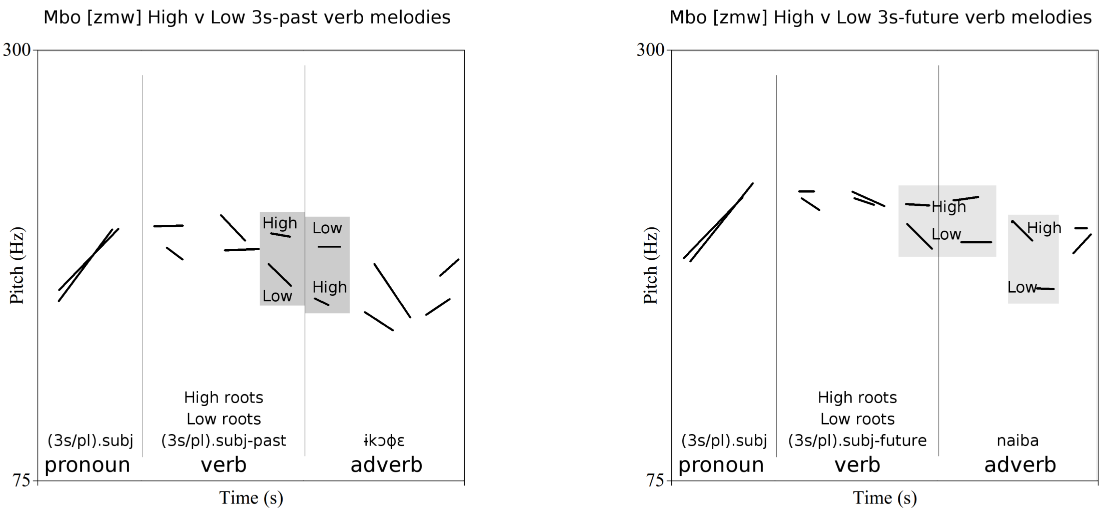


Figure 58. 3s past and future trendlines for Low and High verbs in Mbo [zmw]

The contrast between High and Low in the 3s past forms (on the left) is over the last syllable of the verb and the first syllable of the adverb, with the values opposite in each case: High verbs finish high, with a low initial adverb, whereas Low verbs finish low, with a high initial adverb. The contrast between High and Low in the 3s future forms (on the right) again spreads into the first two syllables of the following adverb, in what may be best analyzed as a register shift over those syllables. Again, while the details may be somewhat unclear, the contrast between the two sets of melodies is clear: High and Low verbs differ in all their infinitive and conjugated tone melodies.

4.3 Mbo [zmw] Rising Verbs

The tone patterns in Sections 4.1 and 4.2 are observed on most verbs where the root consonants are voiceless obstruents, implosives, sonorants, or obstruents modified by sonorants without voicing contrast. That is, verbs in the High and Low verb root classes contain consonants that Bradshaw (1999) predicts would not be depressor consonants. But where voiced obstruents are present in the first root syllable, the rising infinitive tone pattern is observed, as in Figure 59:

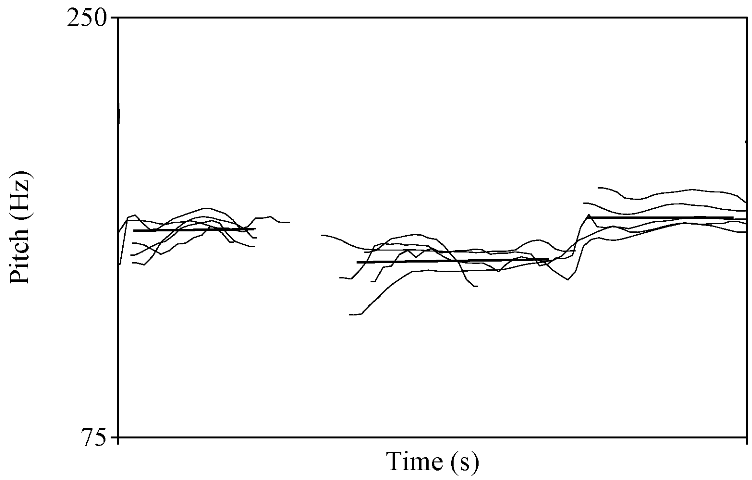


Figure 59. Overlay of infinitive pitch traces for Mbo [zmw] Rising verb root class verbs [kɔzɛnga] [¯ – ¯] 'be drunk', [kɔgʉnda] [¯ – ¯] 'embrace', [kɔgɔa] [¯ – ¯] 'snore', [kɔbanga] [¯ – ¯] 'build' and [kogiso] [¯ – ¯] 'throw away'

As a result, the pitch trace overlay in Figure 59 can be transcribed in an idealized pitch trace as [¯ – ¯]. Without a complete analysis of the underlying forms that result in this melody, one can still note that it is something of an anomaly. That is, if the prefix is high, it is the only high prefix in any closely related language, and thus represents a unique innovation. But if the prefix is low, with the root even lower, then the form in Figure 59 provides evidence of a downstepped low. As observed in Section 4.2.1, there may be an effect in Mbo of low tones causing prior TBU's to be high; this may be one with the contrast observed on the free pronoun in Section 4.2.1, where each high may result from a low tone on the following TBU.

On roots with long vowels, the same HLH sequence observed in Figure 59 covers only the first two syllables:

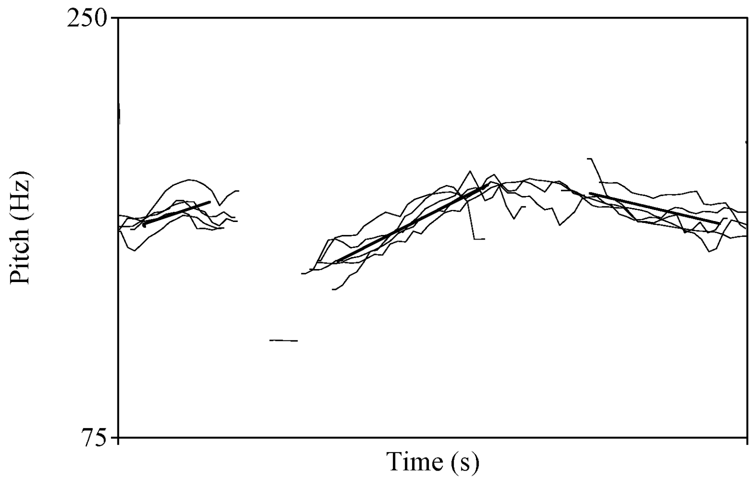


Figure 60. Overlay of infinitive pitch traces for Mbo [zmw] Rising verb root class verbs [kɔdʉːka] [¯ ⁄ ˋ], [kɔdʉːwa] [¯ ⁄ ˋ], [kodʒiːso] [¯ ⁄ ˋ], and [kodʒoːko] [¯ ⁄ ˋ]

I call this verb root class "Rising" to maintain the same verb root class by tone names for Mbo as used for Ndaka, despite the fact that their infinitive forms differ. That is, while the Ndaka Rising infinitive tone pattern rises from low to high on the second syllable (c.f. Table 6), the Mbo Rising infinitive second syllable lowers below the prefix low level (or else the prefix raises), for monomoraic roots. But as will be shown in Chapter 6, the Ndaka and Mbo Rising verb root classes[1], though strikingly different in their infinitive surface melodies, represent the same underlying group of verbs. This can be seen in comparing pairs such as Ndaka [kogiɲo] [– ⁄ –] 'refuse' from (16a) and Mbo [kogiɲo] [¯ – ¯] 'refuse' from (35e) In fact, the verbs in each of the Low, High, and Rising groups correspond for the most part with cognate verbs in the group of the same name in Ndaka, where cognates have been found, despite the fact that the surface form (of at least the infinitive forms) of the groups differs between the two languages. This will be seen perhaps more clearly in Chapter 6.

The 2s past and future tone patterns for Rising verbs are given in Figure 61:

[1] A similar observation can be made for the High verb root classes, where the high contrast spreads to the third syllable in the Mbo infinitive form, but not in that of Ndaka.

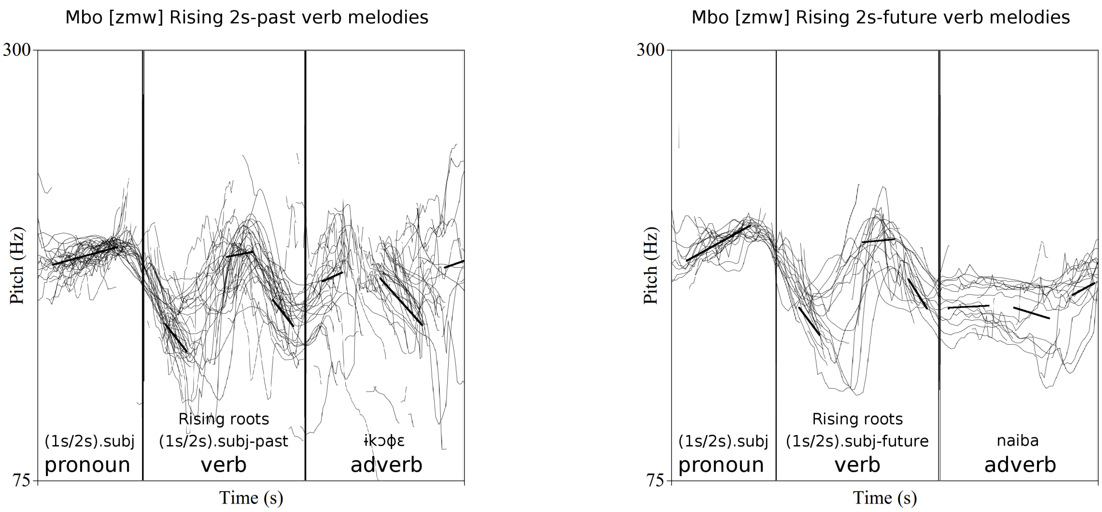


Figure 61. 2s past and future pitch traces for Rising verbs in Mbo [zmw]

Each of the forms in Figure 61 could be described as low-high-low, or perhaps low-high-mid, as the verb final TBU remains higher in pitch than the verb initial low TBU. But for the past form (on the left) the following adverb begins higher than the verb final TBU, whereas the future form (on the right) the adverb begins at the same level as the verb ends.

As a result, the pitch trace overlays in Figure 61 can be transcribed in idealized pitch traces as [⁻|_ ⁻ _|˅] and [˩|_ ⁻ –|˩].

Unlike the forms in Figure 61, the forms in Figure 62 are preceded by a pronoun with a (3s) low-high tone pattern, and begin with a high TBU:

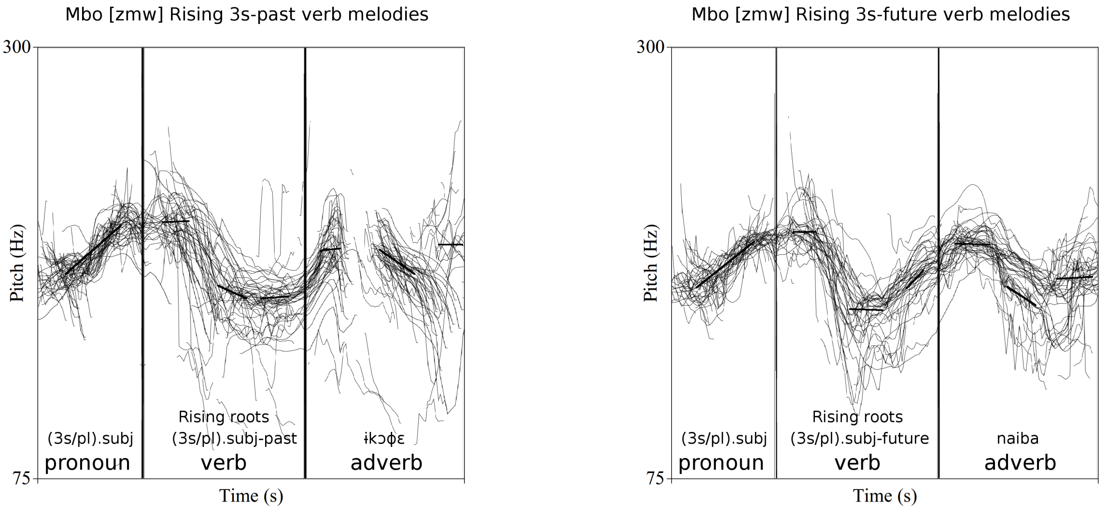


Figure 62. 3s past and future pitch traces for Rising verbs in Mbo [zmw]

The future form (on the right) also begins high and ends low on the last two TBU's, but rises up some on the last TBU before the following initial high adverb.

The past form (on the left) begins high and ends low on the last two TBU's, staying low before the following initial high adverb[2].

As a result, the pitch trace overlays in Figure 62 can be transcribed in idealized pitch traces as [ˊ|¯ _ _|∨] and [ˊ|¯ _ ⟋|∨].

The Mbo Rising tone melodies can be summarized as in Table 26, justifying the third row of Table 23 on page 122:

<table>
<tr><th>inf</th><th>2s-past</th><th>2s-future</th><th>3s-past</th><th>3s-future</th></tr>
<tr><td>[¯ _ ¯]</td><td>[¯|_ ¯ _|∨]</td><td>[⟋|_ ¯ _|⟋]</td><td>[ˊ|¯ _ _|∨]</td><td>[ˊ|¯ _ ⟋|∨]</td></tr>
</table>

Table 26. Mbo [zmw] Rising verb infinitive and conjugated tone melodies

The Mbo Low verbs (c.f. Section 4.1) and the High verbs (c.f. Section 4.2) contain no segments specified for voice, but the Rising verbs in Mbo [zmw] in Figure 61 and in Figure 62 include a number of segments which are specified for voice:

[2] This kind of word final flat pitch has been noted elsewhere (e.g., Section 3.5) as indicating a tone melody with a word final floating tone blocking lowering or assimilation to a following tone (i.e., which would be present in the past form but not in the future form of Figure 62.

(35) Mbo [zmw] Rising verbs (C1 specified for voice)
  a. kɔbɛtɑ [¯ – ¯] open (eye)
  b. kɔbɨjɑ [¯ – ¯] wipe
  c. kɔbɔbɑ [¯ – ¯] winnow flutter
  d. kɔdʉwɑ [¯ – ¯] pound (pestle)
  e. kogiɲo [¯ – ¯] refuse
  f. kogiso [¯ – ¯] throw away
  g. koguɓo [¯ – ¯] bend
  h. kɔgɔɑ [¯ – ¯] snore
  i. kɔgbɔkɑ [¯ – ¯] find
  j. kɔgbɔmɑ [¯ – ¯] bark
  k. kogʷɑjo [¯ – ¯] gather

The following verbs with C1 specified for voice are also taken to be in the Rising verb root class, based on infinitive forms:

(36) a. kɔzɛngɑ [¯ – ¯] be drunk
  b. kɔgʉndɑ [¯ – ¯] embrace
  c. kɔbɑngɑ [¯ – ¯] build
  d. kɔbɨbɑ [¯ – ¯] respect
  e. kɔbɑbɑ [¯ – ¯] carry
  f. kɔbɔbɑ [¯ – ¯] n.d.
  g. kɔgʲɑbɑ [¯ – ¯] n.d.
  h. kɔdʉːkɑ [¯ ˏ ˋ] n.d.
  i. kɔdʉːwɑ [¯ ˏ ˋ] n.d.
  j. kodʒiːso [¯ ˏ ˋ] n.d.
  k. kodʒoːko [¯ ˏ ˋ] n.d.

The presence of contrastive voicing in the C1 of all the data in (35) and (36) aligns nicely with Bradshaw (1999)'s prediction that consonant-tone interaction should be found with contrastively voiced segments, assuming that the Rising tone group is opposed to the High and Low groups on the basis of consonant-tone interaction. Under this assumption, Bradshaw (1999) accounts for a large amount of the Mbo verb tone data, including that in Section 4.1, Section 4.2, and (35) and (36).

Unfortunately for those seeking to make Bradshaw (1999) work for all Mbo data, Mbo's Rising tone group also contains other verbs. These do not have contrastively voiced consonants as the first consonant of the root, as in (37) and (38).

The first includes Mbo Rising verbs which have been verified by conjugated tone patterns:

(37) Mbo [zmw] Rising verbs (C1 NOT specified for voice)
- a. koɗuwo [˥ ˧ ˥] pull (up)
- b. kɔkɔfɑ [˥ ˧ ˥] cough
- c. kɔkpɑgɑ [˥ ˧ ˥] ferment
- d. kɔɸɑnɑ [˥ ˧ ˥] show
- e. kɔmɨjɑ [˥ ˧ ˥] swallow
- f. kɔnɑgbɑ [˥ ˧ ˥] have headache
- g. kɔjɑnɑ [˥ ˧ ˥] cook
- h. kɔwʉmɑ [˥ ˧ ˥] skin (an animal)
- i. konguso [˥ ˧ ˥] remove
- j. kouɸo [˥ ˧ ˥] blow air
- k. kɔʉnɑ [˥ ˧ ˥] sow (seed)

The following verbs are also taken to be in the Rising verb root class, based on infinitive forms:

(38) Other Mbo [zmw] Rising verbs (also with C1 NOT specified for voice)
- a. kɔtʉgɑ [˥ ˧ ˥] n.d.
- b. kosugbɑ [˥ ˧ ˥] emerge
- c. konɑːtɑ [˥ ˏ ˋ] n.d.

The consonants found in C1 position of the verbs in (37) and (38) are implosives (37a), voiceless obstruents (37b-d; 38a-b), nasals (37e-f; 38c), other sonorants (37g-h), or obstruents modified by sonorants without a voicing contrast (37i). Additionally, two verbs lack a consonant in C1 position (37j-k).

In summary, verbs in the Rising verb root class (and no others) include C1 segments which are specified for voice. This restriction on C1 segments argues in favor of the Rising verb root class being a result of consonant-tone interaction (Bradshaw 1999). But it is also the case that a number of Mbo Rising verbs include a C1 segment which is not contrastive for voice, causing something of a problem for a simple [L/Voice] analysis based on Bradshaw (1999). The impact of the data in (37) and (38) will be discussed more fully in Chapter 7.

4.3.1 Comparison of High and Rising Verbs

Having given tone melodies characteristic of the High and Rising verb root classes, I now turn to comparing them, to show how they are distinct one from the other. The trendlines from High 2s forms in Figure 55 and Rising 2s forms in Figure 61 are extracted and superimposed in Figure 63:

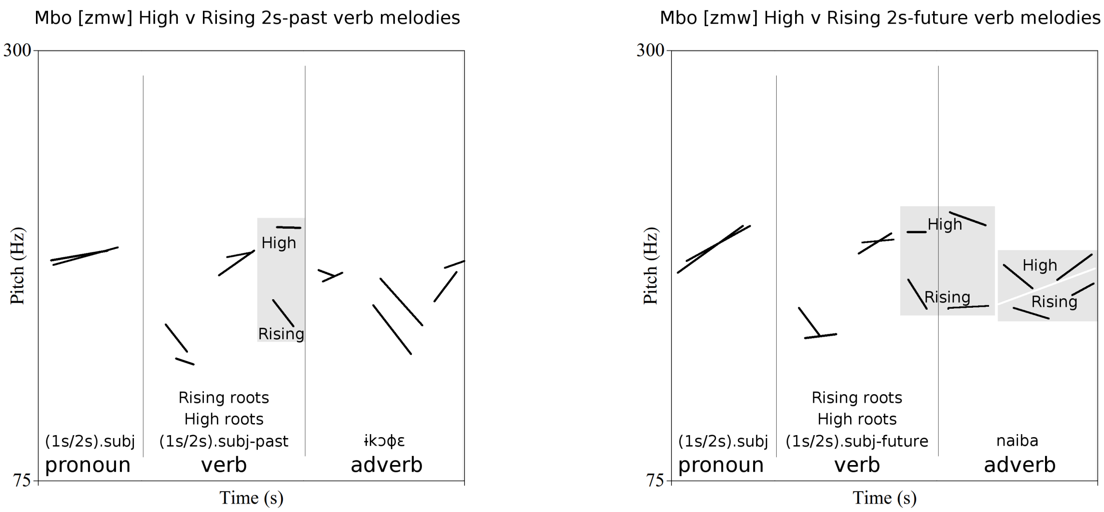


Figure 63. 2s past and future trendlines for High and Rising verbs in Mbo [zmw]

The contrasts in Figure 63 strongly resemble the contrasts between the High and Low classes in Section 4.2.1. The contrast between High and Rising in the 2s-past form (on the left of Figure 63) is essentially only on the final syllable of the verb, where High verbs are higher than Rising verbs. The contrast between High and Rising in the 2s-future form (on the right) spreads into the following adverb, in what may be best analyzed as a register shift over those syllables.

The trendlines from High forms in Figure 56 and Rising forms in Figure 62, which are again preceded by a pronoun with a (3s) low-high tone pattern, and begin with a high TBU, are compared in Figure 64:

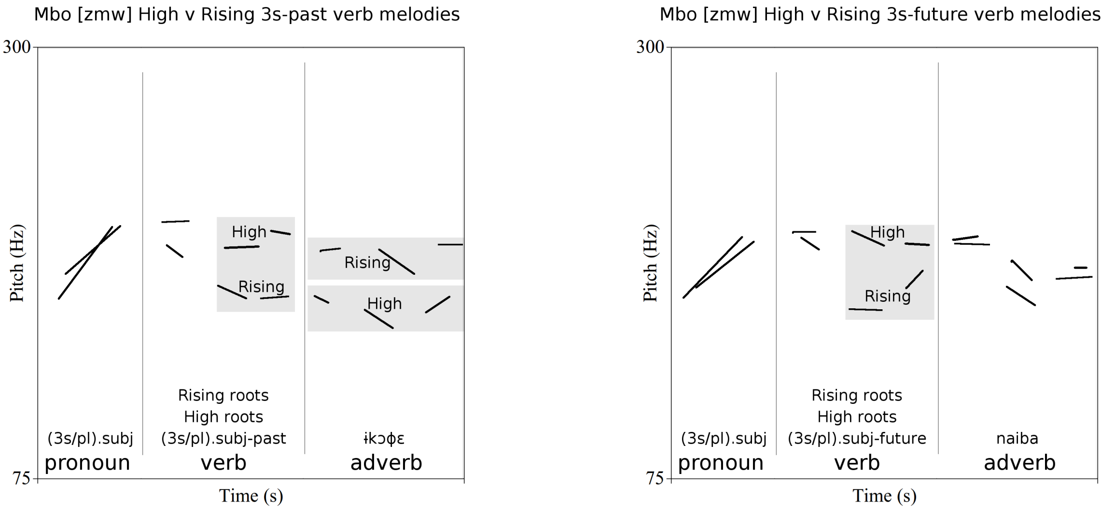


Figure 64. 3s past and future trendlines for High and Rising verbs in Mbo [zmw]

The contrast between High and Rising in the 3s past form (on the left) provides a pitch inversion at the word boundary: High verbs finish high, with a lower register on the adverb, whereas Rising verbs finish low, but with a higher register on the adverb. The contrast between High and Rising in the 3s future form (on the right) covers only the second two syllables of the verb, where High verbs are higher than Rising verbs.

Again, however these details in form may be analyzed, it is clear that the Mbo High and Rising verb root classes have different melodies, for each of the forms so far investigated.

4.3.2 Comparison of Low and Rising Verbs

This section compares the Low and Rising verb root classes, to show how they are distinct one from the other. The trendlines from Low 2s forms in Figure 52 and Rising 2s forms in Figure 61 are extracted and superimposed in Figure 65:

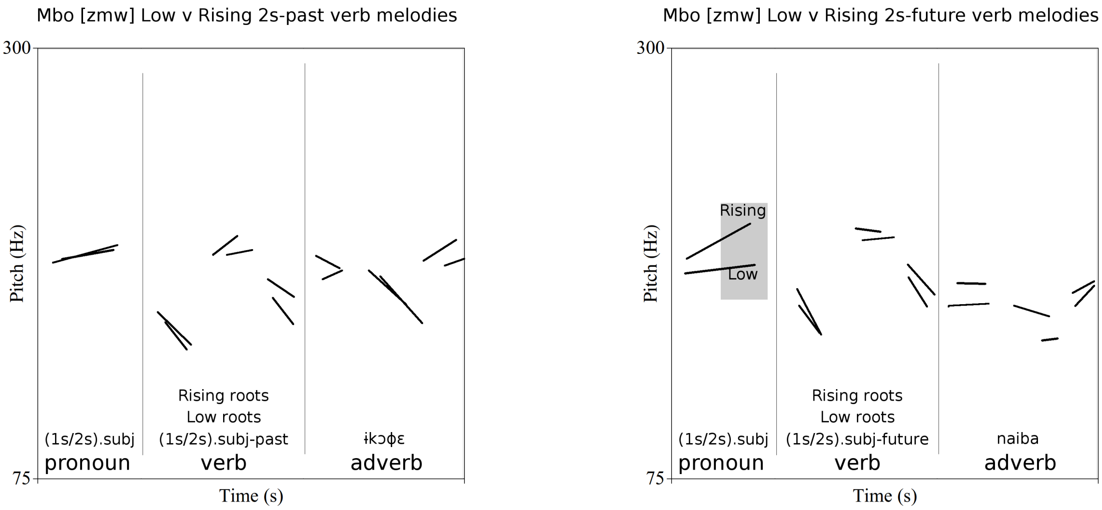


Figure 65. 2s past and future trendlines for Low and Rising verbs in Mbo [zmw]

There is essentially no contrast between Low and Rising in the 2s-past form (on the left, hence the lack of shading); the contrast between Low and Rising in the 2s-future form (on the right) is limited to the rise in tone on the free pronoun, as discussed briefly in Section 4.3.1. In this case, Rising verbs have a free pronoun which ends higher than those which precede Low verbs.

The trendlines from Low forms in Figure 53 and Rising forms in Figure 62, which are again preceded by a pronoun with a (3s) low-high tone pattern, and begin with a high TBU, are compared in Figure 66:

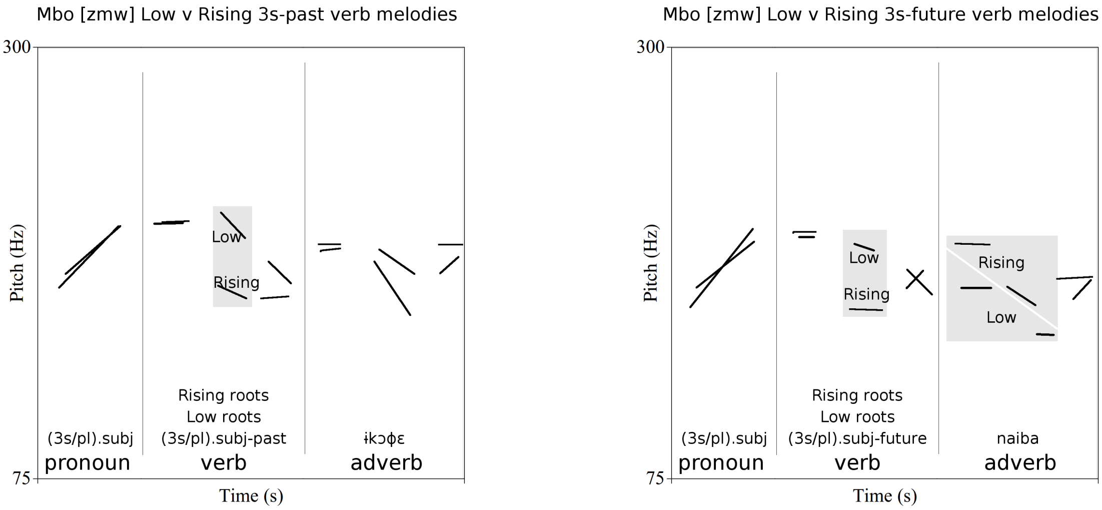


Figure 66. 3s past and future trendlines for Low and Rising verbs in Mbo [zmw]

The contrast between Low and Rising in the 3s-past form (on the left) is essentially only on the second syllable of the verb, the root TBU. The contrast between Low and Rising in the 3s-future form (on the right) again inverts at the word boundary, with Low verbs higher on the root TBU, followed by a lower pitch on the adverb, and with Rising verbs with a lower root TBU, followed by a higher pitch on the adverb.

For three of the conjugated forms, then, there is a clear contrast between the tonal melodies observed on Mbo Low and Rising verbs; this is in addition to the difference in their infinitive tone melodies.

4.4 Summary Comparison of Mbo [zmw] verb root classes by tone

The following figures show a three way comparison between the Mbo verb root classes, for each position in the conjugated data. The Low verb root class data from Figure 52 are superimposed on the High verb root class data from Figure 55 and the Rising data from Figure 61, resulting in Figure 67:

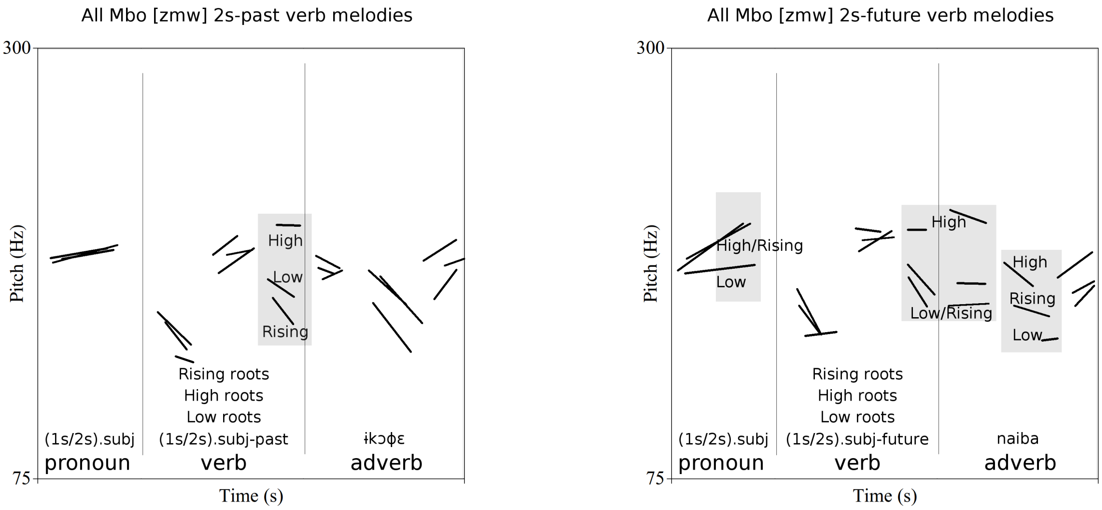


Figure 67. 2s past and future trendlines for Low, High and Rising verbs in Mbo [zmw]

The past forms (on the left) show a contrast for verb root class by tone on the final vowel (FV) only, where High is higher than Low and Rising (which may be lower still). The future forms (on the right) show High and Rising roots opposed to Low roots in the operation of the polar prefix (spread to the pronoun, c.f. Section 4.2.1). At the end of the word, however, High is opposed to Low and Rising. This indicates that the three verb root classes by tone are distinct lexically, however that distinction works out in terms of underlying forms.

The Low verb root class data from Figure 53 are superimposed on the High verb root class data from Figure 56 and the Rising data from Figure 62 (each of which are again preceded by a pronoun with a (3s) low-high tone pattern, and begin with a high TBU) to form Figure 68:

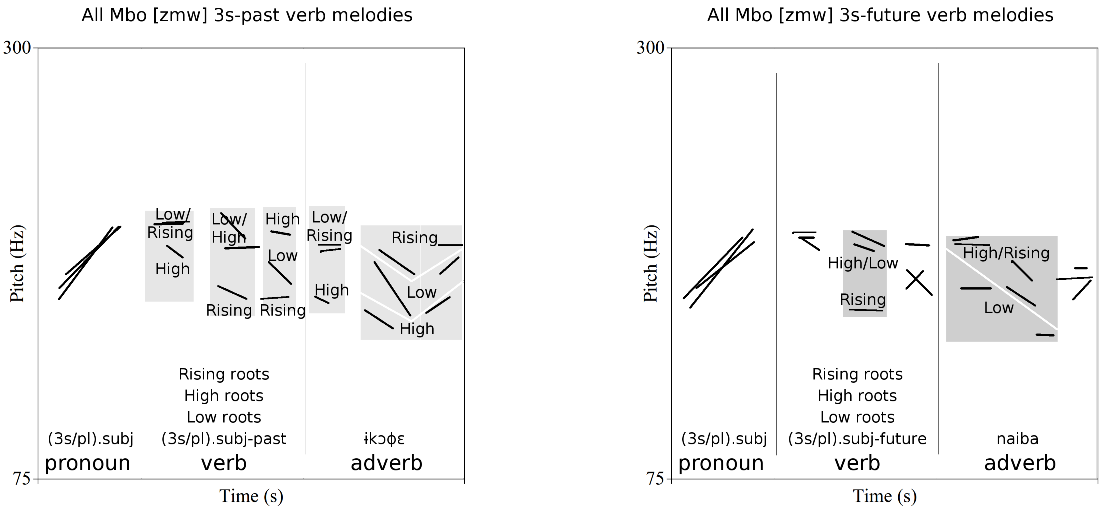


Figure 68. 3s past and future trendlines for Low, High and Rising verbs in Mbo [zmw]

The past form (on the left) shows Low and High verbs higher than Rising verbs on the root TBU, but High verbs alone are higher than Low and Rising verbs on the FV. On the adverb, High verbs are lower than either Low or Rising verbs. The future form (on the right) shows Rising as lower than both High and Low verbs on the root TBU, and Low verbs lower than High and Rising verbs on the following adverb.

Despite the difficulty of finding clear generalizations that apply through all the data in Sections 4.1–4.3, the data clearly show three distinct verb root classes by tone, by both infinitive and conjugated tone patterns.

Furthermore, two of the three verb root classes by tone (High and Low) have only verbs with C1 not specified for voice, while the third (Rising) has verbs with C1 specified for voice, but also verbs with C1 not specified for voice. These fundamental facts mirror exactly the data presented for Ndaka in Chapter 3. The implications of these generalizations will be discussed briefly in Section 4.5, but more fully in Chapter 7.

4.5 Discussion of Mbo [zmw] Data

This section discusses a number of generalizations that come out of the Mbo data presented in Sections 4.1–4.3. Perhaps the most significant is that there are three verb root classes by tone, two of which (High and Low) have verb roots without any C1's which are specified for voicing. The third (Rising) has all the Mbo roots with C1 specified for voicing, which indicates a complementary distribution between one verb root class by tone which derives from [L/Voice] specification and two that lack it (Bradshaw 1999). Such a generalization covers all Mbo High and Low verbs, and 11 of 22 Rising verbs, for a total of 57 of the 68 Mbo verbs for which I have conjugation data (i.e., 84%). But my data set also includes the other 11 of 22 Rising verb root class verbs (i.e., 16% of all Mbo verbs in my conjugated data set) with C1 that is not specified for voicing, indicating that however strong the indication of complementary distribution is, the distribution of C1 types across tone patterns is not in fact complementary, but rather contrastive, as indicated in Table 27:

verb root class	C1 without [L/Voice]	C1 with [L/Voice]
Low	30	-
High	16	-
Rising (CTI)	11	11

Table 27. Summary of distribution of consonants across melodies in Mbo (numbers for conjugated data only)

The presence of the data in the (darker shaded) lower left corner prevents an analysis of complementary distribution, which would be necessary to allow a straightforward and synchronic [L/Voice] analysis to work. This is not only the case when examining larger categories as in Table 27 (i.e., with or without voicing specification). One is similarly forced to analyze the Mbo data as contrastive on the basis of particular segments. There are examples of Low, High and Rising verbs with a variety of C1 segments not specified for voicing, including [ɗ] (31a; 33e; 37a), [k] (31d; 33f-g; 37b), [t] (31c, 32b; 33h-n; 38a), and [s] (31i-j; 34c; 38b). There are also examples of High and Rising verbs with C1 [w] (34d; 37h) and examples of Low and Rising verbs with C1 [kp] (31e-g; 37c), [ɸ] (31h, 32a; 37d), [m] (31k-l; 37e), and [n] (31m; 37f, 38c). So there is no way to segment the data into smaller segment type chunks than done in Table 27, to arrive at a complementary distribution of verb root class by tone based on consonant type.

Yet a generalization expected by a simple [L/Voice] analysis covers a large portion of the Mbo data, so it would seem inappropriate to reject [L/Voice] altogether. In terms of the predictions made by Bradshaw (1999)'s [L/Voice] in (8), the prediction in (8a), that **all** C1's specified for voice are engaged in CTI (the Rising verb root class), is upheld. On the other hand, the prediction in (8b), that **only** C1's specified for voice are engaged in CTI (the Rising verb root class), is not. This last is because some verbs with C1 **not** specified for voice are found in the Rising verb root class. Resolving this tension between the fact that [L/Voice] covers a large part of the Mbo data without covering all of them (or that one [L/Voice] prediction is upheld while the other is not), will be addressed in Chapter 7.

Chapter 5
Nyali-Kilo [nlj]

This chapter presents Nyali-Kilo data collected during the course of my fieldwork. They are organized according to verb root class by tone, according to infinitive forms. Sections also include evidence for the distribution of consonant types in each verb root class by tone.

The data in this chapter are largely taken from Rasmussen (2012), which was itself written from my field notes and recordings made in a series of workshops to help the Nyali-Kilo, Vanuma, and Bɨra communities make advances in the development of their respective writing systems. The workshops were held in Bunia DRC, from December 1–8, 2006 and in Ibambi, DRC from June 9–26, 2009. The workshops were held in a participatory manner (Kutsch Lojenga 1996), to engage the fullest participation of the community in the analysis of their language and development of their writing system. Individual participants are listed in the acknowledgements on page vi. I have unfortunately been unable to maintain productive contact with this language community for follow-up work. As a result, the data presented here are limited to that presented in Rasmussen (2012), and do not include conjugated forms. Because these verb root classes by tone are based entirely on infinitive form data, I refrain from making too many claims based on them.

It may be helpful to note that for related languages Ndaka and Mbo, the conjugated data confirmed, but did not show any significant departures from, the infinitive tone melody groupings. That is, for at least two Bantu D30 languages, the infinitive tone melody seems to be a good predictor of conjugated forms. That being said, Nyali-Kilo has different infinitive tone melodies, so we cannot be certain than infinitive tone melodies absolutely correspond to lexical verb root classes by tone.

The grouping by infinitive verb forms is summarized in Table 28:

Low	[⁻ – ¬]
High	[– ⁻ ⁻]
Low with C1 Depressor	[⁻ – –]
High with C1 Depressor	[– ⁄ ⁻]
Low with C2 Depressor	[⁻ – _]
High with C2 Depressor	[– ⁻ ⁄]

Table 28. Nyali-Kilo [nlj] infinitive tone melodies

While the analysis of these tone patterns hasn't been fully worked out yet, I will use the term Rising to cover the final four rows in Table 28, because they represent the variety of tone patterns due to depressor consonants in Nyali-Kilo. These rows will be justified in the following sections.

5.1 Nyali-Kilo [nlj] Low Verbs

The Ndaka Low infinitive tone pattern is low and level, as for Ndaka and Mbo:

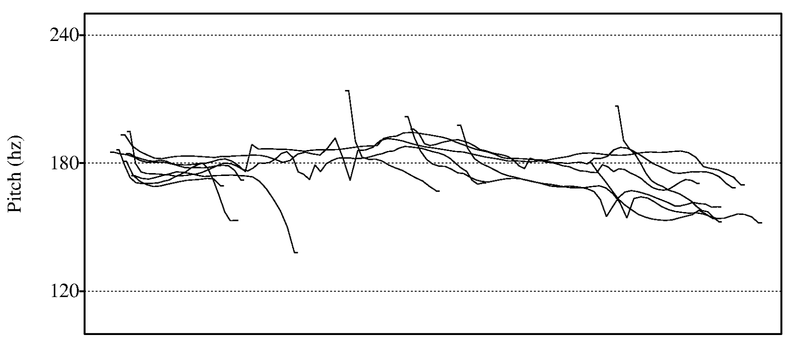


Figure 69. Nyali-kilo [nlj] overlay of pitch traces showing low and (mostly) level tones: [kasama] [⁻ – ¬] 'swim', [kamiɲo] [⁻ – ¬] 'squeeze', [kakʉta] [⁻ – ¬] 'bite', [kafʉla] [⁻ – ¬] 'break wind', [kak^{w}ala] [⁻ – ¬] 'sleep' [kaluso] [⁻ – ¬] 'take, bail out' and [kaliyo] [⁻ – ¬] 'flow'

As a result, the pitch trace overlay in Figure 69 can be transcribed in an idealized pitch trace as [⁻ – ¬]. The following verbs have this Low infinitive tone melody:

(39) Infinitive verb forms with Low roots (infinitive forms with [– – ¬] melody)

kɑɗoɗo	‘carve’
kɑkɨkɑ	‘act’
kɑkɑkɑ	‘ferment’
kɑkɑtɑ	‘be silent’
kɑkuɗo	‘crawl (lizard)’
kɑkuto	‘close’
kɑkʉtɑ	‘bite’
kɑkʉlɑ	‘buy’
kɑkumo	‘fear’
kɑk͡pɑk͡pɑ	‘become jammed’
kɑk͡pɑtɑ	‘ascend’
kɑpɑt͡ʃɑ	‘finish’
kɑsuɓo	‘protect’ ‘steer’
kɑsiɓo	‘care for’
kɑfʉlɑ	‘break wind’
kɑfɑnɑ	‘rest cheek in hand’
kɑfulo	‘spit’
kɑk͡pʉlɑ	‘look for’
kɑk͡peɲo	‘drive away’
kɑtɑmɑ	‘tie’
kɑtɑwɑ	‘cross’
kɑtɨ^{n}dɑ	‘walk’
kɑsɑmɑ	‘swim’
kɑsi^{ŋ}go	‘submerge’
kɑsʉwɑ	‘sing’
kɑmiɲo	‘squeeze’
kɑmɨlɑ	‘swallow’
kɑmejo	‘rain’
kɑmʉmɑ	‘gather’
kɑnʉtɑ	‘give birth’
kɑlɑ^{ŋ}gɑ	‘display’
kɑlijo	‘flow’
kɑlɨtɑ	‘stamp’
kɑluso	‘take’ ‘bail out’
kɑlʉwɑ	‘plait hair’
kɑwʉlɑ	‘spoil’

kɑwolo	‘count’
kɑk^{w}ɑlɑ	‘sleep’ ‘lie down’
kɑk^{w}ejo	‘pack’
kɑk^{w}ɑkɑ	‘cut’ ‘write’
kɑ^{n}dɨlɑ	‘cry’
kɑ^{n}d͡ʒo^{n}d͡ʒo	‘hang up’

The consonants found in C1 position of the verbs in (39) are all implosives, voiceless obstruents, nasals, other sonorants, or obstruents modified by sonorants without voicing contrast. There are no C1 consonants specified for voicing in the Low verbs.

5.2 Nyali-Kilo [nlj] High Verbs

The Nyali-Kilo High infinitive tone pattern resembles that of Mbo:

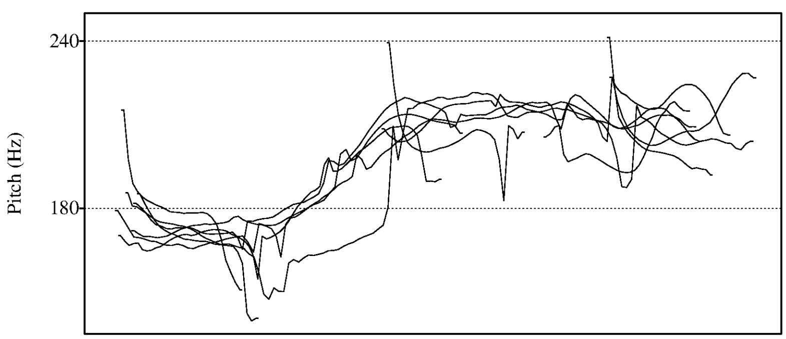


Figure 70. Nyali-kilo [nlj] overlay of pitch traces showing high roots: [katoko] [– ¯ ¯] ‘spit’, [katamba] [– ¯ ¯] ‘play’, [kanʉla] [– ¯ ¯] ‘kill’, [kanʉko] [– ¯ ¯] ‘hear’, [kaɓɨnda] [– ¯ ¯] ‘hit’, [kanɛna] [– ¯ ¯] ‘see’, and [kanɛka] [– ¯ ¯] ‘come’

As a result, the pitch trace overlay in Figure 20 can be transcribed in an idealized pitch trace as [– ¯ ¯]. The following verbs have this High infinitive tone melody:

(40) Infinitive verb forms with High roots (infinitive forms with [– ¯ ¯] melody)

kaɓɨta	‘shell (peanuts)’
kaɓala	‘sprout’
kaɓɨ^{n}de	‘hit’
kaɓolo	‘burst’
kaɓo^{ŋ}go	‘twist’ ‘wring’
kaɓʉla	‘come from’
kaɓu^{n}d͡ʒo	‘break’
kaɗejo	‘lick’
kakɨka	‘cackle’
kakuko	‘surround’
kakʉ^{n}da	‘agree’
katoko	‘spit’
katuto	‘push’
kata^{ŋ}ga	‘hesitate’
katɨja	‘construct’
katolo	‘urinate’
katulo	‘forge’
kata^{m}ba	‘play’
katɨna	‘cut’
kat͡ʃʉ^{ŋ}ga	‘accuse’
kasɨka	‘refuse’
kana^{m}ba	‘obstruct’
kanʉko	‘hear’
kanʉ^{n}da	‘shoot (weapon)’
kanuto	‘pour away’
kali^{n}do	‘avoid’
kaluko	‘knead’
kalʉta	‘indicate’
ka^{ɱ}vʉma	‘eat too much’
kaja^{ŋ}ga	‘be astonished’

The summary of these data, as pertains to this dissertation, is that none of the C1's are contrastively voiced, as with the Low forms in Section 5.1.

The consonants found in C1 position of the verbs in (40) are all implosives, voiceless obstruents, nasals, other sonorants, or obstruents modified by sonorants without voicing contrast. There are no C1 consonants specified for voicing in the High verbs.

5.3 Nyali-Kilo [nlj] Low with C1 Depressor Verbs

One of the C1 depressor verb root classes by tone in Nyali-Kilo has the following melody:

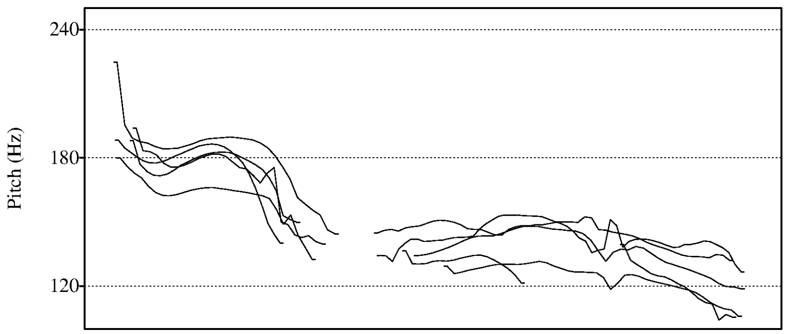


Figure 71. Nyali-kilo [nlj] overlay of pitch traces showing low with C1 depressors: [kabʉba] [¯ – –] ‘flutter’, [kadu$^{ŋ͡m}$gbo] [¯ – –] ‘go back’, [kadʉla] [¯ – –] ‘pound (pestle)’, [kagiɲo] [¯ – –] ‘abandon’ and [kagʉla] [¯ – –] ‘snore’

As a result, the pitch trace overlay in Figure 71 can be transcribed in an idealized pitch trace as [¯ – –]. The following verbs have this Low with C1 depressor infinitive tone melody:

(41) Infinitive verb forms with Low roots and C1 Depressors (infinitive forms with [¯ – –] melody)

a. kɑbʉbɑ flutter
b. kɑd͡ʒid͡ʒo ring
c. kɑdʉlɑ pound (pestle)
d. kɑdu$^{ŋ͡m}$g͡bo go back
e. kɑg͡bʉkɑ rejoin find
f. kɑg͡bejo compromise
g. kɑgiɲo abandon
h. kɑgʉlɑ snore

The summary of these data, as pertains to this dissertation, is that all of the C1's are contrastively voiced (the C2 may or may not be). There are no C1 consonants which are not specified for voicing in this class of verbs.

5.4 Nyali-Kilo [nlj] High with C1 Depressor Verbs

The other C1 depressor verb root class by tone in Nyali-Kilo has the following melody:

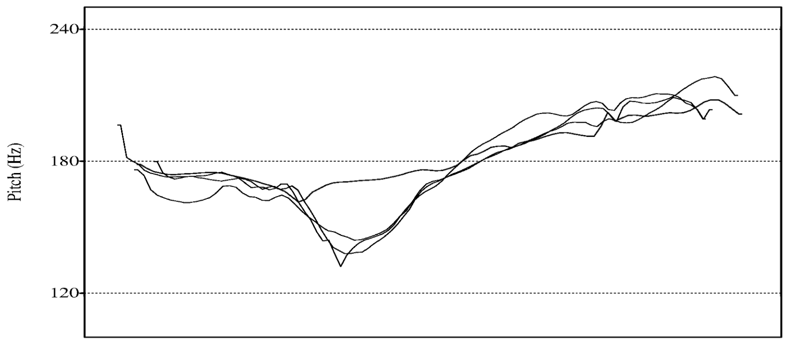


Figure 72. Nyali-kilo [nlj] overlay of pitch traces showing high roots with C1 depressors: [kawʉma] [– ⟋ ¯] ‘skin (an animal)’ and three tokens of [kawolo] [– ⟋ ¯] ‘weed (garden, field)’

As a result, the pitch trace overlay in Figure 72 can be transcribed in an idealized pitch trace as [– ⟋ ¯]. The following verbs have this High with C1 depressor infinitive tone melody:

(42) Infinitive verb forms with High roots and C1 Depressors (infinitive forms with [– ⟋ ¯] melody)

a.	kɑbolo	swing		
b.	kɑg͡bʉmɑ	bark		
c.	kɑguɓo	bend	wind up	
d.	kɑgʉɓɑ	take revenge		
e.	kɑgiso	fall short	fail	lack
f.	kɑvɨlɑ	take forcefully		
g.	kɑwʉkɑ	groan (of pain)		
h.	kɑwʉlɑ	be fully grown		
i.	kɑwʉmɑ	skin (an animal)		
j.	kɑwʉnɑ	sow (seed)		
k.	kɑwʉtɑ	cut		
l.	kɑwulo	blow (mouth)		
m.	kɑwolo	weed		

The summary of these data, as pertains to this dissertation, is that some of the C1's are contrastively voiced, while others (*w* only) are not. None of the C2 are contrastively voiced. There are no C1 consonants which are not specified for voicing other than *w* in this class of verbs.

5.5 Nyali-Kilo [nlj] C2 Depressor Verbs

There are two other tone patterns observed in Nyali-Kilo infinitive verbs, which I will mention only briefly for a number of reasons. First, as mentioned above, my contact with the community has not allowed for much continued investigation, including verification of any of these tone patterns with conjugated forms. But these verb root classes by tone are particularly concerning because they depend on C2 depressors, unlike anywhere else observed in Bantu D30. This being infrequent, there is minimal data in each group (five words each). In addition to the minimal number of words, I have only a single recording in each tone pattern, so it is difficult to show the generalization here, even where it seemed clear in my initial transcription. And finally, each of the tone patterns depends on something missing; the Low group depends on C1 not being a depressor consonant, and the High group depends on there not being a C1 at all. This may well be due to morphological complexity not found in other CVC forms. This last point should disqualify the verb root class by tone from this study altogether, though for the sake of completion I include it here, though with these caveats. Hopefully these tone patterns can be more fully investigated some day, to clarify the number of tone patterns in Nyali-Kilo verbs, and the system by which they are derived.

5.5.1 Nyali-Kilo [nlj] Low with C2 Depressor Verbs

I have just one recorded utterance of a final falling surface melody which is particularly low:

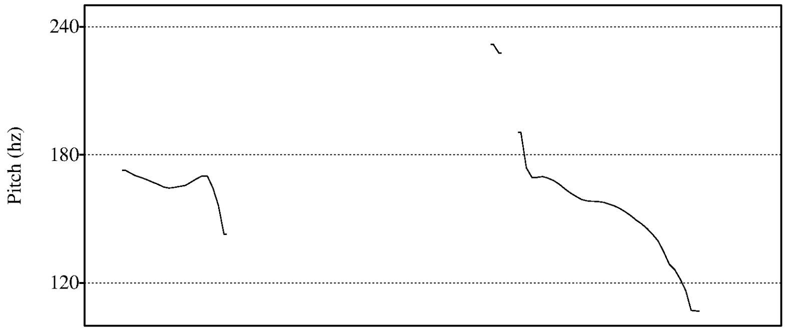


Figure 73. Nyali-kilo [nlj] pitch trace showing low root with C2 depressor: [kakʉwa] [‒ ‒ _] ‘cough’

As a result, the pitch trace overlay in Figure 20 can be transcribed in an idealized pitch trace as [‒ ‒ _]. The following verbs have this Low with C2 depressor infinitive tone melody:

(43) Infinitive verb forms with Low roots and C2 Depressors (infinitive forms with [‒ ‒ _] melody)

a. kɑsʉgɑ plaster smear
b. kɑsugo congratulate
c. kɑɗulo pull (up)
d. kɑkʉwɑ cough
e. kɑɗejɑ oil (something)

The summary of these data, as pertains to this dissertation, is that none of the C1's are contrastively voiced, while C2 may or may not be (including sonorants *l*, *w*, and *j* only).

5.5.2 Nyali-Kilo [nlj] High with VC Depressor Verbs

I have likewise just one recording of a C2 depressor following a High root mora:

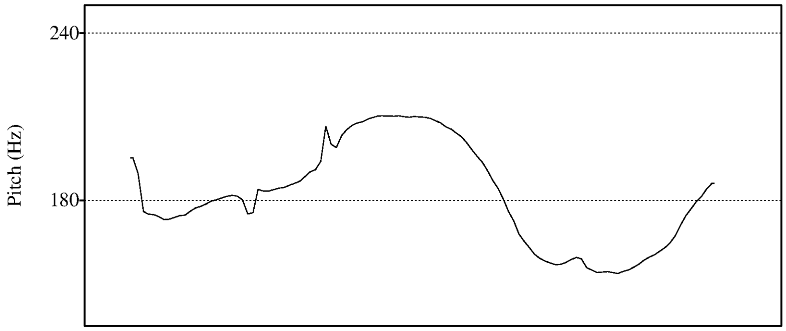


Figure 74. Nyali-kilo [nlj] pitch trace showing high root with VC depressor: [kanawa] [– ¯ ⟋] ‘curse’

As a result, the pitch trace overlay in Figure 20 can be transcribed in an idealized pitch trace as [– ¯ ⟋][1]. The following verbs have this High with C2 depressor infinitive tone melody:

(44) Infinitive verb forms with High roots and C2 Depressors (infinitive forms with [– ¯ ⟋] melody)

a.	kɑ(n)ɑbɑ	kick	ɑbɑ!	kick!
b.	kɑ(n)ɑg͡bɑ	add	ɑg͡bɑ!	add!
c.	kɑ(n)ɑᵑ͡ᵐg͡bo[2]	mix	ɑᵑ͡ᵐg͡bo!	mix!
d.	kɑ(n)ɑwɑ	curse	ɑwɑ!	curse!
e.	kɑ(n)ɨjɑ	prepare food	ɨjɑ!	prepare food!

The summary of these data, as pertains to this dissertation, is that these roots do not have C1's. Furthermore, their C2's may or may not be contrastively voiced (including *w* and *j*).

5.6 Discussion of Nyali-Kilo [nlj] Data

Rather than make claims of much significance on the basis of infinitive forms alone, this section will simply show the pattern of distribution across consonant types for the tone patterns described in this chapter:

[1] A sample of 7 LHH verbs showed a 7.26(±9.33)hz drop due to declination between the two High syllables. In the only recorded example of this LHR tone pattern, the third syllable rise ends 31.12hz lower than the previous high, indicating that this drop is more substantial than that due to declination alone, i.e., downstep.

[2] This does seem to be an anomalous case of an [a]C root causing +ATR harmony, though see appendix D for others in Bantu D30.

verb root class	C1 [L/Voice] without	C1 [L/Voice] with	C2 [L/Voice] without	C2 [L/Voice] with
Low	+	-	+	-
High	+	-	+	-
Low-C1 Depressor (CTI)	-	+	+	+
High-C1 Depressor (CTI)	+	+	+	-
Low-C2 Depressor (CTI)	+	-	+	+
High-C2 Depressor (CTI)	-	-	+	+

Table 29. Summary of distribution of consonants across melodies in Nyali-Kilo

While the pattern of distribution for Nyali-Kilo consonant types across tone patterns differs from that seen in Ndaka and Mbo, similar problems exist. For instance, while the first three rows fit nicely into the complementary distribution expected by Bradshaw (1999)'s [L/Voice] (assuming C1 triggers CTI), the final three rows do not. The tone pattern which would appear to be High with C1 Depressor has both consonant types in C1 position, as do the two tone patterns which appear to have depression in the C2 position (have both consonant types in that C2 position).

As with the data available for Ndaka and Mbo, the vast majority of Nyali-Kilo verb roots fit a pattern expected by Bradshaw (1999)'s [L/Voice] (in lighter shading in Table 29, representing 91 of 104 verbs, or 88%). Nevertheless, there are a number of exceptions in multiple depressor tone patterns, so these require some kind of explanation, if one is to take advantage of the generalization provided by Bradshaw (1999)'s [L/Voice]. An account of this difficulty will be taken up in Chapter 7.

Looking again at the distribution of particular consonants, we find [w], [l], and [j] to be the most problematic in Nyali-Kilo. Bilabial [w] is found in Low verbs, High verbs with C1 depressors, and both C2 depressor groups. Palatal [j] is found in High verbs and both C2 depressor groups. And [l] is found in both High and Low verbs, as well as in one of the C2 depressor groups (Section 5.5.1).

Chapter 6
Regular Bantu D30 Correspondences

This chapter presents data collected during the course of my fieldwork, from multiple Bantu D30 languages. They are grouped by patterns of correspondence across languages. These data provide a foundation of stability against which the less regular data to be presented in Chapter 7 can be evaluated, as well as allowing for provisional genealogical subgrouping of the Bantu D33 languages.

The data in this chapter, beyond what was presented in Chapters 3 and 4, is taken from my fieldnotes[1] and recordings made in a series of workshops to help the Ndaka, Mbo, Nyali-Kilo, Vanuma, and Bɨra communities make advances in the development of their respective writing systems. The workshops were held in four locations in the DRC: Bunia (December 1–7, 2006; June 9–10 2014), Ibambi (June 9–26, 2009), Nya-Nya (December 8–15, 2006; May 20–June 6, 2014; August 2–23, 2016) and Nyankunde (August 24, 2016). Apart from a number of sessions of targeted elicitation, the workshops were held in a participatory manner (Kutsch Lojenga 1996), to engage the fullest participation of the community in the analysis of their language and development of their writing system.

Data are presented in tables, with a row for each cognate set (a set of words in different languages presumed to have originated from a single historical word), and a table for each pattern of correspondence across languages. For the most part, I present here only infinitive forms, though there is some summary comparison of conjugated data as well.

These data tables include proto-Bantu reconstructed forms from the Bantu Lexical Reconstructions 3 (BLR3; Bastin, et al 2002), when a plausibly similar form is found attached to the same (or plausibly similar) gloss. What counts as plausibly similar, for either a form or a gloss, is of course a matter of personal judgement, to some extent. So in the interest of transparency, I have included the BLR3 entry numbers, so the reader can look up the full entry at will. Where I have some doubt of an entry being the historical source of the cognate set, I precede it with

[1] Vanuma and Bɨra data have been supplemented from workshop databases including the contributions of others. For the most part I have confirmed those data myself, but time and access restrictions have not always made that possible.

a question mark (e.g., ?115: **báat* in Table 30). Where multiple BLR3 entries are plausibly similar, I include the most likely (in my judgement) in the table, and others in a footnote (e.g., 'bite', 'cut (tree)' in Table 30). Where a BLR3 gloss differs from the gloss used for the Bantu D30 languages, the D30 gloss is given in the gloss column, with the BLR3 gloss in the BLR3 column (e.g., 282: **búà* 'dog' for 'bark' in Table 34).

Sections 6.1, 6.2 and 6.3 show correspondences for Low, High and Rising verb root classes, respectively. These correspondences show tone values that are essentially stable across the languages compared. The observed stability facilitates the comparison of the languages, as well as confirms one of the basic assumptions the comparative method depends on, that sound change is regular. Taking these verb root class by tone correspondences together with the consonant correspondences presented in Sections 6.4 and 6.5, the data in this chapter shows essentially regular correspondences for Bantu D30. This is not to say that all Bantu D30 correspondences are obviously regular; rather, the correspondences in this chapter will function as a backdrop against which to understand the less obviously regular correspondences to be presented and discussed in Chapter 7.

The correspondence patterns presented in this chapter already provide some advance in our understanding of the history of the Bantu D30 languages. Recall from Section 2.3.2 that the relationship between related languages is established through shared innovations. That is, languages that share an innovation which is understood to have occurred once would have been the same language when the innovation occurred. This logic depends on an understanding of which innovations are unlikely to occur (and therefore likely occurred just once in a subgroup), and which might more likely occur spontaneously, and should therefore not be used as evidence of a more recent shared history. For instance, it seems fair to assume that developing consonant-tone interaction, a somewhat rare phenomenon, happened only once in a subgroup of languages. Similarly, the loss of **d* (c.f., Section 6.5) likely happened once, while Budu, Ndaka and Mbo were one language. On the other hand, bilabial lenition seems to occur rather frequently in Bantu D30 (e.g, **b* > [w] in Nyali-Kilo, **b* > [β] in Vanuma, **b* > [p] in Budu and Ndaka, and **b* > [ɸ] in Mbo, c.f., Section 6.4), so these innovations would be less likely to indicate a shared history, as they were likely developed independently, at least to some extent[2].

[2] It is of course possible to find subgroupings in these bilabial lenitions, e.g, perhaps Nyali-Kilo and Vanuma shared **b* > [w] before Vanuma innovated **w* > [β]. And such subgroupings may well coordinate with other indications of more recent shared history. But because bilabial lenition happened across the board in Bantu D30, and with different results in each language, the argument that any particular phase of bilabial lenition happened only once is harder to make.

Without going into too much detail to justify it, I provide a draft genealogy of the Bantu D30 subgroup in Figure 75, based on a small number of consonant shifts, some of which are shown in this chapter:

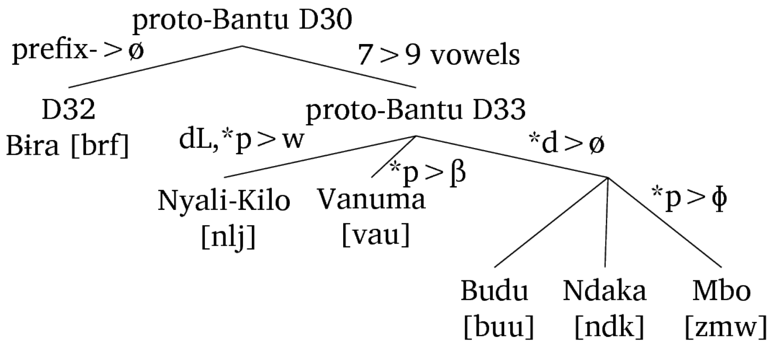


Figure 75. A possible genealogy of the Bantu D32 and D33 languages, based on consonant shifts

But perhaps more important than advances in the genealogy of these particular languages, the cognate sets presented in this chapter form a foundation from which to understand the interaction of consonant changes and tone changes, the subject of Chapter 7. Specifically, when specification for [L/Voice] is lost, what happens to tone patterns, especially to those that seem to (at least historically) depend on specification for [L/Voice]? Ultimately, answering this question leads to a better understanding of how a language may develop new tone melodies (tonogenesis), as well as a way to sensibly analyze the synchronic system without rejecting Bradshaw 1999 completely. To understand different patterns, it is good to start with understanding more straightforward and consistent patterns, as presented in this chapter.

6.1 Low Tone Correspondences

Regular correspondences that indicate a consistent low tone are presented in Table 30. The words in each of Nyali-Kilo [nlj], Vanuma [vau], Ndaka [ndk], and Mbo [zmw] consistently show an almost identical low and level pitch in their infinitive forms:

[nlj]	**[vau]**	**[ndk]**	**[zmw]**	**BLR3**	**gloss**
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]		
kakuto	kakuto	kokuto	kokuto		‘close’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]		‘sleep’, ‘lie down’,
kakʷala	kakʷala	kɔkʷana	kɔkʷana		‘copulate’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]		
kamiɲo	kamiɲo	komijo	komijo		‘squeeze’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]	?115:	
kakpata	kakpata	kɔkpata	kɔkpata	**báat*	‘ascend’, ‘follow’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]		
kamʉma	kamʉma	kɔmɔma	kɔmɔma		‘gather’, ‘keep’, ‘kiss’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]		
kanʉta	kanɔta	kɔnɔta	kɔnɔta		‘give birth’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]	1782:	
kakʉta	kakʷɛta	kɔkʷɛta	kɔkʷɛta	**kèt* ‘cut’[3]	‘bite’, ‘cut (tree)’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]	2663:	
kamejo	kamejo	konejo	konejo	*nì	‘rain’
[− − ¬]	[− − ¬]	[− − ¬]	[− − ¬]	1362:	
katɨⁿda	katɛⁿda	kɔkʲɛⁿda	kɔtɛⁿda	**gènd*	‘walk’
[− − ¬]	[− − ¬]	[− − ¬]			
kakuɗo	kakuɗo	kokuɗo	kokuɗo		‘crawl (lizard)’
[− − ¬]	[− − ¬]	[− − ¬]			
kakumo	kakumo	kakumo			‘fear’ [ndk]:‘run away’
[− − ¬]		[− − ¬]	[− − ¬]		
kakʷejo		kokʷejo	kokʷejo		‘pack’ [ndk]:‘gather’
[− − ¬]	[− − ¬]	[− − ¬]			
kasʉwa	kasʉwa	kɔsʉwa			‘sing’

Table 30. Forms which are consistently Low across languages

The consistency across cognates for each cognate set in Table 30 reflects what has been presented for Low verb root class verbs in Chapters 3, 4, and 5, which is summarized in Table 31:

[3] An alternate source for this cognate set is 719: *cóm* ‘bite’

	inf	2s-past	2s-future	3s-past	3s-future
Ndaka	[− − ¬]	[−|– ˏ ˎ|⋁]	[−|_ _ ¯|˥]	[ˊ|¯ ¯ –|⋁]	[ˊ|¯ – ¯|˥]
Mbo	[− − ¬]	[¯|_ ¯ –|⋁]	[−|– ¯ –|∨]	[ˊ|¯ ¯ –|∨]	[ˊ|¯ ¯ –|⌄]
Nyali-Kilo	[− − ¬]				

Table 31. Low verb root class infinitive and conjugated tone melodies by language

This correspondence is helpful in terms of our capacity to analyze the tone systems of these languages, because for at least these data, low tone in one language corresponds with low tone in the other languages. This gives some basis for the assumption that language change is regular, upon which historical and comparative analysis depends.

6.2 High Tone Verb Correspondences

We also see regular correspondences that indicate a consistent high tone, in Table 32. The words in each of Nyali-Kilo [nlj], Vanuma [vau], Ndaka [ndk], and Mbo [zmw] are consistently high in tone; data for Budu [buu] and Bɨra [brf] are consistent where available.

[nlj]	**[vau]**	**[brf]**	**[buu]**	**[ndk]**	**[zmw]**	**BLR3**	**gloss**
[– ¯ ¯] katoko	[– ¯ ¯] katoko			[– ¯ ˎ] kotoko	[– ¯ ¯] kotoko	3096: **tú*	‘spit’
[– ¯ ¯] katɨna	[– ¯ ¯] katɨna			[– ¯ ˎ] kɔtɨna	[– ¯ ¯] kɔtɨna	?2926: **tínà* ‘root (n)’[4]	‘cut’
[– ¯ ¯] katuto	[– ¯ ¯] katuto			[– ¯ ˎ] kotito	[– ¯ ¯] kotito	2933: **tínd*	‘push’
[– ¯ ¯] kaɓɨ^{n}de	[– ¯ ¯] kaɓɛ^{n}da			[– ¯ ˎ] kɔɓɛ^{n}da	[– ¯ ¯] kɔɓɛ^{n}da	158: **béet*	‘hit’
[– ¯ ¯] katolo	[– ¯ ¯] katoɾo						‘urinate’
[– ¯ ¯] kakʉ^{n}da	[– ¯ ¯] kakʉ^{n}da			[– ¯ ˎ] kɔkʉ^{n}da	[– ¯ ¯] kɔkʉ^{n}da	2044: *kʊ́nd	‘agree’, ‘love’

[4] 2926: **tínà* is also glossed ‘base (n)’ or ‘banana plant (n)’

[– ¯ ¯]	[– ¯ ¯]			[– ¯ ⟍]	[– ¯ ¯]	2708:	
kata^{m}ba	kata^{m}ba			kɔta^{m}ba	kɔta^{m}ba	**tá*	‘play’
[– ¯ ¯]				[– ¯ ⟍]	[– ¯ ¯]	522:	‘cackle’,
kakɨka				kɔkɛka	kɔkɛka	**cèk*	‘laugh’
[– ¯ ¯]	[– ¯ ¯]			[– ¯ ⟍]	[– ¯ ¯]	788:	
kanʉ^{n}da	kanʉ^{n}da			kɛnɨnda	kɛnɨnda	**dác*	‘shoot (weapon)’
[– ¯ ¯]	[– ¯ ¯]	[– –]	[– ¯ –]	[– ¯ ⟍]	[– ¯ ¯]	?841:	
kaɗejo	kaɗejo	ɗɨja	kuɗijo	koɗijo	koɗejo	**dámb*	‘lick’

Table 32. Forms which are consistently High across languages

The consistency across cognates for each cognate set in Table 32 reflects what has been presented for High verb root class verbs in Chapters 3, 4, and 5, which is summarized in Table 33:

<table>
<tr><th></th><th>inf</th><th>2s-past</th><th>2s-future</th><th>3s-past</th><th>3s-future</th></tr>
<tr><td>Ndaka</td><td>[– ¯ ⟍]</td><td>[–|– ⸝ ⸜|⋁]</td><td>[–|_ ¯ ¯|˥˩]</td><td>[ˊ|¯ ¯ ¯|˅]</td><td>[ˊ|¯ ¯ ¯|˥˩]</td></tr>
<tr><td>Mbo</td><td>[– ¯ ¯]</td><td>[¯|_ ¯ ¯|˅]</td><td>[⟋|_ ¯ ¯|˅]</td><td>[ˊ|¯ ¯ ¯|˅]</td><td>[ˊ|¯ ¯ ¯|˅]</td></tr>
<tr><td>Nyali-Kilo</td><td>[– ¯ ¯]</td><td></td><td></td><td></td><td></td></tr>
</table>

Table 33. High verb root class infinitive and conjugated tone melodies by language

Note that the High infinitive melody looks different in Ndaka (without spreading to the final syllable) than in Nyali-Kilo, Vanuma, and Mbo, though that difference corresponds regularly throughout Table 32. The consistent correspondence in Table 32 is helpful in terms of our capacity to analyze the tone systems of these languages, because for at least these data, high tone in one language corresponds with high tone in the other languages. This gives some basis for the assumption that language change is regular, upon which historical and comparative analysis depends.

6.3 Rising Tone Verb Correspondences

While the data in Section 6.1 and Section 6.2 show verb roots without depressor consonants, the data in Table 34 all show contrastively voiced stops in C1 position, for each Bantu D30 language where I have data. There is also a clear correspondence of infinitive tone patterns, though mostly for Ndaka and Mbo; other languages have depressor forms in this table, as well, though they don't necessarily have a straightforward correspondence to the Rising tone patterns in Ndaka and Mbo. Despite the variety of surface patterns involved, I will refer to this group as Rising tone patterns, for the sake of simplicity in the following discussion. For at least the first four, there is reason to believe that the C1 segment remains unchanged from Proto-Bantu:

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[¯ – –] ka**gb**ʉka	[¯ – –] ka**gb**ʉta	zora	[– ¯ –] kʉ**gb**ɔka	[– ⟋ ⟍] kɔ**gb**ɔka	[¯ – –] kɔ**gb**ɔka	266: ****bón***	'find'
[¯ – –] ka**g**ʉla	ka**k**ʉɾa	[– – ¯ ¯] **g**olokana	[– ¯ –] kʉ**g**ʉwa	[– ⟋ ⟍] kɔ**g**ɔwa	[¯ – –] kɔ**g**ɔa	1440: **gòn*	'snore'
[¯ – –] ka**g**iɲo	[¯ – –] ka**g**iɲo	[¯ –] n**g**ana	[– ¯ –] ku**g**iɲo	[– ⟋ ⟍] ko**g**iɲo	[¯ – –] ko**g**iɲo	1394: **gìd*	'refuse', 'abandon'
[_ ⁄ ¯] ka**g**uɓo	[¯ – –] ka**g**ubo	lu**k**uta		[– ⟋ ⟍] ko**g**uɓo	[¯ – –] ko**g**uɓo	6885: **gòb*	'bend'
[¯ – –] ka**d**ʉla	[¯ – –] ka**d**ʉɾa	[– ¯] **ɗ**ʉla		[– ⟋ ⟍] kɔ**d**ʉwa	[¯ – –] kɔ**d**ʉwa	3087: **tòʊt*[5]	'pestle'
[¯ – –] kadu^{ŋm}gbo	[– _ ¯] kadu^{m}bo						'go back'
[¯ – –] kadʒidʒo	[– _ ¯] kadʒidʒo						'ring'
			[– – –] kʉm**b**ɛna	[– ⟋ ⟍] kɔ**b**ɛta	[¯ – –] kɔ**b**ɛta		'open (eye)'
[_ ⁄ ¯] ka**gb**ʉma	ka**gb**ʉma	[– ¯] **gb**ɔma		[– ⟋ ⟍] kɔ**gb**ɔma	[¯ – –] kɔ**gb**ɔma	282: **bʊ́à* 'dog'	'bark'
[_ ⁄ ¯] ka**g**iso	[– – ¯] ka**g**iso			[– ⟋ ⟍] ko**g**iso	[¯ – –] ko**g**iso		'throw', 'fall short'
				[– ⟋ ⟍] kɔbɨja	[¯ – –] kɔbɨja		'wipe'

Table 34. Forms which are consistently Rising across languages

[5] An alternate source for this cognate set is 7174: **to*

So Table 34 shows data with [L/Voice] specification with Rising tone melodies (though recall that there are four Rising tone melodies in Nyali-Kilo, c.f., Section 5).

The consistency across cognates for each cognate set in Table 34 reflects what has been presented for Rising verb root class verbs in Chapters 3, 4, and 5, which is summarized in Table 35:

<table>
<tr><th></th><th>inf</th><th>2s-past</th><th>2s-future</th><th>3s-past</th><th>3s-future</th></tr>
<tr><td>Ndaka</td><td>[– ⟋ ⟍]</td><td>[¯|– ⸝ ⸜|∨]</td><td>[¯|⸜ _ ⸝|˥˩]</td><td>[ˊ|¯ ¯ –|◡]</td><td>[ˊ|¯ – –|˥˩]</td></tr>
<tr><td>Mbo</td><td>[¯ – ¯]</td><td>[¯|_ ¯ _|∨]</td><td>[⟋|_ ¯ –|⸝]</td><td>[ˊ|¯ _ _|∨]</td><td>[ˊ|¯ _ ⟋|∨]</td></tr>
<tr><td>Nyali-Kilo</td><td>[¯ – –]</td><td colspan="4">(Low w/C1 dep)</td></tr>
<tr><td></td><td>[– ⟋ ¯]</td><td colspan="4">(High with C1 dep)</td></tr>
<tr><td></td><td>[¯ ¯ _]</td><td colspan="4">(Low w/C2 dep)</td></tr>
<tr><td></td><td>[– ¯ ⟋]</td><td colspan="4">(High w/C2 dep)</td></tr>
</table>

Table 35. Rising verb root class infinitive and conjugated tone melodies by language

This is helpful in terms of our capacity to analyze the tone systems of these languages, because for the most part in this data, a Rising verb root class in one language corresponds with Rising verb root classes in the other languages. This gives some basis for the assumption that language change is regular, upon which historical and comparative analysis depends.

6.4 Bilabial Correspondences

This section shows both the consistency and correspondence of proto-Bantu D30 bilabials. Different correspondences represent different interaction of voicing specification between the daughter languages and their proto-Bantu sources. For the segment that would reconstruct in proto-Bantu D30 as **w*, given in Table 36, there are a number of different proto-Bantu sources (i.e., **p*, **j*, **u*, and **w*). Despite the variability of sources, we see a consistent reflex of /w/ in the daughter languages today, indicating that these corresponding segments were merged in these cognate sets by the time of proto-Bantu D30:

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
		[– ¯ –]	[– ¯]	[– ¯]	[¯ ⁄]	?2132:	
lu**w**o	lʉ**w**ɔ	nku**w**ɑ	u**h**o	u**w**o	u**w**o	**kú*p*à*	'bone (n)'
			[– – –]	[– – ¯]	[– – ¯]	?1567:	
ʉ**w**ɑnɨ	ʉ**w**ɑnɨ	n.c.	ʉ**w**ɑnɨ	u**w**ɑni	u**w**ɑni	**j*ánì*	'leaf (n)'
		[–]	[¯ –]	[– ¯ –]	[– ¯ ¯]	?2089:	
kɑ**w**o	kɑ**w**o	k**w**ɑ	ku**w**o	keni**w**o	keni**w**o	*k**ú***	'die'
[– –]	[– –]	[– –]	[– –]	[– –]	[– –]		
ʉ**w**ɛ	ʉ**w**ɛ	ʉ**w**ɛ	ʉ**w**ɛ	ʉ**w**ɛ	ʉ**w**ɛ	***u** / ***w**ɛ*	'you.sg (pron)'

Table 36. Bantu D30 [w] from proto-Bantu **p*, **j*, and **u*.

The data in Table 37 show a single proto-Bantu origin of Bantu D30 egressive and implosive stops.

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
	[– ¯ –]					3701:	
lɨ**b**ɑtɑ	lɨ**b**ɑtɑ	lu**b**ɑtɑ	lɨ**b**ɑtɑ	i**b**ɑtɑ	i**b**ɑtɑ	*****b**ààtà*	'duck (n)'
[– ¯ ¯]	[– ¯ ¯]		[– ¯ –]	[– ¯ –]	[– ¯ ¯]	41:	
kɑ**ɓ**ɨtɑ	kɑ**ɓ**ɛtɑ		kʉ**ɓ**ɛtɑ	kɔ**ɓ**ɛtɑ	kɔ**ɓ**ɛtɑ	*****b**ádʊk*	'shell (peanuts)'
						'be split'	
[– ¯ ¯]	[– ¯ ¯]		[– ¯ –]	[– ¯ –]	[– ¯ ¯]	382:	
kɑ**ɓ**u^{n}dʒo	kɑ**ɓ**u^{n}dʒo		ku**ɓ**u^{n}do	ko**ɓ**u^{n}do	ko**ɓ**u^{n}do	*****b**únj*	'break'

Table 37. Proto-Bantu **b* corresponding with D30 [b] and [ɓ]

The data in Table 37 are further interesting in that they show a correspondence between proto-Bantu **b* with **low** tone and Bantu D30 egressive [b] in 'duck', and between proto-Bantu **b* with **high** tone and Bantu D30 implosive [ɓ] in 'shell (peanuts)' and 'break'. While Table 37 doesn't show enough data to form a broad generalization, it is possible that implosive stops in Bantu D30 arose from a reanalysis of the contrast between egressive stops with high and low tone as a contrast for specification of voice (i.e., between egressive and implosive implementation of the stop, which would entail the presence and lack of voicing specification, respectively).

The correspondence in Table 38 represents a novel voicing specification for at least Nyali-Kilo and Vanuma, as these forms have voiced obstruents which correspond with voiceless *p in proto-Bantu. This innovation shows that proto-Bantu stops have become voiced, on at least some occasions in the Bantu D30 languages.

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
kɑ**b**u**b**ɑkɑ	kɑ**b**ʉ**b**ɑkɑ	**p**e**p**ɑ		ko**p**e**p**ɑ	ko**p**o**p**ɑ	2463: **pép*	‘winnow’

Table 38. Proto-Bantu **p* corresponding with D30 [p] and [b]

This innovation of voicing in Bantu D30 will become important as we look at changes in voicing specification more broadly in Bantu D30 in Chapter 7.

The cognate sets up to this point largely show unity across the Bantu D30 languages. Beginning with Table 39, there is variation in bilabial correspondence, with Bantu D30 language reflexes at various stages of lenition. Clearly plosive consonants appear in Ndaka [ndk] and Budu [buu], but fricatives in Mbo [zmw] and Vanuma [vau]. Nyali-Kilo [nlj] has a bilabial semivowel for the corresponding consonant. This correspondence constitutes a nice pattern of lenition across the Bantu D30 languages, which is consistent across each cognate set. The Proto-Bantu reconstruction (on the right) shows **b*, in the first two rows:

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
			[– ¯]	[– ¯]	[– ¯]	1614:	
lɨ**w**ɑ	lɨ**β**ɑ	n.c.	ɨ**p**ɑ	ɨ**p**ɑ	ɨ**ɸ**ɑ	**jʊ́**b**à*	‘sun (n)’
		[– – –]	[– – ¯]	[– – –]	[¯ ¯ –]	134:	
?ɨ**w**elu	ɨ**β**elu	ki**b**ele	ɨ**p**ɛʉ	ɨ**p**ɔɔ	ɨ**ɸ**ɛʉ	*****b**èdò*	‘thigh (n)’
	[– – ¯]			[– – ¯]	[– – ¯]		
liko**w**i	liko**β**i			iko**p**i	iko**ɸ**i		‘stone (n)’
[– ¯ ¯]	[– ¯ ¯]	[¯ –]	[– ¯ –]	[– ¯ ¯]	[– ¯ ¯]	659:	
lusi**w**e	lusi**β**e	usu	ũsi**p**e	usi**p**e	usi**ɸ**e	**có**mb**á*	‘fish (n)’
[– – ¯]	[– – ¯]		[– – ¯]	[– – ¯]	[– – ¯]	?2463:	
ɑ**w**ɑ**w**ɑ	ɑ**β**ɑ**β**ɑ	n.c.	ɑ**p**ɑ**p**ɑ	ɑ**p**ɑ**p**ɑ	ɑ**ɸ**ɑ**ɸ**ɑ	***pép*** ‘fly’	‘housefly (n)’

Table 39. Bilabials corresponding with lenition in Vanuma and Mbo

A similar correspondence, though without the extent of lenition in Mbo (and perhaps Vanuma), is found in Table 40. The comparison between the two correspondences can be seen in ‘housefly (n)’ from Table 39, which is plosive in Ndaka [ndk], but fricative in Mbo [zmw], compared with Mbo ‘spear’ in Table 40, which is not fricative, in addition to differing in both meaning and tone from Mbo ‘housefly’:

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
		[– ¯]		[¯ –]	[¯ –]	?1025:	
ɑ**w**ɨlɑ		golɑ		**p**ɨɑ	**p**ɨɑ	*dì**b**à	‘waterfall (n)’
			[– –]	[– ¯]	[– ¯ –]	?7306:	
u**w**o	u**b**o	n.c.	ʉ**p**o	u**p**o	ɑ**p**ɑ**p**ɑ	*cò**p**[6]	‘spear (n)’

Table 40. Bilabials corresponding without lenition in Vanuma and Mbo

The given proto-Bantu reconstruction for ‘waterfall (n)’ is only plausible on the basis of metathesis some time between the proto-Bantu and Bantu D30 forms, such that the proto-Bantu D30 form would have been something like **bida*. However striking that change may seem, metathesis (**bVd > dVb*) may be seen as only one change to bring the proto-Bantu reconstruction in line with the Bantu D30 forms of the same gloss, so I am accepting it as a plausible source for this Bantu D30 cognate set.

One may question the last two cognate sets of Table 39, along with ‘spear’ in Table 40, as they do not correspond to proto-Bantu **b*. And not only are **mb* and **p* not the same proto-Bantu consonant (i.e. **b*), they are also not specified for voicing[7]. We have two reasons to believe that this apparent issue is not fundamentally a problem.

The first reason the multiplicity of proto-Bantu sources in Table 39 is not problematic derives from the principle that a single set of correspondences across daughter languages implies a single segment in the proto-language. This principle tells us that each of the cognate sets in Table 39 once had the same consonant, as did the cognate sets in Table 40[8]. This principle does not tell us the phonetic value of that consonant, though it should be clearly bilabial, as all of its reflexes are bilabial. This proto-Bantu D30 consonant came from proto-Bantu **b*, **mb*, or **p*, then later developed into [w], [β], [b], [p], and [ɸ]. Given that no modern reflexes have any nasal features, **b* and **p* seem the most likely candidates for a single intermediary between the multiple proto-Bantu sources and the multiple Bantu D30 reflexes.

The second reason the multiple proto-Bantu sources in Table 39 is not problematic is that Table 38 already provides evidence of proto-Bantu consonants gaining

[6] 7306: **còp* is glossed ‘prod (with stick, spear)’ in the BLR3.

[7] The lack of a constructed **mp* for proto-Bantu implies that **mb* should not be analyzed as underlyingly specified for voice, as it is not the voiced member of a pair of sounds minimally different for voice (as **b* does, in opposition to **p*).

[8] We can further say that the two consonants were likely one consonant at some time in their history, as the correspondence patterns are so similar, differing only in Vanuma and Mbo.

voice specification in proto-Bantu D30 (at least in Nyali-Kilo [nlj] and Vanuma [vau]), so it is plausible to imagine that the proto-Bantu D30 bilabial in ‘stone (n)’, ‘fish (n)’, and ‘housefly (n)’ of Table 39 and in ‘spear (n)’ of Table 40 was **b*, though at some time after Proto-Bantu. Further evidence of proto-Bantu **p* which apparently gained voicing specification by Bantu D30 will be presented in Section 7.5, along with a third argument of a voiced intermediary (e.g., **b*) between proto-Bantu **p* and Bantu D30 voiceless consonants (that their tone patterns behave in the same way as others which have voiced ancestors). Further examples of correspondences to voiceless reconstructions in Proto-Bantu, which I assume occurred through a voiced intermediary, will occasionally occur through the rest of this dissertation, but I will not justify that assumption beyond what I have stated here.

Whatever the proto-Bantu D30 source was, the correspondences in Table 39 are particularly important because **b* > [w]/[β]/[p]/[ɸ] and **b* > [w]/[b]/[p] represent a loss of voicing specification, almost across the board in Bantu D30. The implications of this loss of voicing specification will be addressed in further detail in Chapter 7, specifically in Sections 7.5 through 7.7.

6.5 Alveolar Correspondences

This section shows a single correspondence of [l] in Nyali-Kilo and Vanuma[9], to a missing consonant in the other Bantu D30 languages. The data in Table 41 have been left as originally transcribed with regard to the presence or absence of [w] corresponding to **d*. In all likelihood, an epenthetic [w] arises from the loss of **d*, which creates hiatus between the root and final vowels. But whether any of the **d* reflexes in Budu, Ndaka and Mbo are distinct phonetically or phonologically from other $V_{[back]}a$ or $V_{[back]}wa$ sequences has not been investigated. But regardless of the status of this reflex in these languages, there is no voicing specification.

[9] There is an alternation between [l] and [ɾ] in Vanuma, so one may consider this [l] to be a lateral with an allomorph of [ɾ]. In any case, they should be distinguished from [w] and [j], which in the cases given in this section are most likely an epenthetic insertion as a resolution of vowel hiatus.

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[– – –] kafʉ**l**ɑ	[– – –] kɑfu**l**o	ɲʉɑ			[– ¯ ¯] kɔpʉɑ	756: **cù*d***	‘fart’ [10]
[– – –] kɑ**l**ʉwɑ	[– – –] kɑ**l**ʉwɑ		[– –] kʉwɑ	[– – –] kɔʉwɑ	[– – –] kɔʉwɑ	1177: ****d**ʊ̀k*	‘plait (hair)’
[– – –] kɑkʉ**l**ɑ	[– – –] kɑkɔ**l**ɑ		[– – –] kʉkʉwɑ	[– – –] kɔkʉ**w**ɑ	– – – kɔkʉ**w**ɑ	1482: **gʊ̀**d***	‘buy’
mʉ**l**ʉndɑ	[– ¯ ¯] mʉ**l**ʉndɑ	kiindɑ		mʉʉndɑ	mʉʉndɑ	1059: ****dí**mbà* ‘body’	‘corpse (n)’

Table 41. Laterals in Nyali-Kilo and Vanuma corresponding with no consonant, from proto-Bantu **d*

The cognate sets in Table 41 are important because they come from proto-Bantu **d*, which implies a loss of distinctive voicing at some point, either to synchronic [l] or to no (or epenthetic) consonant, for a given language spoken today. It remains somewhat unclear if some of the forms in Table 41 have an epenthetic or contrastive [w], but a contrastive ***w*** could easily have arisen from either an epenthetic [w], or from a loss of place from [l], so in any case the correspondences in Table 41 represent a loss of voicing specification between proto-Bantu and the synchronic forms. The impact of this loss of voicing for words with this correspondence pattern will be taken up in Chapter 7, specifically in Section 7.6.

6.6 Summary of Regular Bantu D30 Correspondences

This chapter establishes two points of background necessary to follow the analysis in Chapter 7, regarding the unity and variety of the Bantu D30 languages.

In terms of unity, I have demonstrated a basic consistency for comparison across the Bantu D30 languages. This provides evidence in support of the assumption that sound change is generally regular. Specifically, Sections 6.1 and 6.2 show a basic unity of low and high tone, respectively, across the Bantu D30 languages. There may be changes in tone, either between two Bantu D30 languages, or between Bantu D30 as a whole and the proto-Bantu forms they correspond with, but there is a large degree of tone stability between these languages.

[10] The gloss ‘fart’applies only to Nyali-Kilo [nlj], Vanuma [vau] and Bɨra [brf]. Mbo [zmw]) has the gloss ‘diarrhea’

One practical implication of the correspondence between verb root classes by tone in these languages is that it argues strongly for the phonological reality of tone in these languages. That is, there is some phonological reality underlying the tone in verbs that is the same across these languages, even though each language may express that reality differently (e.g. the infinitive forms of High Ndaka roots and Rising Mbo roots). Because of this, the correspondence of the verb root classes by tone across languages argues against a thesis that tone is not an underlying phonological reality, though I hope few would argue that thesis today.

In terms of variety, I have demonstrated a range of regular correspondences between sounds in the Bantu D30 languages. These regular segmental correspondence patterns provide a backdrop against which to understand the less regular tone correspondences presented in Chapter 7. Understanding the regularity of these segmental correspondence patterns is important to understanding the different kinds of correspondence patterns found for tonal melodies.

There are two basic consonantal sound correspondences that include a loss of voicing in Bantu D30. For labials, proto-Bantu **b* often corresponds with a specific set of labial sounds in various stages of lenition (i.e., [w], [β], [p], and [ɸ]) across the Bantu D30 languages. None of these synchronic consonants are specified for voicing (despite the fact that [w] and [β] have surface voicing), so each represents a loss of voicing specification from proto-Bantu **b*. Similarly for alveolars, proto-Bantu **d* corresponds with [l] in Nyali-Kilo and Vanuma, and an epenthetic or missing consonant elsewhere in Bantu D30. This sound correspondence also thus represents a loss of voicing specification from proto-Bantu **d* for each language.

This variety is important to understanding the implications of the loss of voicing specification and the development of new tone melodies in Chapter 7. I have shown that at least some of the variety of consonants observed in Bantu D30 is not chaotic, but reflects regular sound correspondences between these languages. This means reader should not be troubled to see [w] in Nyali-Kilo corresponding to [p] in Ndaka in the same cognate set, because this has been established as a regular correspondence (in Section 6.4). But it is similarly critical to understand that each of the two sound correspondences mentioned above represent a loss of voicing specification for each Bantu D30 language, despite their various segmental reflexes. This loss of voicing specification from each of **b* and **d* will be relevant to the examination of voicing specification loss and tone system development in Chapter 7.

Chapter 7
Tonogenesis

> *The first question we are likely to ask when confronted with this vast Babel of tongues is the historical one: How did such linguistic diversity come to exist?* (Greenberg 1959:15)

Having laid out the basics of the verb root class by tone inventory for Ndaka in Chapter 3, for Mbo in Chapter 4, and for Nyali-Kilo in Chapter 5, and having shown some of the basic correspondences between Bantu D30 languages in Chapter 6, this chapter covers one of the fundamental issues arising from those data, namely that the distribution of segments across tone patterns does not align with a straightforward analysis according to Bradshaw (1999)'s [L/Voice]. To resolve this issue, I ultimately claim that these languages have undergone tonogenesis, which "refers to the development of tone systems. It includes both passage from a non-tonal to a tonal stage and also changes from a certain tonal stage to another, generally more complex, tonal stage" (Hombert 1975:4). Some limit the use of the term to Hombert (1975)'s first usage, which I will call tonogenesis *ex nihilo*. My analysis in this chapter will focus on the second usage (developing more complex tone), though my model could be applied to either.[1]

As was discussed in Section 2.1.3, if a language has gone from two verb root classes by tone to three, the means by which the new lexical melody arose must interest us, if we are to understand complex tone systems.

I speak of tonogenesis as the development of **new verb root classes by tone** intentionally, rather than as the development of tone *ex nihilo*, or the first verb root classes by tone in a given language. This usage follows from the understanding that the tonal **melody** distinguishes the verb root class by tone of one morpheme from that of another, rather than any one of a set of "tonemes", e.g., {H,L}, which combine in any number of ways on any given morpheme (c.f., Section 2.1.3 of

[1] I discuss the development of *more* complex tone in Bantu D30 because that is the process which I have relevant data. Given that Proto-Bantu has been reconstructed with tone, applying this model to tonogenesis *ex nihilo* for a Bantu language would require comparing Proto-Bantu with its contemporaries, which is likely beyond our current capabilities.

this dissertation, Snider 1999, 2014, 2018, *inter alia*). From the perspective of verb root classes by tone defined by melodies, a system that moves from {L,H} to {L,H,LH} has gained a new verb root class by tone, having moved from two tone melodies to three, though it may not look like much has changed from a perspective considering simple tonemes only. Because of the importance of tone melodies, throughout this dissertation I use tonogenesis to refer to the development of new tone melodies. However, see appendix F for a discussion of the possible extension of my model towards tonogenesis *ex nihilo*, as well.

Because my thesis regarding tonogenesis in Bantu D30 is somewhat complex, I will take it in stages. In Section 7.1, I will outline the distributions of segments across tone patterns as expected with and without CTI, as well as that actually found in Bantu D30. In Section 7.2 I will show the phonetic loss of voicing which appears to have occurred after the lexicalization of CTI by comparing two cognate verb forms from two closely related Bantu D30 languages. In Section 7.3, I will lay out the theoretical framework for the model I propose. I will describe four language systems which would be derived one from another through a series of minimal reanalyses and the impact of losing voicing specification at each stage. In Section 7.4, I will put hypothetical examples through the model, to show how actual word categories (with or without voicing specification, and with high or low tone) would be reanalyzed at each stage. The hypothetical examples are shown first without loss of voicing specification, then with loss of voicing specification at each stage of the process.

The argument as stated so far is somewhat hypothetical, but will lead to a prediction that there should be different kinds of tone changes observed through the different Bantu D30 languages. Data showing this prediction held will be presented in Sections 7.5–7.7, then summarized in Section 7.8. The diversity of correspondence patterns shown in these sections supports the assertion that while CTI may have arisen once, the processes of lexicalizing CTI and devoicing likely occurred spontaneously and at different times for particular words in each language.

The discussion in Section 7.9 includes a series of topics arising from the model of tonogenesis proposed here, including the meaning of the variation observed in Sections 7.5–7.7, practical implications for the analysis of the Bantu D30 languages, and some theoretical implications about the nature of language and language change more generally. A possible extension to proto-Bantu and tonogenesis *ex nihilo* will be reserved for appendix F.

7.1 Distributions

To see the unexpectedness of the distribution of segments across tone patterns clearly, it may help to first clarify the distribution expected by a straightforward analysis using Bradshaw (1999)'s [L/Voice]. A system without a systematic consonant-tone interaction (CTI) rule would show a distribution like in Table 42 for a binary tone contrast:

\ C1 verb root class \	without [L/Voice] (vl obstruents, sonorants)	with [L/Voice] (vd obstruents)
Low	+	+
High	+	+

Table 42. Complete distribution of consonant types across tone melodies without CTI

That is, there are two verb root classes by tone (High and Low), each of which are fully distributed across consonant types. Each consonant type is found with each tonal melody, and each tonal melody is found with each consonant type, so the distribution is complete across the two binary categories.

But with a CTI rule, those consonants with [L/Voice] specification should produce different surface tone melodies. Because this would be a systematic application to one set of consonants and not the other, the result would be a different set of melodies on the consonants where the rule applies, and the original melodies only with consonants where the rule did not apply. This means that melodies without such changes (i.e., Low and High) would then only be found on words with consonants **without** [L/Voice] specification, and the new melodies would only be found on words with consonants **with** [L/Voice] specification. Assuming for the moment that there is just one CTI melody which we will call “Rising”, one would expect the distribution of consonants across those tone patterns as in Table 43:

\ C1 verb root class \	without [L/Voice] (vl obstruents, sonorants)	with [L/Voice] (vd obstruents)
Low	+	-
High	+	-
Rising	-	+

Table 43. Complementary distribution of consonant types across tone melodies under CTI

That is, there is a strict complementary distribution between words without C1 which are specified for voicing[2] in the **Low** and **High** verb root classes (and only there) on the one hand, and words with C1 which **are** specified for voicing in the **Rising** verb root class (and only there) on the other. This kind of complementary distribution reflects a straightforward phonological rule, as in the rightmost column of Figure 13 (from p41), repeated in Figure 76:

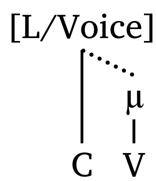


Figure 76. CTI rule; my interpretation of Bradshaw (1999)

The result of this rule is that all consonants specified for voicing contribute an extra low tone to the lexical tone melody, and those which are not thus specified do not.

But the distribution in Table 43 is not strictly found in the data in Chapters 3, 4 and 5. For each of Ndaka, Mbo, and Nyali-Kilo[3], the Low and High verb root classes have no words with C1 which are specified for voicing, which falls into line with the distribution in Table 43. However, while the verb root classes by tone

[2] I continue to assume Bradshaw (1999)'s [L/Voice], but I will use the term "voicing" to keep my discussion as theory neutral as possible. At no point should I be taken to propose a contrast between a feature [L/Voice] and another feature [voice] or [±voice]. For those that accept that my data indicate a correlation between contrastive voicing (however the reader understands that, in terms of features) and low tone, I believe these data argue for Bradshaw (1999)'s conception of a single, privative feature [L/Voice].

[3] For the sake of simplicity in this discussion, I will assume that the Nyali-Kilo infinitive tone patterns presented in Chapter 5 reflect underlying verb root classes by tone. Furthermore, and again to facilitate this discussion, I will treat the four Nyali-Kilo verb root classes by tone which arose due to the activity of depressor consonants (i.e., Sections 5.3–5.5) under a single label, "Rising".

which *prima facie* belong to depressor consonants (i.e., Rising) contain words with C1 which are specified for voicing, as expected by an [L/Voice] analysis, they also contain words with C1 which are **not** specified for voicing, which is unexpected by an [L/Voice] analysis. This observed distribution is summarized in Table 44:

**\ C1 verb root class **	**without [L/Voice] (vl obstruents, sonorants)**	**with [L/Voice] (vd obstruents)**
Low	+	-
High	+	-
Rising (CTI)	+	+

Table 44. Contrastive distribution of consonant types across tone melodies observed in Ndaka, Mbo, and Nyali-Kilo

Given that there is a Rising lexical verb root class which largely, but not exclusively, consists of verbs with C1 containing voice specification, there are a limited number of analytical options to choose from. Setting aside Bradshaw (1999)'s hypothesis of [L/Voice] would require some other explanation for the distributions observed in the Bantu D30 languages on two counts. First, one would need to account for the fact that all words with contrastively voiced C1 are entirely in one particular verb root class by tone. And second, one would need to account for the fact that 84%-88% of the available data in each language fits the distribution in Table 43, as detailed in Sections 3.6, 4.5, and 5.6. As a result, any analysis without [L/Voice] would need some other account for the gap in the upper right corner of Table 44, even if one is not bothered by the strong majority of the data corresponding with the distribution in Table 43, as would be predicted by [L/Voice].

Alternatively, one may look for a model that allows for the essential claims of Bradshaw (1999) and [L/Voice], while at the same time allowing for the presence of verbs both with and without contrastively voiced C1's to appear with the only tone pattern available to those with contrastively voiced C1's.

I propose such a model in this chapter, to maintain the essential claims of Bradshaw (1999), while also accounting for data from the Bantu D30 languages. In this model, proto-Bantu D30 had consonant-tone interaction (CTI), complete with a CTI rule such as that in Figure 76, but the CTI rule was at some point reanalyzed as lexical specification. Subsequently (at least for some lexemes) loss of voicing specification occurred, producing verbs with C1's which are no longer specified for voicing, but which retain Rising tone melodies nonetheless.

This model accounts for the strong bias in consonant distributions across tone patterns today as an artifact of a period in recent history, where consonant-tone interaction was active by rule (c.f., Figure 76), which at the time resulted in a distribution as in Table 43.

Additionally, this model accounts for the numerous exceptions to the distribution in Table 43 by understanding the CTI rule to have been lexicalized in the Bantu D30 languages today.

By itself, the reanalysis of the CTI rule would not change the distribution in Table 43, as the lexicalized distribution would be based on the distribution resulting from the CTI rule. But once the surface tone melodies were reanalyzed as lexical specification (rather than applied by rule), they were no longer conditioned by consonants, so changes to those consonants would no longer impact the tone pattern. In this way, the combination of the lexicalization of a CTI rule, followed by loss of voicing specification, resulted in a new distribution: words without voicing specification, but nonetheless with Rising tone patterns (i.e., the lower left corner of Table 44).

This model then provides for the exact pattern of consonant distribution we find in Ndaka and Mbo, and summarized in Table 44. Not only does the analysis I present in this chapter provide for a way of understanding the synchronic distribution of consonants across verb root classes by tone, it also provides for a model of tonogenesis, the development of new verb root classes by tone.

7.2 Phonetics

Rather than simply ask the reader to accept all my transcriptions at face value, this section provides an example of the loss of voicing in the obstruents of the infinitive forms of 'flutter' in Ndaka [ndk], as compared to its Mbo [zmw] cognate.

To see consonant-tone interaction (CTI) and loss of [L/Voice] specification working out in two of the Bantu D30 languages, consider the following two cognates, first in Mbo in Figure 77, which has voiced root consonants:

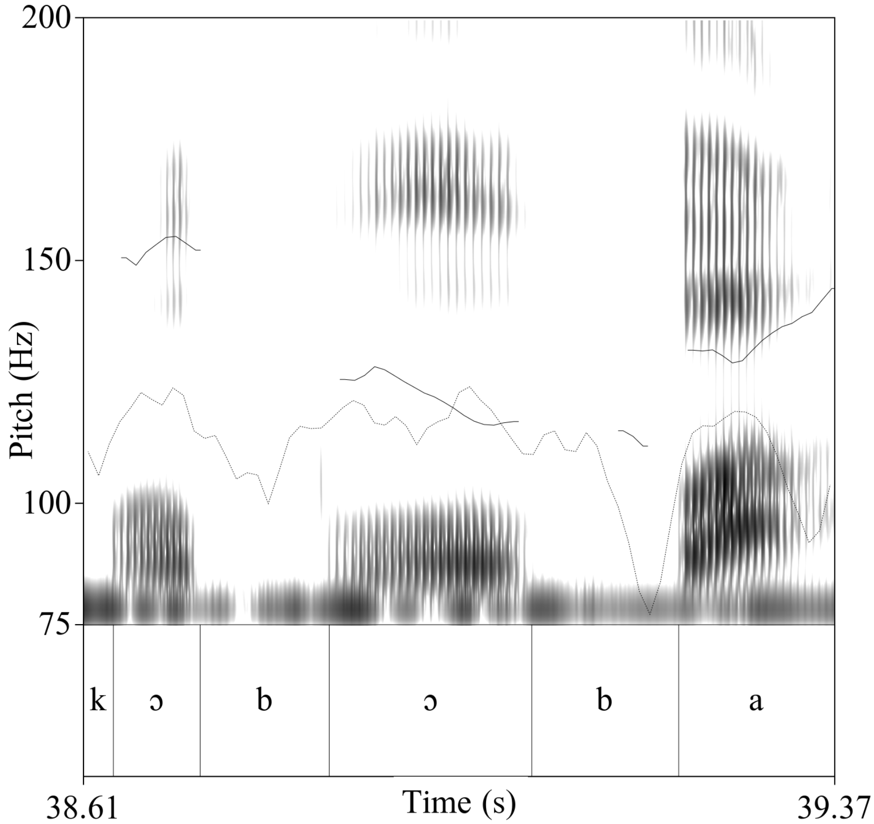


Figure 77. Mbo [zmw] ‘flutter’, with Rising tone melody (c.f. Section 4.3)

Figure 77 shows Mbo [zmw] [kɔbɔba] ‘flutter’ with Rising tone melody ([¯ — ¯] in the 120-150hz range). There is periodic signal in the F_0 of the spectrogram (<85hz on the pitch trace), corresponding with consonants transcribed as voiced [b]. The spectrogram F_0 is somewhat obscured by wind noise, but the periodicity remains visible, especially in contrast to Figure 78:

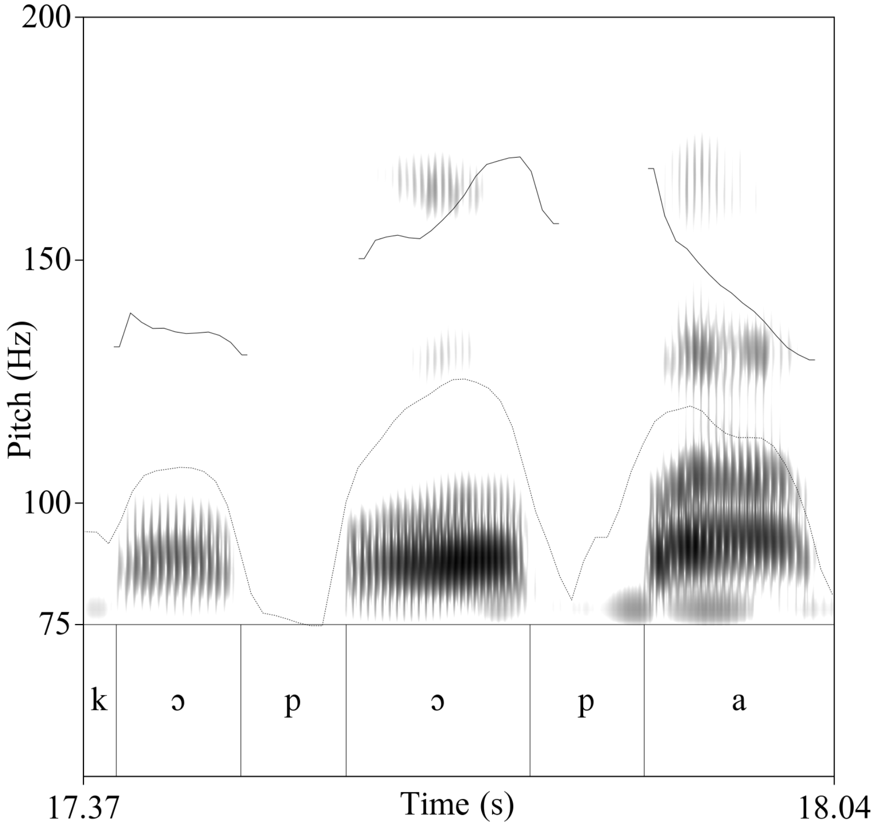


Figure 78. Ndaka [ndk] ‘flutter’, with Rising tone melody (c.f. Section 3.3)

Figure 78 shows Ndaka [ndk] [kɔpɔpa] ‘flutter’ with Rising tone melody ([_ ⟋ ⟍] in the 130-180hz range). Recall that the Rising pattern looks different than in Mbo, c.f., Section 4.3, so this melody represents the same underlying verb root class (i.e., Rising) as that in Figure 77. The lack of F_0 in the spectrogram ($<$ 85 hz on the pitch axis) corresponds to consonants transcribed as voiceless [p].

The spectrograms in Figure 77 and Figure 78 thus confirm the original transcription of the root consonants in Figure 77 as voiced, while those in Figure 78 are not.

Despite the difference in [L/Voice] specification, each word has the corresponding Rising tone pattern for its language. That is, the difference between tone patterns in Figure 77 and Figure 78 is a difference between the two languages, which holds across all rising forms, with or without [L/Voice] specification (c.f., Section 6.3).

The cognate words in Figures 77 and 78 are thus minimally different in segmental form, and have the same meaning and verb root class by tone, according to a systematic difference that operates across all Rising forms in Ndaka and Mbo. What this means is that the verb root class by tone of the verb did not change when the Ndaka form lost voicing specification on its consonants (**b* > [p]). These are the same basic data which will be repeated in Sections 7.5–7.8, showing that forms lose voicing specification without a change in lexical verb root class by tone. This observation matches the expectation that the Rising surface tone represents a lexical verb root class by tone, and is not applied by rule.

This interplay between CTI lexicalization and loss of voicing specification is critical to the theoretical framework I propose in Sections 7.3–7.4.

7.3 Theoretical framework

Having laid out the basic Bantu D30 tone groups in Chapters 3–5, including the difficulty posed by the Rising verb root class(es) in each language, and having shown briefly the presence of voicing loss in Bantu D30 without affecting verb root class by tone (Section 7.2), this section will present the theoretical aspects of the model itself.

The basic thrust of the model I present in this chapter is that somewhere in the history of the Bantu D30 languages, they developed a consonant-tone interaction (CTI) rule. At some point, the CTI rule was reanalyzed as lexical specification. Finally the CTI rule is dropped. This flow from one language system to another through reanalysis is outlined in (45) and summarized graphically in Figure 79:

(45) a. Develop a CTI rule (c.f., Figure 76; W → X in Figure 79)
b. Reanalyze CTI rule as lexical specification (X → Y in Figure 79)
c. Drop CTI rule (Y → Z in Figure 79)

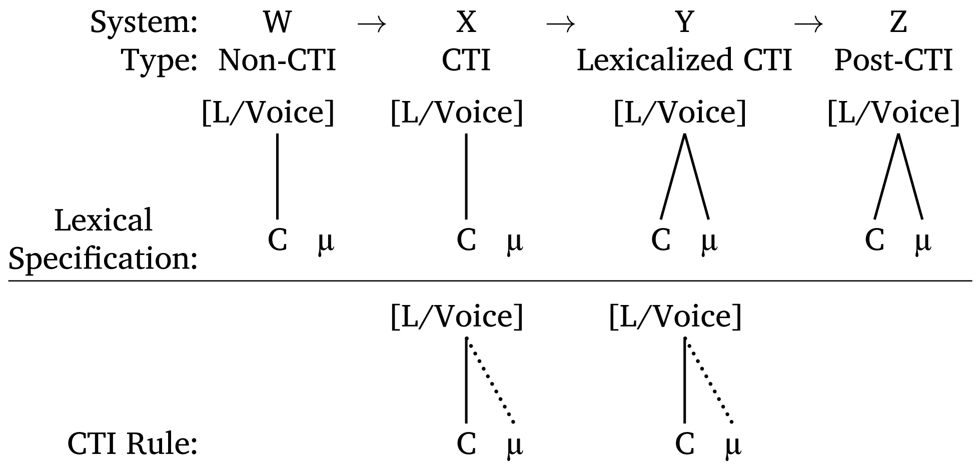


Figure 79. Historical development through four language systems

Figure 79 shows a three step process of reanalysis, with the first column (W) representing a language without CTI. The development of CTI is shown by autosegmental rule in column (X), as in Figure 76. Column (Y) shows that same surface form, but now as a result of lexical specification, rather than by rule. Finally, column (Z) shows a language where the CTI rule has been dropped, perhaps because it was judged redundant and unnecessary. For each language (W-Z) in Figure 79, I have provided a lexical representation for a voiced consonant followed by a mora (above) and a possible CTI rule (below); together, the lexical representation and presence or absence of the CTI rule distinguish the four systems in Figure 79.

Note that none of the above denies the presence of a universal, privative feature [L/Voice] (Bradshaw 1999); the difference between the systems in Figure 79 is whether [L/Voice] associates by rule or not, and how [L/Voice] is associated to consonants and/or mora in the lexicon. In other words, [L/Voice] is present throughout each of these systems, but it systematically associates to a following mora in languages X and Y only, in what is called consonant-tone interaction (CTI).

But the reanalysis of a CTI rule as lexical specification is only the first part of the path to modern day Bantu D30 languages. The first reanalysis (W→X) in Figure 79 adds a new surface tone pattern (in language X). The second reanalysis (X→Y) reinterprets this surface tone pattern (from language X) as a lexical tone pattern (in language Y). Up to this point, the correlation between voicing and tone melodies would largely remain as the CTI distribution in Table 43[4] from p176

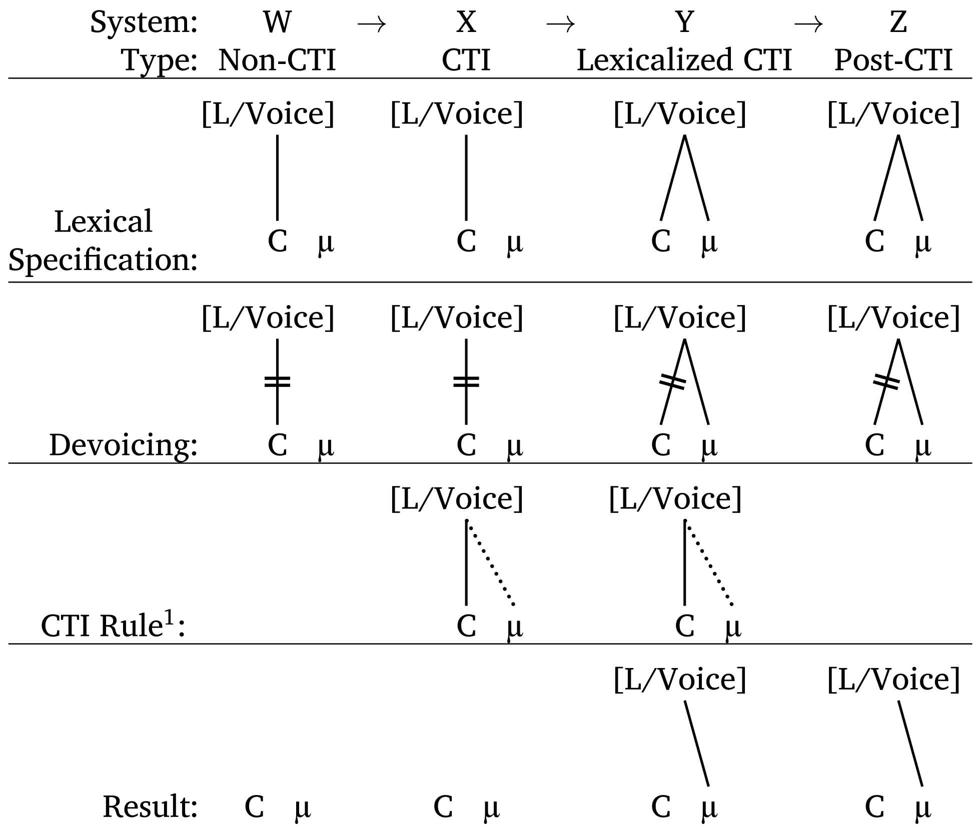
[4] The distribution as in Table 43 would of course not apply to new words, which at this point should be assignable to any lexical tone pattern regardless of consonant types.

without some other change. But when combined with a loss of [L/Voice] specification, the correlation between consonant types and tone melodies changes, as the new tone patterns are found on words with consonants which would not have conditioned the CTI rule. Finally, a third reanalysis (Y→Z) drops the CTI rule altogether, making way for exceptions on the other side, words with consonants which should trigger a CTI rule (i.e., voiced obstruents), but which do not have a Rising tone pattern, because the rule no longer exists in the language.

Figure 80 shows the same systems W-Z from Figure 79, but for words which undergo a loss of voicing specification. In addition to the lexical representation and CTI rule, there is a devoicing rule, which applies for these data across each language. The final line shows the result of the two rules applying to the lexical representation for each language. In this way, Figure 80 shows diachronic change from left to right, and synchronic derivation from top to bottom; for each language system (column). It is critical to my model that the reanalyses presented in Figure 79 and the devoicing of particular words in a given language represent two independent dimensions of language change. Because of this, the rules in Figure 80 apply to the lexical forms of that system (at the top of that column), rather than to an output of a previous system. The resulting model thus allows for words from any one of four stages of language change in Figure 79, plus another four types of words, derived from those same four languages, but with loss of voicing specification, for a total of eight types of words in this model (from any one of W-Z, each with or without devoicing).

Regarding the specifics of Figure 80, The loss of voicing specification in language W results in a consonant unspecified for voicing, without other impact on the system. I assume that the CTI rule in language X would be bled by loss of voicing specification, making words which lose voicing specification identical to their cognates which lost voicing specification from language W (i.e., as if the system had never developed CTI in the first place).[5] For both languages Y and Z, the loss of voicing specification on the consonant does not impact the presence of [L/Voice] in the lexical representation, which remains associated to the mora.

[5] One could of course take the opposite view, that loss of voicing specification in X would apply after the CTI rule, in which case X forms which lose voicing specification would be pronounced the same as their cognates in language Y which lost voicing specification. But in any case, there are only two surface forms which result from loss of voicing specification in Figure 80, based on whether the loss of voicing specification occurs before the CTI rule, or after.



1: N.B.: The CTI rule is not applied without [L/voice] on a consonant.

Figure 80. The impact of devoicing on four language systems

In the case where loss of voicing specification occurs before the lexicalization of the CTI rule (i.e., in W or X), segments and tone melodies continue to correlate as in Table 43. But where loss of voicing specification occurs after the lexicalization of the CTI rule (i.e., Y or Z), segments and tone melodies lose complementary distribution, and now correlate as in Table 44, resulting in a new, contrastive Rising tone melody on words with consonants which are not specified for voicing. This difference arises from the fact that the additional [L/Voice] in the surface representation is now found on a word following a consonant which is not specified for voice.

Word final devoicing is argued to be a natural process that applies in completely unrelated languages for similar physiological reasons (Blevins 2004). It may be that there is something generally natural about spontaneous devoicing in other contexts as well, as in Bantu D30 verb C1's. In any case, it will be observed in the data in Sections 7.5–7.7 that devoicing must occur at multiple times for different words in these languages.

Figures 79 and 80 can be summarized together in Figure 81, which shows a genealogical progression through four language systems (W-Z), and the impact of loss of voicing specification at each stage. Language W has L and H tone melodies[6] for all words, regardless of voicing specification on segments. Loss of voicing specification makes no change in that distribution, only changing words specified for voicing into those that were not.

Language X differs in that it has surface, predictable LL and LH tone melodies for words with consonants specified for voicing, but the distribution is again unchanged by loss of voicing specification.

Language Y has the same surface distributions as Language X, but they are lexical, rather than derived by rule. Because the language Y tone patterns are lexical, words in language Y with consonants which lose specification for voicing do not lose those tone patterns. The result of loss of voicing specification in some language Y words, then, is language Y words with consonants which are not specified for voicing both with H/L tone patterns (inherited from language X words without voicing specification) and with LH/LL tone patterns (inherited from language X words with voicing specification, which subsequently lost voicing specification).

Language Z has the same lexical tone melodies as Language Y, and the same impact of losing voicing specification on words with segments that have voicing specification. But because the voicing specification on consonants is now completely independent of the voicing specification on TBU's, the language is free to develop exceptions in both directions: continuing to lose voicing specification without changing tone pattern (to only L or H), but also adding voicing specification without changing tone pattern (to only LH or LL), in addition to borrowing or inventing words with any type consonant bearing any of the lexically possible tone patterns (i.e., H/L/LH/LL).

[6] I assume H v L for the sake of moving forward with the model. Of course the initial state could be otherwise, e.g., H v ∅, as has been claimed for much of Bantu (Hyman 2000).

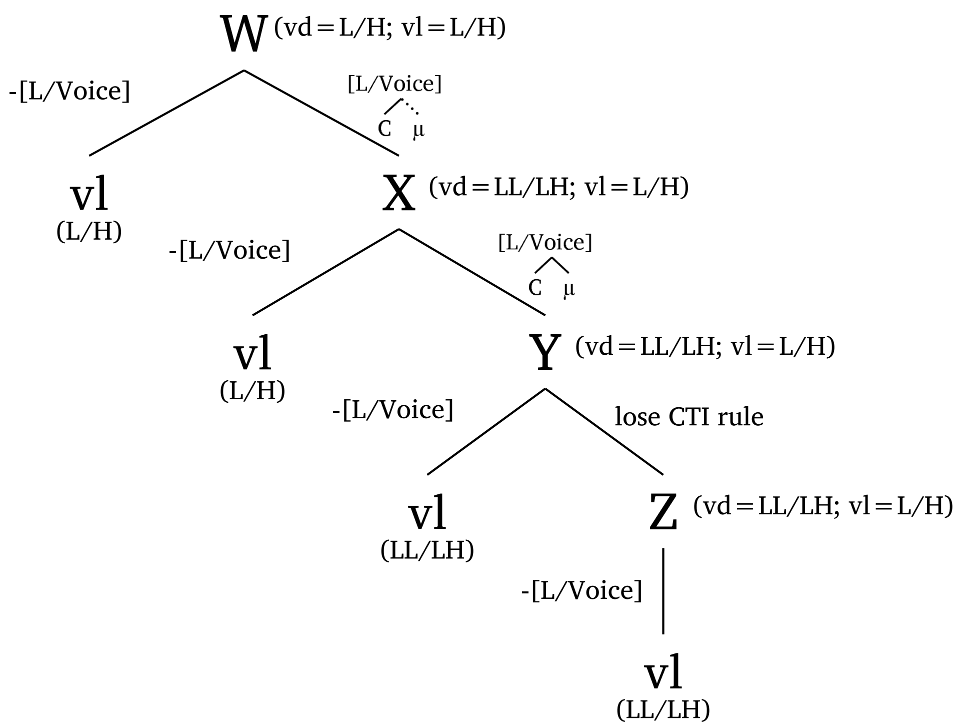


Figure 81. Historical development of four language systems with loss of [L/Voice] specification for some words at each stage

The analysis presented here makes the following predictions:

(46) a. Consonants **without** specification for voice which correspond with consonants **without** voicing specification in all related languages (including sisters and ancestors) should lack Rising tone patterns exclusively, since there is no evidence that the word in question was specified for voicing while CTI was in effect.
b. Consonants **without** specification for voice which correspond with consonants **with** voicing specification in at least one related language (including sisters and/or an ancestor) should exist either

1. **without** Rising tone patterns, indicating a loss of voicing specification from a language system like W or X, **before** CTI was reanalyzed as lexical, or
2. **with** Rising tone patterns, indicating a loss of voicing specification from a language system like Y or Z, **after** CTI was reanalyzed as lexical

c. If loss of voicing specification is relatively frequent and spontaneous (as the data in Section 7.5 will strongly indicate), we should see large number of patterns showing the interaction of consonants who have lost voicing specification and Rising tone melodies. Those patterns should reflect the historic unity of these languages, as well as their recent departures, assuming the voicing loss occurred throughout each stage of the languages' history.

d. If loss of voicing specification is infrequent and non-spontaneous (which is not indicated by my data), we should see a limited number of patterns showing the interaction of consonants which have lost voicing specification and Rising tone melodies.

e. Words without cognates in related languages may appear with any tone pattern, for languages Y and Z.

f. Any number of consonants specified for voicing with non-CTI tone patterns indicate a language Z system. This is because such consonants should trigger a CTI rule, if it were present. The lack of a CTI tone pattern on words with C1 voicing specification would thus indicate that the CTI rule had been dropped.

The following section will describe this model further, through the use of possible categories of example words, following each through the changes described in the model.

7.4 Hypothetical examples

Having addressed the transitions between four language systems in terms of general principles, I now apply those transitions to logically possible example types, before dealing with specific examples in Sections 7.5–7.7. Because of the potential complexity of the data to be presented in Sections 7.5–7.7, I show in this section how the model presented in Section 7.3 works out for each of a set of hypothetical examples, which are simplified to ignore distracting details, yet cover important categories. The goal in this section is not to avoid dealing with real data (which will be done in Sections 7.5–7.7), but rather to prepare the reader for those data by thoroughly examining the implications of the model presented in Section 7.3 for particular categories of data.

The figures in this section are organized according to four logical possibilities, for each language system. The four logical possibilities are organized by two binary options: according to the presence or absence of [L/Voice] on the initial root consonant (C1; rows *a/c* or *b/d*, respectively), and according to low or high tone on the vowel (rows *a/b* or *c/d*, respectively). The first set of examples, in Figure 82, represent these four possible consonant-tone possibilities, across the four language systems of Figure 79 without (or before) any devoicing:

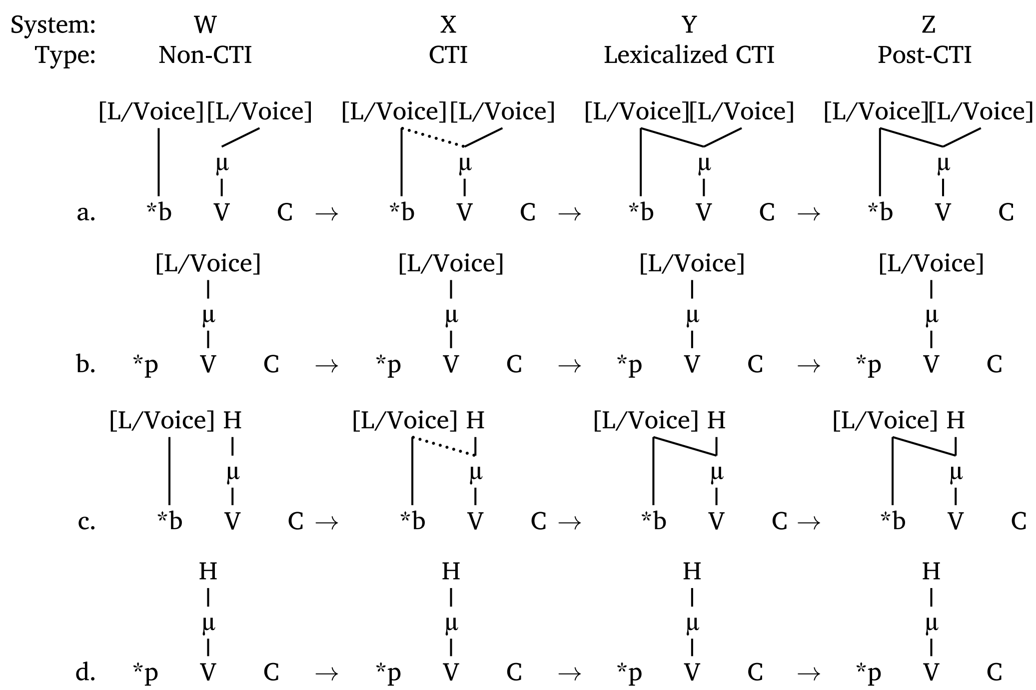


Figure 82. Examples of possible word types in four language systems **before or without** losing [L/Voice] specification

In Figure 82, row *a* represents a root with an onset with [L/Voice] specification and a low root tone; row *b* represents a root with an onset **without** [L/Voice] specification, but also with a low root tone. Row *c* represents a root with an onset with [L/Voice] specification again, but now with a **high** root tone, and row *d* represents a root with an onset **without** [L/Voice] specification, but again with a high root tone. The columns remain the same as in Figures 79 and 80, moving from a non-CTI system (W) to a post-CTI system (Z).

In terms of surface tone patterns, Figure 82 shows that words in groups W*a* and W*b* are the same (L), as are W*c* and W*d*, which are H. X*a* is LL, X*b* is L, X*c* is LH, and X*d* is H (though recall that the difference between X*a* and X*b* is predictable on the basis of a CTI rule, as is that between X*c* and X*d*). In this way, language X has two new surface tone melodies, derived from its two lexical melodies. Languages Y and Z have the same surface melodies as X, the only difference being that those surface forms result from lexical specification, rather than by rule.

The distribution of surface tone melodies in Figure 79 is summarized in Table 45:

	C1	**lex T**	**W**	**X**	**Y**	**Z**
Figure 82a	*b	*L	[L]	[LL]	[LL]	[LL]
Figure 82b	*p	*L	[L]	[L]	[L]	[L]
Figure 82c	*b	*H	[H]	[LH]	[LH]	[LH]
Figure 82d	*p	*H	[H]	[H]	[H]	[H]

Table 45. Surface melodies by word type and language type (without devoicing)

The lexical tone patterns across these systems are shown in Table 46. While X has four surface tone patterns (i.e., L, H, LL and LH in Table 45), it only has two lexical tone patterns (i.e., L and H in Table 46). Languages Y and Z, however, have four lexical tone patterns (i.e., L, H, LL and LH), each of which are also expressed on the surface (as in Table 45).

	C1	**lex T**	**W**	**X**	**Y**	**Z**
Figure 82a	*b	*L	/L/	/L/	/LL/	/LL/
Figure 82b	*p	*L	/L/	/L/	/L/	/L/
Figure 82c	*b	*H	/H/	/H/	/LH/	/LH/
Figure 82d	*p	*H	/H/	/H/	/H/	/H/

Table 46. Lexical melodies by word type and language type (without devoicing)

So far this model has shown a new set of surface tone patterns (in Table 45), and a new set of lexical tone patterns (in Table 46), but it is important to note that there is no change in consonant-tone distribution other than the addition of new melodies in the reanalysis of W → X (45a, on p181). That is, the distribution of surface tone patterns established in language X is maintained in Y and Z for the examples in Figure 82, i.e., for the data which have not undergone loss of voicing specification. Despite the fact that the tone patterns of languages Y and Z are lexicalized, their surface forms are still predictable on the basis of consonant types, as LL and LH (lexical and surface) melodies only go with **b*, and L and H (lexical and surface) melodies only go with **p*, without exceptions. So up to this point, the distribution of surface tone patterns across consonant types remains as in Table 43, as would be expected by a simple [L/Voice] analysis: consonants with voicing specification have an extra initial low tone, and others do not.

But devoicing can happen to a word of any word type (a-d), in any language type (W-Z), resulting in another dimension with possible combinations of consonants and tone patterns, based on whether or not a consonant loses voicing specification. Example categories showing this devoicing are given in Figure 83. The rows and columns in Figure 83 represent the same categories as in Figure 82, the only difference being that the words in Figure 83 have lost [L/Voice] specification through merger of **p* and **b* to [w].

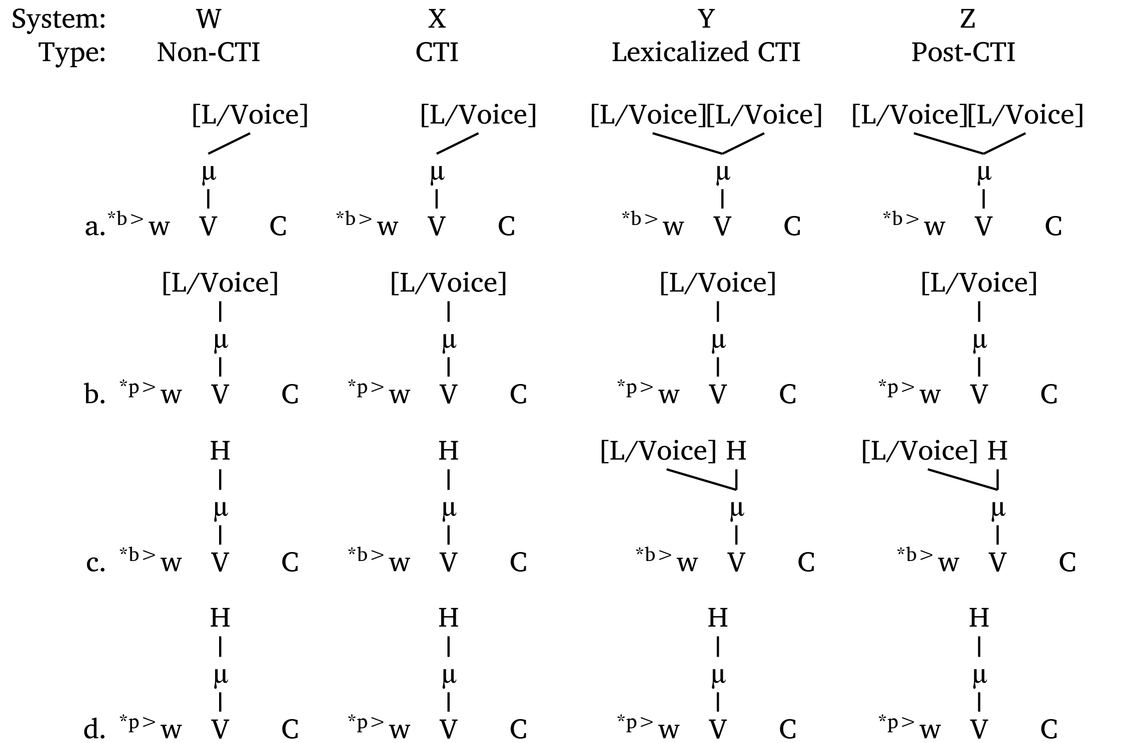


Figure 83. Examples of possible word types in four language systems **after** losing [L/Voice] specification

The exact consonant shift in Figure 83 is not important; it could be **b* > [p] (or at some other point or manner of articulation), so long as it involves a loss of voicing specification. I exemplify with a merger of **p* and **b* to [w] to make clear that the presence of surface voicing is irrelevant to the loss of voicing specification. In the Bantu D30 languages, [b] has both surface and underlying voicing, and [p] has neither, while [w] has surface voicing only (as there is no [w̥] in any Bantu D30 language, so [w] should not be analyzed as underlyingly voiced). This means that **b* > [w] represents a loss of voicing specification without change in surface voicing, while **p* > [w] involves a gain of surface voicing, though without adding underlying voicing specification.

In terms of surface tone patterns, Figure 83 shows that words which lose voicing specification in W and X each have either H or L, depending only on their lexical verb root class by tone. In fact, none of the examples in Figure 83 trigger the CTI rule, because the historically voiced consonants have lost their voicing specification (i.e., in rows *a* and *c*). This results in no extra surface tone melodies in Figure 83, where the surface forms depend only on their lexical tone melodies. The surface forms of the examples in Figure 83 are summarized in Table 47.

	C1	lex T	W	X	Y	Z
Figure 83a	*b > [w]	*L	[L]	[L]	[LL]	[LL]
Figure 83b	*p > [w]	*L	[L]	[L]	[L]	[L]
Figure 83c	*b > [w]	*H	[H]	[H]	[LH]	[LH]
Figure 83d	*p > [w]	*H	[H]	[H]	[H]	[H]

Table 47. Surface melodies by word type and language type (with devoicing)

While the CTI rule is likewise not triggered in languages Y and Z, they do have lexical tone patterns not in W and X, as indicated in Table 48. These new lexical tone patterns arose in the X → Y reanalysis of the CTI rule as lexical specification (45b on p181; crucially before these forms lost voicing specification). The lexical tone melodies in Table 48 are the same as the surface melodies in Table 47 because there is no application of a CTI rule, given the lack of voicing specification for all these categories. The lexical forms of the examples in Figure 83 are summarized in Table 48.

	C1	lex T	W	X	Y	Z
Figure 83a	*b > [w]	*L	/L/	/L/	/LL/	/LL/
Figure 83b	*p > [w]	*L	/L/	/L/	/L/	/L/
Figure 83c	*b > [w]	*H	/H/	/H/	/LH/	/LH/
Figure 83d	*p > [w]	*H	/H/	/H/	/H/	/H/

Table 48. Lexical melodies by word type and language type (with devoicing)

The variety of surface and lexical melodies in Table 47 and Table 48 are not much different from those in Table 45 and Table 46. In fact, the only difference is that in Table 47 language X is missing two surface forms (i.e., LL and LH, as the CTI rule doesn't apply there).

But there is a critical change in the distribution of those tone melodies across consonant types in Table 47 and Table 48. Because of the loss of voicing specification after the lexicalization of the CTI rule (i.e., in Y and Z), we now see words with C1 not specified for voice (i.e., [w]) with LL and LH melodies.

As noted earlier, W and X have the same surface patterns after loss of [L/Voice] specification in Table 47, with *a* and *b* groups L and *c* and *d* groups H. The case where W-X *a-d* all lack [L/Voice] specification does not impact the distribution of consonants and tone patterns, as words with consonants not specified for voicing (either [p] or [w]) bear H and L tone melodies, as they do without any loss of voicing specification (i.e., all of these words follow a distribution according to [L/Voice], as in Table 43).

In languages Y and Z, however, loss of [L/Voice] specification provides for all tone patterns (LL, L, LH and H) to appear on roots with consonants without [L/Voice] specification. The result of voicing specification loss in languages Y and Z, then, is a change from a CTI distribution like that in Table 43 (on p176) to the contrastive distribution observed in Bantu D30 languages today (in Table 44, on p177), where CTI tone patterns exist on words with and without voicing specification.

This change in distribution may seem trivial, but it is essential proof that a language has made the X → Y reanalysis lexicalizing the CTI rule (45b, on p181). As noted above, Table 45 and Table 46 do not provide distributional evidence of the lexicalization of their tone patterns, since the lexicalized patterns in Y and Z follow the same distributions as those derived by rule in W (because the X → Y lexicalization was based on the CTI rule). On the other hand, Table 47 and Table 48 show four tone melodies (lexical and surface) with the same C1 (i.e., [w]).

In the same way, evidence that a language has completed the Y → Z reanalysis would require a change in distribution. If a language has dropped a CTI rule, there should be at least some exceptions to such a rule. Evidence of a Z language, then, would be a more complete distribution than that in Table 44 (on p177), as in the one in Table 49:

\ C1 **verb root class **	**without [L/Voice]** **(vl obstruents, sonorants)**	**with [L/Voice]** **(vd obstruents)**
Low	+	?
High	+	?
Rising (CTI)	+	+

Table 49. Distribution of consonant types across tone melodies; to show language Z, at least one of ? must exist

Because the distribution to prove language Z goes beyond the observed distribution for Bantu D30 languages in Table 44 (on p177), we do not have evidence that these languages have lost their CTI rule, and may well remain at the language Y stage of this model. If this is the case, modern Bantu D30 languages maintain a CTI rule, while at the same time maintain lexical representations based on it.

This section concludes with a summary of all the example categories discussed so far. Putting the word types in Figure 82 (without devoicing) together with those in Figure 83 (with devoicing) provides eight word types: with historically H or L tone, with _*b_ or _*p_ as C1 source, and merged into [w] or not. The lexical and surface tone melodies for these word types are summarized in Table 50. The first two rows (above the line) are synchronically specified for [L/Voice]; the remainder are not.

	C1	**lex T**	**W**		**X**		**Y**		**Z**	
Figure 82a	b	*L	/L/	[L]	/L/	[LL]	/LL/	[LL]	/LL/	[LL]
Figure 82c	b	*H	/H/	[H]	/H/	[LH]	/LH/	[LH]	/LH/	[LH]
Figure 83a	$^{*b>}$w	*L	/L/	[L]	/L/	[L]	/LL/	[LL]	/LL/	[LL]
Figure 83c	$^{*b>}$w	*H	/H/	[H]	/H/	[H]	/LH/	[LH]	/LH/	[LH]
Figure 83b	$^{*p>}$w	*L	/L/	[L]	/L/	[L]	/L/	[L]	/L/	[L]
Figure 83d	$^{*p>}$w	*H	/H/	[H]	/H/	[H]	/H/	[H]	/H/	[H]
Figure 82b	p	*L	/L/	[L]	/L/	[L]	/L/	[L]	/L/	[L]
Figure 82d	p	*H	/H/	[H]	/H/	[H]	/H/	[H]	/H/	[H]

Table 50. Melodies by word type and language type (new melodies highlighted)

To summarize the distributions of these data across tone melodies and consonant types, those that are specified for [L/Voice] show new melodies LL and LH (only), beginning with language X. While a more complex system than that in language W which only has H and L), the new melodies introduced in language X are predictable on words with C1 which are specified for voicing (the first two rows in Table 50). The next two rows in Table 50, however, show these same new tone melodies on words with C1 which are not specified for voicing, where they are not predictable, in languages Y and Z. This is because they bear the additional low tone historically derived from [L/Voice] on C1, though they are preceded by a consonant which is not specified for voicing (i.e., [w]). Together with the final four rows, Languages Y and Z have four tone melodies (i.e., LL, LH, L, and H) on words with C1 not specified for voicing, whereas languages W and X only have two (i.e., L and H).

The end result of the combination of CTI lexicalization and loss of [L/Voice] specification is that a system which had a simple two way tone split, both underlyingly and on the surface (i.e., W) now has new tone melodies, which are completely independent from the consonants on the words which carry them.

In terms of distribution of consonants across surface tone melodies, the last six rows of Table 50, which lack [L/Voice] specification, show all four tone melodies, while the two rows with [L/Voice] specification only have two of the available tone melodies, i.e., LH and LL. This lopsided distribution is because there is no corresponding process to add [L/Voice] specification, which would be required to give [L/Voice] specification to those that historically lacked it, and would therefore have the L and H tone melodies. The result is precisely the distribution of consonants and tone melodies that we see in Table 44 (on p177). As mentioned above, the complete lack of exceptions to this distribution (i.e., no voiced C1 H or L) argues for Y as opposed to Z in contemporary Bantu D30 languages, as the CTI rule is likely still in effect. Were this not the case, exceptions would eventually enter the system.

In order to move from the initial evidence of a Z system (as in Table 49) to a more complete distribution of consonants and tone melodies (as in Table 51), I would expect one of two conditions to hold. One possibility would be a new voicing rule (e.g., intervocalic voicing). With a systematic source of new voicing contrasts (in addition to the devoicing already described for Bantu D30), a language would have a robust context for gaining as well as for losing voicing specification. The result would be that words in the High and Low verb root classes would be given [L/Voice] specification in large numbers, resulting in the complete distribution in Table 51.

\ C1 verb root class \	without [L/Voice] (vl obstruents, sonorants)	with [L/Voice] (vd obstruents)
Low	+	+
High	+	+
Rising (CTI)	+	+

Table 51. Complete distribution of consonant types across tone melodies

Alternatively, a complete distribution of consonant types across tone melodies should eventually be achieved even without a systematic voicing rule, by the passage of time. With sufficient time since the loss of the CTI rule, the normal course of a language adding borrowings, coined words, and sporadic voicing would eventually provide enough High and Low verbs with contrastive voicing so as to appear systematic. One way or another, I expect that any language passing through my model to language Z (i.e., having lost a CTI rule), will ultimately show a complete distribution as in Table 51.[7]

Finally, given the eight classes of words across the four language systems, the model presented in this chapter predicts that there would not be a simple and straightforward change of categorization, but rather that as each language undergoes loss of voicing specification and CTI lexicalization, each word may or may not lose [L/Voice] specification at any point, resulting in a large number of different patterns of correspondences across related languages. Yet there are predicted to be constraints on those variations, as mentioned in Section 7.3. And correspondence patterns varying within certain parameters is just what we find, as will be described in Sections 7.5–7.7.

7.5 Bilabial Correspondences with some loss of voicing specification

The data in this section show bilabial consonants in words with tone patterns not predicted by a straightforward analysis using Bradshaw (1999)'s [L/Voice]. That is, they lack voicing specification, yet appear in words with Rising tone melodies. As the data to be presented in this section will show, these consonants correspond with consonants which are specified for voice in related languages (other D30 languages and/or in their Proto-Bantu sources). Furthermore, the observed patterns of correspondence are varied, as predicted by my model.

Throughout this section, two types of reflexes which have lost voicing specification will be presented. Some retain Rising tone patterns, which I take to mean the loss of voicing occurred after CTI was lexicalized for that word in that language. That is, the loss of voicing occurred while the language was in stage Y or Z on Figure 83. Other reflexes will show loss of voicing before CTI was lexicalized for that word in that language. These do not have Rising patterns, and reflect the loss of voicing while the language was in stage W or X of Figure 83.

Furthermore, the cognate sets presented in this and following sections will show that a word in a given language may be of one type (e.g., with W/X devoicing, before CTI lexicalization), and a cognate word in another language may be of the other type (e.g., with Y/Z devoicing, after CTI lexicalization). In fact; this is most

[7] For some time after achieving a complete distribution (as in Table 51), I expect that the relative number of verbs in each category would still indicate which category were more recently empty.

frequently the case; I have few cognate sets where devoicing and CTI lexicalization happened in the same order across all Bantu D30 languages. Because of this diversity across languages from cognate set to cognate set, a large part of the discussion of these data is taken up with the significance of the fact that one language devoiced a particular word before CTI lexicalization, while a sister language devoiced a cognate word after CTI lexicalization. This issue is taken up throughout this section, in figures showing words from each of two languages, both derived from the same proto-Bantu source, but showing devoicing before and after CTI lexicalization.

Recall from the introduction to Chapter 6 that proto-Bantu reconstructed forms are included when a plausibly similar form is attached to the same (or plausibly similar) gloss in the BLR3 (Bastin, et al 2002). Where I have some doubt of an entry being the historical source of the cognate set, I precede it with a question mark (e.g., ?139: **bèga* in Table 56). Where multiple BLR3 entries are plausibly similar, I include the most plausibly similar (in my judgement) in the table, and others in a footnote (e.g., 'father (n)' in Table 52). Where a BLR3 gloss differs from the gloss used for the Bantu D30 languages, the D30 gloss is given in the gloss column, with the BLR3 gloss in the BLR3 column (e.g., 1632: **jòngʊ́* 'cooking-pot (n)' for 'cook' in Table 64).

Unfortunately, I have not been able to collect complete (i.e., segmental and tonal) forms for all the languages in each cognate set in this chapter. Where data are lacking, I provide what I have (e.g., Vanuma 'father (n)', which lacks tone data in Table 52), but do not include those data in my discussion of that cognate set.

The first cognate sets presented will include sources that fairly certainly derive from proto-Bantu **b*. In Table 52, we see loss of [L/Voice] specification with a clearly maintained Rising tone pattern in Nyali-Kilo [nlj] only (shaded):

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[¯ ⟋]		[– –]		[– –]	[– ¯]	12:	
ɑ**w**e	ɑ**β**ɑ	**b**ɑ**b**ɑ		**b**ɑ**b**ɑ	**b**ɑ**b**ɑ	***b**à**b**á*[8]	'father (n)'

Table 52. Historical depression in Nyali-Kilo only (*b)

The historical derivation of the Nyali-Kilo [nlj] form is compared with that of the Mbo [zmw] form in Figure 84. Note that the innovation **b* > [w] follows the

[8] An alternative source for this cognate set is ?9428: **pàpá* also glossed 'father (n)'

loss of the **b* in C1 position, an innovation in Nyali-Kilo presumably shared with Vanuma[9]:

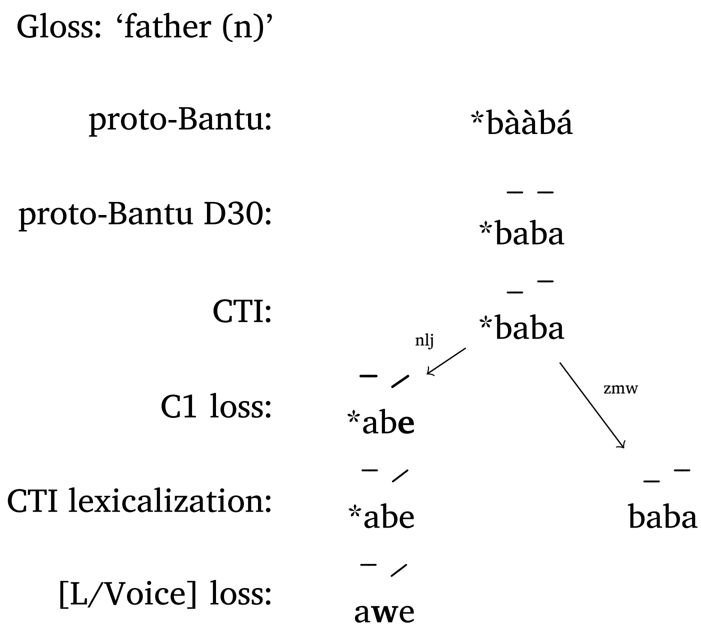


Figure 84. Comparative derivation of 'father (n)' in Nyali-Kilo [nlj] and Mbo [zmw]

Figure 84 shows C1 loss in Nyali-Kilo [nlj] after developing CTI. Because CTI normally applies to C1 (at least in the Bantu D30 languages), the loss of C1 in Nyali-Kilo apparently makes the CTI rule apply to the next syllable, causing a change to a tone pattern with depression on the second syllable. After CTI is lexicalized in each language, voicing specification[10] is lost in Nyali-Kilo, but not in Mbo. This results in synchronic forms which correspond with the data presented in Table 52.

[9] There is a potentially suspicious change from *LH > L/H > LH from proto-Bantu to the CTI form in Figure 84. As previously stated, I do not defend my proto-Bantu D30 forms at this time, though this is the form that would result in the ultimate derivation of the current surface forms. I do not believe that it can be known with much certainty how CTI operated, if at all, in proto-Bantu or its peers, though I do find it suspicious that this LH proto-Bantu noun has a contrastively voiced C1. Further speculation will be reserved for Appendix F.

[10] Recall that **b* > [w] represents a loss of voicing specification.

In Table 53, all languages (except D32 Bɨra [brf]) lose voicing specification coming from proto-Bantu **b*. All this devoicing was before CTI was lexicalized, as no Rising tone patterns result from the C1 bilabial without voicing specification.[11]

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[– – –]	[– – –]	[– ¯]	[¯ ¯ –]	[– – –]	[– – –]	14:	
kɑ**w**olo	kɑ**β**oɹo	**b**ɑɾɑ	ku**p**oo	ko**p**oo	ko**ɸ**oo	*****b****àd*	‘count’

Table 53. Data without Rising patterns due to **b in any language*

The data in Tables 52 and 53 are important because they show the two different orderings of CTI lexicalization and loss of voicing specification, in two different words of the same language. In Table 52, the Nyali-Kilo form lost voicing specification **after** CTI lexicalization, as the Rising form is maintained synchronically. But in Table 53 the Nyali-Kilo form lost voicing specification **before** CTI lexicalization, as the synchronic tone melody is not Rising, but Low. This difference in ordering happened despite the fact that the C1's of the two Nyali-Kilo words derive from the same proto-Bantu consonant **b*, and have the same reflex, [w].

The presence of forms reconstructed to the same consonant, yet behaving differently in terms of their tone, argues against a categorical devoicing and lexicalization processes across (at least) Nyali-Kilo. That is, even if one of the two processes (devoicing and CTI lexicalization) were argued to have happened categorically across these languages, the other must have occurred both before and after it (but see reasons in Section 7.8 to doubt that either process applied categorically).

Table 54 shows the cognate sets for ‘bark (n)’ and ‘calabash (n)’, with loss of voicing specification after CTI lexicalization in Mbo [zmw] only (shaded):

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
	[– – –]	[– ¯ ¯]	[– – –]	[– – –]	[¯ – –]	?4790:	
ɨ**w**ɑ**w**i	ɨ**β**ɑ**β**i	kikobɑ	ɨ**p**ɑ**p**ɨ	ɨ**p**ɑ**p**ɨ	ɨ**ɸ**ɑ**ɸ**ɨ	*****p****ú*	‘bark (n)’
[– – –]	[– – –]				[¯ – –]	?746:	
ʉ**w**ɑtɑ	ʉ**β**ɑtɑ			ʉ**p**ɑtɑ	ʉ**ɸ**ɑtɑ	*cʉ́**p**à	‘calabash (n)’

Table 54. Historical depression in Mbo ($*p > *b$)

[11] There is depression in Budu ‘count’, but it is due to the **d* in C2 position, not the **b* in C1 position, as the depression falls on the third, not second syllable (see Section 7.6).

The cognate sets in Table 54 are each likely reconstructed with **p* in proto-Bantu, but appear to have passed through a voiced intermediary (i.e., **p* > **b* in proto-Bantu D30 or a more recent ancestor). I consider this and other cognate sets to have passed through **b* at some point for multiple reasons, as mentioned in Section 6.4. First, they produce the same segmental correspondence patterns across the Bantu D30 languages (i.e., [w]:[β]:[p]:[ɸ]), and should therefore be reconstructed as one consonant at some point, as discussed in Section 6.4. This section adds another (though lesser) argument: their tone patterns behave the same as those reconstructed with **b*. That is, the Mbo forms in Table 54, though likely coming from proto-Bantu **p,* each act as if they had voiced ancestors, in that their tone patterns are depressed[12].

The cognate set for ‘winnow, flutter’ in Table 55 is also reconstructed with **p* in proto-Bantu, and also appears to have passed through a voiced intermediary, though with a different correspondence pattern. This correspondence pattern shows proto-Bantu **p* with voicing in Nyali-Kilo [nlj], Vanuma [vau], and Mbo [zmw], but without voicing in Bɨra [brf], Budu [buu], and Ndaka [ndk]. The Rising tone pattern only remains without voicing in Ndaka (shaded):

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[¯ — —]	[— _ ¯]	[¯ ¯]	[– ¯ –]	[– ⁄ –]	[¯ – ¯]	2463:	‘winnow’,
kɑ**b**ʉ**b**ɑ	kɑ**b**ʉ**b**ɑ	**p**u**p**ɑ	kʉ**p**ɔ**p**ɑ[13]	kɔ**p**ɔ**p**ɑ	kɔ**b**ɔ**b**ɑ	***pép***	‘flutter’

Table 55. Rising tone patterns with and without devoicing, plus late loss of [L/Voice] specification in Ndaka (**p* > **b*)

The data in Table 55 add an additional argument to the thesis that there was some voicing between proto-Bantu and some ancestor of the Bantu D30 languages. The relevant points are a proto-Bantu reconstructed form with **p* for each of C1 and C2, with corresponding consonants voiced in reflexes for Nyali-Kilo, Vanuma, and Mbo. One must on the face of it propose either the implausible fact that this voicing occurred three times, or else that the voicing happened while at least Nyali-Kilo, Vanuma, and Mbo were one language. One could propose that it was an

[12] That is, of course, unless the thesis of this chapter is entirely wrong, and there is some other reason apparently depressor tone patterns should appear on words with voiceless segments. But given that my thesis already works for the rest of the data (as this chapter will show), and along with the stronger arguments repeated here from Section 6.4, I will continue on the assumption that the forms in question had a voiced ancestor.

[13] This Budu [buu] form is glossed ‘fan fire’

ancestor of closely related Nyali-Kilo and Vanuma only, but one must still account for voicing in synchronic Mbo, so voicing must have happened at least twice, if not for all Bantu D30 at one time (i.e., in proto-Bantu D30). The hypothesis that Mbo is more closely related to Nyali-Kilo and Vanuma than to Ndaka is not borne out by even the most casual glance at the languages, so at least Ndaka must have a voiced ancestor. But the simpler analysis is that proto-Bantu D30 innovated the voicing on consonants in Table 55, when it was one language, after which Bɨra [brf], Budu [buu], and Ndaka [ndk] each lost voicing. This analysis also accounts for the Rising tone melody on the Ndaka form, without a voiced ancestor of this form, one would need to develop another analysis to explain the origin of this Rising melody.

The derivation of ‘winnow, flutter’ from Table 55 is given in Figure 85 for Ndaka [ndk] and Budu [buu] :

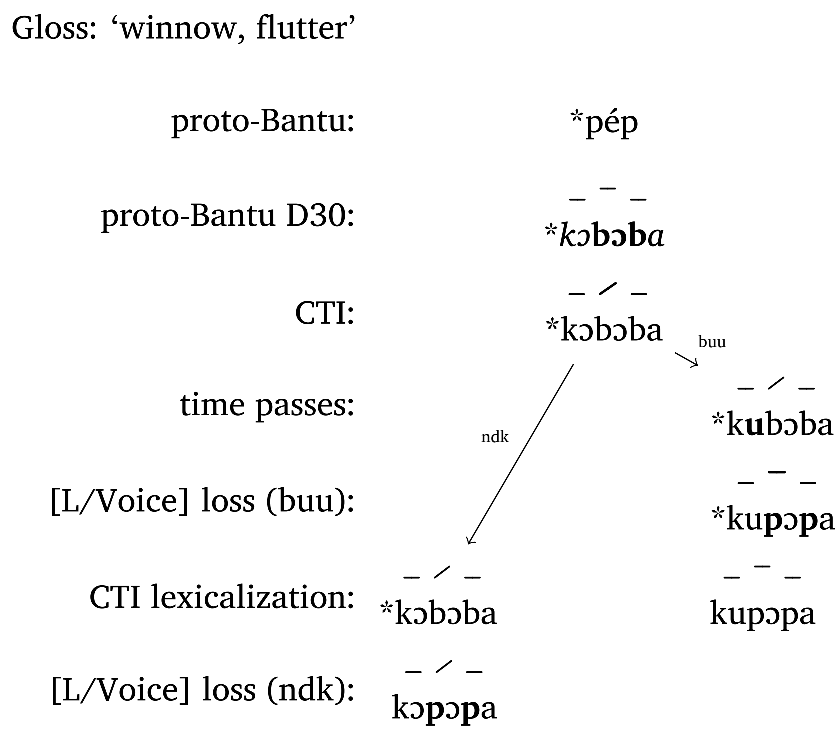


Figure 85. Comparative derivation of ‘winnow, flutter’ in Ndaka [ndk] and Budu [buu]

In Figure 85, we see the loss of voicing specification in Budu [buu] before CTI lexicalization, but in Ndaka [ndk] after CTI lexicalization, resulting in synchronic forms which correspond with the data presented in Table 55. Recall that ‘winnow, flutter’ is presented in a spectrogram in Figure 78 in Section 7.2. The Mbo data, from Figure 77, show no loss of voicing at all.

The data in Tables 52, 54, and 55 show late (YZ) devoicing in only one language, for each of Nyali-Kilo, Mbo, and Ndaka, respectively. This again speaks to the nature of devoicing as a spontaneous occurrence, which can happen in one language at a different time than for very closely related languages. These data also provide some evidence of the variety of correspondence patterns expected by my model of tonogenesis, where devoicing may happen either before or after CTI lexicalization, for each word in each language.

This spontaneity of devoicing does pose some difficulty for the reconstruction of the history of these languages. For instance, the data in Table 56 show forms in Budu [buu] and Ndaka [ndk] which lost voicing specification **after** CTI lexicalization, as tonal depression remains synchronically (shaded). On the other hand, the Mbo [zmw] forms lost voicing specification **before** CTI lexicalization, as they do not show tonal depression:

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
			[¯ – – –]	[¯ – –]	[– ¯ –]	290:	
lɨ**w**o**w**i	lɨ**β**ɔ**β**ɨ	n.c.[14]	ɨ**p**o**p**ɨɑ	jɑ**p**ɔ**p**ɨ	jɑ**ɸ**ɔ**ɸ**ɨ	****b**ʊ̀**b**ì*	‘spider (n)’
				[¯ – –]	[– – –]	?139:	
ʉ**w**ɑkɑ	ʉ**β**ɑkɑ**β**ɑkɑ	n.c.	ʉ**p**ɑkɑ	ʉ**p**ɑkɑ	ʉ**ɸ**ɑkɑ	****b**ègà*[15]	‘shoulder (n)’

Table 56. Historical depression in Ndaka and Budu, but not Mbo (*b)

The data in Table 56 are somewhat problematic because even a cursory glance at the Bantu D30 languages shows a large degree of similarity between Ndaka and Mbo, beyond the similarity of either to Budu. Yet in Table 56 Ndaka and Budu have the same ordering of voicing specification loss and CTI lexicalization, and different than that of Mbo. But rather than argue that Ndaka and Budu share an innovation which is not shared by Mbo (i.e., that they remained one language after they and Mbo split), it seems more plausible that voicing specification loss happened independently in Ndaka and Budu, such that Ndaka and Budu share similar timing for voicing specification loss in ‘spider (n)’ by coincidence. This coincidence does require that devoicing be a somewhat spontaneous, natural process.

[14] There is perhaps a cognate form in *bakambalimbali*[– – ¯ ¯ – –]

[15] another possible source for this cognate set is 8708: **panga*.

The derivation of ‘shoulder (n)’ from Table 56 is given in Figure 86 for Ndaka [ndk] and Mbo [zmw]:

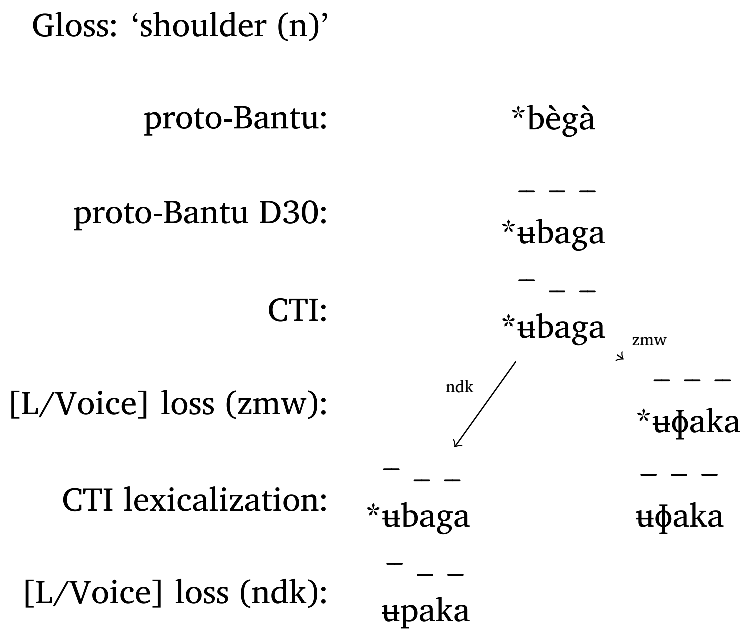


Figure 86. Comparative derivation of ‘shoulder (n)’ in Ndaka [ndk] and Mbo [zmw]

In Figure 86, loss of voicing occurred in Mbo [zmw] before CTI lexicalization, while loss of voicing occurs after CTI lexicalization for Ndaka [ndk], resulting in synchronic forms which correspond with the data presented in Table 56. Recall that despite the difference in derivation shown in Figure 86, Ndaka and Mbo almost certainly share a more recent history with each other than either does with any other language.

Spontaneous devoicing should not only produce late (YZ) devoicing in a single language, as shown above. For instance, Table 57 shows late (YZ) devoicing across all D33 languages for ‘blow’, ‘show’, and ‘cough’ (shaded). The forms for ‘blow’ are given as originally transcribed, though I have no real evidence to distinguish the presence of a semivowel onset (as in Nyali-Kilo and Vanuma) from the lack of a root onset (as in Ndaka and Mbo). Assuming the presence of an onset phonetically, it is further unclear if such a segment would be phonemic or epenthetic, as a resolution of vowel hiatus.

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[_ ⁄ ¯] ka**w**ulo	[¯ _ ¯] ka**w**ulo	[– –] huwɑ		[– ⁄ –] kouo	[¯ _ ¯] kou**ɸ**o	2672: ****púd***[16]	‘blow’
kɑ**w**ɑnɑkiso	kɑ**w**ɑnɑkiso	tɑndɑ		[– ⁄ –] kɔ**p**ɑnɑ	[¯ _ ¯] kɔ**ɸ**ɑnɑ		‘show’
[¯ ¯ _] kakʉ**w**ɑ	kɑkʉ**β**ɑ	[– –] k**p**olɑ	[¯ ¯ _] kʉkʉɑ	[– ⁄ –] kɔkɔ**w**ɑ	[¯ _ ¯] kɔkɔ**f**ɑ	1958: *kó**p**ʉd[17]	‘cough’

Table 57. Rising tone patterns due to (**p* > **b*)

Because these languages are known to be related on other grounds, it should not be surprising to find words where they all agree on the order of devoicing, in this case after CTI lexicalization. In this case, it may well be that these words lexicalized the CTI rule while these were all one language, followed by a loss of voicing specification, either as individual languages or before separation into smaller groups of languages (as they do not all have the same reflex today).

In summary, this section has shown evidence of three more general principles. The first is the spontaneity of devoicing, at least for these languages. This is attested to by data showing three languages acting independently of their sister languages in late (YZ) devoicing, as well as by different ordering for different words in the same language. Because of this spontaneity, similar ordering between languages for a given word does not necessarily imply a history more recently shared than that shared with languages who differ on that word.

The second general principle attested to by the data in this section is that it is plausible to assume a voiced intermediary between the proto-Bantu reconstructed form and its reflexes in the Bantu D30 languages of today. This is attested to by data from proto-Bantu sources which are not specified for voicing, but which nonetheless have voiced reflexes today (evidencing that voicing has emerged, on at least these occasions), or else Rising tone patterns (which require some sensible origin).

The third principle is that there are a variety of patterns, as expected by my model of tonogenesis. In addition to the above mentioned data where one language disagrees with her sisters, this section also shows data with late (YZ) devoicing shared across all Bantu D33 languages. This is expected in my model, because if devoicing is spontaneous it should be expected to have occurred both while these

[16] Another possible source for this cognate set is 2660: **pùʊp*, also glossed ‘blow’
[17] Another possible source for this cognate set is 1868: **kɔ́c*, also glossed ‘cough’

languages were one language, as well as at each later stage in their development. This point will be taken up again in Section 7.8.

7.6 Alveolar correspondences with some loss of voicing specification

The data in this section show alveolar consonants with tone patterns not predicted by a straightforward analysis using Bradshaw (1999)'s [L/Voice]. That is, they lack voicing specification, yet appear in words with Rising tone melodies. As the data to be presented in this section will show, these consonants correspond with consonants which are specified for voice in related languages (other D30 languages and/or in their Proto-Bantu sources).

Cognate sets also show devoicing of proto-Bantu **d* at different times in different languages, as will be detailed in this section. Recall that the correspondence pattern from Section 6.5 has current reflexes of [l] in Nyali-Kilo and Vanuma, but the segment has been lost in Budu, Ndaka, and Mbo. But interestingly for the historical validity of the model presented in this chapter, a historical depressor effect can be seen even where there is no consonant observed today.

The data in Table 58 show a loss of voicing specification after CTI lexicalization in Mbo [zmw] only (shaded). The Ndaka and Mbo forms are given as originally transcribed, though I have no real evidence to assert a contrastive C2 root segment (as in Nyali-Kilo and Vanuma). That is, [j] may well be epenthesized to resolve the hiatus caused by the loss of **d*, rather than the result of an implausible **d* > [j] innovation. In any case, the voicing specification originally present in **d* has been lost.

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[⁻ ⁻ ⁻]	[⁻ ⁻ ⁻]	[⁻ ⁻ ⁻]		[⁻ ⁻ ˥˨]	[⁻ ˍ ⁻]	2184:	
kamɨla	kamɨla	kamɨɾa [18]		komijo [19]	kɔmɨja	**mì**d***	‘swallow’

Table 58. Rising tone pattern in Mbo only

The derivation of ‘swallow’ is given for Mbo [zmw] and Nyali-Kilo [nlj] in Figure 87:

[18] This form is glossed ‘squeeze, press’

[19] c.f., *kɔmɨja* – ⟋ – ‘throw’

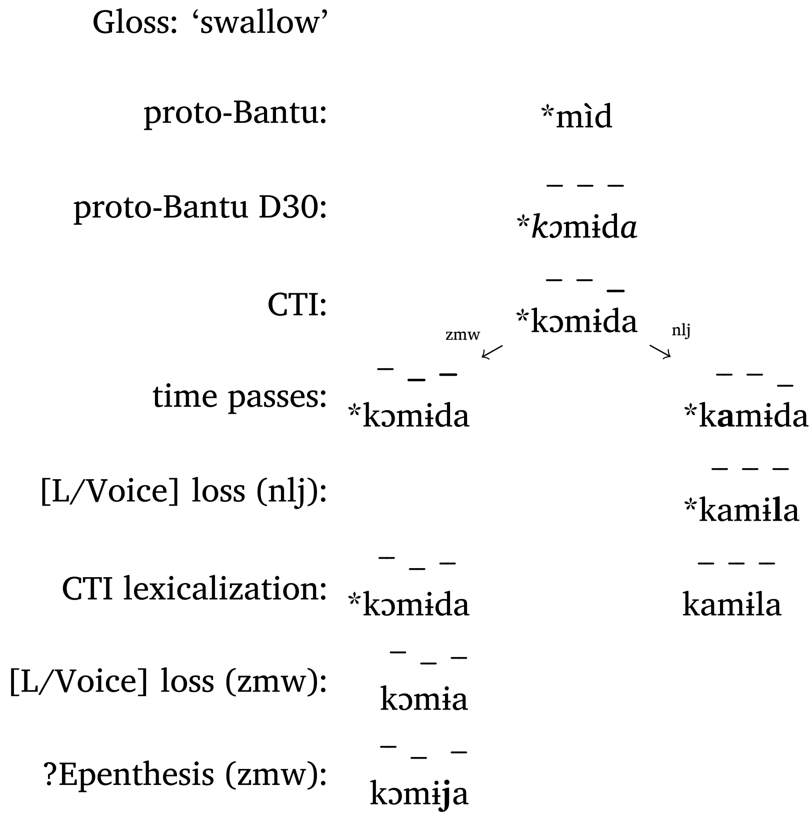


Figure 87. Comparative derivation of 'swallow' in Mbo and Nyali-Kilo

The derivation in Figure 87 shows the differential ordering of loss of voicing specification and CTI lexicalization. That is, the loss of voicing specification occurred before CTI lexicalization for Nyali-Kilo [nlj], but after CTI lexicalization for Mbo [zmw], resulting in synchronic forms which correspond with the data presented in Table 58.

As with the bilabial data presented in Section 7.5, the data in Table 58 show *d devoiced late (YZ) in one language only (i.e., Mbo). The data in Table 59 show a loss of voicing specification after CTI lexicalization in Budu [buu] only (shaded):

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[− − −]	[− − −]	[− ⁻]	[⁻ ⁻ –]	[− − −]	[− − −]	14:	
kɑwo**l**o	kɑβoɹo	bɑɾɑ	kupoo	kopoo	koɸoo	***bàd***	'count'

Table 59. Rising tone pattern in Budu only

The derivation of ‘count’ is given for Budu [buu] and Ndaka [ndk] in Figure 88:

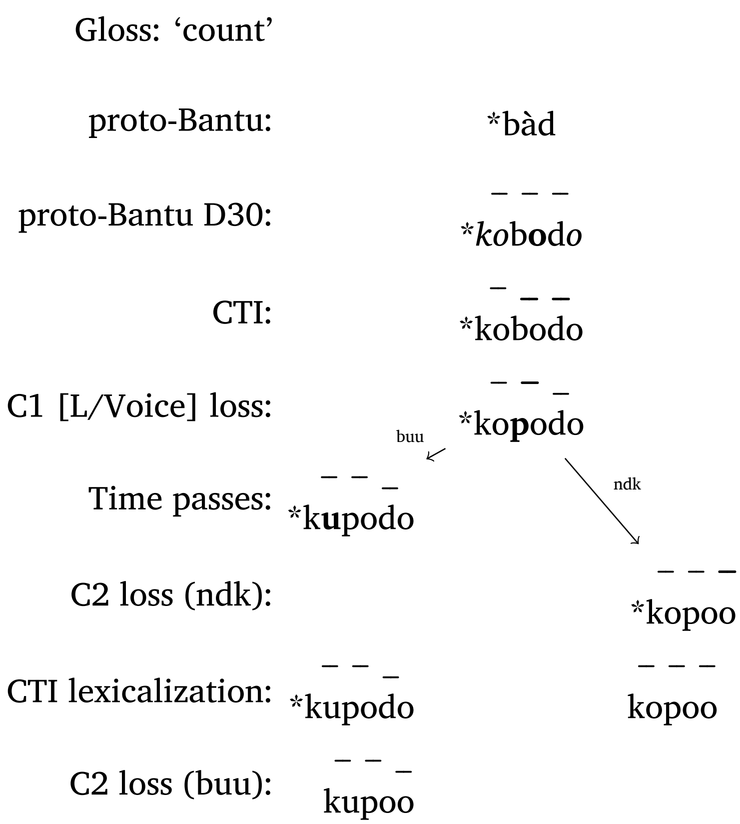


Figure 88. Comparative derivation of ‘count’ in Budu and Ndaka

The derivation in Figure 87 shows loss of voicing specification (with the rest of the consonant) **before** CTI lexicalization for Ndaka, but **after** CTI lexicalization for Budu, resulting in synchronic forms which correspond with the data presented in Table 59.

The cognate sets ‘swallow’ and ‘count’ are interesting in that they both derive from **d* in C2 position, yet the order of loss of voicing specification and CTI lexicalization is not the same across languages for each. The differential ordering of voicing specification loss and CTI lexicalization in Table 58 and Table 59 further strengthens the idea that loss of voicing specification and/or CTI lexicalization occurred spontaneously and potentially independently for each language.

But it is not the case that all data show one language with late (YZ) devoicing, in opposition to her sisters. The cognate set in Table 60 shows loss of voicing specification on **d* after CTI lexicalization for Budu [buu], Ndaka [ndk], and Mbo [zmw] only (shaded):

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[⁻ ⁻ ⁻]		[– ⁻]	[⁻ ⁻ –]	[⁻ ⁻ –]	[⁻ ⁻ –]	253:	
kɑwʉ**l**ɑ	yɑibo**l**e	hɔlɑ	kʉpʉwɑ	kɔpɔwɑ	kɔɸɔwɑ	**bò**d***	'rot'

Table 60. Depressed Low tone patterns due to **d* in Budu, Ndaka, and Mbo

The data in Table 60 are interesting because they are evidence of a possible innovation (*d > ∅ after CTI lexicalization, followed by [w] epenthesis as transcribed in Table 60) shared by Budu, Ndaka, and Mbo. This will be taken up again in Section 7.8.

The difference in derivation of 'rot' between Ndaka [ndk] and Nyali-Kilo [nlj] is given in Figure 89:

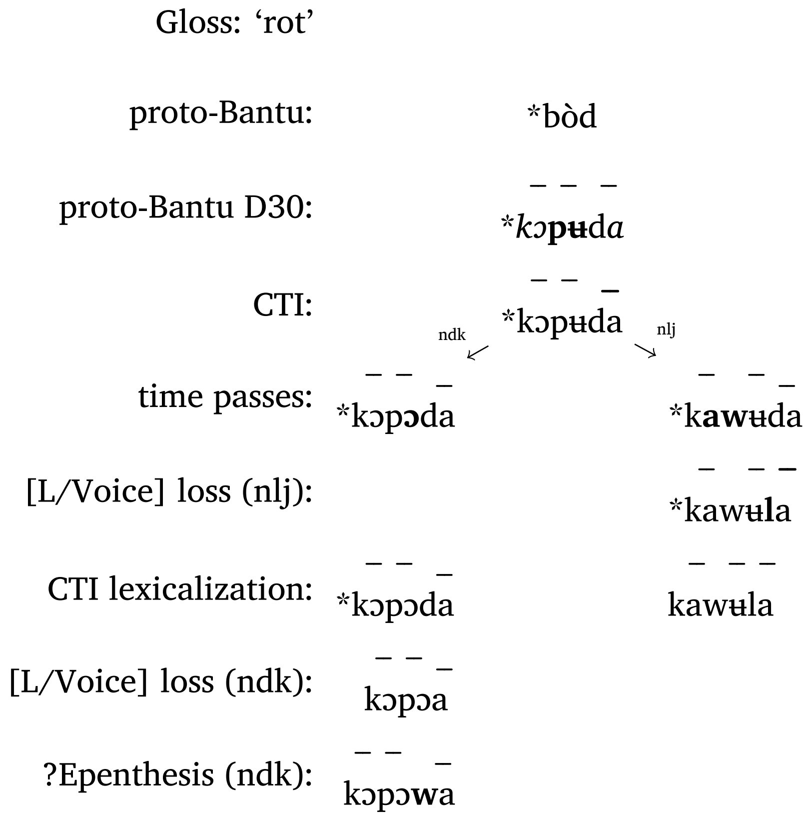


Figure 89. Comparative derivation of 'rot' in Ndaka [ndk] and Nyali-Kilo [nlj]

Figure 89 is interesting in that it speaks to the transition from **d* > [l] and **d* > ∅ in Bantu D30. I have in the past assumed the existence of one chain (**d* > **l* > ∅), but the necessary ordering in Figure 89 to derive 'rot' in Table 60 indicates that the **d* > [l] shift in Nyali-Kilo occurred before CTI lexicalization, whereas Ndaka maintained [L/Voice] in C1 of this word until after CTI lexicalization. So whether the Ndaka transition from **d* > ∅ occurred with an **l* intermediary or not, it must have begun after (i.e., at a different time than) the Nyali-Kilo shift away from **d*[20].

The cognate sets in Table 61 show loss of voicing specification on **d* before CTI lexicalization for all languages:

[20] This is true unless, of course, CTI in Ndaka 'rot' was lexicalized independently of (either before or in place of) Nyali-Kilo 'rot'. In either case, as the languages would not have shared CTI lexicalization for 'rot', they would necessarily have been independent of each other before the Nyali-Kilo shift away from **d* (which occurred after CTI lexicalization for Nyali-Kilo 'rot'), and so Ndaka and Nyali-Kilo could not have shared **d* > [l], at least for 'rot'.

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[_ ⁄ ¯]	[¯ – ¯]	[− −]		[– ⁄ –]	[¯ – ¯]	2672:	
kɑ(w)ulo	kɑ(w)ulo	huwɑ		kouo	kouɸo	**púd**[21]	‘blow’

Table 61. No depressed Low tone patterns due to *d*

The data in Tables 59-61 are interesting because even with bilabial C1 and alveolar C2 plosives in each case, the proto-Bantu **d* does not result in the same reflex tone patterns across these words. That is, ‘count’ from Table 59 shows depression from **d* in Budu only, while ‘rot’ from Table 60 shows depression from **d* in Budu, Ndaka, and Mbo, and ‘blow’ in Table 61 shows no depressor tone pattern in any language (none that are likely due to the **d* in C2 position[22], in any case). This speaks strongly to the spontaneity of devoicing in these languages, as the difference in devoicing patterns cannot obviously be accounted for by the phonological environment.

The cognate set in Table 62 indicates a loss of voicing specification after CTI lexicalization for most Bantu D33 languages (i.e., Nyali-Kilo [nlj], Budu [buu], Ndaka [ndk], and Mbo [zmw], each shaded):

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[¯ ¯ –]	[− − −]	[– ¯]	[¯ ¯ –]	[– ⁄ –]	[¯ – ¯]	1267:	
kɑɗulo	kɑɗulo	ɓʉlɑ	kuɗuo	koɗuwo	koɗuwo	**dù**t***	‘pull (up)’

Table 62. Rising tone patterns in most languages (**t* > **d*)

Arguments for a voiced intermediary between the proto-Bantu C2 **t* and the reflexes given in Table 62 are similar to those presented for **p* > **b* in Section 7.5. First of all, the correspondence pattern for **t* in Table 62 is the same as for **d* in Tables 58-61, as well as for the data coming from **d* in Section 6.5, i.e., [l] in Nyali-Kilo and Vanuma, and consonant loss in Budu, Ndaka, and Mbo.

The presence of a C2 depressor (i.e., not the C1 [ɗ], which is not specified for voicing) is somewhat puzzling here, though it is clear that the Nyali-Kilo [nlj] and Budu [buu] forms are depressing on the third syllable, rather than the second.

[21] Another possible source for this cognate set is 2660: **pʊ̀ʊp*

[22] The apparent metathesis in the Mbo form indicates that there may be something else going on between the forms in the ‘blow’ cognate set in Table 60. Unfortunately given that the reflex of **d* is the lack of a consonant for Budu, Ndaka, and Mbo, and given that Ndaka has lost the **b* as well, any metathesis in Ndaka and Budu would be difficult to see.

Ndaka and Mbo do not typically have tone patterns showing depression on the third syllable of the verb, though see Table 60 for an example. In any case, it would appear that the historical C2 depressor triggered the more typical depressor tone patterns for Ndaka and Mbo in Table 62.

The final cognate set in this section shows the same correspondence pattern from a **n* proto-Bantu source, this time with a loss of voicing specification after CTI lexicalization in Budu [buu] only (shaded):

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[– – –]		[– ¯]	[¯ ¯ –]	[– – –]	[– – –]	3547:	
kɑwʉ**l**ɑ		hɔ**l**ɑ	kʉ**p**ʉɑ	kɔpɔ**w**ɑ	kɔ**ɸ**ʉɑ	**jó**n***	‘spoil’

Table 63. Rising tone patterns in Budu [buu] only (**n* > **d*)

Arguments for a voiced intermediary between the proto-Bantu C2 **n* and the reflexes given in Table 63 are similar to those presented for **p* > **b* in Section 7.5 and for Table 62. First of all, the correspondence pattern for **n* in Table 62 is the same as for other data coming from **d*, i.e., [l] in Nyali-Kilo, and consonant loss in Budu, Ndaka, and Mbo. Additionally, there is the depression in Budu after the C2 slot, which would be unaccounted for without a voiced consonant in C2 position. The shift from proto-Bantu **j* to these Bantu D30 bilabials remains unclear at this point.

In summary, this section has shown data coming from **d* (though perhaps after **t* or **n*) with late (YZ) devoicing in only one language (for each of Budu and Mbo), for a group of three languages (Budu, Ndaka and Mbo), for no languages, and for most of the D33 languages. This variety is expected by my model of tonogenesis, reflecting the sporadic devoicing of these segments over the history of these languages. This variety will be taken up again in Section 7.8.

7.7 Other correspondences with some loss of voicing specification

The data in this section show other consonants with tone patterns not predicted by a straightforward analysis using Bradshaw (1999)'s [L/Voice]. That is, they lack voicing specification, yet appear in words with Rising tone melodies. These consonants also for the most part correspond with consonants which are specified for voice in related languages (other D30 languages and/or in their Proto-Bantu sources).

I have not presented a broader sampling of palatal correspondences in Chapter 6, and my data showing unexpected tone patterns with segments which are neither bilabial nor alveolar is minimal, so the data in this section must be taken somewhat lightly. Yet along with the data in Sections 7.5–7.6, these data are consistent with Rising tone patterns being the result of historically active consonant-tone interaction (CTI) followed by loss of voicing specification. Given the evidence of voiced intermediaries between proto-Bantu and Proto-Bantu D30 sources for both bilabials and alveolars, it is interesting to find, even in the limited data presented here, Rising tone patterns on verbs coming from both **j* and **c*, voiced and voiceless palatal stops reconstructed for Proto-Bantu.

Before looking at those data, a note on transcription is in order. The phonetic content of proto-Bantu **j* seems to be a matter of some debate, with some evidence of it being a semivowel, rather than an obstruent (Odden 2014). But the BLR3 legend has **j* listed in the consonants, in opposition to **c*, and in parallel with **b/*d/*g*. There is also a note after the consonant chart: “Guthrie's **j* and **y* have been merged into **j*. The problems regarding **j/*y/*zero are far from being resolved” (Bastin, et al 2002). So I take **j* to be a voiced obstruent in opposition to **c*, in line with what is at least implied in most presentations of the proto-Bantu consonant inventory (e.g., Bastin, et al 2002, Hyman 2003a, Odden 2014). Furthermore, the position that proto-Bantu **j* was specified for voicing is supported by the evidence of its Bantu D30 reflexes bearing Rising tone patterns. The transcription of **j*, then, while standard among those who study Bantu, is somewhat at odds with the IPA transcription of consonants, as in my transcriptions of Bantu D30 language data. The result is that the IPA palatal semivowel [j] is observed in each of the Bantu D30 languages with data in Table 64, corresponding with the proto-Bantu original consonant **j*, which was likely a voiced obstruent.

The implication of this fact for this study is that the data in Table 64 represent loss of voicing for each of the Bantu D30 languages with data, as proto-Bantu **j* likely existed in opposition to voiceless **c*, while the the Bantu D30 language sonorant [j] does not exist in opposition to a voiceless alternate. So Table 64 shows loss of voicing specification from a proto-Bantu form (**j*) to forms currently lacking voicing; all reflexes are semivowels, and therefore not specified for [L/Voice], yet so far as can be known all reflexes also bear a Rising tone pattern (shaded).

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[_ ¯ ⁄]	[– ¯ –]			[– ⁄ –]	[¯ – ¯]	1632: **jòŋgʊ́*	
kɑnijɑ	kɑnɨjɑ			kɔjɑnɑ	kɔjɑnɑ	'cooking-pot (n)'	'cook'

Table 64. Rising tone melodies without [L/Voice] specification for all languages with data (*j)

The data in Table 64, therefore, represent a loss of voicing specification after CTI lexicalization for all Bantu D30 languages for which I have data. The **c* reflexes in Table 65 show a different pattern of correspondence. While the proto-Bantu consonant is reconstructed as voiceless (*c), I assume there is a voiced intermediary between proto-Bantu and proto-Bantu D30, such that an ancestor of the languages in Table 65 innovated **c* > **j* for this word, which subsequently devoiced in at least Nyali-Kilo [nlj], Vanuma [vau], Ndaka [ndk] and Mbo [zmw] (but not in Bɨra [brf]).

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[– – –]	[– – –]	[– ¯]		[– ⁄ –]	[¯ – ¯]	395:	
kɑ**k**ɑkɑ	kɑ**k**ɑkɑ	**z**ogɑ		kɔ**kp**ɑgɑ	kɔ**kp**ɑgɑ	*****c****àc*	'ferment'

Table 65. Rising tone melodies without [L/Voice] specification in Ndaka [ndk] and Mbo [zmw] only (*c > **j*)

The data in Table 65 are consistent with an early (W or X, before CTI lexicalization) loss of [L/Voice] specification for 'ferment' in Nyali-kilo and Vanuma, followed by a later (Y or Z, after CTI lexicalization) loss of [L/Voice] specification in Ndaka and Mbo (shaded), since only Ndaka and Mbo forms have Rising tone patterns.

The difference in ordering of loss of voicing specification and CTI lexicalization coincides with a difference in consonantal reflex as well. Ndaka and Mbo, which lexicalized their tone pattern before devoicing in Table 65, indicate a total consonant shift from proto-Bantu of **c* > [kp]. On the other hand, Nyali-Kilo and Vanuma, which lost contrastive voicing before/without CTI lexicalization show a total consonant shift from proto-Bantu of **c* > [k]. So while the extent of intermediaries along those whole consonant shifts remains unclear, the loss of voicing at different times corresponds with different changes in point of articulation.

The data in Table 66 are similarly consistent with my model, assuming a voiced intermediary as discussed above. Under this analysis, the Bantu D30 data in Table

66 underwent **k* > **g* > ∅, perhaps followed by *w* epenthesis, as indicated by the transcriptions for Nyali-Kilo [nlj] and Vanuma [vau]. Each of Nyali-Kilo [nlj], Vanuma [vau], Ndaka [ndk], and Mbo [zmw] would have lexicalized their tone melodies (shaded) before losing the intermediary **g*.

[nlj]	[vau]	[brf]	[buu]	[ndk]	[zmw]	BLR3	gloss
[_ / ¯]	[¯ – –]			[– ⁄ –]	[¯ – –]	2041:	
kawʉnɑ	kawʉnɑ	bhunjɑ		kɔʉnɑ	kɔʉnɑ	**kʊ́n*	'sow'

Table 66. Rising tone melodies with loss of [L/Voice] specification in Nyali-Kilo [nlj], Vanuma [vau], Ndaka [ndk], and Mbo [zmw] (*k > **g*)

In summary, while the data in this section are clearly minimal, they are consistent with the data presented in Sections 7.5–7.6, showing Rising tone patterns which might be taken to be exceptional or unexpected, in a historical context which accounts for their exceptionality, according to the model presented in this chapter.

7.8 Summary of data showing some loss of voicing specification

Sections 7.5–7.7 presented cognate sets showing the non-exceptional sources of modern day exceptional tone patterns. This section will summarize those cognate sets to draw some generalizations from them.

Because there are only two options for the ordering of devoicing, either before or after consonant-tone interaction (CTI) lexicalization, and because that devoicing is taken to be spontaneous, occurring at multiple times throughout the history of these languages, one might expect that there is little to be said about the history of these languages on the basis of the ordering of devoicing and CTI lexicalization. But as I hope to show in this section, apparently random data over time will show some patterns that indicate the history of these languages. This is because both devoicing and CTI lexicalization are innovations, which would be shared by daughters of the innovating language. As a result, while there would certainly be some coincidental agreement between languages which happened to innovate devoicing and CTI lexicalization in the same order, there should also be agreement as a result of shared innovation, leading to more agreement between languages with a more recently shared history (i.e., a closer genetic relationship).

To enable this type of generalization, I have summarized the data in Sections 7.5–7.7 here in a series of tables, which show for each cognate set which languages evidence devoicing before CTI lexicalization, and which evidence devoicing after

CTI lexicalization. I also include in each table a column to indicate languages where devoicing did not occur for that cognate set, which amounts to a lack of the devoicing innovation. Finally, I include a column to indicate where the data I have is unclear, either because I have no cognate for that cognate set in that language, or because I lack some aspect of that data (e.g., the tone pattern). This last is important because I don't want to make arguments from silence haphazardly; reflexes for which I have unclear data may well fit into a generalization formed on the other data, or they may not.

The bilabial correspondence data presented in Section 7.5 are summarized in Table 67, where there are patterns to be observed, despite the appearance of chaos. The presence of some patterning among relative chaos is to be expected, if loss of voicing specification truly is spontaneous, because not only should devoicing have occurred at different times in different words in different languages, but it also should have applied before and after languages diverged, meaning that it would at times be an innovation shared between languages (c.f., Section 2.3.2), while at other times it would have occurred after languages diverged. Unfortunately, the difference would not be observable today; an event that occurred once for a pair of languages (which later split) would look the same as the same event occurring in each language after they split.

	before CTI Lexicalization	**After CTI Lexicalization**	**None**	**unclear data**
Table 52		[nlj]	[ndk]-[zmw], [brf]	[vau], [buu]
Table 53	[nlj]-[vau]-[buu]-[ndk]-[zmw]		[brf]	
Table 54	[nlj]-[vau], [brf], [buu], [ndk]	[zmw]		
Table 55	[buu]	[ndk]	[nlj]-[vau], [zmw]	[brf]
Table 56	[zmw]	[buu], [ndk]		[nlj], [vau], [brf]
Table 57		[nlj]-[vau]-[buu]-[ndk]-[zmw]		[brf]

Table 67. Summary of **b* data from Section 7.5 by loss of [L/Voice] specification

But despite the apparent chaos, languages which we elsewhere have reason to believe are closer to one another than to the other Bantu D30 languages end up in the same position in Table 67. For instance, Nyali-Kilo [nlj] and Vanuma [vau] (which are *prima facie* similar) both lost voicing specification before CTI lexicalization in Table 53 and Table 54, and after CTI lexicalization in Table 57. Neither lost voicing specification at all in Table 55. So everywhere we have clear data in Table 67, Nyali-Kilo and Vanuma agree. Assuming the relative ordering of lost voicing specification and CTI lexicalization represents the relative timing of two innovations, the coincidence of this ordering across languages builds a case for at least some of these innovations being shared between those languages. That is, if Nyali-Kilo and Vanuma shared the above orderings entirely by coincidence, that would be a highly unlikely coincidence. This likelihood is increased when one considers that not only are the above four pairings common to Nyali-Kilo and Vanuma, but there are also no cognate sets in Table 67 for which we have data in both languages, and they differ. As a result, it would be more likely that at least some of these orderings arose from either shared CTI lexicalization, shared loss of voicing specification, or both, at least for the words in question.

A similar case can be made for Ndaka [ndk] and Mbo [zmw], which also have noticeable *prima facie* evidence of close relationship. They both lost voicing specification before CTI lexicalization in Table 53 and after CTI lexicalization in Table 57. Neither lost voicing specification in Table 52. So Ndaka and Mbo have the same behavior in half of the examples in Table 67 for which I have data on both languages. This agreement between Ndaka and Mbo is consistent with the hypothesis that some devoicing happened while Ndaka and Mbo were one language, and some after they split.

One might further compare these four languages, observing that synchronic Nyali-Kilo and Vanuma agree more completely than Ndaka and Mbo, at least for the data in Table 67. We have therefore reason to believe that Nyali-Kilo and Vanuma diverged from each other later, at least in comparison to the period in which they lost voicing specification, whereas Ndaka and Mbo diverged from each other earlier, again at least in comparison to the period in which they lost voicing specification. There is no obvious reason to assume that Nyali-Kilo, Vanuma, Ndaka and Mbo all went through just one period of voicing specification loss, but there do seem to be fewer obviously language specific examples of loss of voicing specification for Nyali-Kilo and Vanuma, indicating that they have spent less time as separate languages. And a more recent unity of Nyali-Kilo and Vanuma matches the available ethnographic information, in that those who speak Nyali-Kilo and Vanuma consider themselves one ethnic group (or at least with a currently acknowledged shared history), while those that speak Ndaka and Mbo consider themselves to be distinct ethnic groups[23].

Less can be said about other languages from the data in Table 67. Budu aligns with no other language in any particular way, though wherever Nyali-Kilo and Vanuma agree with Ndaka and Mbo, Budu is also there. That is, all five languages lost voicing specification before CTI lexicalization in Table 53 (and perhaps Table 54[24]), and they all lost voicing specification after CTI lexicalization in Table 57. The unity of these five languages indicates a third level of voicing specification loss, which occurred before any of the Bantu D33 languages split off. This is in addition to the devoicing which occurred while Nyali-Kilo and Vanuma were one language, and while Ndaka and Mbo were one language. And all the above devoicing occurred before divergence into the five Bantu D33 languages which are vigorous today, where we have seen further devoicing.

Given that Budu does not share any particular pairing with another language, unless it is shared with all the Bantu D30 languages, Budu may have diverged from proto-Bantu D33 before any of the other languages, as indicated in Figure 90. This divergence would have been followed by Ndaka and Mbo (which seem to share a smaller proportion of known innovations, indicating more time as separate languages), which was again followed by Nyali-Kilo and Vanuma (which seem to share a larger proportion of known innovations, indicating more time as one language). The data in Table 67 would thus result in a family tree something like that in Figure 90:

[23] The question of what counts as an ethnic group is of course to some degree subjective, and might be argued to be circular, to whatever extent it depends on linguistic distinctions. But regarding these statements, I am depending solely on what has been said by speakers of these languages about themselves. For the first week working together the Nyali-Kilo and Vanuma groups referred to themselves as Nyali-North and Nyali-South, respectively. They also called each other dialects of one language, despite the fact that a survey of about 2,000 words showed only about 60% lexical similarity.

Regarding the Ndaka and Mbo, on the other hand, I cannot recall any statements implying common heritage, apart from a few references to them as "brothers", a term they would also use for speakers of Budu, Vanuma, and Nyali-Kilo, perhaps among others. This might be due to the distinct contrast between the two groups socioeconomically, or due to the conflict between the DRC army and a militia led by a member of the Mbo. While they seemed happy to work together, I had the continual sense that they were two different groups working together, rather than two sections of one group.

[24] The data in Table 54 could, of course be coincidental, though they could also represent a loss of voicing in Bantu D30, followed by a later return of voicing in Mbo only, all of which preceded CTI lexicalization in these languages.

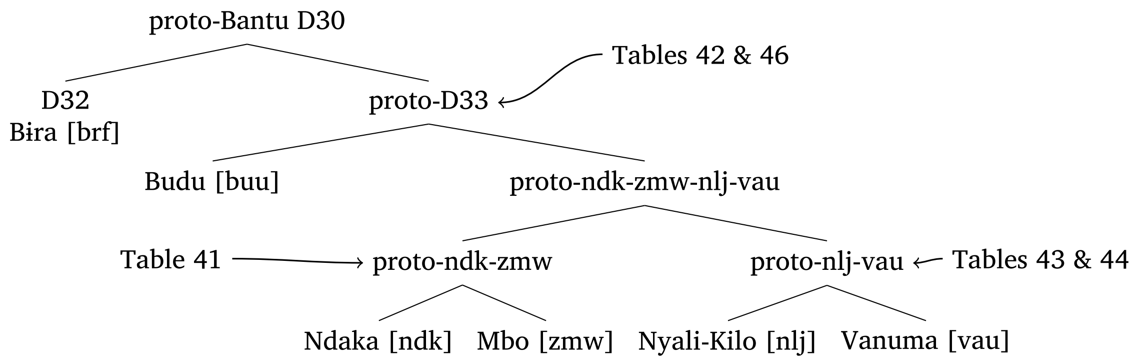


Figure 90. Genetic affiliations of the Bantu D30 Languages, based on **b* correspondences

The alveolar correspondence data presented in Section 7.6 are summarized in Table 68:

	before CTI Lexicalization	**After CTI Lexicalization**	**None**	**unclear data**
Table 58	[nlj]-[vau], [brf], [ndk]	[zmw]		[buu]
Table 59	[nlj]-[vau], [ndk]-[zmw]	[buu]	[brf]	
Table 60	[nlj], [brf]	[buu]-[ndk]-[zmw]		[vau]
Table 61	[nlj]-[vau], [ndk]-[zmw]			[brf], [buu]
Table 62	[vau], [brf]	[nlj], [buu]-[ndk]-[zmw]		

Table 68. Summary of **d* data from Section 7.6 by loss of [L/Voice] specification

The data in Table 68 provide for generalizations similar to those drawn from Table 67. That is, it remains difficult to say with much certainty, but there is again a strong correspondence between Nyali-Kilo [nlj] and Vanuma [vau], as well as between Ndaka [ndk] and Mbo [zmw]. What may indicate a difference in the data summarized in Table 68 is the fact that Budu [buu] agrees with Ndaka and Mbo more often than not where there are data for all three, indicating that it is possible that Budu separated from the ancestor of Ndaka and Mbo, rather than from D33, as in Figure 91[25]:

[25] Budu aligning with Ndaka and Mbo also aligns with data regarding the loss of **d* as described in Section 6.5, which seems to be an innovation shared by Ndaka, Mbo, and Budu.

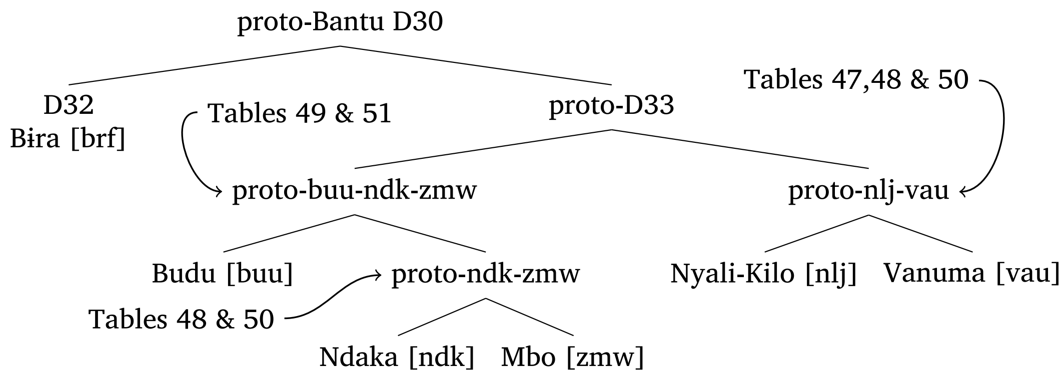


Figure 91. Genetic affiliations of the Bantu D30 Languages, based on **d* correspondences

The other consonant correspondence data presented in Section 7.7 are summarized in Table 69:

	before CTI Lexicalization	**After CTI Lexicalization**	**None**	**unclear data**
Table 64		[nlj]-[vau], [ndk]-[zmw]		[brf],[buu]
Table 65	[nlj]-[vau]	[ndk]-[zmw]	[brf]	[buu]
Table 66		[nlj]-[vau], [ndk]-[zmw]	[brf]	[buu]

Table 69. Summary of **c* data from Section 7.7 by loss of [L/Voice] specification

The correspondences in Table 69 show the same basic trends as those in Table 67 and Table 68 do, that Nyali-Kilo and Vanuma align, as do Ndaka and Mbo. Given the lack of Budu [buu] data, it is difficult to tell exactly what to make of Table 64 and Table 66. They are at least evidence of unity between Nyali-Kilo and Vanuma, and between Ndaka and Mbo, as indicated in Figure 92. But if Budu agreed with the other four languages on these cognate sets, it would be evidence of devoicing by a single proto-Bantu D33, as again indicated in Figure 92.

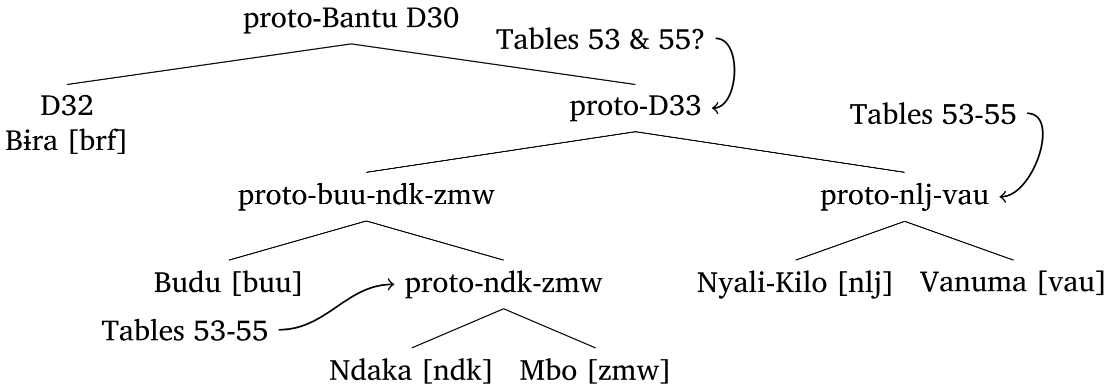


Figure 92. Genetic affiliations of the Bantu D30 Languages, according to other correspondences

The generalizations drawn in this section, along with the genealogical observations based on them, must remain tentative given the limited amount of data they are based on. However, I find it interesting that the data so far presented would indicate what is already suspected from other quarters. So while we could conclude that the data merely show chaos, we don't have to. There are patterns in the loss of voicing specification. And some of those patterns seem to indicate shared innovations between more closely related languages. These patterns exist alongside other correspondence patterns which seem to indicate loss of voicing specification and/or CTI lexicalization which was not shared. All of this together is as would be expected by a pair of processes that operated spontaneously throughout the history of the Bantu D30 languages.

Figure 75 contains a summary of the genealogical evidence presented in this chapter, along with segmental evidence from Figure 75.

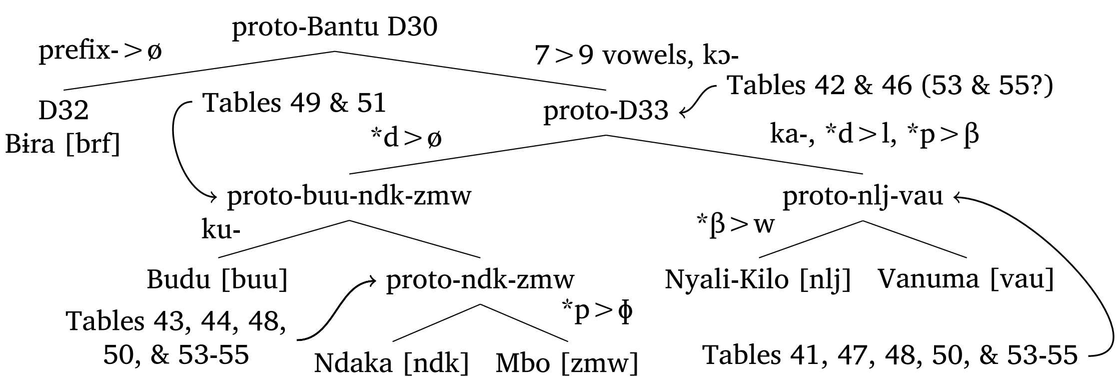


Figure 93. Genetic Affiliations of the Bantu D30 Languages (summary)

7.9 Discussion

The model of tonogenesis presented in this chapter accounts for Bantu D30 data unaccounted for by Bradshaw (1999) alone. Rather than view a distribution that includes words with Rising tone melodies without [L/Voice] specification as exceptional, my model provides a way to see forms unexpected by Bradshaw (1999) as arising through a historical consonant-tone interaction (CTI) rule which was subsequently lexicalized, followed by spontaneous loss of [L/Voice] specification. The result is a principled analysis that includes one prediction made by Bradshaw (1999), i.e., that CTI effects come with consonants specified for [L/Voice], while not requiring that CTI tone melodies only appear with consonants specified for [L/Voice].

I have shown in Sections 7.5–7.7 a large amount of data which would be taken as exceptional with Bradshaw (1999) alone. While these data contain consonants not specified for [L/Voice] with CTI tone melodies, they correspond with consonants which are specified for [L/Voice] in cognate forms in other Bantu D30 languages, and/or in proto-Bantu reconstructed forms.

The remainder of this chapter will be committed to discussing the meaning of variation, practical implications for the analysis of the Bantu D30 languages, and theoretical implications for our understanding of language more generally.

7.9.1 The meaning of variation

The correspondence patterns in Sections 7.5–7.7 show that a large number of words have lost voicing specification, but they vary as to whether their tone patterns would be expected by a synchronic analysis of CTI according to [L/Voice] specification (Bradshaw 1999). This variation represents a distinction between words with consonants that lost voicing specification before CTI was lexicalized, and those with consonants that lost it after CTI was lexicalized.

The analysis presented in this chapter assumes, but does not critically depend on, a single development of a CTI rule in proto-Bantu D30 or an ancestor, which then subsequently became lexicalized. It is less important to this model exactly when this CTI rule lexicalized, either for a given language, or for a given word within a given language. That is, the data presented in Sections 7.5–7.7 speak only to the ordering of CTI lexicalization and loss of voicing specification for each word in each language. It is possible, though by no means necessary, to claim that CTI lexicalization occurred once for all Bantu D30 languages in proto-Bantu D30 (as I assume for the development of the CTI rule). Under such an analysis, the ordering of CTI lexicalization and loss of voicing specification would depend on when loss of voicing specification occurred[26]. I am personally inclined toward the idea that if CTI occurs infrequently in the world's languages, it may be unstable or dispreferred in some way. If this is true, then the phonologization of the correlation between voice and pitch into a CTI rule should be understood to have occurred once per group of languages where it is observed, while CTI lexicalization (or loss of a CTI rule by some other means) may well occur spontaneously (as perhaps does voicing specification loss).

The spontaneous and unpredictable ordering of CTI lexicalization and voicing specification loss is perhaps particularly evident in the data summaries presented in Section 7.8. These summaries in large part show unpredictability from language to language and from word to word, even with words derived from the same proto-Bantu consonants. There remains, however, some correlation between languages which otherwise appear to be more closely related, such as between Nyali-Kilo and Vanuma, and between Ndaka and Mbo. That is, it may be that the number of different correspondence patterns observed in Sections 7.5–7.7 seems to arise from the interplay of two spontaneously occurring processes throughout the history of each word in each language, resulting in more agreement between languages with more shared history, as they would have shared more of those instances of CTI lexicalization and/or loss of voicing specification, before diverging into separate languages. This agreement, however, is contained in a great deal of observed variation in ordering from language to language and from word to word.

One final note on variation: this model includes a reanalysis of a CTI rule as lexical specification, followed by the loss of that CTI rule in a second step, though I have not provided evidence of such a rule loss (i.e., Language Z). My analysis of the Bantu D30 languages is that they remain as Language Y, having lexicalized the CTI rule, but without having such lexicalization replace the rule, i.e., with the rule remaining in effect. I maintain this analysis because there are no exceptions to the rule that words with C1 specified for voice bear CTI tone patterns. This argument is, unfortunately, an argument from silence, however, as it is possible that the CTI rule has already been dropped, but exceptions have not yet arisen for some reason. To show the rule remaining in effect, one would need evidence of recent voicing, which would trigger a CTI tone melody. Such evidence would

[26] Alternatively, it is possible to posit that devoicing occurred once, where it occurred at all, and individual lexemes were lexicalized either before or after that one point in time. This seems less plausible to me, given that not all words show devoicing, whereas it is harder to show that a word's surface tone pattern has not been lexicalized. As a result, a single, categorical lexicalization seems more plausible than a single devoicing operation, as the latter would necessarily only cover certain lexemes. But in no case is either required by my analysis; both CTI lexicalization and loss of voicing specification may have occurred spontaneously across time and lexical space.

include words that became recently voiced gaining a CTI (Rising) melody, such that words with C1 synchronically specified for voice, corresponding with (and reconstructing to) words with C1's which are not specified for voicing (the result of which is comparative evidence of a gain of voicing specification), would have a CTI melody nonetheless. Unfortunately, I do not have any such data in my corpus[27], so I state this caveat, but without further discussion. Despite the tentative nature of this argument, the complete lack of voiced exceptions to the distribution in Table 44 leads me to believe that the CTI rule remains in effect. I would not be surprised, however, to see exceptions arise (or be found), indicating that the rule had been dropped, and that the language in question had transitioned to a system like Language Z in this model of tonogenesis.

7.9.2 Practical implications

The model of tonogenesis presented in this chapter is not just of value to a general understanding of how depressor consonants contribute to tonogenesis. If I have truly modeled the history of the Bantu D30 languages (and what may well be happening in other languages), then this analysis should help understand data in these languages which would otherwise be problematic. There are two points in particular, where this model clarifies outstanding issues in the analysis of the Bantu D30 languages.

The first outstanding issue which is resolved by this model is the presence of a sonorant in Nyali-Kilo which appears both to depress and to not depress tone. Bradshaw (1999) predicts that sonorants should always either depress tone or not in a given language, based on the analysis of their underlying specification of [L/Voice] in that language. There should not, then, be a single sonorant in a single language which depresses tone in some cases, but not in others (assuming it has one [L/Voice] specification wherever it appears, of course). But such a case was found in Nyali-Kilo with [w], which occurs with and without a CTI melody:

(47) **with** a CTI melody and corresponding to the following in other Bantu D30 languages:
  a. the obstruent [b], in Table 52
  b. a fricative ([f] or [ɸ]) or deleted consonant, in Table 57

[27] This is true unless of course you count examples like ‘winnow, flutter’ in Table 55. This correspondence shows voicing specification added before CTI lexicalization (for at least four languages), but also subsequent loss of voicing specification (in Ndaka), indicating that the addition of voicing specification would not be particularly recent.

(48) **without** a CTI melody and corresponding to the following in other Bantu D30 languages:
  a. [w], in Table 36
  b. bilabial obstruents ([β], [p], and [ɸ]), in Table 39 and in Tables 53 and 54

So according to the available data as summarized in (47) and (48), there are four historical sources of Nyali-Kilo [w], only two of which bear CTI melodies (47a,b). While it would be possible to propose an analysis where Nyali-Kilo has two *w* phonemes, one of which is specified for [L/Voice], and the other of which is not (the two being neutralized on the surface, and only distinguishable by their impact on tone patterns), it is simpler and more principled (at least in terms of claims made about underlying geometries) to say that the language has developed new lexical tone patterns on the basis of historically active consonant-tone interaction, which was then followed by a loss of voicing specification in some cases. The result is the merger of forms with and without CTI melodies into a single *w* phoneme.

Setting aside the need for a synchronic CTI rule enables us to see that some cases of Nyali-Kilo [w] come from depressors who lost voicing specification after their word tone melodies were lexicalized, while other cases of Nyali-Kilo [w] come from other sources, e.g., depressors who lost voicing specification before a CTI rule was lexicalized, or else historic **w*. This allows us to see the patterns in the language resulting from historical association of [L/Voice], i.e., that consonants with established history as depressors do this, while others don't. And this is done without needing to maintain a currently active [L/Voice] association rule, which would be untenable given the available Nyali-Kilo data regarding [w].

My model similarly accounts for another outstanding issue, the presence of a voiceless obstruent in Ndaka which apparently acts as a depressor consonant. This is an even more serious issue than Nyali-Kilo [w], because Bradshaw (1999) predicts that voiceless obstruents should never depress tone in any language, as they should never be understood to be specified underlyingly for [L/Voice]. But Ndaka [p] has similar behavior to Nyali-Kilo with [w], as discussed above. Ndaka [p] occurs with and without a CTI melody:

(49) **with** a CTI melody and corresponding to the following in other Bantu D30 languages:
  a. bilabial obstruents ([w], [β], [p], and [ɸ]), in Table 56
  b. bilabial plosives ([b] and [p]), in Table 55
  c. bilabial fricatives ([w] and [ɸ]), in Table 57

(50) **without** a CTI melody and corresponding to the following in other Bantu D30 languages:
  a. bilabial obstruents ([w], [β], [p], and [ɸ]), in Table 39 and in Tables 53 and 54
  b. bilabial plosives and a sonorant ([w], [b], and [p]), in Table 40

So (49) and (50) show potentially five different historical sources for Ndaka [p] —three of which have CTI melodies, and two of which do not. They are each very similar, though, so it would be implausible to trace all of these back to clearly distinct proto-Bantu D33 consonants. In particular, the correspondence with bilabial obstruents [w], [β], [p], and [ɸ] occurs both with and without the CTI melodies, in (49a) and (50a). Based on their segmental correspondences, these should be reconstructed as the same consonant. But to do so would require claiming that a single Ndaka consonant (i.e., [p]) comes from a single proto-Bantu D33 consonant, both with and without CTI melodies. As a result, there is no obvious conditioning factor to motivate a historical innovation which would distinguish the correspondences in (49a) and those in (50a). As a result, apart from an analysis that includes spontaneous devoicing throughout history (such as that presented in this chapter), it would be difficult to use [L/Voice] to account for a single phoneme, with a single historical source, appearing both with and without CTI melodies.

The alternative analysis for Nyali-Kilo [w] above (that the segment exists as two phonemes, one of which is specified for [L/Voice], and the other of which is not) is not plausible for Ndaka [p], as such an analysis would require a /p/ phoneme which is underlyingly specified for voice, but which somehow surfaces voiceless. As implausible as that would be, the difficulty is aggravated by the fact that the system would also require a /p/ phoneme which is underlyingly **not** specified for voice, but which somehow also surfaces voiceless –and each of these would need to be distinct from the phoneme /b/, which is voiced both underlyingly and on the surface.

But the analysis I propose above for Nyali-Kilo [w] works also for Ndaka [p]: a loss of voicing merges forms with CTI melodies and forms without them into forms with a single synchronic [p], which then appears to depress or not, in accordance with its lexical tone pattern. This lexical tone pattern reflects the history of the consonants, rather than a synchronic CTI rule.

This analysis of Nyali-Kilo [w] and Ndaka [p] holds for each consonant despite the fact that [w] has surface voicing and [p] does not, because they both evidence historical loss of [L/Voice] specification, which is the critical distinction for this analysis.

Finally, and perhaps at the risk of hyperbole, a third problem, similar to Nyali-Kilo [w] and Ndaka [p], presented itself during the course of this inquiry: an empty consonant slot with the same behavior. The presence of depressor tone patterns apparently triggered by no consonant at all (c.f. Section 7.6) must be nonsensical when only considering synchronic [L/Voice], since there should be no [L/Voice] where there is no consonant. Assuming transcribed semivowels in these data arise from vacated consonant slots, empty consonant slots in Ndaka and Mbo occur with and without CTI melodies:

(51) **with** a CTI melody and corresponding to the following in Nyali-Kilo and Vanuma:
  a. lateral [l] from proto-Bantu **d*, in Table 58 (Mbo only), Tables 60 and 61
  b. lateral [l] from proto-Bantu **t*, in Table 62
  c. a deleted consonant (or epenthetic [w]) from proto-Bantu **k*, in Table 66

(52) **without** a CTI melody and corresponding to Nyali-Kilo and Vanuma [l]:
  a. from proto-Bantu **d*, in Table 58 (Ndaka only) and Table 59
  b. from proto-Bantu **n*, in Table 63

So this third practical difficulty for the analysis of the Bantu D30 languages shows CTI melodies apparently triggered by an empty consonant slot (or perhaps an epenthetic consonant). This difficulty contains two problems for any attempt to bring these data under a synchronic analysis based on Bradshaw (1999)'s [L/Voice]. First, since the [L/Voice] feature is necessary to trigger CTI melodies, one would have to posit it as floating in an otherwise empty consonant slot. And second, this would have to contrast with another empty consonant slot, which would not result in CTI melodies. The difficulty posed for Ndaka [p] above is aggravated here, in that there is a minimal difference between the data in (51a) and (52a), coming from the same source, and with the same synchronic reflex. But for the data in (51) and (52), there is no segmental content on which to hang any kind of dual phoneme hypothesis, even if one were to try to make something of the kind work for Nyali-Kilo [w] and Ndaka [p]. But the data in (51) and (52) fit nicely into an analysis which shows **d* > ∅ after CTI lexicalization for the data in (51), and **d* > ∅ before CTI lexicalization for the data in (52), as given in Section 7.6.

Apart from an analysis similar to that presented in this chapter, the presence of a consonant apparently engaging in consonant-tone interaction (CTI) in some words of a language, but not in others in the same language, must be problematic for any analysis based on Bradshaw (1999)'s [L/Voice]. Even more problematic is this kind of behavior in voiceless obstruents, which should not engage in CTI at any point, in any language (Bradshaw 1999). And the presence of empty consonant slots behaving in the same manner stretches to absurdity the possibility of a synchronic [L/Voice] analysis for these languages. Yet if one were to abandon Bradshaw (1999) altogether, with it one would lose a number of very powerful generalizations, covering some 80% of the data in each language (c.f., Sections 3.6, 4.5, and 5.6).

Given the practical necessity to describe observed patterns in a language, while at the same time understand apparent exceptions to those patterns in a rational manner, my model provides a coherent way of understanding both the large generalizations caused by historically active CTI in Bantu D30, and the apparent exceptions to those generalizations. And this is done without rejecting Bradshaw (1999), and in a manner that has already been of practical use, specifically in understanding previously unaccounted for Nyali-Kilo [w] and Ndaka [p].

7.9.3 Theoretical implications

This section briefly touches on a number of remaining implications for how we think about language more generally, including tonogenesis as shifting functional load from consonants to tone, the impact of distributions on phonological analyses, and the expectation of regularity in sound change.

While not an original goal of my research, the analysis presented in this chapter also provides a model by which we can understand the creation of new tone melodies (in the Bantu D30 languages, but perhaps also in other languages). In addition to the kind of split *w* forms described in Section 7.9.2, which may depress (and can clearly be seen as deriving from a historical depressor **b*) or not, one may find any number of consonants associated with apparent depressor tone patterns, where no synchronic [L/Voice] specification is indicated on that consonant.

In addition to not requiring a synchronic CTI rule to make sense of the Nyali-Kilo data, this analysis shows that the functional load for the relevant distinctions has shifted from the consonants (e.g., [b] vs [w], etc) to lexical tone melodies. This is the case because words which once had a predictable tone melody on words with distinctive consonants lost the distinction of these consonants, and the tone melodies became unpredictable. In this way, the phonemic distinction lost to the consonants as a result of **b* > **w* was taken up by the tone system. Such a reanalysis maintained the same distinctions on the same words, but it was now another difference in tone melody that signaled the distinction, rather than a difference in consonants (e.g., between [b] and [w]). This shift of functional load from consonants onto the tone system is the mechanism for tonogenesis, the creation of new tone melodies. From this perspective, Nyali-Kilo [w] and Ndaka [p] signal innovation in the lexical tone melodies, rather than a problem for Bradshaw (1999)'s [L/Voice]. The result, then, is that Ndaka and Mbo (and likely also the other Bantu D30 languages) are now systems with three true verb root class by tone, with roots in a depressor consonant system, but no longer strictly attached to it.

Additionally, the analysis presented in this chapter predicts the fact that the majority of verbs will follow consonant and tone melody distributions as expected by Bradshaw (1999). But it also allows for the fact that not all verbs will follow expected distributions. And furthermore, it predicts that those exceptions will fall mostly in one specific direction[28]. Exceptions to Bradshaw (1999) in the data presented here are all in the direction of CTI melodies on words without [L/Voice] specification; no words with [L/Voice] specification are shown without CTI melodies. This kind of strong but not absolute generalization must be difficult for a strictly synchronic phonological analysis, where a rule either applies or it doesn't. If it applies, there should be no exceptions. But if it doesn't apply, why is there such a strong generalization in the data? Evidence of a rule which applied historically, and which subsequently provided a source for lexical associations, provides a rationale for the strong generalizations. The fact that the rule no longer applies provides a means of understanding the exceptions to those generalizations.

In terms of how languages change, the analysis presented in this chapter affirms the central thesis maintained throughout the Handbook of Historical Sociolinguistics (Hernández-Campoy & Conde-Silvestre 2012), that language change is not as categorical as our assumptions may lead us to believe (c.f. Section 2.3.2). Of particular interest is the fact that **d* does not seem to follow the same pattern of devoicing as does **b*, nor **c*. Differential behavior of **b*, **d* and **c* is consistent with language change starting at one part of a language (e.g., place of articulation) and moving to others, even where that change may appear to be systematic. If systematic changes start in one place of a language, or in one language of a family, then we should expect the kind of variation we have seen in this chapter.

[28] The propensity for exceptions in one direction works only for languages of type Z. For Y languages, the exceptions would fall exclusively in one direction, as the CTI rule remains in effect, acting in parallel with CTI lexicalization and devoicing.

Chapter 8
Conclusion

This dissertation has presented a new model of tonogenesis, based on lexicalization of a historically active consonant-tone interaction system. This model arose in response to an otherwise problematic distribution of consonant types across tone patterns for three Bantu D30 languages. In this distribution, more than 80% of the data indicates a consonant-tone interaction system, yet the distribution of consonant types across verb root classes by tone is clearly contrastive. I have further supported this model by providing comparative data showing patterns predicted by this model. These data show a significant amount of variation, but also a number of patterns indicating a closer relationship for some language pairs.

In Chapter 2, I provided a review of literature on three broader topics, as foundation and background for this dissertation. The first is an overview of background research necessary to a principled analysis of tone, including how to understand the critical differences between fundamental frequency (F_0), pitch, and tone. This included the relative nature of tone, as well as how to establish tone contrasts. The relative nature of tone leads to two other topics surveyed: the relationship between tone and register, and the importance of using a relative transcription system to present phonetic information, rather than the *de facto* standard method with diacritics, which is generally suitable only for phonemic transcription.

The second section of the literature review gave an overview of more specific research relevant to consonant-tone interaction (CTI), the more particular focus of this dissertation. I addressed the phonetic correlation between voicing and lower pitch, as well as how Bradshaw (1999) accounts for that correlation with a single, privative feature [L/Voice]. This feature interacts in a number of geometries (i.e., to each of vowels and consonants), and predicts CTI on certain consonants but not others, namely only those with phonologically contrastive voicing.

The third section of the literature review gave an overview of the Bantu D30 languages, the source of the majority of the data in this dissertation, in terms of their genealogy, geography, and segmental phonology. Also included is background on the comparative method and typological classification, two methods of comparing languages. Also included is specific information relevant to the Bantu

D30 data in this dissertation, including verb structure, paradigmatic categories I used as tone frames, and segmental inventory.

In Chapters 3–6 I provided original data from the Bantu D30 languages. These data were collected in community workshops, with the object of helping those communities develop their orthographies.

Chapter 3 walked through the three verb root tone categories for Ndaka: Low, High, and Rising. Each one is identified on the basis of infinitive forms, as well as the four conjugated forms used as tone frames. Each of the verb root classes by tone was compared in a pairwise fashion as well as all three together, to show their distinctiveness. In addition to showing the clear distinctions between these verb root classes by tone, the consonants found in C1 position for each was also provided and summarized. Low and High verbs lack contrastively voiced C1 consonants, while some Rising verbs have C1 consonants. But other Rising verbs do not have C1 consonants, so these tone patterns do not represent a complementary distribution of consonant types across tone patterns, as voiceless C1 consonants exist in each of the three verb root classes by tone.

In addition, I provided an initial autosegmental analysis of rules and underlying forms, which accounts for the available Ndaka data. 2s pronouns and High verb roots are unspecified, while Rising roots bear another low tone. A doubling of the prefix tone indicates past TAM, and future TAM is represented by a high toned suffix, which initially associates to the leftmost empty TBU of the stem.

Chapter 4 showed fundamentally the same data, but for Mbo. There are again three lexical verb root classes by tone, though their infinite and conjugated forms differ. They are again compared pairwise and *en masse*, showing three distinctive classes. The distribution of these verb root classes by tone across consonants is the same as for Ndaka, namely that High and Low lack contrastively voiced C1 consonants, and Rising verbs have both C1 consonants with contrastive voicing and others without it.

The third set of data was presented in Chapter 5, which covered infinitive forms of Nyali-Kilo. There are again the two High and Low verb root classes, which again lack contrastively voiced C1 consonants. But Nyali-Kilo distinguishes itself in having four tone patterns impacted by depressor consonants, which together I call "Rising". They represent both High and Low tone with a contrastively voiced C1 consonant, as well as High and Low with a contrastively voiced C2 consonant. The latter pair of tone patterns only arise in cases without a contrastively voiced C1 consonant; in the case of High verbs with a contrastively voiced C2 consonant, their C1 consonants are all epenthetic.

Chapter 6 brought the previous three chapters together, along with data from other Bantu D30 languages, in sets of cognate sets showing consistent correspondences for the High, Low, and Rising verb root classes. Additionally, a number of bilabial and alveolar correspondence patterns are provided, as a frame of reference for devoicing in Bantu D30, which ultimately impacts the analysis of tonogenesis in Chapter 7.

Chapter 7 presented a new model of tonogenesis as a resolution of the contrastive distributions observed in the Bantu D30 languages, which largely appear to be consonant tone interaction systems. This model involves the historical development of a CTI rule, followed by the lexicalization and eventual dropping of that rule. As a given consonant in a given word is devoiced anywhere along that process, the tone pattern of that word is either lexicalized with the effect of the CTI rule (if devoicing took place after lexicalization) or without it (if devoicing took place before or without lexicalization). This model is exemplified first using four categories of data (with H or L tone, and with voiced or voiceless C1), to show how the system evolves over time, eventually developing a new tone melody which is not predictable on the basis of consonant types. This model predicts the distribution observed in Bantu D30 languages.

This model is further supported by cognate sets like those in Chapter 6, but showing devoicing either with or without Rising tone patterns for a given language. The various patterns show the spontaneity of devoicing across the history of the Bantu D30 languages, as different words in each language devoiced either before or after the tone pattern of that particular word became lexicalized.

The data presented in support of this model also show a number of patterns. These patterns confirm what is otherwise known about these languages, i.e., that Nyali-Kilo and Vanuma are closer to each other than either is to any other Bantu D30 language, and that Ndaka and Mbo are closer to each other than either is to any other Bantu D30 language.

This chapter closes with a discussion on the meaning of the observed variation, as well as the resolution of a number of outstanding issues in the analysis of the Bantu D30 languages. It concludes with a discussion of the theoretical impact of this model, including the value of interpreting consonant-tone interaction as historically but not synchronically active, with the added benefit of understanding how new tone melodies may be formed.

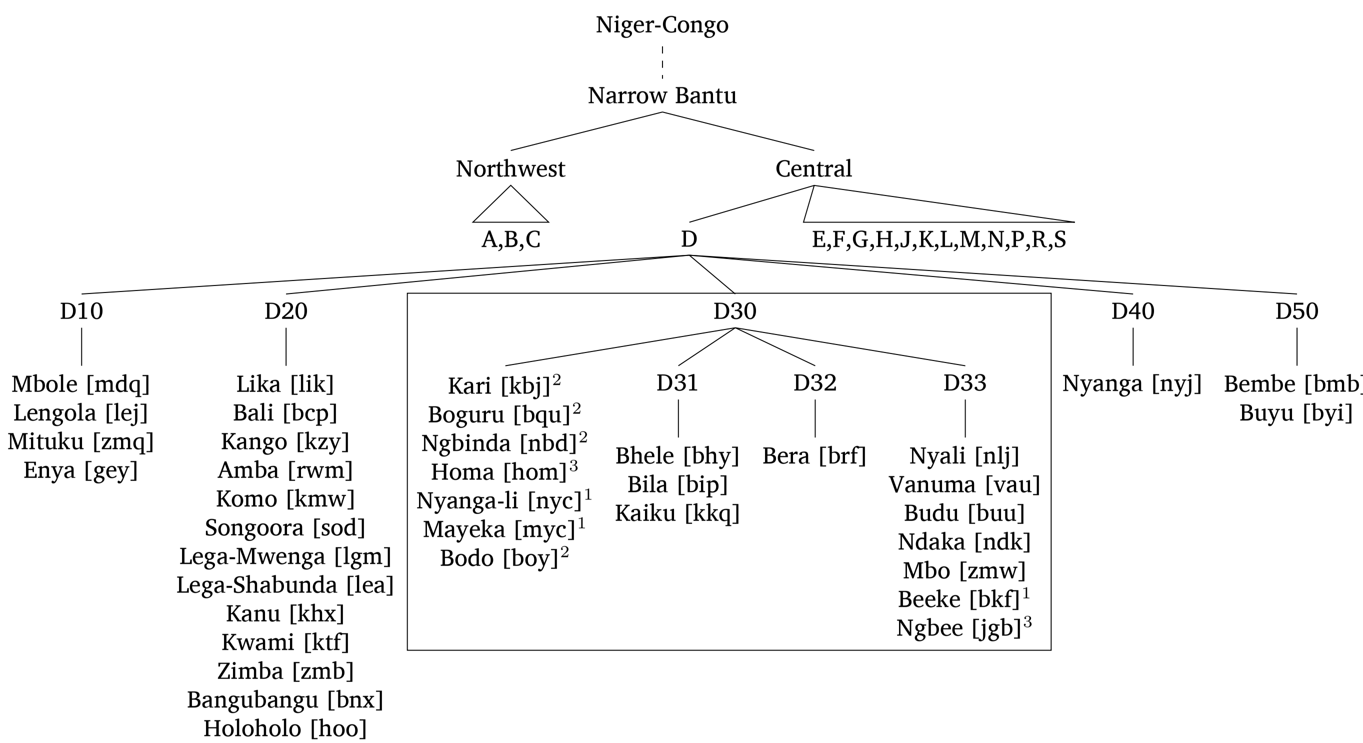
Taken together, this dissertation has established contrast between at least three lexical verb root classes by tone for three Bantu D30 languages, as well as correspondences with data from other languages in that subgroup. Furthermore, those three lexical verb root classes by tone prove problematic, in that they strongly indicate a consonant-tone interaction system, yet the distribution of consonant types across verb root classes by tone is clearly contrastive. This problem is resolved by a new model of tonogenesis, whereby the strong indication of consonant-tone interaction can be seen as indication of a historically active system, even though there is contrast in the modern day languages. Furthermore, this model provides a path to the observed contrastive systems, which includes the historical consonant-tone interaction, but also lexicalization of that rule and subsequent devoicing for some forms. Finally, I have provided data showing the patterns predicted by this model, namely a significant amount of variation, but also a number of patterns indicating a closer relationship for some language pairs.

In this way, the data provided in this dissertation demand an explanation, which is provided in a new model of tonogenesis, and which is further illustrated and defended on the basis of other comparative data. This new understanding allows for a more insightful study of these languages, as well as a better understanding of how languages develop new tone patterns generally.

APPENDICES

APPENDICES

APPENDIX A
The Bantu D30 Languages (family tree)



1. moribund; 2. nearly extinct; 3. extinct; source: Lewis, Simons, & Fennig (2015)

The Bantu D30 Languages (family tree)

APPENDIX B
Processing Pitch Trace Data

Producing the kind of pitch trace overlays I present in this dissertation is typically not a straightforward process, so I state here what I have and haven't done to make the generalizations I see in those pitch traces clear visually. Given the relative nature of tone (c.f. Section 2.1.2), it is important to preserve and highlight the relative contrasts which are consistent across utterances of a particular tone pattern. At the same time, it is not necessary to preserve elements of utterance pitch traces which are not consistent across those utterances; in fact, attempting to do so may ultimately prove distracting to the reader.

Before continuing with the technical aspects of the creation of these images, I should remind the reader of what is written in Section 2.1.1. The fact that pitch is abstracted from F_0, and that tone is abstracted from pitch, means that to get at the categories relevant to language (i.e., tone), one must necessarily transition from the objective to the subjective. The thicker bands in these images are intended to be a subjective expression of the general trend in that set of pitch traces, rather than an objectively defensible trend line. This is not from laziness, or an inability to do statistics, but rather because the categories we are looking to understand are inherently subjective. It is my hope that these lines allow one human to communicate to another, and that most of us would agree that the lines are correct in their essences. But I also hope that if the reader's subjective impression of the trend of a set of pitch traces differs in any substantial way from mine, then these images will provide a starting point for discussing those differences —and that such a discussion would be mutually beneficial for our understanding of tone.

The more obvious variation in pitch height has been discussed in Section 2.1.2, but in addition to individual (and especially gender) variation causing pitch changes in the vertical dimension, a number of factors impact the rate of speech, and therefore the duration of an utterance, in the horizontal dimension. This variability can be observed even across tokens of the same words spoken by the same speaker.

Because of the relative need to preserve the above features, I have made a number of modifications to the pitch traces before overlaying them. In order to show the pitch trace generalizations more clearly, the individual pitch traces have been shifted and expanded or contracted in the horizontal dimension (time), in order to align the word breaks across tokens. This assures that the differences between an utterance spoken faster or slower do not cause visual chaos, but that the overall pattern in each can be seen. Similarly, I have shifted (but not expanded or contracted) some of the pitch traces vertically, to account for the differences in voices between speakers. In this sense the frequency axis on most of these figures is not entirely correct, in terms of the frequencies for each line within the figure. But the overall scale of each figure is correct, and the numbers do give the reader an idea of the frequencies involved in these utterances.

Finally, there are a number of jumps in F_0, which do not mean much in terms of the pitch contrasts in a given utterance, but which are very visually distracting when 50 pitch traces are laid one upon another. Some of these amount to problems with Praat's algorithm for calculating F_0, as can be seen by jumps to twice or half the previous F_0 value. At other times, disturbances in the recording itself results in F_0 calculations that do not accurately reflect the pitch contour of an utterance. Because these recordings took place in a public space, during a public workshop, recordings can include wind, coughing, children playing in the background, chickens, and other distracting noises. As a result, I have removed these glitches using layer masks in GIMP (www.gimp.org), applied to each utterance individually (more details below). I have generally left sharp F_0 changes due to (voiceless) consonants alone, though occasionally (when they seem particularly distracting to the overall trend) they have been removed.

It is important to the integrity of the data to not make changes to the F_0 traces extracted from Praat, other than what I mentioned above.

With this basic overview of modifications made, the following provides a more detailed account of the process I used to create the pitch trace overlay images in this dissertation.

The first step on developing these images is annotating the recordings in Praat (Boersma & Weenink 2014). In order to align the tone frames, I used a point tier to mark the beginnings of word breaks before and after each verb.

With the utterances annotated, I used the pitchtraces.praat script (currently available at http://omega.uta.edu/~rasmussenk/praat/) to draw images based on those annotations. While the script is for general use and rather configurable, for pitch trace overlays I included only the pitch trace (i.e., no spectrograms or textgrids), along with vertical lines where the point tier has marked word breaks.

The individual images in this script are then imported into GIMP for organization and further manipulation. For my convenience, images were organized as layers in a layer group hierarchy (according to verb root class by tone, TAM group, and subject pronoun). The garnishing of one image was cut out to make a mask for all the images, and the garnishing (axes, labels, etc) on all the images was cut off by cropping the entire hierarchy, to allow for the horizontal manipulations without making multiple garnishings visible.

The horizontal shifting and scaling was done with the exact-x-aligner.scm GIMP script (currently available at http://omega.uta.edu/~rasmussenk/gimp/), which I derived from the Exact Aligner (©2009 Dr. Volker Tries ; volker.tries@kfopraxis-oberursel.de). The original script uses a path with four points, to align the first two points on one layer with the second two points on the second layer. My variation removes the y dimension modifications, so only the x coordinates of the four path points are used. To use the script to align pitch traces, I placed the first two points of a path on some standard x coordinates, then the second two points on the vertical lines in the image created in Praat (defining the beginning of word breaks). At this point all that remains is to select the layer to modify and run the script. This process continues for each layer, leaving the first two path points in place, and moving the second two to align to the wordbreaks on the next layer, then running the script. At some point in this process, the white of these layers is made transparent, so the pitch traces of multiple layers can be seen at the same time. With the wordbreak lines from each layer aligned in this way, they are directly on top of each other and appear as one line, as in the images in this document.

Images were also annotated in GIMP with textual notifications for subject, TAM, and verb root class by tone. I also included a layer with a thicker line indicating the general trend of each set of pitch traces overlaid in an image. These trendlines also allowed the creation of images comparing trendlines between verb classes, as seen in Section 3.2.1 and elsewhere. Finally, additional layers added shading and annotations to highlight the difference between the various trendlines.

APPENDIX C
Orthographies of Bantu D30 and the functional load of tone

Do we need to write tone? —Anonymous

One of the questions that confronts anyone working on a writing system for a tone language is whether tone needs to be written. Social and political arguments on either side can and are made, some perhaps better than others. But regarding linguistic factors, it is my strong impression that most people charged with orthography development in a tone language feel at a complete loss as to how to evaluate linguistic factors that impact the need for tone marking in the writing system. That is, given that languages use forms (e.g., consonants, vowels, and pitch) to encode information differently, and given that writing systems typically cannot include all meaningful differences expressed in the associated spoken language, it makes sense to evaluate the relative importance of the various forms, to better ensure that the writing system includes the most important forms, whichever they be.

One example of the evaluation of which elements of spoken language should be written is the decision to not mark stress in the English writing system (as is done in Spanish, for instance), despite the fact that there are words which as a result are spelled the same, while having different pronunciations and different meanings (e.g., the nouns *desert* [ˈdɛsəɹt] 'barren land (n)' and *desert (n)* [dəˈsəɹt] 'thing deserved (n)'). The decision to not write stress in English is justified, however, given the small number of pairs of words that could be confused, along with the fact that most of them would never occur in the same syntactic environment (e.g., the verb *desert* [dəˈsəɹt] 'abandon' is much more common than the noun 'thing deserved (n)' cited earlier, but verbs and nouns occur in the same places in sentences).

So a way to more or less objectively evaluate how much a language uses tone would be helpful to advise a writing system committee in terms of linguistic factors that should be considered in designing a writing system. The more important tone is in a particular language, the less likely leaving it out of the writing system makes sense. On the other hand, the less important tone is in a language, the more sense it makes to leave it out of the writing system, if that is desired by the community.

The Bantu D30 languages have similar consonant and vowel inventories, and similar numbers of lexically contrastive tones. The question remains, then, how to evaluate the functional load of tone in the actual usage of these consonants, vowels, and tones in these languages? Despite their similarities in terms of inventory, we find a variety of usage between these languages, in terms of which language features distinguish which conjugated verb forms.

(53) Free pronouns in four Bantu D30 languages (with Kiswahili [swc] for reference)

[ndk]		[zmw]		[buu]		[brf]		[swc]	
ɨmɨ	[— —]	ɨmɨ	[— —]	ɨmɨ	[— —]	ɨmɨ	[— —]	mimi	‘ 1s ’
ʉwɛ	[— —]	ʉwɛ	[— —]	ʉwɛ	[— —]	ʉwɛ	[— —]	wewe	‘ 2s ’
ɨjɛ	[_ ¯]	ɨjɛ	[_ ¯]	ɛjɔ	[_ ¯]	jɛɨ	[¯ _]	yeye	‘ 3s ’
isu	[_ ¯]	isu	[_ ¯]	isu	[_ ¯]	ɓesu	[¯ ¯]	sisi	‘ 1p ’
inu	[_ ¯]	inu	[_ ¯]	inu	[_ ¯]	ɓenu	[¯ ¯]	ninyi	‘ 2p ’
ɨɓɔ	[_ ¯]	ɨɓɔ	[_ ¯]	ɨɓɔ	[_ ¯]	ɓoli	[¯ _]	wɑo	‘ 3p ’

The free pronouns provided in (53) show a great deal of similarity between Ndaka [ndk], Mbo [zmw], and Budu [buu], with less for Bɨra [brf], and even less for Congo Swahili [swc]. Yet the relatedness of the language can be seen across the board, with common consonants and/or vowels in each row.

Tonally, Ndaka, Mbo, and Budu have the same system, which opposes 1s and 2s against all other forms, in a simple binary split between simple speech act participants versus others. Bɨra, on the other hand, includes 1p and 2p, creating a binary split between those that include a speech act participant and those that do not (i.e., 3s and 3p). Congo Swahili differs even further, in not using tone at all.

Beginning with Congo Swahili as a baseline of comparison, we see the following conjugated past and future forms:

(54) Past conjugation of *kukata* 'cut' by person in Congo Swahili [swc]

nilikata	'I (did, have) cut.'
ulikata	'You (did, have) cut.'
alikata	'He (did, has) cut.'
tulikata	'We (did, have) cut.'
mlikata	'You all (did, have) cut.'
walikata	'They (did, have) cut.'

(55) Future conjugation of *kukata* 'cut' by person in Congo Swahili [swc]

nitakata	'I will cut.'
utakata	'You will cut.'
atakata	'He will cut.'
tutakata	'We will cut.'
mtakata	'You all will cut.'
watakata	'They will cut.'

In non-tonal Congo Swahili, person and TAM are exhaustively and exclusively marked with consonants and vowels. Because there is no tone at all in the system, the entire weight of the distinctions in (54) and (55) is carried by the consonants and vowels.

The following shows corresponding conjugated forms in Bɨra:

(56)
past conjugation of *kʉtina* [— ¯ —] 'cut' by person in Bɨra [brf]
nɨtinɨ [— ¯ ¯] 'I (pst) cut.'
ɔtinɨ [— ¯ ¯] 'You (pst) cut.'
ɑtinɨ [— ¯ ¯] 'He (pst) cut.'
kɛtinɨ [— ¯ ¯] 'We (pst) cut.'
ɓʉtinɨ [¯ ¯ ¯] 'You all (pst) cut.'
ɓɑtinɨ [¯ ¯ ¯] 'They (pst) cut.'

(57)
future conjugation of *kʉtina* [— ¯ —] 'cut' by person in Bɨra [brf]
nɨtinɨ [— ¯ —] 'I will cut.'
ɔtinɨ [— ¯ —] 'You will cut.'
ɑtinɨ [— ¯ —] 'He will cut.'
kɛtinɨ [— ¯ —] 'We will cut.'
ɓʉtinɨ [¯ ¯ —] 'You all will cut.'
ɓɑtinɨ [¯ ¯ —] 'They will cut.'

Bɨra shows full segmental specification of bound pronouns, the majority of which are marked by two segments (CV). Nothing depends on tone by person or number; past and future TAM is distinguished only by tone, as is the case for all the Bantu D30 languages.

The following is a corresponding set of conjugated forms in Budu:

(58)
past conjugation of *kutoko* [— ¯ —] ‘spit’ by person in Budu [buu]
mɑtoko [— ¯ _] ‘I (pst) spit.’
wɑtoko [— ¯ _] ‘You (pst) spit.’
ɑtoko [¯ ¯ —] ‘He (pst) spit.’
kɑtoko [¯ ¯ —] ‘We (pst) spit.’
nɑtoko [¯ ¯ —] ‘you all (pst) spit.’
ɓɑtoko [¯ ¯ —] ‘They (pst) spit.’

(59)
future conjugation of *kutoko* [— ¯ —] ‘spit’ by person in Budu [buu]
mɑtoko [— ¯ ¯] ‘I will spit.’
wɑtoko [— ¯ ¯] ‘You will spit.’
ɑtoko [¯ ¯ ¯] ‘He will spit.’
kɑtoko [¯ ¯ ¯] ‘We will spit.’
nɑtoko [¯ ¯ ¯] ‘You all will spit.’
ɓɑtoko [¯ ¯ ¯] ‘They will spit.’

For Budu, bound pronouns are all unique, as was the case for Bɨra. However, they maximally consist of one (consonant) segment, and one of them is an empty default (3s). Because of this, the Budu system depends less on consonants (and much less on vowels), to distinguish person and number, than does Bɨra. One might argue that in each case the tone is not required to distinguish the person and number of the subject, and this is true. However, the amount of redundancy in Budu is much less than that in Bɨra. Because of this, if each were to lose some of that segmental information, Bɨra would more likely be able to maintain those contrasts solely on the basis of its segments, as it has more segments currently carrying person/number information today. If Budu were to lose some of these bound subject pronoun segments, it would run into danger of not being able to maintain those contrasts, as it has fewer person/number segments to lose.

The following show the corresponding conjugations in Mbo:

(60)
past conjugation of *kotoko* [– ¯ ¯] ‘spit’ by person in Mbo [zmw]
motoko [– ¯ ¯ (–)] ‘I (pst) spit.’
otoko [– ¯ ¯ (–)] ‘You (pst) spit.’
otoko [¯ ¯ ¯ (–)] ‘He (pst) spit.’
totoko [¯ ¯ ¯ (–)] ‘We (pst) spit.’
notoko [¯ ¯ ¯ (–)] ‘You all (pst) spit.’
ɓotoko [¯ ¯ ¯ (–)] ‘They (pst) spit.’

(61)
future conjugation of *kotoko* [– ¯ ¯] ‘spit’ by person in Mbo [zmw]
motoko [– ¯ ¯ (¯)] ‘I will spit.’
otoko [– ¯ ¯ (¯)] ‘You will spit.’
otoko [¯ ¯ ¯ (¯)] ‘He will spit.’
totoko [¯ ¯ ¯ (¯)] ‘We will spit.’
notoko [¯ ¯ ¯ (¯)] ‘You all will spit.’
ɓotoko [¯ ¯ ¯ (¯)] ‘They will spit.’

For Mbo, bound subject pronouns are no longer segmentally unique. Pronouns for 2s and 3s are segmentally identical (both ∅), for each of past and future forms. Hence, there exist systematic minimal tone quadruplets (two persons x 2 TAM categories: “He/you will/did <verb>”). Note that this is not very much different from the Budu data presented in (58) and (59); the loss of an initial consonant *w* on the 2s form makes the 3s default no longer unique. As a result, Mbo cannot use segmental material to distinguish 2s and 3s forms; this can only be done with tone.

The following show the corresponding conjugations in Ndaka:

(62)
past conjugation of *kotoko* [ˍ ˉ ⟍] 'spit' by person in Ndaka [ndk]
notoko [— — — (ˉ)] 'I (pst) spit.'
otoko [— — — (ˉ)] 'You (pst) spit.'
otoko [ˉ ˉ ˍ (ˉ)] 'He (pst) spit.'
kotoko [ˉ ˉ ˍ (ˉ)] 'We (pst) spit.'
notoko [ˉ ˉ ˍ (ˉ)] 'You all (pst) spit.'
ɓotoko [ˉ ˉ ˍ (ˉ)] 'They (pst) spit.'

(63)
future conjugation of *kotoko* [ˍ ˉ ⟍] 'spit' by person in Ndaka [ndk]
notoko [— ˉ ˉ (ˉ)] 'I will spit.'
otoko [— ˉ ˉ (ˉ)] 'You will spit.'
otoko [ˉ ˉ ˉ (ˉ)] 'He will spit.'
kotoko [ˉ ˉ ˉ (ˉ)] 'We will spit.'
notoko [ˉ ˉ ˉ (ˉ)] 'You all will spit.'
ɓotoko [ˉ ˉ ˉ (ˉ)] 'They will spit.'

Ndaka has a system which is very similar to that of Mbo, as Mbo's system is similar to Budu's. The difference here is that 1s *m-* in Mbo is *n-* in Ndaka, making Ndaka's 1s form segmentally identical to its 2p form[1]. The result is another set of systematic minimal tone quadrulplets (i.e., "I/you.all will/did <verb>"), though the 1s/2p combinations begin with *n-*, whereas the 2s/3s combinations lack an initial consonant.

[1] This is apparently also true of Budu Nita, which is geographically separated from Ndaka by Budu Akoya, the dialect of reference here.

The above differences can be summarized as in (64):

(64) Bound subject pronouns in Bantu D30 languages

[ndk]	[zmw]	[buu]	[brf]	
n- [L]	m- [L]	m- [L]	nɨ- [L]	'I'
∅- [L]	∅- [L]	w- [L]	ʉ- [L]	'2s'
∅- [H]	∅- [H]	∅- [H]	ɑ- [L]	'3s'
k- [H]	t- [H]	k- [H]	kɨ- [L]	'we'
n- [H]	n- [H]	n- [H]	ɓʉ- [H]	'you.pl'
ɓ- [H]	ɓ- [H]	ɓ- [H]	ɓɑ- [H]	'they'

To summarize the above, minimal changes from one language to another imply a shift of the balance of information carried on segments, as opposed to that carried by tone. That is, as even one consonant is lost in a language, that language must at times either lose the relevant distinction, or else maintain it through the use of tone. This kind of comparison should be helpful in not understanding only whether a given language uses tone, but how, and to what extent it depends on tone to maintain important grammatical and lexical distinctions.

In addition to the above systematic differences between these languages' use of person/number in bound pronouns and TAM, Ndaka has a number of lexical minimal pairs between verb roots:

(65) Ndaka [ndk] High/Rising tonal minimal pair between verb roots
  a. kɔjɑnɑ [_ / _] prepare food
  b. kɔjɑnɑ [_ ¯ _] be tired

(66) Ndaka [ndk] Low/High tonal minimal pair between verb roots
  a. komijo [— — —] swallow
  b. komijo [_ ¯ _] squeeze press

Hence, there exist in Ndaka systematic minimal tone octuplets (two persons x 2 TAM categories x two roots: "He/you will/did cook/be tired"). While one might expect the same situation in Mbo, it doesn't exist. Some verb root minimal pairs in Ndaka seem to not have cognate forms (n.c.) in Mbo, as in (67):

(67) Mbo [zmw] High/Rising tonal minimal pair between verb roots (not there)
  a. kɔjɑnɑ [¯ _ —] prepare food
  b. n.c. be tired

Other Ndaka verb root minimal pairs have cognate forms in Mbo, but their forms are different enough that they are not minimally different in Mbo, as in (68):

(68) Mbo [zmw] segmentally distinct cognates to Low/High tonal minimal pair between verb roots in (66)
   a. kɔmɨjɑ [¯ — —] swallow
   b. komiɲo [— — —] squeeze press

The result of this is that, just as Ndaka relies on tone more than Mbo does to express person and number of subject, it also relies on tone more than Mbo to distinguish one root from another.

feature\language	Swahili	Bɨra	Budu	Mbo	Ndaka
past/future	CV	T	T	T	T
2s/3s	V	V	C/∅	T	T
1s/2p	CV/C	CV	C	C	T
roots	CV	CVT	CVT	CVT	CVT/T

Table 70. Summary of the functional load of tone by feature and language (T = distinction carried by tone; CV = distinction carried by Cs and Vs, etc)

What is interesting about the comparison in Table 70, in terms of orthography development for the Bantu D30 languages, is that they all use tone to distinguish at least two TAM categories (unlike Kiswahili). But apart from distinguishing TAM, they have different functional loads on tone, as opposed to the amount of functional load placed on consonants and vowels. Budu has had a minimal diacritical marking of tone in their orthography[2] for over 20 years now, and it seems to work for them. The apparent success of a diacritical marking system in Budu makes sense given that their person marking is completely distinctive with just segments, even if Budu doesn't use segments as heavily as Bɨra does.

Based on the initial analysis in this chapter, Bɨra *may* be able to use a similar diacritical TAM marking system, *if* they want to write tone minimally. On the other hand, Mbo, and most certainly Ndaka, rely much more heavily on tone, making

[2] The Budu orthography manual indicates an equal sign (" = ") after the first vowel of a verb for future, and a colon (":") for past (Projet Budu & Bamata-Subama 1997). This system is diacritical in that these symbols do not indicate surface tone, nor even underlying tonal units, but rather a particular TAM function, which may manifest in a number of different surface tone patterns, depending on whatever other morphology is on the verb.

it much more likely that an orthography that attempts to zero or minimally mark tone will ultimately fail.

One further point should be made regarding the development of writing systems in the Bantu D30 languages. I have used the term consonant-tone interaction (CTI) perhaps somewhat loosely in this dissertation, but I should be clear that I have observed no application of a rule by which tone patterns are predictable on the basis of consonants in these languages as spoken today. If such were found, it could simplify the need for tone marking, as one melody (or set of melodies) would be predictable, and therefore not need to be written. But on the basis of available data in these languages, while there is evidence of **historical** consonant-tone interaction, there is also evidence of **lexicalization** of that interaction, such that a number of exceptions have arisen in the Bantu D30 languages. I have documented such exceptions for Ndaka, Mbo, and Nyali-Kilo; if the claim is made that Vanuma has none, I would encourage the community to look carefully for them. If the thesis of this dissertation is correct, such exceptions should be there already, or else be ready to arrive at any time.

What the presence of exceptions to a CTI rule means for a community developing a writing system is that they will most likely need to either not write tone at all, or else come up with a tone marking system which is more complex than they would need if their language did not have those additional tone melodies. But for Ndaka and Mbo, at least, I think I have shown that tone is so decidedly important in these languages, that the community will need to write it in a manner than accounts for all existing words, and I trust that the community will be able to see that. As far as specific ideas about how to write tone, others have written on this at length, e.g., Bird (1999a) and (1999b); Cahill (2001), Roberts & Walter (2012), Roberts, Borgwaldt & Joyce (2013), Kutsch Lojenga (2014), and Snider (2018). Other citations are available at https://www.sil.org/orthography/tone.

APPENDIX D
(V)C roots and the permanence of ATR

This dissertation only deals with canonical CVC roots generally, though it is worth noting that there are also a number of (V)C[1] roots in these languages, as well. These data do not directly impact the question of tonogenesis in Bantu D30, yet is very interesting for what it says about ATR in these languages. The presence of [ATR] as a feature making up the entirety of a grammatical morpheme is discussed in Roberts (1994), but I am not aware of another description of [ATR] making up part of a grammatical morpheme without moraic value, i.e., with consonants, but not syllabic nasals or vowels.

In Ndaka [ndk] and Mbo [zmw], verb roots lacking a full CVC syllable structure have their missing segments made up with an ATR harmonizing *kɛn(ɨ)-* prefix in the infinitive (the prefix final [ɨ] is present on roots missing both a consonant and a vowel):

(69) Ndaka [ndk] *s* 'go'
- a. kɛnɨsɑ [− − −] go
- b. ɔsɑ [¯ ¯] he went
- c. sɑ! go!

(70) Mbo [zmw] *s* 'go'
- a. kɛnɨsɑ [– ¯ ¯] go
- b. ɔsɑ [¯ ¯] he went
- c. sɑ! go!

On some occasions the loss of segmental material differs between languages. Example (71) also shows the +ATR harmonized *ken(i)-* on 'dip.in.water', a +ATR root:

[1] As will be noted in the following data, none of these consonants are contrastively voiced (depressor consonants).

(V)C roots and the permanence of ATR

(71) a. kenino [_ ¯ _] dip.in.water *[ndk]*
b. ino! dip.in.water! *[ndk]*
c. kenino [_ ¯ ¯] dip.in.water *[zmw]*
d. no! dip.in.water! *[zmw]*
e. kɑnuno [_ ¯ ¯] dip.in.water *[nlj]*
f. uno! dip.in.water! *[nlj]*

Note that Nyali-Kilo [nlj] does not use the *kɛn(ɨ)-* prefix, but rather *kan-*. Compare (71ef) with (72):

(72) Canonical *ka-CVC-a/o* in Nyali-Kilo [nlj]
a. kɑnuto [_ ¯ ¯] pour away
b. nuto! pour away!

It is also noted that the only examples of VC roots I have where V is not [ɨ] or [i] do not use this prefix:

(73) Normal *kɔ-* prefix on V(C) roots where V ≠{ɨ,i}
a. kɔʉnɑ [– ¯ –] sow ɔʉnɑ [⌄ ⟋ ⌄] He sowed (seed) *[ndk]*
b. kɔʉnɑ [¯ – –] sow ɔʉnɑ [— ¯ —] He sowed (seed) *[zmw]*
c. kouo [– ⟋ –] blow (air) ouo [⌄ ⟋ ⌄] He blew (air) *[ndk]*
d. kouɸo [¯ – –] blow (air) ouɸo [— ¯ —] He blew (air) *[zmw]*

These roots are also interesting in that they are the only VC roots in the Rising verb root class. All other (V)C roots are analyzed as Low, according to current data. This amounts to a lack of lexical tone contrast for (V)C roots where V is [ɨ] or [i], at least for Ndaka.

The non-canonical data presented in this appendix are interesting in that they provide evidence of ATR stability (or persistence). The ATR harmonization on minimal single consonant (C) roots can be seen in the Mbo data in (74):

(74) C roots in Mbo [zmw]
a. kɛnɨtɑ [– ¯ ¯] ask ɔtɑ [¯ ⟋] he asked
b. kenito [– ¯ ¯] pour out oto [¯ –] he poured (it) out
c. kɛnɨkɑ [– ¯ ¯] come ɔkɑ [¯ –] he came
d. keniko [– ¯ ¯] hear oko [¯ –] he heard
e. kɛnɨnɑ [– ¯ ¯] see ɔnɑ [¯ –] he saw
f. kenino [– ¯ ¯] dip.in.water ono [¯ –] he dipped (it in water)

These verbs correspond with Ndaka verbs that show an extra vowel, for the most part, which appears to contain ATR information[2].

(75) (V)C roots in Nyali-Kilo [ndk]

a.	kenito	[– ¯ –]	pour out	oto	[¯ ¯]	he poured (it) out
b.	kɛnɨtɑ	[– ¯ –]	ask	ɛɨtɑ	[¯ ¯ –]	he asked
c.	kɛnɨkɑ		come	ɛɨkɑ	[¯ ¯ –]	he came
d.	keniko	[– ¯ –]	hear	eiko	[¯ ¯ –]	he heard
e.	kɛnɨnɑ	[– ¯ –]	see	ɛɨnɑ	[¯ ¯ –]	he saw
f.	kɛnino	[– ¯ –]	dip.in.water	eino	[¯ ¯ –]	he dipped (it in water)

The vowel slot is lost in Ndaka disproportionately before consonants toward the front of the mouth. That is, it remains on velars ([k]), palatal stops ([ndʒ]), alveolar nasals ([n]), and some alveolar oral stops ([t]). There is no root initial vowel before other consonants: labials, prenasalized and sibilant alveolars ([nd] and [s]), and palatal semivowel ([j])[3].

So not only do the Mbo data in (74) show an ATR contrast for roots which otherwise seem to contain only a single consonant, those ATR values correspond with what is visible in a root vowel slot in Ndaka. The conclusion I draw then is that while the vowel slot has been lost in the Mbo roots in (74), those roots maintain an ATR feature nonetheless, which is somehow attached to either the consonant or some other part of the root feature geometry, or else remains somehow floating until the rest of the word is built.

[2] An alternative analysis could posit an independent ATR feature in the root. It is somewhat suspicious that VC roots in Ndaka are (?almost) all iC or ɨC, calling into question the amount of information contained in that vowel slot.

[3] As the only palatal consonant which is not preceded by *i\ɨ* in these Ndaka roots, an alternative analysis would posit the *j* root as an underlying V slot. This alternative analysis would, however, require that the segment get syllabified before the prefix is applied, as the surface infinitive form comes out *kɛnɨ-j-a*.

APPENDIX E
Integrating [L/Voice] (Bradshaw 1999) and RTT (Snider 1999)

Given that my analysis assumes both Register Tier Theory (RTT) from Snider (1999) and [L/Voice] from Bradshaw (1999), but neither account assumes the other, I address here how their potentially competing claims may be integrated.

RTT describes four tones on the basis of two binary features, one on a tonal tier, with values of high (H) and low (L), and the other on a register tier, with values of higher (𝒽) and lower (ℓ). This is modeled in Figure 94, repeated from Figure 6 on page 21:

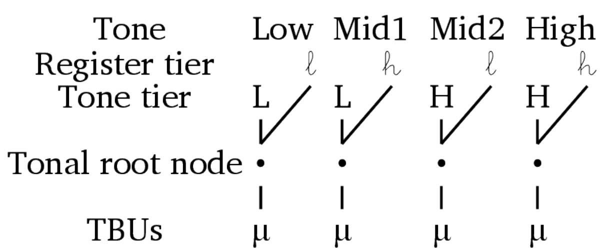


Figure 94. Review of tonal feature geometry in RTT (Snider 1999)

The tone and register tiers of RTT are each associated to a node on the tonal root tier, which is then in turn associated to TBUs. What is typically thought of as a low tone has a low (L) feature on the tone tier associated to a tonal root node that is also associated to a lower (ℓ) register feature on the register tier, as in the leftmost column of Figure 94. Alternatively, the same low (L) feature may be associated to a tonal root node which is also associated to a higher (𝒽) register feature, producing one kind of mid tone, as in the second column of Figure 94. A high (H) tone feature may be associated to a tonal root node which is also associated to a lower (ℓ) register feature, producing another kind of mid tone, as in the third column of Figure 94. Finally, what is typically thought of as a high tone has a high (H) feature on the tone tier, and a higher (𝒽) register feature on the register tier, as in the rightmost column of Figure 94.

The RTT geometry allows us to understand tone and register not only as distinct features, each with different associations to the TBU, but also as different kinds of things —namely the relative domain for tones (register) and the tones which are expressed on that domain.

Bradshaw (1999) makes other claims about the relationship between features. The theoretical explanation for why tone and contrastive voicing are so consistently correlated is given in Bradshaw (1999) as a privative (single valued, or non-binary) feature which contains both low tone and voice, which she calls [L/Voice]. That is, either the feature is present, or it is not; in any case, one gets either both contrastive voicing and low tone, or neither, but never only one or the other. Surface voicing for consonants not specified for voicing (e.g., most voiced sonorants) would be dealt with in the phonetic implementation without the use of distinctive features, since there is no phonemic contrast.

Looking again at Figure 13 from page 41, Bradshaw (1999)'s [L/Voice] feature may be associated in multiple ways, as shown in Figure 95:

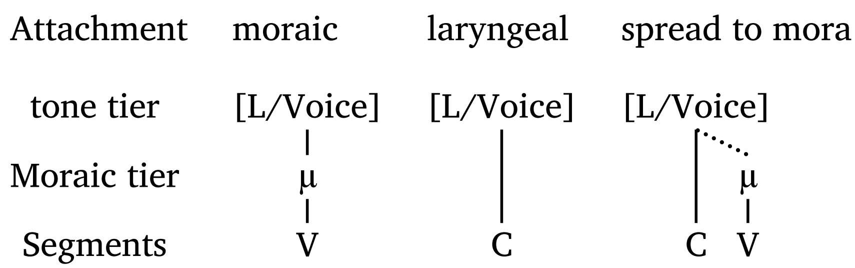


Figure 95. Review of feature geometry in Bradshaw (1999); (=Figure 13)

The left column of Figure 95 shows low tone on a vowel, with superfluous voicing. The center column shows contrastive voicing on a consonant, with an unrealized low tone. The right column shows a depressor consonant, a contrastively voiced consonant in a language where [L/Voice] systematically spreads to a TBU, thereby impacting the tone system.

Given the above summaries of the feature geometries in Snider (1999) and Bradshaw (1999), the potential conflict is that Snider (1999) separates what some consider one feature into two (i.e., tone and register), while Bradshaw (1999) joins a tone with a feature (i.e., voice) which is not elsewhere considered part of the tone system. Given my analysis depends on both Snider (1999) and Bradshaw (1999), it is important to understand how these theories, and the geometries they imply, may coexist. Figure 96 shows five logically possible options for understanding the geometric relationship between tone, register and voice:

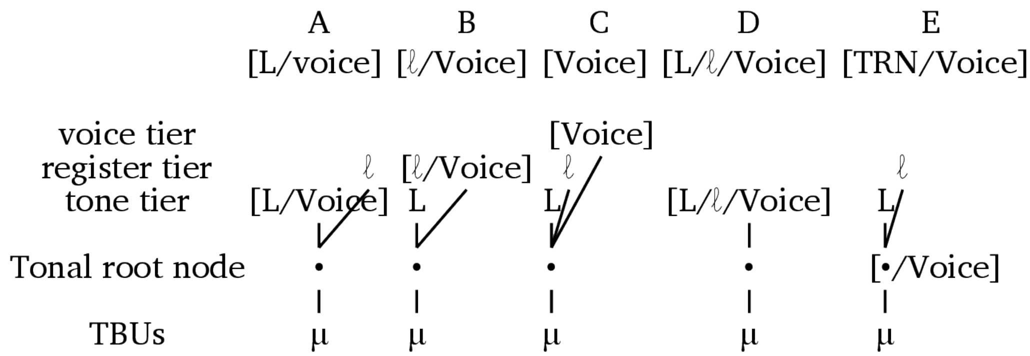


Figure 96. Five possible reconciliations of [L/Voice] (Bradshaw 1999) and RTT (Snider 1999) feature geometries

From left to right, the five geometries combine tone, register, and voice in different manners: tone and voice in one feature/tier, then register and voice on one feature/tier, then each of the three on different features/tiers, then all three in one feature/tier, and finally voice joined to the tonal root node (TRN), with both tone and voice on daughter tiers[1].

Having laid out the logical possibilities for joining Snider (1999)'s RTT and Bradshaw (1999)'s [L/Voice] in Figure 96, I argue that the two columns on the left (A&B) are the only real possibilities for reconciling the two theories. Column C, in separating voice and low tone, denies the fundamental thesis of Bradshaw (1999), which is that tone and voice is contained in a single, privative feature. Column D similarly denies the fundamental thesis of Snider (1999), which is that tone and register operate as independent features in the tone system.

The far right column (E) represents an attempt to join voice to the tone in a manner that leaves tone and register independent, while keeping voice bound to low tone in some intrinsic manner (i.e., through the identity of features, not through association). While this logical possibility ought to be considered (and while it does allow for the spreading of voice with both tone and register features through spreading of the TRN), it is difficult to imagine how binding voice intrinsically to the tonal root node would bind it intrinsically to either of the low tone features (as would be done by either [L/Voice] or [ℓ/Voice]), since either, both, or neither may be associated to a given TRN.

The difficulty of binding voice to a TRN becomes particularly problematic when a TRN is associated to a H tone and/or 𝒽 register feature. In this case, voicing would

[1] I include the geometry in Figure 96 column E as a logical possibility, not because I think it is a plausible geometry in either (Snider 1999) or (Bradshaw 1999).

be intrinsically bound to a TRN with no intrinsic connection or autosegmental association to either L or ℓ. With a TRN bound to a high tone (and not a low tone), the connection between voice and anything that might be called low tone is lost, along with the advances of Bradshaw (1999)'s [L/Voice].

Furthermore, it is not obvious that there is any provision in RTT for the TRN itself to bear any featural value. It is a structural node which allows spreading and copying of tones in part (between tones/features and another TRN) or as a whole (between a TRN and another mora), but it does not bear any distinctive featural value in and of itself.

Even if one were to modify the theory to include the proposal of voice intrinsically bound to a TRN, that proposal would pose other problems. For instance, while it would not typically be a problem for each TBU to have voicing by definition, it would be problematic for a TRN to spread to or from a consonant, as would be required to enact CTI as in Figure 95. As a result, while columns C and D set aside fundamental elements of Bradshaw (1999) and Snider (1999), respectively, column E denies fundamental elements of each.

In consequence of the above, the question at hand is whether, when Bradshaw (1999) discusses the unity of tone and voicing in a privative feature, the feature joined with voice is either tone or register. Because tone and register are shown to operate independently (Snider 1999), voice may be intrinsically bound to either low tone or lower register, but not both. If each of low tone and lower register were intrinsically bound to voice, they would necessarily be intrinsically bound to each other (i.e., in a single feature [L/ℓ/voice]).

Having laid out the theoretical possibilities, it remains to ask which geometry fits the data better, first for a given language, and then cross-linguistically. As a practical matter, at some point the distinction between [L/Voice] and [ℓ/Voice] may be necessary or helpful to an analysis of tone in a language with CTI. That is, in the analysis of a tone system which is theorized to have (or have had) depressor consonants, the question ought to be asked as to whether the consonant-tone interaction shows a correlation between voice and either a tone feature or a register feature.

Practically, this means that we should see either tone values or register effects correlating with voicing specification. In an analysis where voice is one with a tone feature (i.e., [L/Voice] from column A of Figure 96 above), one should expect that contrastive voicing would correlate with a low tone, rather than with any particular change in register. There may yet be a connection between voice and register, but it should be indirect, at most. On the other hand, if voice is one with a register feature (i.e., [ℓ/Voice] from column B of Figure 96 above), one should expect that contrastive voicing would correlate with a lower register, and only indirectly correlate to a lower position on the same register. Finally, any analysis requiring both tone and register to correlate directly with voicing must break either Bradshaw (1999) or Snider (1999), or both.

Another complication arises from the understanding that "each feature defines a tier", which may have origins in Clements & Hume (1995). If H and L are not on the same (i.e., tonal) tier, then a floating Low tone must either include a TRN or timing node (i.e., be a floating feature bundle), or else it cannot be reliably placed in its temporal ordering. That is, if some feature which dominates both H and [L/Voice] is needed to maintain timing, then that feature must float with the low, in order for that floating low to be reliably before one tone and after another, in terms of its impact on the system. For example, if a floating low tone is supposed to lower the second of two high tones, then there must be some mechanism to establish that the low tone is between the two high tones, since that ordering would not be provided by adjacency on a single tier (e.g., a "tone tier", with both high and low tone features). Such a timing mechanism cannot, however, involve being associated to morae, as the floating low should not be associated to a TBU.

Furthermore, in three dimensional space two different features may be associated to a third (e.g., a TBU) without having any particular ordering with regard to each other. Yet the difference between L ordered before H, each of which is associated to a given TBU (which would amount to a rising tone), is critically different from H ordered before L each associated to a given TBU (which would be a falling tone). In addition to the critical distinction between rising and falling tones, there would be the additional logically possible case where L and H on the same TBU would neither precede the other (since they would not occupy the same geometric space, being on two different tiers), and it is not clear how such an occurrence should be interpreted.

However the specifics of the readers assumptions about tone geometry work out, it remains clear that we should expect either tone or register, but not both, to correlate with underlying voicing specification.

APPENDIX F

Extending the model to proto-Bantu and tonogenesis *ex nihilo*

Some readers may take offense at the title of Chapter 7, and perhaps that of the entire work, claiming that I have not shown tonogenesis at all, but merely an addition of a new tone melody. At another time, we might discuss the importance of adding a new tone melody to a language, but here I hope to begin to show that this model can be extended to include languages which did not originally have tone, in any case.

While the model presented in this chapter is justified solely on the basis of languages adding tone melodies on the basis of losing [L/Voice] specification after consonant-tone interaction (CTI) lexicalization, the model presented in this dissertation should be easily extensible to cover languages developing tone without having had tone contrasts initially (tonogenesis *ex nihilo*). This is because there is nothing in this model which requires the input language (W) to have an original tone specification; in fact, such specification, while present in Section 7.4 to make the tonal categories more easily applicable to Bantu D30, is absent from most of Section 7.3; initial tone specification is not necessary to the development of new tonal categories through CTI lexicalization and loss of [L/Voice] specification.

Unfortunately for one trying to reconstruct tonogenesis *ex nihilo* for the Bantu D30 languages, these languages have been reconstructed as having tone at least as far back as proto-Bantu. So any attempt to show tonogenesis *ex nihilo* for Bantu D30 would require arguing for tonogenesis *ex nihilo* in proto-Bantu itself. Fortunately, applying this model to proto-Bantu is not a completely lost cause. It seems still out of reach to point to the first tone contrasts in proto-Bantu, but I hope in this section to make at least some headway toward showing how new tone melodies could have arisen in proto-Bantu through CTI lexicalization and loss of [L/Voice] specification. I will address only one generalization, leaving others for future work.

Any attempt to see the development of tone in proto-Bantu must begin with an acknowledgement that the system is not composed of roots in just two tone melodies, e.g., high and low. Given that there are at least four tone melodies (i.e., H, L, LH and HL), the presumably more recent developments of LH and HL need to be dealt with before any attempt can be made to see the language developing H v L.

If new tone melodies arise from a distribution like that in Table 43 to one like that in Table 44 through loss of contrastive voicing, one would expect that the distribution initially would look exactly like that in Table 44, but with enough passage of time slots which were categorically empty would eventually be filled, either by the kind of devoicing I suggest, or other processes, however random and spontaneous or systematic they may be. Still, one should see a disproportionately large number of lexical items in those categories that were originally filled, and a disproportionately small number of lexical items in those categories that were originally empty.

According to my model, one would expect that the initial distribution of LH tone, for instance, would be entirely of voiced consonants, if it originated as a group of H verbs with an initial depressor consonant. At the same time, one would expect a smaller number of words constructed with voiceless C1 and LH tone, given that the historical lexical bias is toward words constructed with voiced C1 and LH tone. And a distribution of LH words heavily skewed to voiced C1 is precisely what we find, all the way down Table 71:

	voiced	**voiceless**
b/p	36	9
d/t	13	5
j/c	22	11
g/k	11	7

Table 71. LH tone melody by C1 obstruents in BLR3 (Bastin, et al 2002)

This data is far from conclusive. But I find the striking difference in proportions interesting, and worth further consideration. To show this case more clearly, one would need evidence that the data in the right column of Table 71 were originally voiced, though this would require comparison with proto-languages contemporary to proto-Bantu.

References

Alege, Way. 2012. Signes de souffrance des OMD (1-7). Journée Porte Ouverte Faculté de Developpement. Bunia, DRC.

Anderson, Stephen R. 1978. Tone features. In Fromkin (ed), 133-175.

Asangama, Natisa. 1983. *Le budu: esquisse phonologique et grammaticale.* Paris: Université de Sorbonne (Paris III) Doctoral dissertation.

Bastin, Yvonne, André Coupez, Evariste Mumba, & Thilo C. Schadeberg. 2002. Bantu Lexical Reconstructions 3 / Reconstructions Lexicales Bantoues 3. Tervuren: Royal Museum for Central Africa. http://linguistics.africamuseum.be/BLR3.html. (November 6th 2005 version.)

Beach, Douglas M. 1923. The science of tonetics and its application to Bantu languages. *Bantu studies* 2(1). 75-106.

Bird, Steven. 1999a. When marking tone reduces fluency: an orthography experiment in Cameroon. *Language and Speech* 42(1). 83-115.

Bird, Steven. 1999b. Strategies for representing tone in African writing systems. *Written Language & Literacy* 2(1). 1-44.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns.* New York: Cambridge University Press.

Boersma, Paul & Weenink, David. 2014. Praat: doing phonetics by computer [Computer program]. 5.4 edn. http://www.praat.org/.

Bradshaw, Mary M. 1995. Tone on verbs in Suma. In Akinbiyi Akinlabi (ed), *Theoretical approaches to African linguistics,* 255-71. Africa World Press.

Bradshaw, Mary M. 1997. A phonology-phonetics mismatch:[voice] in consonant-tone interaction. *Studies in the Linguistic Sciences* 27(1). 17-31.

Bradshaw, Mary M. 1999. *A Crosslinguistic Study of Consonant-Tone Interaction.* Ohio State University Doctoral dissertation.

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

Cahill, Mike. 2001. Avoiding tone marks: a remnant of English education? *Notes on Literacy* 26.1. 1-11.

Casali, Roderic F. 2003. [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7(3). 307-382.

Casali, Roderic F. 2008. ATR harmony in African languages. *Language and Linguistics compass* 2(3). 496-549.

Cassimjee, Farida. 1998. *Isixhosa tonology: an optimal domains theory analysis.* Munich: Lincom Europa.

Cassimjee, Farida, Charles Kisseberth. 1998. Optimal domains theory and Bantu tonology. In Larry M Hyman & Charles W. Kisseberth (eds.), *Theoretical Aspects of Bantu Tone,* 33-132. Stanford, California: Center for the Study of Language and Information.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* Cambridge, MA: MIT Press.

Clements, George N. & Annie Rialland. 2008. Africa as a phonological area. In Heine & Nurse (eds), 36-85.
Clements, George N., Elizabeth V. Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed), *The Handbook of Phonological Theory,* 245-306. John Wiley & Sons.
Clements, George Nick, Alexis Michaud & Cédric Patin. 2011. Do we need tone features? In Goldsmith, Hume & Wetzels, 3-24.
Connell, Bruce. 2001. Downdrift, downstep, and declination. *Typology of African Prosodic Systems Workshop,* 3-12. Bielefeld University, Germany.
Cope, Anthony T. 1970. Zulu tonal morphology. *Journal of African Languages* 9(3). 111-152.
Dalby, David. 1970. *Language and history in Africa: a volume of collected papers presented to the London Seminar on Language and History in Africa (held at the School of Oriental and African Studies, 1967-69).* Africana Pub. Corp.
Downing, Laura J. 2009. On pitch lowering not linked to voicing: Nguni and Shona group depressors. *Language Sciences* 31(2). 179-198.
Downing, Laura J., Bryan Gick. 2001. Voiceless tone depressors in Nambya and Botswana Kalang'a. *Proceedings of BLS,* 11-11. Citeseer.
Fromkin, Victoria A. (ed). 1978. *Tone: A Linguistic Survey.* New York: Academic Press.
Gafos, Adamantios & Amanda Dye. 2011. Vowel harmony: transparent and opaque vowels. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice*The Blackwell Companion to Phonology: Suprasegmental and Prosodic Phonology,* 2164-2189. John Wiley & Sons.
Gary F. Simons, & Charles D. Fennig (eds). 2018. *Ethnologue: Languages of the world, twenty-first edition.* Dallas, Texas: SIL International. online version: http://ethnologue.com/.
Geoghegan, Tracy. 2013. *State of the World's Mothers.* London: Save the Children International.
Givón, Talmy. 1971. On the verbal origin of the Bantu verb suffixes. *Studies in African linguistics* 2(2). 145-163.
Goldsmith, John A. 1979. *Autosegmental phonology.* MIT Doctoral dissertation.
Goldsmith, John A., Elizabeth Hume, & Leo Wetzels. 2011. *Tones and Features: Phonetic and Phonological Perspectives.* Berlin, Boston: De Gruyter Mouton.
Good, Jeff. 2013. A (micro-) accretion zone in a remnant zone? *Language Typology and Historical Contingency: In honor of Johanna Nichols* 104. 265–282.
Greenberg, Joseph Harold. 1955. *Studies in African Linguistics Classification.* Compass Publishing Company.
Greenberg, Joseph Harold. 1959. Africa as a linguistic area. In William Russel Bascom & Melville J. Herskovits (eds), *Continuity and Change in African Cultures,* 15-27. Chicago: University of Chicago Press.
Greenberg, Joseph Harold. 1960. A quantitative approach to the morphological typology of language. *International journal of American linguistics* 26(3). 178-194.
Greenberg, Joseph Harold. 1970. Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics* 36(2). 123-145.
Güldemann, Tom. 2008. The "Macro-Sudan belt": Towards identifying a linguistic area in northern Sub-Saharan Africa. In Heine & Nurse (eds), 151-185.
Halle, Morris. 1995. Feature geometry and feature spreading. *Linguistic inquiry* 26(1). 1-46.
Hartell, Rhonda L. 1993. *Alphabets of Africa.* UNESCO, Dakar regional Office.
Heine, Bernd & Derek Nurse (eds). 2000. *African languages: An introduction.* Cambridge University Press.

Heine, Bernd & Derek Nurse. 2000a. Introduction. In Heine & Nurse (eds), 1-10.

Heine, Bernd & Derek Nurse (eds). 2008. *A linguistic geography of Africa.* Cambridge: Cambridge University Press.

Heine, Bernd & Zelealem Leyew. 2008. Is Africa a linguistic area? In Heine & Nurse (eds), 15-35.

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

Hernández-Campoy, Juan Manuel & Juan Camilo Conde-Silvestre. 2012. *The handbook of historical sociolinguistics.* Malaysia: John Wiley & Sons.

Hobley, June. 1964. A preliminary tonal analysis of the Bassa language. *Journal of West African Languages* 1(2). 51-55.

Hombert, Jean-Marie. 1975. *Towards a theory of tonogenesis: an empirical, physiologically and perceptually based account of the development of tonal contrasts in languages.* University of California, Berkeley Doctoral dissertation.

Hombert, Jean-Marie. 1978. Consonant types, vowel quality, and tone. In Fromkin (ed), 77-111.

Hombert, Jean-Marie, John J Ohala, William G Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55(1). 37-58.

House, Arthur S., Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25(1). 105-113.

van der Hulst, Harry & Keith L. Snider (eds). 1993. *The Phonology of tone: the representation of tonal register.* Walter de Gruyter.

Hyman, Larry M. 1973. *Consonant types and tone,* vol. 1 (Southern California Occasional Papers in Linguistics). Los Angeles: USC.

Hyman, Larry M. 1976. On some controversial questions in the study of consonant types and tone. In Victoria A. Fromkin (ed), *Studies on Production and Perception of Tones* (Working Papers in Phonetics), vol. 33, 90-98. University of California, Los Angeles.

Hyman, Larry M. 1993. Conceptual issues in the comparative study of the Bantu verb stem. In Lioba Moshi Salikoko S. Mufwene (eds). *Topics in African Linguistics* (Amsterdam studies in the theory and history of linguistic sciences), vol. IV. Amsterdam/Philidelphia: John Benjamins.

Hyman, Larry M. 2000. Privative Tone in Bantu. Symposium on Tone. ILCAA, Tokyo.

Hyman, Larry M. 2003a. Segmental phonology. In Nurse & Philippson (eds), 42-58.

Hyman, Larry M. 2003b. Basaá (A43). In Nurse & Philippson (eds), 257-282.

Hyman, Larry M. 2011. Do tones have features. In Goldsmith, Hume & Wetzels, 50-80.

Hyman, Larry M., Russell G. Schuh. 1974. Universals of tone rules: Evidence from West Africa. *Linguistic inquiry* 5. 81-115.

Kutsch Lojenga, Constance. 1994. Kibudu: a Bantu language with nine vowels. *Africana Linguistica* XI. 127-133. Tervuren, Belgium: Royal Museum for Central Africa.

Kutsch Lojenga, Constance. 1996. Participatory Research in Linguistics. *Notes on Linguistics* 73(2). 13-27. Dallas, TX: SIL.

Kutsch Lojenga, Constance. 2006. Bila (D 32). In Nurse & Philippson (eds), 450-474.

Kutsch Lojenga, Constance. 2014. Orthography and tone: a tone system typology with implications for orthography development. In Michael Cahill, Keren Rice (eds), *Developing orthographies for unwritten languages,* 49-72. SIL International.

Lea, Wayne A. 1973. Segmental and suprasegmental influences on fundamental frequency contours. In Larry M. Hyman*Consonant types and tone* (Southern California Occasional Papers in Linguistics 1). 15-70. Los Angeles: USC.
Leben, William Ronald. 1973. *Suprasegmental phonology*. Massachusetts Institute of Technology Doctoral dissertation.
Leben, William Ronald. 1978. The Representation of Tone. In Fromkin (ed), 177-219.
Lehiste, Ilse, & Gordon E Peterson. 1961. Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America* 33(4). 419-425.
Lewis, M. Paul, Gary F. Simons. 2010. Assessing endangerment: expanding Fishman's GIDS. *Revue roumaine de linguistique* 2. -.
Lewis, M. Paul, Gary F. Simons, & Charles D. Fennig (eds). 2013. *Ethnologue: Languages of the world, seventeenth edition.* Dallas, Texas: SIL International. online version: http://ethnologue.com/.
Lewis, M. Paul, Gary F. Simons, & Charles D. Fennig (eds). 2015. *Ethnologue: Languages of the world, eighteenth edition.* Dallas, Texas: SIL International. online version: http://ethnologue.com/.
Louw, JA. 1971. The tonal paradigm of the verb in Xhosa. *Taalfasette* 13(1). 33-44.
Maddieson, Ian. 1974. A note on tone and consonants. In Ian Maddieson (ed), *The tone tome: studies on tone from the UCLA Tone Project* (Working Papers in Phonetics), vol. 27, 18-27. University of California, Los Angeles.
McPherson, Laura. 2015. Tone features revisited: evidence from Seenku (Mande, Burkina Faso). Annual Conference on African Linguistics (ACAL) 46. Eugene, OR.
Meeussen, Achiel Emiel. 1970. Tone typologies for West African languages. *African Language Studies* 11. 266-271.
Meyer, Ernst A. 1937. *Die Intonation im Schwedischen. T. 1: Die Sveamundarten.* Fritzes.
Mohr, Boudewijn. 1968. Intrinsic fundamental frequency variation. *Monthly Internal Memorandum. University of California, Berkeley* June. 23-32.
Mohr, Burkhart. 1971. Intrinsic variations in the speech signal. *Phonetica* 23(2). 65-93.
Morgan, Laurel & Karen Van Otterloo. 2009. Nyali Tchabi VAU Phonology Sketch. SIL Eastern Congo Group, ms.
Morgan, Paul & Johnny Walker. 2009. Overview of Kibira Phonological Features: Results of Participative Orthography Workshop. SIL Eastern Congo Group, ms.
Mous, Maarten. 2003. Nen (A44). In Nurse & Philippson (eds), 283-306.
Newman, Paul. 2000. Comparative linguistics. In Heine & Nurse (eds), 123-160.
Nurse, Derek & Gérard Philippson (eds). 2006. *The Bantu languages,* 2nd edn. Routledge.
Odden, David. 2011. Features impinging on tone. In Goldsmith, Hume & Wetzels, 81-107.
Odden, David. 2014. Bantu Phonology. *To appear in Oxford Handbooks Online in Linguistics,* http://www.ling.ohio-state.edu/~odden/BantuPhonology.pdf.
Projet de Traduction de la Bible et d'Alphabétisation en Langue Budu et Mr. Bamata-Subama. 1997. *Manuel d'Orthographe Budu.* Isiro, DR Congo: le Projet de Traduction de la Bible et d'Alphabétisation en Langue Budu.
Rasmussen, Kent. 2012. Lexicalized Consonant-Tone Interaction in Nyali-Kilo [nlj] Infinitive Verbs. World Congress of African Linguistics 7. Buea, Cameroon.
Rasmussen, Kent. 2014. Report on Nia-Nia 2014 Alphabet Workshop: Summary of Linguistic Facts. SIL East Congo Group, ms.

Rasmussen, Kent. 2015a. A First Look at Mbo [zmw] (D. R. Congo). Annual Conference on African Linguistics 46. Eugene, OR.
Rasmussen, Kent. 2015b. Ndaka [ndk] (D.R. Congo): Another Bantu D30 Language with Nine Vowels and Depressor Consonants. 7th Annual Metroplex Linguistics Conference. University of North Texas.
Rasmussen, Kent. 2016. Ndaka [ndk] (D.R. Congo): Another Bantu D30 Language with Nine Vowels and Depressor Consonants. Annual Conference on African Linguistics 47. University of California, Berkeley.
Roberts, David, S Borgwaldt, T Joyce. 2013. A tone orthography typology. *Typology of writing systems* 51. 85-85.
Roberts, David, Stephen L Walter. 2012. Writing grammar rather than tone: An orthography experiment in Togo. *Written Language & Literacy* 15(2). 226-253.
Roberts, James S. 1994. Nontonal floating features as grammatical morphemes. *Work Papers-Summer Institute of Linguistics, University of North Dakota Session* 38. 87-99.
Rose, Sharon, & Rachel Walker. 2011. Harmony Systems. In John A. Goldsmith, Jason Riggle & C. L. Alan (eds), *The handbook of phonological theory,* 240-290. John Wiley & Sons.
Schuh, Russell G. 1978. Tone rules. In Fromkin (ed), 221-256.
Snider, Keith. 1999. *The geometry and features of tone* (Publications in Linguistics 113). USA: Summer Institute of Linguistics and The University of Texas at Arlington.
Snider, Keith. 2014. On establishing underlying tonal contrast. *Language Documentation & Conservation* 8. 707-737.
Snider, Keith. 2018. *Tone Analysis for Field Linguists.* Dallas, TX: SIL International.
Snider, Keith & Harry van der Hulst. 1993. Issues in the representation of tonal register. In van der Hulst & Snider (eds), 1-27.
Starwalt, Coleen Grace Anderson. 2008. *The acoustic correlates of ATR harmony in seven-and nine-vowel African languages: A phonetic inquiry into phonological structure.* University of Texas at Arlington Doctoral dissertation.
Stewart, John M. 1983. Downstep and floating low tones in Adioukrou. *Journal of African Languages and Linguistics* 5(1). 57-78.
Stewart, John M. 1993. Dschang and Ebrié as Akan-type total downstep languages. In van der Hulst & Snider (eds), 185-244.
Stewart, John Massie. 1964. *The typology of the Twi tone system.* Institute of African Studies, University of Ghana.
Stewart, John Massie. 1968. The typology of the Twi tone system. In John Massie Stewart, Paul Schachter, & William Everett Welmers, *The Typology of the Twi Tone System: With Comments by P. Schachter & WE Welmers. (preprint)* (Bulletin of the Institute of African Studies 1). 1-27. Institute of African Studies, University of Ghana.
Stirtz, Timothy M. 2009. [ATR] vowel harmony in Gaahmg–a six-vowel system. *Journal of African Languages and Linguistics* 30(5). 73-79.
UNDP. 2011. *Human Development Report 2011: Sustainability and Equity: A Better Future for All.* New York: The United Nations Development Program (UNDP).
UNDP. 2013. *Human Development Report 2013: The Rise of the South: Empowered lives. Resilient nations. Human Progress in a Diverse World.* New York: The United Nations Development Program (UNDP).
Welmers, William E. 1962. The phonology of Kpelle. *Journal of African Languages* 1(1). 69-93.

References

de Wit, Gerrit. 2010. Lika Phonologie. *SIL Electronic Working Papers* 2010(002). http://www.sil.org/silewp/abstract.asp?ref=2010-002.

Wright, Richard, Aaron Shryock. 1993. The effects of implosives on pitch in SiSwati. *Journal of the International Phonetic Association* 23(1). 16-23.

Yip, Moira. 1989. Feature geometry and cooccurrence restrictions. *Phonology* 6(02). 349-374.

Yip, Moira. 2002. *Tone.* Cambridge: Cambridge University Press.

Glosses Index

Entries are verbs unless otherwise indicated.

Segmental Forms Index

CVC verb roots follow kɔ/ko/ka- prefix;
(V)C verb roots follow kɛn(ɨ)-/ken(i)-/kan- prefix.

Segmental Forms Index

Segmental Forms Index

Segmental Forms Index