Consonant-tone interaction in the Khoisan language Tsua

by

Timothy K. Mathes

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

New York University

January, 2015

Christopher T. Collins

Gillian Gallagher

© Timothy K. Mathes

All Rights Reserved, 2015

'...the notion that one can capture some aspect of reality by making a description of it using a symbol and that to do so can be useful seems to me a fascinating and powerful idea.'

-- *D. Marr*

DEDICATION

For Ma Lil and Aunt Matt

ACKNOWLEDGEMENTS

I am deeply grateful to my committee, Chris Collins, Gillian Gallagher, Lisa Davidson, Andy Chebanne and Doug Whalen, for their invaluable advice and support. My committee co-chairs, Chris and Gillian, provided the guidance and expertise I needed to turn a language documentation project in Botswana into a full-fledged dissertation. I would like to thank Lisa, Andy and Doug for their numerous and helpful comments throughout the process.

I am very appreciative for having Chris Collins as a mentor. Chris has provided me with a tremendous number of opportunities to expand my fieldwork experience and knowledge of linguistics since we first met. What began as a casual conversation about the documentary "The Language You Cry In" has grown into a fruitful collaboration on many levels that was essential for my success at New York University. Chris's suggestion that I apply to the first-ever African Linguistics School (ALS) held in Accra, Ghana, in the summer of 2009 and right before my first semester of graduate school, led to one of the greatest experiences of my life. Not only was it a wonderful opportunity to learn about African linguistics from the most eminent scholars in the field, but it was amazing to study in an environment with highly motivated and welcoming classmates.

I would like to give a warm thanks to the organizers of ALS Ghana, John Singler, Chris Collins, Enoch Aboh and Akinbiyi Akinlabi, as well as the rest of the faculty, Hilda

Koopman, Malte Zimmermann, Ana Deumert, Amanda Miller, Violeta Vasquez Rojas, Eno-Abasi Urua, Pieter Muysken, Evershed Amuzu and Firmin Ahoua. Warm wishes go to my classmates Tolulope Odebunmi, Gbenga Ajileye, George Iloene, Emmanual Tonukari, Beatrice Bruku, Ebitare Obikudo, Reginald Duah, Golden Ekpo, Ekeme Ukpong, Adeola Agoke, Atikonda Mtenje, Dorothy Adusei, Jean-Claude Dodo, Claire Halpert, Khady Tamba, Theophile Nasse, Johnson Ilori, Juliet Oppong-Asare, Kamal Tagba, Beatrice Koffi, Musah Anthony Agoswin, Margot van den Berg, Cosmas Amenorvi, Regina Caesar, Ezekiel Bolaji, Blibli Serge, Afua Blay, and Elvis Rescue Yevudey. I am particularly indebted to Tolulope Odebunmi for discussions ranging from sociolinguistics, tone, and the Yoruba language to Nigerian pidgin. I am so blessed to have met such a fine group of people.

I would like to acknowledge the faculty and students from the 2nd African Linguistics School held in Porto-Novo, Benin during the summer of 2011. Everything I learned from ALS Ghana and Benin has heavily influenced my way of thinking about natural languages. I would like to thank Mary Paster in particular for taking the time to read my Masalit tone paper and providing so many helpful comments and encouragement to study tone languages.

My fieldwork in Botswana would not have been possible without Chris Collins and Andy Chebanne's knowledge of and enthusiasm for Khoisan languages. I am deeply thankful to Andy and his family for offering me a place to stay in Gaborone for not one but *two* summers in a row. It was a very generous and kind act that I will never forget. I am grateful to the Tsua consultants, speakers S, B, M, and the Kua consultant, speaker O, for their patience and wisdom during the elicitation sessions. I will always remember their warmth and good humor.

Many thanks go to my professors and classmates at New York University. I have learned a great deal from everyone during my time there. I am especially grateful to Txuss Martin, Violeta Vasquez Rojas, Tricia Irwin, Vincent Chanethom, and my co-horts Emily Nguyen, Luiza Newlin-Lukowicz, Meera Saeed Al Kaabi, Carina Bauman, Sang-Im Lee, Tim Leffel, Kim Leiken, Sean Martin, Neil Myler, and Sangjin Hwang for sharing the ups and downs of graduate school in one way or another.

I would like to express my gratitude to the audiences at the 4th International Symposium on Tonal Aspects of Languages, Nijmegen, The Netherlands, and the 5th International Symposium on Khoisan Languages and Linguistics, Riezlern, Austria, for their excellent comments and advice. Thanks also go to Kristine Yu for organizing the UMass Tonal Spaces Workshop, where I decided to incorporate Yi Xu's superb ProsodyPro Praat script into the analysis of the Tsua tonal melodies. The timing of these three events during the summer of 2014 couldn't have been better! I was able to improve my research many times over because of all I learned and experienced. In particular, a warm thanks goes to Hirosi Nakagawa at the Khoisan symposium for taking the time to listen to numerous Tsua sound files and suggesting appropriate transcriptions for some of the trickier clicks and non-clicks.

Finally, I would not have finished this dissertation without the constant support from family, friends and my wife Irene. Irene's unwavering support was a godsend. Everyone's confidence in me bolstered my voyage through graduate school from the beginning, through the invariable frustrations, up until the very end. I would not have made it without numerous people being there every step of the way. This dissertation is dedicated to Ma Lil and Aunt Matt, whose kind words and belief in me as a child would positively influence the rest of my life.

PREFACE

The neuroscientist and psychologist David Marr once wrote that any machine conducting an information processing task must be understood at three different levels before it could be understood completely (1982). The three levels of description were presented as an approach in developing a theory of perceptual information processing. These levels can be summarized as:

Table 1 Marr's three levels of description

Computational Theory	Representation and Algorithm	Hardware Implementation
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?	How can the representation and algorithm be realized physically?

With respect to developing a computational theory of perception based on the three levels, Marr writes:

> At one extreme, the top level, is the abstract computational theory of the device, in which the performance of the device is characterized as a mapping from one kind of information to another, the abstract properties of this mapping are defined precisely, and its appropriateness and adequacy for the task at hand are demonstrated. In the center is the choice of representation for the input and output and the algorithm to be used to transform one into the other. And at the other extreme are the details of how the algorithm and representation are realized physically - the detailed computer architecture, so to speak (p. 72).

My journey toward understanding the Tsua tonal system, which constitutes the heart of this dissertation, has been guided by Marr's intuition. The fundamental research question in this dissertation can be observed at the Hardware Implementation level: How are tones in Tsua realized acoustically? Marr held the view that the three levels are rather loosely coupled; nevertheless, in carrying out the current inquiry, I have given a considerable amount of thought to the Representation and Algorithm level in an attempt to further inform the Hardware Implementation level. Simply put, how could Tsua depressed High tones be represented? What process plausibly accounts for the acoustic properties of their surface realization?

One goal of this research is to increase our understanding of Khoisan tonal systems by carefully and systematically considering instrumental evidence while pondering the levels outlined above. Although many instrumental studies in Khoisan linguistics have focused on the production of the typologically rare click consonants, it is my hope that researchers will find the realization of tone in Khoisan equally fascinating. It is in this spirit, guided by Marr's intuition, that has given me the impetus to forge ahead.

ABSTRACT

This dissertation examines the acoustic phonetic properties of consonant-tone interaction in the critically-endangered and little studied Khoisan language Tsua, using speech production data from original field research in Botswana. Tsua exhibits an extreme tonal depressor effect that lowers a post-consonantal, root-initial High tone's production by 50 Hz or more when it is followed by a non-High tone and preceded by a voiced obstruent, aspirated obstruent or the glottal fricative /h/, a typologically rare pattern.

The root-initial High tones in the HM and HL tone melodies with onset consonants /b, d, dz, ɟ, g, ɢ, p^h, t^h, ts^h, c^h, k^h, q^h, h, gǀ, gǂ, gǁ, ǀɢ, ǁɢ, ǀ^h, ǂ^h, ǁ^h, ǀq^h, ǂq^h, ǁq^h/ are depressed. However, this inquiry finds that 19.3% of the root-initial High tones that are expected to be lowered in the post-depressor context are not. It is argued that the tonal depression exceptions involving voiced obstruents correspond to historically nasal sonorants or nasalized clicks, consonants which are not depressors in Tsua synchronically or diachronically. Thus, voiced obstruents function as depressors in lexical items that have not undergone this historic sound change, an explanation referred to as the Tsua Depression Exceptions Hypothesis in this thesis.

To tease apart the acoustic effects by depressor type, F0 curves are evaluated using Smoothing Spline Analysis of Variance (SS ANOVA), a domain-general statistical tool for

the rigorous comparison of curves along multiple reference points (Gu 2002, 2013). Three main conditions in the implementation of depressed High tones are tested via SS ANOVA on normalized F0 z-scores: (i) when the depressed High tone follows voiced versus aspirated obstruents and is followed by a Mid tone; (ii) when the depressed High tone follows voiced versus aspirated obstruents and is followed by a Low tone; and (iii) when the depressed High tone is followed by a Mid versus Low tone regardless of depressor type. The effect on Mid tones preceded by depressors is examined as well.

The results indicate that the F0 curve shapes are almost identical regardless of whether a root-initial High tone is produced following a root-initial aspirated obstruent or a root-initial voiced obstruent and when it precedes a Mid tone (i.e. DH-M). The F0 curve shapes are also quite similar when a root-initial depressed High tone is produced following aspirated obstruents as opposed to voiced obstruents and when it precedes a Low tone (DH-L). The F0 curve shapes from normalized time point 30 to the F0 peaks are similar when a depressed High tone precedes a Mid versus Low tone regardless of depressor type. There is no evidence to suggest that Mid tones undergo extreme depression in Tsua.

Tsua tonal depression may be the result of a Low tone insertion rule given the nature of the F0 curves. The purported rule is not completely isomorphic as the Low tone insertion rule for depression in Ikalanga proposed by Hyman and Mathangwane (1998), as Tsua depression is more context dependent. Moreover, the Tsua depression pattern is not fully accounted for by the [L/voice] feature in Bradshaw's Multiplanar model of consonant-tone interaction. Formally, I account for the Tsua depression pattern via the Low tone insertion rule in (1).

(1) Tsua tonal depression rule

$\varnothing \rightarrow$ L / [-sonorant, +slack] ___ H [-H] #

The rule in (1) proposes that [+slack] is the key feature that unifies the depressor types pending further acoustic and articulatory supporting evidence. What is important is that a Low tone insertion account is plausible given modifications to previous proposals for other languages.

It is unclear whether the rule in (1) is present in Tsua's phonology or if the current depression pattern is a remnant of Tsua's phonology from long ago. As a first step towards determining whether the rule is synchronically active, a two-part tone Wug Test is proposed for a future experiment. If the rule is found to be productive, this raises the interesting possibility that click replacement, a sound change process whereby a click consonant is replaced by another click or non-click consonant diachronically, may interact with rule application in some way, given the tonal depression exceptions for historically nasal clicks that were replaced by voiced clicks or voiced non-clicks. Such a possibility has not been explored in the Khoisan literature. The findings reported in this dissertation expand our knowledge of tonal phonetics by showing what is possible in a typologically rare tone system and highlight the importance of statistical methods in phonetic fieldwork.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Tone production varies considerably in terms of how complex F0 curves can be cross-linguistically. A complex F0 curve is defined here as excursions between the F0 onset, F0 maximum, F0 minimum and F0 offset that are not “flat” and may even be extreme over the course of a tone or tone melody. The East Asian tone language Thai is a good example based on plots from several acoustic studies (see Gandour 1974; Abramson 1979; Morén and Zsiga 2006; Zsiga and Nitisaroj 2007). The Khoisan language Tsua resembles Thai in this respect: original field research reveals tone melodies that have complex F0 curves. Fig. 1 introduces the Tsua tonal space.

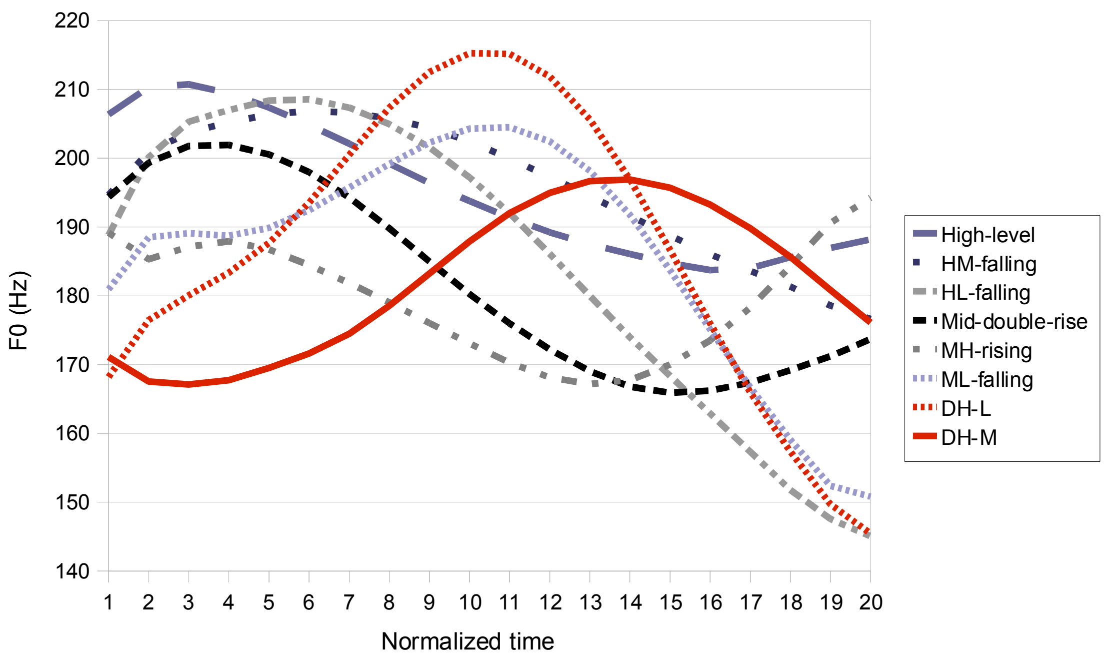


Fig. 1 The Tsua tonal space with its complex F0 curves.

The central aim of this dissertation is to explain why the Tsua tonal space is the way it is and which historical factors may have led to its synchronic state. I show that the Tsua tonal melodies with F0 shapes having the most extreme excursions (i.e., the red curves in Fig. 1) are the result of a rare consonant-tone interaction pattern with depressor types found in both African and East Asian tone languages. The Tsua depressor types are voiced obstruents, aspirated obstruents and the glottal fricative /h/. The pattern's rarity can be observed when we consider cross-linguistic data on depressor consonants in Bradshaw (1999) and Tang (2008). It is quite common for voiced obstruents to be tone depressors, particularly in African languages. Aspirated obstruents are reported to be depressors in many East Asian languages. It is not common for the glottal fricative /h/ to be a depressor in either African or East Asian languages. It is rare for all three depressor types to be found in one language as in Tsua.

Table 2 has examples of African and East Asian tone languages and their respective depressor consonant types. Citations are provided in the table. Yaka is analyzed as having Low tone insertion after root-initial voiced obstruents. Bassa voiced obstruents and the glottal fricative /h/ most often co-occur with Low tones. Kalang’a [Ikalanga] has voiced obstruents and aspirated obstruents as depressor consonants. Voiced obstruents and aspiration have an affinity for Low tone in the Wujiang dialect of Chinese, while in Jingpho, voiced obstruents and /ħ/ have an affinity for Low tone. Aspiration has an affinity for Low tone in Manange.

Table 2 Consonant-tone interaction in African and East Asian languages

(based on Bradshaw 1999 and Tang 2008)

African Languages		
Language	**Location**	**Effect**
Yaka	Central African Republic, Rep. of Congo-Brazzaville	Low inserted after root-initial voiced obstruents, Low triggers voicing (Kutsch Lojenga 1998)
Bassa	Liberia	Voiced obstruents and /h/ have an affinity for Low (Hobley 1964)
Kalanga'a	Botswana	Voiced obstruents and aspiration have an affinity for Low tone (Downing and Gick 2005)
East Asian Languages		
Wujiang dialect	China	Voiced obstruents and aspiration have an affinity for Low tone (Shen 1994)
Jingpho	Myanmar	Voiced obstruents and /ħ/ have an affinity for Low tone (Maran 1971)
Manange	Nepal	Aspiration has an affinity for Low tone (Hildebrandt 2003)

Tsua's rare depression pattern provides a unique opportunity to investigate the acoustic phonetic properties of consonant-tone interaction by tone and depressor type. The research questions motivated by this rare depression pattern include whether all tones are depressed and if the effect on F0 varies by depressor type. Statistical analysis via SS ANOVA pairwise comparisons reveals two striking generalizations: (i) only root-initial High tones followed by a non-High tone are depressed, and (ii) the overall shapes of the F0 curves are more important for identification than whether they are produced at a slightly higher or lower Hertz value, even if the difference is statistically significant. The first finding reflects the heavily context dependent nature of Tsua depression. The second finding stands in contrast to studies of other tone systems that suggest the relative Hertz differential between adjacent tones is more important for identification, a differential that has been reported to be at least 10 Hz for tones to be reliably interpreted (Rietveld and Gussenhoven 1985).

The dissertation is organized as follows. Chapter 1 presents the societal context and evidence for Tsua's critical endangerment, because these factors may be partially responsible for the sound changes that have affected the interaction between Tsua's consonants and tones. Chapter 2 describes the consonant and vowel inventories, including the click and non-click consonants. The description of the inventories lays the groundwork for the consonant-tone interaction analysis later in the thesis. Chapter 3 details the tonal system through a rigorous acoustical analysis. Chapter 4 provides evidence for the rare tonal depression pattern found in Tsua. A hypothesis explaining exceptions to tonal depression is put forth with supporting evidence for three predictions made by the hypothesis. Chapter 5 incorporates SS ANOVA for statistical comparison of the tonal melodies by depressor type. Theoretical issues in consonant-tone interaction are addressed based on the results of the statistical analysis. The dissertation concludes with a summary of the most important findings in Chapter 6.

CHAPTER 1

1.1 Introduction

Every natural language has a societal context, in part because of the historical, social and economic factors that influence the daily lives of the speech community. These key factors affect language use, change and endangerment, among other things. In the case of Tsua, the sociolinguistic situation has led to changes in aspects of its phonological system; namely, language-contact induced click replacement. This chapter presents the societal context that has created the conditions conducive for click replacement, which has had a direct consequence on Tsua lexical tone and tonal depression, as will be shown in subsequent chapters. The relationship between click replacement, lexical tone and tonal depression has not been investigated in any Khoisan language as far as I know.

I begin by giving a brief history of southern Africa and the impact outside groups have had on the indigenous societies when they migrated to the area. Background information on modern-day Botswana with statistics is presented followed by Tsua's classification and evidence of its endangerment. The topics of click replacement and its causes are explored. The chapter concludes with technical information about the data collection process that serves as the basis for my analysis throughout the rest of the dissertation.

1.2 A brief history

Indigenous groups have inhabited southern Africa for thousands of years, where they practiced hunting and foraging before the arrival of the Bantu over a millennium ago, living on fruits, roots and prey. There is archaeological, historical and linguistic evidence that these groups occupied a much greater area in southern Africa compared to today (Andersson and Janson 1997:107). The indigenous groups spoke a variety of languages, which were often only distantly related to each other and unrelated to any others in Africa (Westphal 1962; Vossen 1984; Childs 2003; Güldemann 2006).

The Bantu may have originated in the area that is now Cameroon and spoke a language referred to as Proto-Bantu before their migration eastwards about 2,500 years ago to the region around Lake Victoria (Andersson and Janson 1997:14). Approximately 500 years later, there were waves of migration from the Lake Victoria area, with various Bantu groups moving southwards along the eastern coast of Africa and others migrating to the central and western parts of southern Africa over a period of several hundred years, populating vast swaths of the southern African continent, bringing with them cattle and their knowledge of farming (ibid).

It is important to emphasize the intricate patterns of movement that occurred as the new populations came into contact with the indigenous peoples and each other in southern Africa, for this has had profound consequences on the current sociolinguistic situation. The migration of the Bantu groups into what is now the country of Botswana took place at different periods and from different directions, taking at times complex routes from the west, north, east and southeast in wave after wave over a long period of time, perhaps a thousand years. The earliest Bantu-speaking groups most likely arrived about 1,700 years ago, although their descendants may not have stayed in the area. The Bakalanga, on the other hand, arrived in modern-day western Zimbabwe and northeastern Botswana around 1000 A.D. and remained in the area, the result of migration from the north or northwest, speaking a language of the Shona group. Not long after, the Bakgalagadi first settled the south-eastern part of modern-day Botswana, speaking a Southern Bantu language belonging to the Sotho group that also includes Setswana, Sesotho and Sepedi (Andersson and Janson 1997:19).

Around the 16th century, the Bakgalagadi may have been pushed further west into the desert by several groups of Batswana arriving from the south-east, which included the people known as Bakwena. The Batswana eventually dominated the region, expanding throughout the area in a complex series of splits and movements: two large groups, the Bangwaketse and the Bangwato, seceded from the Bakwena and formed independent chiefdoms. The Bangwato moved northward and extended their power over a large area that is now the Botswana Central District, with the Bakalanga later becoming their subjects (Andersson and Janson 1997:23). While the Batswana increased their power and influence over the region, other Bantu groups continued to arrive from the north, including the Basubiya, the Wayeyi and the Hambukushu.

Meanwhile, the encounter between European immigrants and the Khoisan-speaking populations began in earnest in 1652 with the establishment of a supply post by Jan van Riebeeck on behalf of the Dutch East India Company at Table Bay near present-day Cape Town, South Africa (Vossen 2013a:1). The Cape region was inhabited by reportedly "light-skinned" non-Bantu speaking people who had herds of sheep or kept cattle, lived permanently in small villages, and later became referred to as the "Hottentots", while the adjacent hinterland was inhabited by small groups of non-Bantu hunter-gatherers with "brownish complexion" who appeared to be ethnically and culturally distinct from the pastoralists. The term "Bushman", derived from "Bosjesman" which was in turn derived from "Bosmanneken" was used to designate the second group (ibid). Linguistically, both groups spoke languages that were characterized by a high proportion of click consonants.

As time passed, Dutch frontier farmers (mostly Trek Boers) from the Cape Colony pushed deeper into the hinterland, in search of pasture for their livestock. The Dutch settlers came into contact with an increasing number of hunter-gatherer societies, whom they despised and drove away in tremendous numbers. The destruction of the indigenous societies as a result of disease and genocidal persecution by the settlers took only three hundred years (Childs 2003:3). Many of the indigenous societies were assimilated by Bantu groups as unfree laborers or forced to migrate to the most inhospitable areas of southern Africa, although there is evidence to suggest that contact with the Bantu groups was not always disharmonious (Chebanne 2008).

The sociocultural distinction made between the two aforesaid groups *Khoi*, formerly “Hottentot” and meaning 'person' in the Nama language, and *San*, formerly “Bushman” and meaning 'gatherer, forager' in Nama as well, along linguistic and ethnic lines was first discussed in historical records in the early 20th century. The terms Hottentot and Bushman fell out of favor as they became considered offensive and inappropriate as linguistic descriptors.

The compound term *Khoïsan* was coined in 1928 by Leonard Scholtze for linguistic and anthropological classification, although this nomenclature has not been without its problems either. Referring to “Khoisan”, Chebanne (2008:93) states: “This term has to non-linguists therefore taken a pejorative connotation, even relegating these people to a subhuman class.” Nevertheless, the term gradually lost its pejorative connotation, as more comprehensive research on the Khoi and San communities gave them validity, perhaps beginning with Isaac Schapera's 1930 publication “The Khoisan Peoples of South Africa: Bushmen and Hottentots” (ibid).

1.3 Modern-day trends

Recently, it has been estimated that there may have been more than 100 languages spoken by the indigenous populations in southern Africa before contact with outside groups (Güldemann and Vossen 2000:99). Today, the term Khoisan (also spelled *Khoesan*[1])

---

[1] See Nienaber's 1990 proposal for a general spelling change from *oi* to *oe* to harmonize with Khoisan orthographies, since these vowel sequences can be contrastive in Khoisan languages. I have chosen to

is used as an umbrella term that embraces about 30 languages spoken in the southern African countries of Botswana, Namibia, Angola, Zambia, Zimbabwe and South Africa, and two isolates, Hadza and Sandawe, spoken in Tanzania. Many of them are bound for extinction. A significant number are spoken in Botswana, a sparsely-populated country that was peacefully transformed to an independent nation from the British Protectorate of Bechuanaland in 1966 and is home to the arid conditions of the Kalahari Desert.

Fig. 2 shows the population density in 2007, with most residents living in the eastern part of the country as illustrated by the darker shading, presumably influenced by the geographical location of the Kalahari desert. The blue arrow north of the country's capital Gaborone points to the approximate location of the Tsua consultants' home village Moralane. Additional background information on the consultants is presented in §1.7. Botswana's population density may be relevant for sound changes observed in some Khoisan languages because of the increased inter-group contact in the more densely populated areas, particularly in the eastern part of the country where the languages are less isolated compared to the western areas. These processes will be defined and examined in detail in §1.4 of this chapter and in subsequent chapters.

---

adopt the spelling *Khoisan* throughout this dissertation, which is familiar among Khoisanists and non-Khoisanists alike.

© Copyright 2007 by World Trade Press. All Rights Reserved.

Fig. 2 Botswana population density in 2007. The blue arrow points to the approximate location of the Tsua consultants' home village Moralane.

The latest 2011 census lists a total population of 2,024,904 in Botswana.[2] According to an earlier 2006 Botswana census, 36,580 individuals or 2.2% of the population at that time spoke *Sesarwa* at home, a slight increase compared to the 2001 census numbers given in Chebanne and Nyati-Ramahobo (2003:396), of about 30,037 people, or 1.9% of the population. Sesarwa is a generic label used in Botswana to designate the Khoisan communities. Thus, official reports overlook the ethnic and linguistic identities of the Khoisan communities (Chebanne 2008). The very low percentage of individuals who speak Khoisan languages in Botswana is one indicator of their minority status. Table 3 presents the 2006 census data of languages most often spoken at home in Botswana:[3]

---

2 http://www.cso.gov.bw/index.php?option=com_content1&id=2&site=census

3 A list of languages most often spoken at home based on the 2011 census data was not available at the time of writing.

Table 3 Number and percentage of individuals aged 2 years and over by language most often spoken at home (Botswana Census Data 2006)

Language	Number	Percentage
Setswana	1,221,595	72.6
Ikalanga	142,077	8.4
Shekgalagadi	48,029	2.9
Sebirwa	43,126	2.6
Setswapong	37,201	2.2
Sesarwa	**36,580**	**2.2**
English	34,772	2.1
Mbukushu	33,077	2.0
Other	23,109	1.4
Shona	22,029	1.3
Ndebele	11,275	0.7
Herero	8,226	0.5
Afrikaans	8,215	0.5
Shiyeyi	7,578	0.5
Subiya	6,721	0.4
Total	**1,683,611**	**100**

It can be observed in Table 3 that Botswana contains a diverse range of Bantu languages despite its small population, likely due to its geographical location at the crossroads of several regions with different histories and cultures (Andersson and Janson 1997:15). Setswana is spoken by the vast majority of the Botswana population (72.6%), followed by Ikalanga at 8.4%. English is the declared official language and Setswana is the national language, although Setswana serves as a de facto co-official language, having more prestige than any other indigenous language (Smieja 2002:212-213). Another factor contributing to the prestige enjoyed by Setswana is the fact that it is spoken not only in Botswana but also in parts of Namibia, South Africa and Zimbabwe, with a grand total of about 4.6 million speakers and functioning as a supranational means of communication (ibid).

In spite of the elaborate web of turmoil that resulted from the historical contact between disparate groups in southern Africa, modern-day Botswana has made significant progress in terms of economic development, benefiting from democracy, peace and stability in the decades since independence from the United Kingdom in 1966. Botswana's Gross Domestic Product (GDP) was $14.41 billion USD in 2012 according to the World Bank. The 2012 Ibrahim Index of African Governance (IIAG), which provides an independent, annual assessment of governance performance in Africa based on 88 indicators, ranks Botswana third overall out of 52 African nations with a score of 77.2 out of 100, only to be outranked by Cape Verde (#2; 78.4/100) and Mauritius (#1; 82.8/100).[4] The indicators are meant to track the general categories of (a) Safety and the Rule of Law, (b) Participation and Human Rights, (c) Sustainable Economic Opportunity, and (d) Human Development.

Nonetheless, regional economic development has had its drawbacks. Life has drastically changed for the people of Khoisan origin, with some moving to the cities in search of employment opportunities and others working on farms or herding cattle. This

---

[4] http://www.moibrahimfoundation.org/botswana/

state of affairs reflects the perpetual economic dependency that some Khoisan communities have had on Bantu groups historically, such as the Bakalanga in the north, the Batawana in the south and the Bakgalagadi in the central Kalahari, through the *mafisa* system, which essentially means that one herds cattle in return for milk and meat, sometimes including the right to keep the calves (Wilmsen and Vossen 1990:10; Andersson and Janson 1997:138). The long tradition of economic dependence on the Bantu is mainly due to the fragile and unpredictable socioeconomic life of the Khoisan communities, who had always relied on hunting and gathering in the past, but found it no longer feasible to do so (Batibo 2003:285).

The Botswana government appears to espouse the view that with development and modernity, there is no compelling reason to maintain the rural languages and cultures of the Khoisan communities, and would rather assimilate them into modernity (Chebanne 2003a, 2008). From colonial times until the present, the Botswana administration's language use policy has sought to promote Setswana's dominance in all areas of social life and as the unifying vehicle of communication, penetrating the remote areas as deeply as possible (Chebanne 2002b; Smieja 2002; Nyati-Ramahobo 2001; Batibo 2005). For many of the indigenous groups, assimilation has meant very low wage jobs as cattle post workers, goat herders or domestic laborers, with the net effect of continued ethnic and linguistic marginalization and stigmatization, which has led to increased levels of language shift and endangerment in Botswana (Chebanne 2002a; Sommer and Widlock 2013). To the indigenous groups, Setswana is viewed as a language of wider interaction and social integration — it is a language associated with high status, socioeconomic advancement, modern living and education (Batibo 2002:197). To the majority group, speakers of the minority languages are often regarded as hampering national unity and economic growth, and are pushed aside as an impediment to faster development (Smieja 2002:211).

1.4 Tsua classification and endangerment

Greenberg (1963) divides Khoisan languages into Northern, Central and Southern groups, although this arrangement continues to be under debate. Fig. 3 gives the geographic location of the Kua-Tsua dialect continuum as mapped in Güldemann and Vossen (2000:100). Tsua is classified by Güldemann (2006) and Güldemann and Vossen (2000) as a Kalahari Khoe East language of the Central Khoisan branch and part of the Tshwa subgroup (Fig. 4). The Kalahari Khoe East languages occupy the eastern part of Botswana, southeast of the Central Kalahari Game Reserve and eastward towards the Western Sandveld of the Central District. These languages are spoken mainly around the Makgadikgadi Salt Pans, as well as towards the Shashe river, Serule and Mabesekwa areas in eastern Botswana, Serowe and Shoshong (Barnard 1992). The languages in Figs. 3 and 4, particularly the ones in the Khoe family that subsumes the Khoekhoe, Kalahari Khoe East and Kalahari Khoe West branches, are referred to throughout this dissertation.

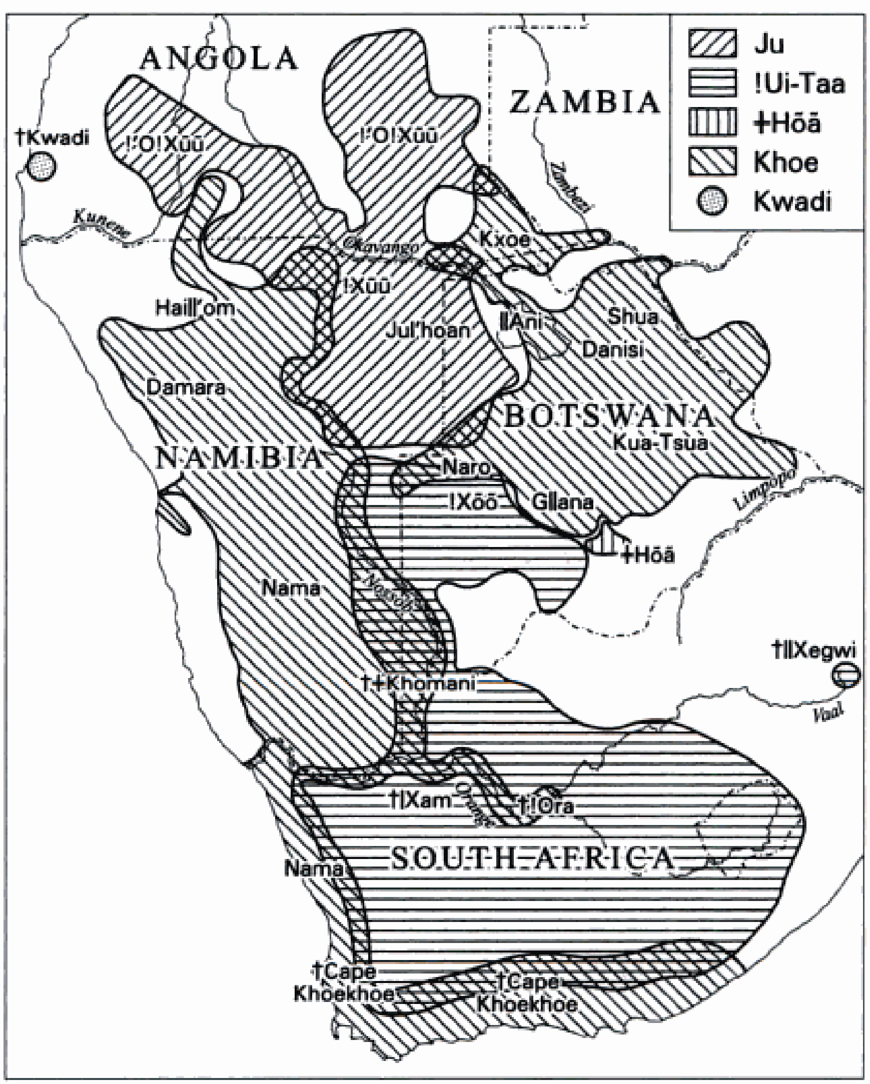


Fig. 3 Kua-Tsua geographic location in Eastern Botswana (Güldemann and Vossen 2000:100).

Lineages and branches	**Language(s) or dialects**
(1) Khoe-Kwadi (*Central Khoisan*)	
Khoe	
Khoekhoe	
North:	Eini†, Nama-Damara, Hai||’om
South:	!Ora†, Cape varieties†
Kalahari	
East	
Shua:	Cara, Deti†, |Xaise, Danisi, Ts’ixa, etc.
Tshwa:	**Kua**, Cua, **Tsua**, etc.
West	
Kxoe:	Khwe, ||Ani, Buga, G|anda, etc.
G||ana:	G||ana, G|ui, ǂHaba, etc.
Naro:	Naro, etc.
Kwadi	single language†
(2) Ju-ǂHoan (*Northern Khoisan*)	
Ju (DC)	
Northwest:	!’O!Xuun, !Xuun
Southeast:	Ju|’hoan, ǂKx’au||’e
ǂHoan	ǂHoan, Sasi
(3) Tuu (*Southern Khoisan*)	
Taa-Lower Nossob	
Taa (DC):	West !Xoon, ’N|ohan, N|amani†, East !Xoon, Kakia†
Lower Nossob (?DC):	|’Auni†, |Haasi†
!Ui:	N||ng (DC); ǂUngkue†; |Xam† (DC); ||Xegwi†

Note: DC = dialect cluster, † = presumed extinct

Fig. 4 Classification of non-Bantu lineages in southern Africa (based on Güldemann 2006).

Current research suggests that Kua is closely related to G‖ana, and not a language in a completely separate subgroup (Chebanne and Collins 2012). Moreover, Chebanne and Collins (2012) contends that Tsua, Cua, Shua may be more closely related to G‖ana than G‖ana is related to Naro, and Naro may be more closely related to Khoekhoe than to G‖ana. In light of these claims stemming from recent fieldwork, it's likely the genetic relationships of the Kalahari Khoe East and West languages need continued reassessment.

UNESCO's Atlas of the World's Languages in Danger designates Tsua as 'definitely endangered,' specifying a total of approximately 900 speakers living in two settlements east and southeast of Serowe, although one must consider how many people are still *fluent* in a language versus residing in a settlement.[5] It may be the case that the number of fluent Tsua speakers is much lower than 900, a number that could be verified through a future sociolinguistic survey of the area. The framework for determining the vitality of a language was developed by an international group of linguists at the behest of UNESCO in 2002 and 2003. The Ad Hoc Expert Group on Endangered Languages established the following nine factors in assessing the sociolinguistic situation of a language.[6]

5 http://www.unesco.org/new/en/culture/themes/endangered-languages/atlas-of-languages-in-danger/

6 The 9 factors are *not* listed in order of importance. The Ad Hoc Expert Group suggests taking the factors into account together depending on the local context.

(2) The UNESCO Ad Hoc Expert Group language endangerment criteria

a. Intergenerational Language Transmission
b. Absolute Number of Speakers
c. Proportion of Speakers within the Total Population
d. Shifts in Domains of Language Use
e. Response to New Domains and Media
f. Availability of Materials for Language Education and Literacy
g. Governmental and Institutional Language Attitudes and Policies, Including Official Status and Use
h. Community Members' Attitudes towards Their Own Language
i. Type and Quality of Documentation

Perhaps the most striking evidence of Tsua's endangerment is the advanced age of the remaining speakers and the very low intergenerational language transmission rates, which is the most commonly used factor in evaluating language vitality (Fishman 1991). The results of a survey conducted in November 1997 indicate the Kua and Tshwa languages as endangered varieties, with language use shifting to Tswana [Setswana] (Hasselbring, Segatlhe and Munch 2000:124). Survey data on language use by the Tsua population itself is unavailable; therefore, data from the Tshwa subgroup, which includes Tsua, will be discussed. The Kua and Tshwa language use by children to parents numbers (to the left) and percentages (to the right) from the survey are presented in Table 4 (Hasselbring, Segatlhe and Munch 2000:152).

Table 4 Language spoken by children to parents

Language from Parents	Own language only	Tswana only	Own language + Tswana	Other language only	Total
Kua	7 36.8%	9 47.4%	3 15.8%		19 100%
Tshwa subgroup	4 8.7%	36 78.2%	5 10.9%	1 2.2%	46 100%

The Kua language use situation seems to be somewhat less vulnerable among the youth compared to the Tshwa subgroup languages: 36.8% spoke Kua exclusively to their parents, while 47.4% spoke Tswana only at the time of the survey. Out of 46 Tshwa (including Tsua) participants, only 4 stated that their children spoke a Tshwa language to them exclusively (8.7%). 36 out of 46 participants (78.2%) indicated that their children spoke Tswana [Setswana] exclusively to them, even when the parents were speaking a Tshwa language to the children. Supplemental survey data adopted from Hasselbring and Segatlhe (2000:83) of the language spoken by children when playing are in Table 5.

Table 5 Language spoken by children when playing

Language from Parents	Own language only	Tswana only	Own language + Tswana	Tswana + regional language	Total
Kua	13 39%	12 36%	8 24%	0 0%	33 100%
Tshwa subgroup	2 4%	37 76%	9 18%	1 2%	49 100%

The Kua youth spoke their own language while playing more than the Tshwa-speaking children at 39% for Kua only versus 36% for Tswana only. A paltry 4% of Tshwa-speaking youth used their own language while playing as opposed to 76% who spoke

Tswana only. The Tshwa statistics can be partially attributed to the exclusive use of Setswana and English as the medium of education at all school levels, including preschool. In some places, teachers insist, even during playtime or break, that students speak the school languages to the exclusion of their mother tongues (Chebanne 2002b:205). It is not surprising, then, that Tsua is considered endangered by UNESCO and perhaps should be designated as critically-endangered, especially if the trend of low intergenerational language transmission rates have continued unabated since the survey was conducted over 15 years ago.

1.5 Sound changes involving click replacement

The dwindling use of some Khoisan languages has led to the rapid erosion or simplification of structures with respect to their phonological systems, morphological forms and lexicons according to some linguists (e.g. Batibo 2002:197). In terms of their phonological systems, many show evidence of systematic changes termed click replacement. I have opted to use the concept "click replacement" to mean the substitution of a non-click sound segment in place of a click segment following Traill (1986).[7]

---

7 With respect to the concepts "click replacement" and "click loss", Traill and Vossen (1997:24) diverges from Traill (1986) by stating: "We apply the concept of "click loss" to languages where a click...has been replaced with a non-click consonant and "click replacement" to cases in which one of the [click] influxes...has been replaced by another [click] influx...Obviously the term "replacement" is not altogether felicitous because both "loss" and "replacement" involve replacement of one sound by another." I have elected to use "click replacement" as the substitution of a non-click sound segment in place of a click segment following Traill (1986) and discarded the term "click loss" unless it refers to Traill and Vossen's (1997) paper.

Attributing click replacement as part of a process of language shift and death stands in stark contrast to earlier hypotheses predicated upon the intrinsic linguistic instability of clicks in *all* Khoisan languages, rather than as an index of language death for some.

The earliest unpublished vocabulary of an eastern dialect of the Central Khoisan languages was collected by the explorer David Livingstone as he traveled through the Ngwato territory via Serowe and north of the Makarikari Pan [Makgadikgadi Salt Pans] around 1851 (Maingard 1961:114). A few of the relevant examples demonstrate the language under investigation by Livingstone had replaced clicks with non-click consonants in a pattern typical of eastern Khoisan (Traill and Vossen 1997:25). Moreover, Dornan (1911) supplies an abbreviated description of the language spoken by the Hiechware in eastern Botswana, with examples showing evidence of click replacement (ibid).

Several scholars claimed that clicks were intrinsically unstable, giving reasons from phonetic license to imprecise "Bushman" speech to phonological entropy of the primitive mind (see, for example, Doke 1937; Maingard 1937 and Bleek 1939-40). These claims were at odds with mounting evidence of click *stability* in other Khoisan languages, e.g. Nama, Zhu and !Xóõ (Beach 1938; Snyman 1975; and Traill 1985, respectively). It was not until Traill (1986) that the systematic nature of click replacement was carefully studied. He found an implicational relationship in which alveolar and palatal clicks were most often replaced by their velar and palatal non-click oral or nasal stop counterparts, with loss of the palatal clicks implying loss of the alveolar clicks. Citing Dressler (1972), Traill agreed with the observation pointing to decaying languages as the target of click replacement, most often occurring in disintegrating communities (Traill 1986:303).

One reason for click replacement is because click consonants are often viewed as "peculiar" by the non-Khoisan speaking groups, sentiments which are in part fueled by the historically low status of the Khoisan populace (Wilmsen and Vossen 1990:22-24). Thus, click replacement may have come about by the desire for these groups to distance themselves from an association with a permanent underclass while attempting to give their languages increased legitimacy (ibid). The situation is, however, complex. The intricate link between click usage and socioeconomic status is aptly elucidated in the following passage:

> The southern African situation is, of course, complicated by the presence of Bantu languages (Sindebele, Kizulu, Isixhosa, and Sesotho) which are respected means of communication in spite of their large phonetic-phonological inventories of clicks. These languages may have retained respectability because they are spoken by peoples who were hegemonically dominant in their geographical areas well into the nineteenth century and because Khoisan-speakers, who may in the later nineteenth century have presented a subdued minority, low-status contrast there, had by then long been absorbed into dominant Bantu click-speakers in those areas. Only when Bantu and Khoisan click-bearing languages coexist geographically in a context of hierarchical dominancy would a differentiation in their respective social statuses be expected (Wilmsen and Vossen 1990:22-23).

The passage's essence is applicable to the Tsua: they have been the subjects of the dominant Bantu communities in a more populated area of eastern Botswana (Fig. 2) in which these communities have coexisted for a long time, living in an area of close economic relations with the agro-pastoralists (Vossen 1991:372; Traill and Vossen 1997:51). In the next chapter, I will supply examples of click replacement, which support Traill's claim of an implicational relationship between deleted clicks. Additionally, vis-à-vis click replacement, Traill and Vossen (1997:25) claims: "...tone plays no part in the phonetic processes we are dealing with." While the statement may be generally true, I will argue these processes obscure the Tsua tonal system, as exceptions to tonal depression are best explained by looking at comparative data, taking into account click replacement. I believe future research on Khoisan tonal systems found in eastern Botswana would greatly benefit from this approach.

1.6 Researching the Kua-Tsua dialect continuum

It is important to research the Kua-Tsua dialect continuum given the language endangerment situation explored in §1.4 before the last remaining speakers pass away, because "documenting differences among the world's most disparate languages is of central importance to the field of linguistics and to the language community's heritage" (Whalen et al. 2011:40). Little is known about the languages of the Khoisan communities because few have been systematically studied (Köhler 1971; Tanaka 1980; Güldemann and Vossen 2000; Chebanne 2003a). In particular, the Tshwa subgroup languages have not attracted the same interest as other Khoisan languages (Andersson and Janson 1997:138). Tsua has a typologically rare tonal depression pattern that has been overlooked until now.

It may be that other Tshwa subgroup languages have the same interesting pattern. Tsua tonal depression is described in detail in Chapter 4. Next, the documentation methodology employed during the summers of 2012 and 2013 for eliciting the language data referenced in this dissertation will be discussed, which was essential for studying the lexicon and eliciting phrases and sentences.

1.7 Data collection methodology

The dissertation field research was divided into two parts: Phase 1 (summer 2012) and Phase 2 (summer 2013). I worked closely with three female Tsua consultants during Phase 1, designating them as speakers S, B and M in the elicitation sessions, and the same three consultants plus one male Kua consultant, designated as speaker O, for Phase 2, relying heavily on the fieldwork output found in an earlier version of Chebanne and Collins (2014). Information about each consultant is summarized in Table 6.

Table 6 Consultant information

Speaker	Reported Age (2012)	Home Village	Spoken Languages
S	78	Moralane	Tsua; Setswana
B	57	Moralane	Tsua; Setswana
M	80	Moralane	Tsua; Setswana
O	64	Khekhenye	Kua; Setswana; Shekgalagadi

In July 2012, the three female Tsua consultants were transported 230 kilometers from their home village Moralane to the Phakalane neighborhood in Gaborone to begin

Phase 1. One reason why the research was not conducted in their village was to minimize noise and distractions (e.g. wandering children, curious community members, etc.) while avoiding the frequently reported problem of consultants missing data collection appointments (Bowern 2008:135). All elicitation sessions were conducted in Setswana and took place in a quiet living room with a short reverberation time. I used the Zoom H4n handheld audio recording device to capture lexical items, phrases and sentences produced by each consultant, proceeding in order according to the steps listed below.

(3) Data elicitation procedure

a. The lexical item, phrase or sentence was presented to the consultants in Setswana by a native speaker and an appropriate context given.
b. The first consultant spoke the Tsua translation three times into the Zoom recorder, holding it at a distance of 6-8 inches (phrases and sentences were spoken once to avoid consultant fatigue).
c. The first consultant handed the recorder to the second, who then spoke the translation into the Zoom.
d. The third consultant did the same.
e. Breaks were taken once every hour or so.[8]

Language consultants have a tendency to get further away from the microphone or recording device if it is mounted on a stand as time goes by. In my experience, particularly from working with the Tsua over the course of two summers, having them hold the Zoom

---

[8] Frequent breaks were essential during the sessions since the elderly consultants would tire easily. It should also be noted here that the consultant recording order was varied during sessions to dissuade them from potentially relying on mimicry.

and passing it along was effective in maintaining the high quality of the recordings by keeping the microphone distance relatively consistent. Even when they became more relaxed in their chairs during the session, the consultants still kept the Zoom's stereo microphone at an appropriate distance.

The Zoom was interfaced with a laptop computer equipped with the phonetic software Praat version 5.1.25 (Boersma and Weenink 2010) via the USB port. The advantage of this setup was the laptop powered the Zoom through the USB port; batteries were not necessary. A second advantage was in situations when the power went out: rolling blackouts in Botswana were not uncommon at the time. The elicitation sessions continued in spite of the lack of power as long as the laptop battery lasted, which was usually a couple of hours. The digital output of the Zoom's high quality, stereo condenser microphones fed Praat's recording function set for a 44.1 kHz sampling rate at 16-bit quantization in .wav format, allowing me to segment each consultant's utterance for the duration of all elicitation sessions by pressing record and stop throughout. This approach was crucial for efficiency: without on-the-fly segmentation, it would have taken many hours to edit the audio files later, taking valuable time away from transcription and acoustic analysis.

Phase 2 occurred in June-July 2013, with all elicitation sessions taking place in a small, two-room building in the Mogoditshane neighborhood of the Greater Gaborone area. While the corrugated roof dispersed the sound waves due to its curvy surface, the first challenge was to reduce the new location's standing waves in the recording room because an audible slap could be heard when I clapped my hands in the middle of the room. The culprits were two highly reflective walls facing each other. I decided to hang a shipping blanket from the rafters to cover one wall and prop up a foam mattress on a sofa to cover the other. The objectionable echo disappeared and the elicitation sessions commenced from there starting with the Kua consultant at the beginning of Phase 2. Subsequent to his departure, the same three Tsua consultants from Phase 1 agreed to return for Phase 2 after a meeting in their village.

There were two technical issues that arose during Phase 1 which needed to be addressed before the start of Phase 2. First, the lack of an adequate set of sonically precise headphones for checking transcriptions after elicitation sessions had ended for the day was problematic. I purchased a pair of Sony MDR-7506 professional, large diaphragm headphones for Phase 2. Second, the Zoom recorder was sensitive to bursts of air generated by plosives or aspiration produced at close proximity to the condenser microphone diaphragms. The result was occasional signal distortion or low frequency "popping". To mitigate the signal distortion from air bursts, I employed the Røde Dead Kitten fur windscreen to cover the Zoom's foam windscreen. Fur windscreens are widely used in production audio recording as a way to counter the detrimental effect of wind in outdoor shoots while accurately capturing the sonic detail of dialog. The Røde windscreen did an excellent job of controlling the low frequency popping so that the fine phonetic detail of the Tsua language could be recorded for analysis.

1.8 Summary

This chapter has presented the societal context for Tsua and its relevance to click replacement. Understanding Tsua's sociolinguistic state of affairs is important because it directly impacts the phonological structure of the language, a circumstance often reported in the eastern Central Khoisan languages. What has not been previously reported is the effect click replacement has on tone in eastern Central Khoisan languages. It is argued in subsequent chapters that nasal consonants and nasalized clicks that have been replaced by voiced consonants in Tsua are not followed by depressed H tones. First, the consonant and vowel inventories are established in Chapter 2 along with a description of Tsua phonotactics.

CHAPTER 2

2.1 Introduction

The starting point for addressing the interaction between segments and tones and the nature of conditioning environments in Tsua is to establish the consonant and vowel inventories. The subsections that follow begin by outlining a classification of the non-click consonant inventory with examples from the Tsua data, and subsequent subsections present the click consonant and vowel inventories. Comparisons will be made to the Tshwa subgroup consonant and vowel inventories in Vossen (1997, 2013b). The current Tsua data set comprises ~1,300 roots. One goal is to allow a clearer picture to emerge vis-à-vis the less frequently occurring segments in the lexicon. For instance, there are pharyngealized vowels that have not been previously reported by Vossen.[9]

Establishing the segmental inventory will lead into a discussion of the phonotactic constraints on Tsua lexical morphemes, which includes three major topics: (i) the distribution of consonants in C_1 and C_2 positions, (ii) the phonotactics of consonants in the C_2 position with respect to the vowels in V_1 and/or V_2 position, and (iii) the restrictions on vowels in V_1 position in $C_1V_1C_2V_2$ structures, known as the Khoisan-specific Back Vowel Constraint (BVC). These topics are descriptive in nature and important for the

---

9 The identical analyses in Vossen (1997) and Vossen (2013b) are based on the same data set (Rainer Vossen, personal communication).

interpretation of the consonant-tone interactions put forth in Chapters 4 and 5.

The Tsua click consonants are especially notable due to a number of inventory gaps, especially concerning the alveolar clicks. It may be that the process of click replacement has *asymmetrically* reduced the Tsua inventory, a process which has been reported in Kalahari Khoe East languages as reviewed in Chapter 1 (see also Vossen 2013b:72). I use the term "asymmetrical" to refer to the phenomenon where certain click place/manner series possess the full set of accompaniments while others do not. For example, the dental and lateral clicks have contrastive aspiration but the alveolar click does not in the current data set. These phonological changes have been observed for quite some time, as Traill (1978) writes: "One such detail involves the erosion of clicks that takes place as one moves eastwards. Kua retains more clicks or reflexes of them than its neighbours" (p. 258). It is likely that the Tsua click inventory gaps emerged as a consequence of the sociolinguistic situation and extensive linguistic contact in eastern Botswana with languages such as Setswana and Ikalanga as discussed in Chapter 1.

2.2 Segmental inventories

2.2.1 Non-click consonants

This subsection presents the non-click consonant inventory derived from the field research results in Mathes and Chebanne (2013). Lexical morphemes are referred to as "roots" in keeping with a Khoe terminological tradition since Beach (1938). Transcriptions were produced with the help of spectrograms and waveforms rendered in Praat, supplemented by cognate transcriptions from: (a) an earlier edition of Nakagawa et al.'s (2014) trilingual G|ui-English-Japanese dictionary; (b) an earlier version of the Kua lexicon in Chebanne and Collins (2014); and (c) Phase 2 field research results on Kua during the summer of 2013. Several hundred phrases and sentences were consulted when necessary to aid in the transcriptions.

In keeping with inventories across Khoisan languages, the non-click as well as click consonant inventories are large and complex. It was not possible to conduct static palatography and linguography during fieldwork to help determine place of articulation for the Tsua segments. Therefore, the elicitation sessions required extra careful work with the consultants to ascertain the place of articulation for some of the most challenging segments such as the ejective /tsqχ'/. This cluster commences with the release burst of the affricate [ts], followed by the burst from the fricated [qχ'], a cluster also found in G|ui (see Nakagawa 2006:130-131 for a discussion).

Nakagawa (2006) is a highly detailed description of the articulatory and acoustic properties of the segment inventory in G|ui. This resource was crucial for working out the Tsua inventory.[10] Since Nakagawa (2006) and subsequent papers contain precise and detailed descriptions of the acoustic and articulatory characteristics of G|ui segments with an abundance of spectrograms, waveforms and palatograms, the interested reader may

[10] Many thanks to Hirosi Nakagawa for listening to numerous Tsua sound files and suggesting appropriate transcriptions for some of the non-clicks and clicks in Tables 7 and 10, respectively.

consult his excellent work. These details are not repeated here.

Table 7 provides the non-click consonant inventory and Table 8 gives examples of the corresponding consonants in root-initial position. Depressor consonants are indicated in red. Tonal depression is analyzed in Chapters 4 and 5.

Table 7 Tsua non-click consonant inventory (n = 40). Depressors are indicated in red.

	Labial	**Alveolar**	**Palatal**	**Velar**	**Uvular**	**Glottal**
voiceless stops	p	t ts	c	k	q	ʔ
voiced stops	b	d dz	ɟ	g	ɢ	
aspirated stops	pʰ	tʰ tsʰ	cʰ	kʰ	qʰ	
ejectives		t' ts' tsqχ'	c' cqχ'	k'	qχ'	
stop + /χ/		tχ tsχ	cχ			
nasals	m	n	ɲ	ŋ		
fricatives		s			χ	h
non-nasal sonorants	w	r l[11]	j			

Table 8 Examples of non-click consonants in root-initial position

	Labial	**Alveolar**	**Palatal**	**Velar**	**Uvular**	**Glottal**
voiceless stops	pee 'to jump'	turu 'rat' tsao 'tail'	coo 'heart'	kae 'to tie'	qãã 'Kalahari cucumber'	ʔũũ 'to hunt'
voiced stops	boo 'axe'	dum 'voice'	ɟue 'to deny'	gui 'string; rope'	ɢoba 'hornbill bird'	

11 [r] and [l] are only attested in root-medial position, a common restriction in Khoisan languages. The set of permissable, root-medial consonants is described in §2.3.1.

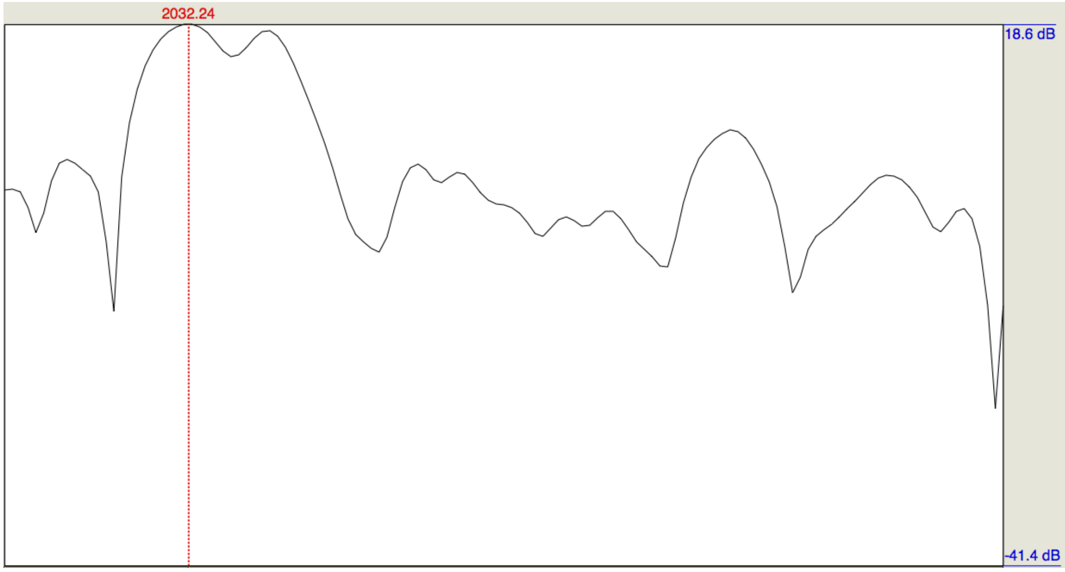
		dzam 'whip'					
aspirate stops	p^haa ta 'a hide to sleep on'	t^hebe 'to burn' ts^huu 'charcoal'	c^huri 'tree stump'	k^han 'to crawl'	q^han 'news'	
ejectives		t’ana 'to hatch' ts’ãũ 'to do' tsqχ’ae 'sky; blue; green'	c’oa 'to come out' cqχ’ae 'to spit'	k’are 'to cross a road'	qχ’am 'mouth; beak'	
stop + /χ/		tχara 'torn' tsχara 'to hurry'	cχaro 'to plead'			
nasals	mũũ 'to see'	nai nai 'crowned plover'	ɲũũ 'to eat'	ŋole 'to stir'		
fricatives		sam 'to suckle'			χun 'to grind'	hube 'to step'
non-nasal sonorants	woo tsuri 'to play'	dzara 'bird' tsala 'foolish'	juu 'skin spot'			

The current data set has seven egressive consonants not reported in Vossen (1997:127, 2013b:74) — voiceless aspirates /p^h/, /q^h/; the voiced uvular stop /ɢ/; voiceless ejectives /tsqχ’/, /cqχ’/, /qχ’/; and the voiceless uvular fricative /χ/. The voiceless bilabial aspirate /p^h/ is marginal in that it only occurs in the root /p^haa ta/ 'a hide to sleep on' and Vossen reports the voiceless velar fricative /x/ as opposed to the current inquiry's voiceless uvular fricative /χ/ for corresponding roots.

The egressives reported here but not in Vossen's analysis pattern together by having a cluster final uvular place of articulation for the instances concerning ejectives. Evidence for the uvular place of articulation can be observed acoustically in tokens for the minimal pair /kãã/ 'to put in smoothly' versus /qãã/ 'Kalahari cucumber', both with the HL tonal contour. Tone is covered in the next chapter. Fig. 5 displays example spectra of [k] and [q] burst frequencies for speaker M. We would expect the highest concentration of burst energy to be lower for the voiceless uvular stop [q] compared to the voiceless velar stop [k] because of the longer portion of the vocal tract in front of the uvular stop's more posterior constriction location (Ladefoged and Maddieson 1996:36; Reetz and Jongman 2009:193), an expectation confirmed by the mean burst frequency measurements in Table 9.

Table 9 Mean velar and uvular burst frequencies of [k] and [q] for /kãã/ 'to put in smoothly' versus /qãã/ 'Kalahari cucumber' (*n* = 3 repetitions per word for each speaker). Burst frequencies are measured via the spectral slice function in Praat at the point of highest burst amplitude.

Speaker	**[k]**	**[q]**
S	1251 Hz	972 Hz
B	1608 Hz	1268 Hz
M	2367 Hz	1237 Hz



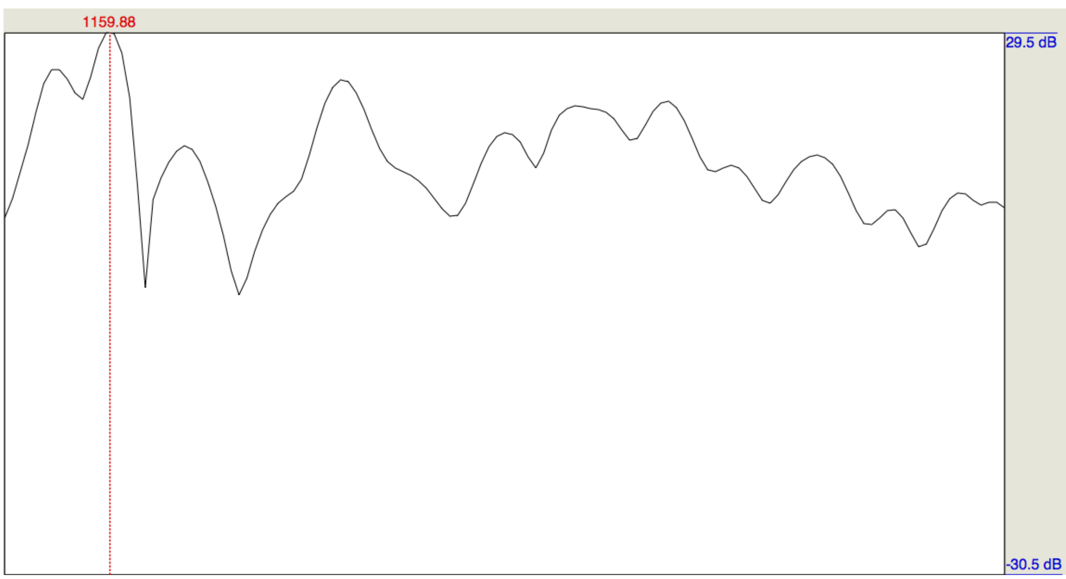


Fig. 5 Spectral slice plots of /kãã/ 'to put in smoothly' (top) versus /qãã/ 'Kalahari cucumber' (bottom) as produced by speaker M. The concentration of burst energy around 1159 Hz for [q] (bottom) is lower compared to 2032 Hz for [k] (top). Burst frequencies are measured via the spectral slice function in Praat at the point of highest burst amplitude.

2.2.2 Click consonants

The click consonant inventory presented in this section follows Ladefoged and Traill (1994) by using the terms “click type” and “accompaniment” to describe the clicks. Click production proceeds as a sequence that starts with the formation of an anterior and posterior closure in the oral cavity by the tongue. The tongue body is then lowered, creating a cavity of rarefied air between the two closures, and once the anterior closure is released, atmospheric air rushes into the mouth followed by the posterior closure release, a spatio-temporal organization of gestures that causes the salient click sound.

Fig. 6 exemplifies the basic click production process — the location and release manner of the anterior closure determines the click type and the velar or uvular posterior closure release is referred to as the accompaniment here (but see Miller et al. 2009 for

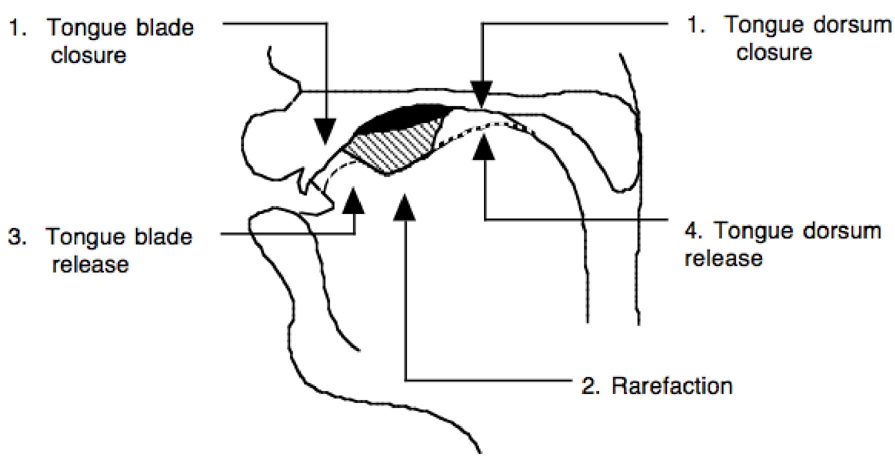


Fig. 6 Basic click production mechanism in which the striped lines represent the cavity of rarefied air after the tongue is lowered from the position denoted by the solid black area (taken from Thomas-Vilakati 2010:10, based on Ladefoged and Maddieson 1996).

arguments against the term accompaniment). As observed in Tables 10 and 11, the accompaniment consists of (a) voicing, aspiration and ejection timed with the posterior closure, (b) posterior closure place of articulation, (c) the oro-nasal setting, and (d) the plosion or friction in the uvula or glottis following the release of the anterior closure (Nakagawa 2006:138). Tsua has four click types symbolized as: [|] dental; [ǂ] palatal; [!] alveolar; and [ǁ] lateral, with voicing and nasality transcribed before the click symbol, a convention commonly employed in Khoisan linguistics. For instance, [ɡ|] represents the voiced dental click and [ŋ|] is the nasalized dental click. A [k] before the click symbol to indicate voicelessness is not used in the transcriptions to reduce redundancy and to make the forthcoming Tsua click replacement patterns easier to discern.

Table 10 Tsua click consonant inventory (*n* = 34). Depressors are indicated in red.

	Dental	Palatal	Alveolar	Lateral
voiceless	|	ǂ	!	ǁ
voiced	ɡ|	ɡǂ		ɡǁ
aspirated	|ʰ	ǂʰ		ǁʰ
ejective	|’			ǁ’
glottalized	|ʔ	ǂʔ	!ʔ	ǁʔ
clusters	|χ			ǁχ
	|qχ’		!qχ’	ǁqχ’
	|q	ǂq	!q	ǁq
	|ɢ			ǁɢ
	|qʰ	ǂqʰ		ǁqʰ
nasals	ŋ|	ŋǂ	ŋ!	ŋǁ

Table 11 Examples of click consonants in root-initial position

	Dental	**Palatal**	**Alveolar**	**Lateral**
plain	ǀam 'sun'	ǂea 'valley'	!ao 'tall'	ǁaa 'black korhaan'
voiced	gǀao 'caracal'	gǂãã 'to deceive'		gǁuu 'to build a fire'
aspirated	ǀ^{h}ãã 'ant species'	ǂ^{h}ere 'pimple'		ǁ^{h}au 'animal'
ejective	ǀ'ũũ 'to break'			ǁ'ui 'to cough'
glottalized	ǀʔũũ 'hair; body hair'	ǂʔujo 'mamba'	!ʔuu 'white'	ǁʔaa 'bat-eared fox'
clusters	ǀχoo 'gemsbok'			ǁχãã 'light'
	ǀqχ'uri 'dirt'		!qχ'ao 'neck'	ǁqχ'ari 'scorpion'
	ǀqãã ti 'Sasi people'	ǂqai 'springbok'	!qãũ 'cheetah'	ǁqore 'claw; fingernail'
	ǀɢoe 'yellow mongoose'			ǁɢai 'maggot'
	ǀq^{h}unu 'walk in dry bush'	ǂq^{h}uru 'dust'		ǁq^{h}am 'spider web silk'
nasals	ŋǀii na 'earring'	ŋǂau 'to chase'	ŋ!au 'frog'	ŋǁau 'to spray liquid'

Khoisanists may deem only one type of aspiration in the tables above as noteworthy because previous work on Khoisan languages has found two different types of click consonant aspiration. Fig. 7 illustrates the two types in Nama as reported in Traill (1991). In the first aspiration type, voiceless aspirated clicks like [ǁ^h] have an absence of any nasal airflow before, during and after the click. The air pressure rises in the pharynx leading up to the anterior closure release with a rapid acceleration of oral airflow, co-occurring with an audible release of the posterior closure. In the second aspiration type, click with delayed aspiration, nasal airflow commences before the click closure and is sustained during and

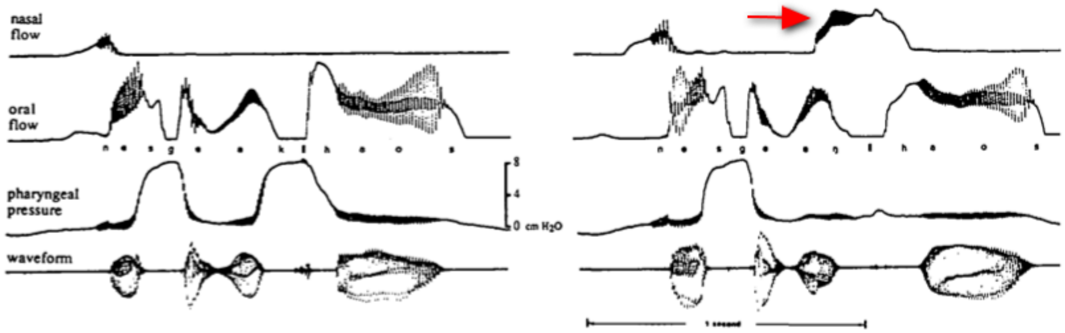


Fig. 7 Aerodynamic measurements and waveforms of two types of aspiration in Nama lateral clicks — aspiration in the Nama word *kǁhaos* 'striking' that lacks nasal airflow during the lateral click (to the left) and delayed aspiration co-occurring with nasal venting in the Nama word *ŋǁhaos* 'traditional cooking place' (to the right). The red arrow points to the nasal airflow during the lateral delayed aspirated click (from Traill 1991).

after click release. The air pressure is vented through the nasal cavity. The red arrow in Fig. 7 points to aerodynamic evidence of nasal venting during the delayed aspirated click. Due to this passive nasal venting, the oral airflow upon click release is far slower compared to the first aspiration type. As a consequence, aspiration slowly increases in amplitude over time for delayed aspirated clicks, as opposed to reaching full amplitude right after click release as in aspirated clicks, quickly followed by an *inaudible* posterior closure release.

The current data set does not provide conclusive evidence of delayed aspirated clicks in the Tsua inventory, although Chebanne and Collins (2014) lists them as part of the Kua inventory. Honken (2008:188) states that most Central Khoisan languages only have the type of aspiration found in aspirated non-clicks, while the contrast between aspiration and delayed aspiration is common in Non-Central Khoisan languages. There is one example in which a Kua cognate has a root-initial delayed aspirated lateral click, transcribed as [ǁ'h] in Chebanne and Collins (2014), that corresponds with a Tsua root-initial nasalized click, evidence of a potential sound change involving the passive nasal venting of the Kua delayed aspirated click being produced as a nasalized click in Tsua.

(4) Kua delayed aspirated click to Tsua nasal click correspondence

Kua	**Tsua**	**Gloss**
ǁ'haˤo	ŋǁaˤo	'to throw nimbly'

It is tempting to conclude that delayed aspiration was a casualty of click replacement (i.e. a delayed aspirated click replaced by a nasalized click in Tsua) brought about in part by the sociolinguistic factors from Chapter 1, but this would involve a change in the conventional view on click replacement and click loss, as Traill and Vossen (1997:24) states: "Neither process involves the accompaniment." The expected click replacement patterns are described in §2.2.3.

Honken (2008:188-192) gives examples of a correspondence between delayed aspirated clicks in Khoekhoe and nasal clicks in Naro, G|ui and ||Ani. Nevertheless, the one example in (4) does not supply ample evidence of a sound change that explains the lack of Tsua clicks with delayed aspiration but is included here as one possibility to be explored in future field research. All that can be said at the moment is there is no evidence for or against the existence of delayed aspiration in the Tsua click consonant inventory, so this issue must be left for future study.

Another general property of the Tsua click consonant inventory revealed by Tables 10 and 11 is the palatal and alveolar click types have accompaniment gaps (the blank cells) while the dental and lateral click types have the full set of accompaniments. One question is how such an asymmetrical inventory could be possible. These gaps are considered in the next subsection framed within the context of Traill's observations regarding click replacement. It is particularly crucial to notice in Table 11 that the click consonants only occur in initial position and not in word-medial or word-final position, an observation that will be important for Chapter 4 as Tsua tonal depression targets the initial H tone in a root.

2.2.3 Traill's click replacement patterns

Traill (1986) marked the first attempt to organize the facts about click replacement into a systematic phenomenon using personally collected data from the Khoe languages. Traill makes the observation that certain Khoe languages but not all replace clicks with non-clicks, and when the phenomenon occurs, palatal [ǂ] and alveolar [!] clicks may be replaced by cognate palatal and velar stops, respectively. In addition, there are Khoe varieties that replace at least some alveolar [!] clicks while preserving all palatal [ǂ] clicks, or varieties that replace both, but never a language that has replaced [ǂ] while keeping [!]. For instance, Nama, Naro and G|ui are conservative in that they preserve all of their clicks as phonetically and phonologically stable elements at one end of the spectrum, but Shua, Teti and Tyire are non-conservative varieties that replace both alveolar and palatal clicks at the other end of the spectrum. The Khoe click-series systems can be summarized as in (5) in diachronic succession according to Traill (1986:305).

(5) Khoe click-series systems

Stage 1	**Stage 2**	**Stage 3**
ǀ ǁ ǂ !	ǀ ǁ ǂ k	ǀ ǁ c k

but

* ǀ ǁ c !

Recall that "click type" is defined as a click plus the full set of accompaniments in a given language. For example, the "alveolar click type" consists of the voiced alveolar click [ɡ!], the nasal alveolar click [ŋ!], the aspirated alveolar click [!ʰ], and so on. A Khoe language in Stage 1 preserves all four click types. A Stage 2 language has replaced at least some alveolar clicks with velar stops. For instance, the aspirated alveolar click [!ʰ] has been replaced with the velar stop [k] and is no longer in a language's inventory, although alveolar clicks with other accompaniments may still be present. A Stage 3 language has replaced at least some alveolar clicks with velar stops and at least some palatal clicks with palatal stops. Dental and lateral click types are always preserved. There is not a situation in which at least some palatal clicks have been replaced by palatal stops while retaining all alveolar clicks.

One question that immediately arises concerns why the palatal [ǂ] and alveolar [!]

clicks are the target of replacement. Traill and Vossen (1997) appeals to the single factor "weakening" of what they categorize as the "abrupt" clicks (i.e., the palatal and alveolar clicks) as the phonetic basis for these sound changes, arguing that it provides a unitary account. Acoustic data reveal the peak amplitudes, spectra and intensities of weakened palatal and alveolar clicks are less like damped impulses, being noisier and more diffuse, compared to the compact peak of the unweakened forms (p. 46-48). These weakened clicks are correlated with languages that have click replacement (e.g. G‖ana) while the unweakened versions most commonly occur in languages that do not (e.g. G|ui).

The mechanism underlying this acoustic attenuation is articulatory undershoot of the abrupt clicks, in which the degree of rarefaction is reduced due to the formation of a larger click cavity prior to click release, a lingual gesture that compromises the compact peak normally associated with these clicks. Traill and Vossen (1997:49) provides evidence of articulatory undershoot in palatographic form to support their argument, concluding that the reduction of articulatory complexity ends with replacement of a click by a non-click.

The click consonant inventory gaps illustrated in Tables 10 and 11 are a reflection of the extent of Tsua click replacement, with some click consonants being replaced by non-clicks in accord with Stage 3. The clicks absent from Tsua are: g!, !^{h}, ǂ', !', ǂχ, !χ, ǂqχ', ǂɢ, !ɢ, and !q^h. Further evidence of Stage 3 click replacement in Tsua comes in comparative form with vocabulary based on Traill (1986:305), seen in Table 12.[12]

---

12 Traill's transcription system is somewhat different from the system used in this dissertation. However, the vocabulary is still illustrative of the click replacement patterns under discussion.

Table 12 Click replacement comparative data for [!] and [ǂ] clicks

Stage 1	Stage 2	Stage 3		Gloss
G|ui[1]	G‖ana[1]	Ts’ixa[1]	Tsua[2]	
!naro	ŋaro	ŋgaro	garo	'chameleon'
!ganee	ganee	gaani	gani	'chin'
!hae	khae	khae	kʰae	'pierce'
!xan	kxan	xan	χan	'sew'
ǂnu	ǂnuu	ɲɟuu	ɟuu	'black'
ǂgoa	ǂgoa	ɟua	ɟoa	'ash'
ǂxoa	ǂxoa	cxoa	cχoa	'elephant'
ǂhuni	ǂhuni	chuni	cʰuni	'elbow'

[1]Traill (1986:305) [2]Mathes and Chebanne (2013)

Table 12 shows a few examples of click replacement patterns for alveolar and palatal clicks from Traill (1986:305). Dental and lateral click types have not been replaced in these languages for reasons that have already been discussed. G|ui has preserved all four click types, thereby being a Stage 1 language. G‖ana is a Stage 2 language because at least some alveolar clicks have been replaced with velar stops. For instance, [!ganee] 'chin' in G|ui is the clickless [ganee] in G‖ana. Ts’ixa is a Stage 3 language since at least some alveolar clicks have been replaced with velar stops and at least some palatal clicks with palatal stops in its lexicon. The Tsua data are most similar to Ts’ixa: at least some alveolar clicks have been replaced with velar stops and at least some palatal clicks with palatal stops in its lexicon. Table 12 gives [!xan] 'sew' in G|ui versus the clickless [χan] in Tsua.

The click replacement comparative data in Table 12 gives insight into the lack of certain clicks in the Tsua inventory, namely, ɡ!, !ʰ, !χ, and ǂχ, as they have undergone replacement by ɡ, k^h, χ, and cχ, respectively (see Traill and Vossen 1997:29 for identical findings in other Khoe varieties). There are extant clicks in the inventory that are replaced in some cases and not in others. [ŋ!abe] 'giraffe' in G|ui is [ɡabe] in Tsua, but Tsua has the alveolar click in [ŋ!au] 'small frog'. Click replacement operates on a continuum in terms of the degree to which a language has undergone a particular change, i.e., the degree to which it has spread through the lexicon (Traill 1986:306). The click replacement patterns found across Khoe languages help explain the Tsua click inventory gaps since they have been replaced by non-clicks.[13] The implications of click replacement will be pursued in Chapter 4, because this process is relevant for tonal depression exceptions in the current data set.

2.2.4 Oral and nasal vowels

Tsua has five phonemic oral vowels /i e a o u/ and three phonemic nasal vowels /ĩ ã ũ/, an inventory that is consistent with what has been reported in numerous Khoisan languages and Vossen's analysis of the Tshwa subgroup languages (1997:128, 2013b:73). Table 13 lists the 8 vowels. Table 14 illustrates the permissible V_1V_2 sequences found in mono-syllabic CV_1V_2 roots, with examples in Table 15. The V_1V_2 sequences are given for descriptive purposes. They are intended to assist future cross-Khoisan comparative studies

13 The gaps may be due to insufficient data, a situation that will be addressed in future field research.

on segmental phonotactics, as data from Kalahari Khoe East languages are often scant.

Table 13 Tsua oral and nasal vowel inventory

	Front	**Central**	**Back**
High	i, ĩ		u, ũ
Mid	e		o
Low		a, ã	

Table 14 Tsua V_1V_2 sequences arranged by initial vowel

	Front	**Central**	**Back**
High	ie, ia, ĩã		ui, ũĩ, ua, ũã, ue
Mid	ea		oe, oa
Low		ai, ãĩ, au, ãũ, ae, ao	

Table 15 Examples of Tsua V_1V_2 sequences arranged by initial vowel

	Front	**Central**	**Back**
High	ʔie 'hole; burrow' ʔia tii 'to take foot off' ʔĩã 'wind'		ʔui 'afternoon' cũĩ 'to be seated' gua 'duiker' ǁʔũã 'to return' ʔue 'to answer'
Mid	ǂea 'valley'		χoe 'to flee' cχoa 'elephant'
Low		ʔai 'take!' gãĩ 'steenbok' ʔau 'to call' tsʼãũ 'to do; to make' ʔae 'village' ǀʔao 'blood; money'	

2.2.5 Pharyngealized vowels

Pharyngealized vowels, which occur primarily in Caucasian and Khoisan languages, are articulated with the active retraction of the tongue root, causing a pharyngeal narrowing to a greater degree than what is normally found in [-ATR] vowel production (Ladefoged and Maddieson 1996:306). Pharyngealized vowels in Tsua are not reported in Vossen (1997, 2013b), but acoustic, auditory and comparative evidence point to two such vowels appearing in V_1 position: /aˤ/ and /uˤ/, although it should be noted there was inter-speaker variability in the consultants' productions.

Speakers S and M had the most salient pharyngealization, with a consistent lowering of the third formant from the vowel midpoint, often but not always accompanied by a raising of the first and second formants and noisy voice quality, acoustic properties typical of these vowels. Speaker B's production was more subtle impressionistically. Examples are /haˤma/ 'daylight' with the pharyngealized Kua cognate /haˤma/ supplying comparative evidence (Chebanne and Collins 2014), and /suˤmi/ 'to drink from the palm of the hand' supported by the G|ui cognate /sṵmi/ 'to slurp' (Nakagawa et al. 2014:223).[14]

Fig. 8 displays spectrograms of /suˤmi/ 'to drink from the palm of the hand' with the pharyngealized high back vowel /uˤ/ next to /sum/ 'shade; shadow' produced with the oral high back vowel /u/. The third formants are marked by the red arrows — auditorily, the

---

[14] Nakagawa et al.'s dictionary symbolizes pharyngealization with a tilde underneath the vowel. I have elected to use the IPA diacritic [ˤ] after the vowel.

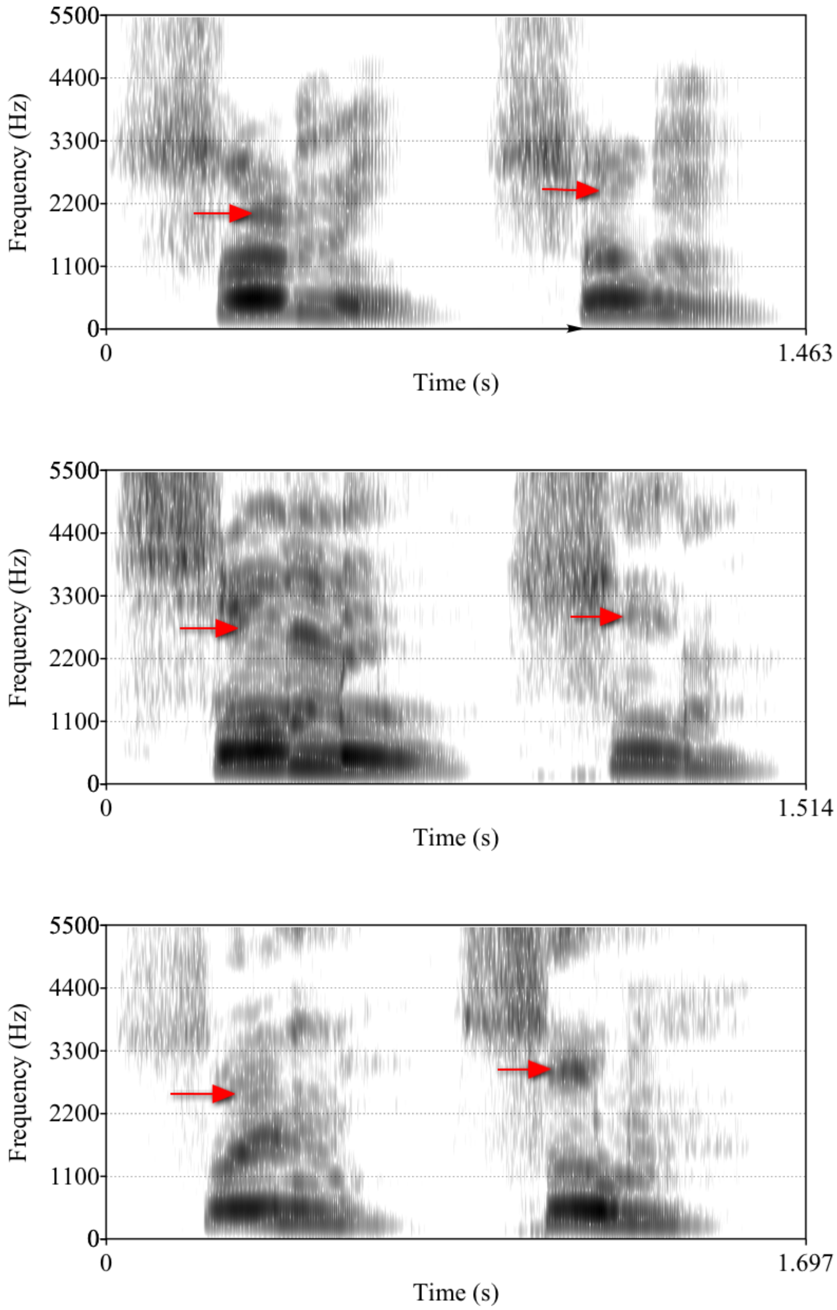


Fig. 8 Spectrograms of /suˤmi/ 'to drink from the palm of the hand' (to the left) versus /sum/ 'shade; shadow' (to the right) for speakers S, B and M from top to bottom, respectively. F3 is indicated by the red arrows.

lowered third formant gives the vowels a “choked” quality reminiscent of [-ATR] vowels but a bit more extreme as Ladefoged and Maddieson suggest (see Casali 2008 for a discussion of auditory distinctions and the descriptive literature vis-à-vis [ATR] vowels).

2.3 Tsua phonotactic generalizations

2.3.1 Canonical roots

Güldemann and Vossen (2000:107) claims that $C(C)_1$-V_1-C_2-V_2 is the template for canonical roots in Khoisan languages.

(6)	$C(C)_1$	$-V_1$	$-C_2$	$-V_2$
	strong stop	short/oral/back	sonorant or anterior voiced stop	short/oral

According to the template, a canonical root has a strong stop in C_1 position, which may be a cluster as indicated by the optional consonant (C). In Tsua, C_1 does not necessarily have to be a strong stop, as the glottal fricative can occur — for instance, /hube/ 'to step'. The vowel in V_1 position is short, oral and [+back] when preceded by a strong stop. A Tsua counterexample is /ǂere/ 'tick' with /e/ in V_1 position. C_2 is a sonorant or anterior voiced stop, yet Tsua has the fricative /s/ in C_2 position for /bisi/ 'spirit'. The final vowel is short and oral in Güldemann and Vossen's template. For ease of reference, the structure in (6) will be referred to as CVCV. There are two other common syllable structures across

Khoisan languages: CVV and CVN. However, since they are related to or derived from the CVCV structure historically and/or synchronically, the CVCV structure can be regarded as the basic structure of Khoisan lexical morphemes (Nakagawa 2010:23; Haacke 2013:96).

While there are counterexamples, the disyllabic Tsua data largely conform to the structure in (6). Monosyllabic lexical morphemes have the syllable structure of CVV, where the V_1 and V_2 positions may have identical vowel qualities. The phonemic nasal vowels are restricted to the syllable structure CVV. Lexical morphemes of the shapes CVVCV, CVCVCV and CVCVCVCV are attested but quite uncommon. They are loan words from Setswana or Ikalanga in many cases. The monomoraic CV structure only occurs in grammatical morphemes, a phenomenon reported in other Khoisan languages (see, for example, Miller-Ockhuizen (1998:222) for evidence of this in !Xóõ).

The C_2 position in Tsua lexical morphemes is restricted to the labial voiced stop /b/, the fricative /s/ and the sonorants /m, n, ɲ, r, l, j/. It appears the C_3 and C_4 positions are restricted to the same set of consonants as the C_2 position. For example, /baramu/ 'concubine' and /sebelelo/ 'cape glossy sterling.'[15] Furthermore, when the final vowel V_2 is a high vowel /i, u/, the C_2 consonant is restricted to /b, m, n, r/.[16] There are many cases when the consultants produce [d] in C_2 position, but the evidence suggests that it is underlying /r/ in this position. The production of [d] is subject to inter-speaker variability. Speaker S

---

15 /sebelelo/ is a Setswana loan word.

16 Except for /s/ in C_2 position with a final high vowel /i/ in the aforementioned /bisi/ 'spirit'. Additionally, /k/ occurs in C_2 position for the onomatopoeic /kuku qχ'ao/ 'rooster' and several lexical items associated with chickens.

produces it as a flap [ɾ] for some tokens and at times as [d] for others; speakers B and M produce the stop [d] for almost all of their corresponding tokens. Most importantly, when a Kua cognate has the trill [r] in C_2 position followed by a high vowel /i, u/, the Tsua lexical item is almost always produced with the voiced alveolar stop [d]. The stop closure can be seen in the spectrograms for Tsua [kʰadi] 'liquor' produced by all three consultants (Fig. 9), which corresponds to [kʰari] 'liquor' in Kua. Spectral slice measurements of the burst frequencies were calculated as 3712 Hz, 3423 Hz and 2754 Hz for consultants S, B and M, respectively, measurements that are consistent with burst frequencies for alveolar stops (Reetz and Jongman 2009:193).

The trill [r] in C_2 position co-occurs with /e, o, a/ in V_2 position while the voiced alveolar stop [d] in C_2 position co-occurs with /i, u/ in V_2 position. Given the complementary distribution of [r] and [d], the predictability of [d], and evidence from Kua,

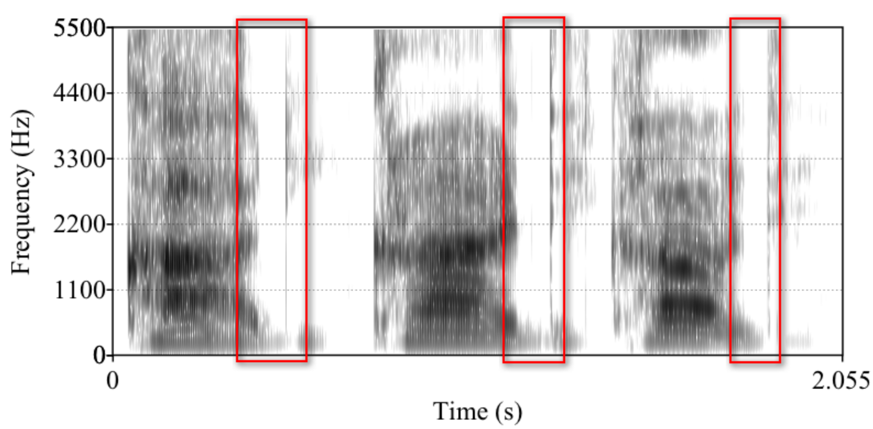


Fig. 9 Stop closures for [kʰadi] 'liquor' spoken in citation form by speakers S, B and M displayed from left to right.

[r] and [d] are allophones of the phoneme /r/ in C_2 position. Moreover, when /r/ occurs intervocalically between /a/, the trill [r] is at its most salient, particularly for speaker S.

There is a restriction on the lateral /l/ as well — it only occurs in the environment of non-high vowels /a, e, o/ and in C_2 position. The result is a limited number of permissible vowel-lateral l-vowel sequences in roots that are observed in the /-V_1-C_2-V_2/ positions: /-alo/, /-ale/, /-ala/, /-olo/, /-ole/, and /-ola/. All of the consonants, whether click or non-click, can occur in C_1 position with the exception of /l, r/. The limited number of consonants that appear in C_2 position are further restricted by the quality of surrounding vowels in V_1 and/or V_2 positions. Thus, it is the root-initial consonantal position that is crucial for contrast, which is consistent with the Khoisan literature.

2.3.2 Back Vowel Constraint

The Back Vowel Constraint is a Khoisan-specific phonotactic generalization proposed in Traill (1985) that captures the co-occurrence restrictions on the phonological shape of the majority of lexical items in the Tuu language !Xóõ. Traill (1985:90) claims a vowel following a [+back] consonant, which includes all of the click consonants in his interpretation, must be /a/, /o/, or /u/ as expressed in the following form.

(7) Back Vowel Constraint (Traill 1985:90)

If:	C_1	V_1
	[+back]	
then:	C_1	V_1
	[+back]	[+back]

Traill (1985:91-92) claims there are very few exceptions across Khoisan languages, suggesting the constraint directly reflects a phonological property of languages with large *native* click inventories rather than back consonants *per se*. He notes that the Nguni languages Xhosa and Zulu have well-developed click systems but show a number of exceptions to the BVC. Nevertheless, he argues that certain assimilatory changes in !Xóõ yield a well-defined class of words that violate the BVC, words that have a long front vowel [iː] or [eː] following the dental [|] and palatal [ǂ] clicks. Traill's singular and plural examples with dental and palatal clicks are given in (8) — tones are omitted for clarity.

(8) BVC exceptions in !Xóõ

Singular	**Plural**	**Gloss**
|ʔi.i	|ʔa.ba + te	'lover'
|qhi.i	|ha.ba + te	'buffalo'
ǂi.i	ǂa.ba + te	'steenbuck'
ǂʔi.i	ǂʔa.a + sa (nom.)	'shoot it!'
ǂqhi.i	ǂha.ba + te	'dog'
ǂqhe.e	ǂqha.m	'sp. bush (terminalia sericea)'

His argument is that the initial stem vowel in the singular has completely assimilated to the Class 1 noun suffix -i and should be seen as deriving from an underlying /a/, as dialectal forms in Khakhea such as [|ʔəi] 'lover', [ǂai] 'steenbuck' and [ǂhai] 'dog' reflect partially assimilated or unassimilated stems (ibid). Thus, Traill claims these exceptions still conform *phonologically* to the constraint, although he did find a handful of genuine, non-alternating exceptions with the dental [|] and palatal [ǂ] clicks preceding /i/ or /e/, but asserts that they likely arose by the same assimilatory process since they have the same phonetic shapes as the examples in (8). For instance, the copulative or stative morpheme [|ii] 'be' is one of the non-alternating exceptions.

Khoisan scholars since then have reported divergent results regarding the BVC in comparison to Traill's study of !Xóõ. Miller-Ockhuizen (2003) and Miller (2010) contend the BVC applies to the uvular consonants [χ], [kχ’] and the alveolar clicks [!], [ǁ] but not [k] in Ju|’hoansi and N|uu, arguing the BVC has its basis in tongue root retraction. A similar line of argumentation is found in Miller, Namaseb and Iskarous (2007) and Miller et al. (2009).

Examining how the BVC could be applicable to G|ui, Nakagawa (2006:226) states it as a set of two constraints:

(9) Nakagawa's version of the BVC

a. * Cq $V_{[\text{-back}]}$

b. * !/ǁ $V_{[\text{-back}]}$

In (9a), Cq represents either the non-click uvular consonants /q ɢ qʰ q’ qχ’ χ/ or consonant clusters with a uvular offset /tχ tsχ tqχ’ or tsqχ’/; $V_{[-back]}$ are the vowels /i e ĩ/; and !/ǁ in (9b) stands for the alveolar and lateral clicks which he categorizes as apical. Nakagawa argues the BVC is not applicable to |/ǂ or Ck in G|ui (the non-click velar consonants), citing exceptions in which the [-back] vowels do not derive from underlyingly [+back] vowels: *n* = 17 for Ck and *n* = 120 for |/ǂ (p. 229). Otherwise, the BVC applies.

Brugman (2009:73) writes that Khoekhoe is one of the Khoisan exceptions to the BVC, as front vowels and [+back] consonants can co-occur, albeit the distribution is skewed in the direction predicted by the BVC when the number of roots containing the consonant-vowel combinations are counted. Moreover, [i] manifests itself freely after all click types in the east African isolates Hadza and Sandawe (Miller 2012:4).

Of interest here is whether there is evidence of a variety of the BVC present in Tsua. Following Nakagawa, Cq represents the non-click uvular consonants /q ɢ qʰ qχ’ χ/ and non-click consonant clusters with a uvular offset /tsqχ’ cqχ’ tχ tsχ cχ/. Ck stands for the non-click velar consonants /k g kʰ k’ ŋ/. $V_{[-back]}$ are the vowels /i e ĩ/. The four click types with all of their respective accompaniments, voicing and nasality are subsumed under the general headings [|] dental, [ǂ] palatal, [!] alveolar, and [ǁ] lateral for simplicity.

An examination of the Tsua roots found numerous violations of the BVC for the [|] dental and [ǂ] palatal categories by all three consultants (*n* = 35 roots). This finding is in line with G|ui. Interestingly, there are violations of the BVC in the Cq, Ck and [ǁ]

categories depending on the speaker. Table 16 gives a list of these violations by speaker.

Table 16 BVC violations in Tsua by speaker in the Cq, Ck and [ǁ] categories. Speaker S violates the BVC in all of these cases (*n* = 9). Speaker B has 1 violation. Speaker M does not violate the BVC.

speaker S	speaker B	speaker M	Gloss
ǁχee	ǁχae	ǁχae	'to scrape hide'
cqχ'ee	cqχ'ae	cqχ'ae	'to spit'
ǂqee	ǂqae	ǂqae	'bone marrow'
qχ'ee	qχ'ae	qχ'ae	'to cry'
baramu ǀqχ'ee	baramu ǀqχ'ae	baramu ǀqχ'ae	'to visit the opposite sex'
qχ'ee ǀʰunu	qχ'ae ǀʰunu	qχ'ae ǀʰuni	'to sob'
gǁee	gǁae	gǁae	'to count'
gee	gae	gae	'to pass by'
gǁee	gǁee	gǁae	'to read'

The front vowel [e] is the cause of speaker S's violations in all of the cases. Speaker B largely adheres to the BVC in the Cq, Ck and [ǁ] categories with the exception of [gǁee] 'to read'. Speaker M conforms to the BVC. Moreover, speakers S and B produce [ki̲] '1SG.POSS', a BVC violation in the Ck category. Speaker M produces [ci] '1SG.POSS', which does not violate the BVC. The fact that speakers S and B produce [ki] and speaker B [ci] may be analogous to what Traill found in !Xóõ. The grammatical particles /ki/ and /ke/ in !Xóõ that violate the BVC are reported by Traill as being produced by native speakers with alternative pronunciations [ti] and [te], owing to the power of the BVC (1985:90). The results presented here suggest that a strong version of the BVC is untenable for Tsua. Nonetheless, the distribution of front vowels and [+back] consonants is skewed in the direction predicted by the BVC.

2.4 Summary

One goal of this chapter was to present a clearer picture of the Tsua segmental inventory. There were two primary findings. First, a significant part of the inventory has a uvular place of articulation for singletons and clusters. Second, there is plausible evidence in favor of including pharyngealized vowels in the inventory. Moreover, click replacement has left the click inventory asymmetrical by reducing the contrasts among alveolar [ǃ] and palatal [ǂ] clicks while leaving the dental [ǀ] and lateral [ǁ] click contrasts intact, a Stage 3 click-series system (i.e., ǀ ǁ c k) in line with click replacement patterns in Khoe languages (Traill 1986). Now that the segmental inventory has been described, the next chapter establishes the Tsua tonal system in detail through a rigorous acoustical analysis.

CHAPTER 3

3.1 Introduction

As much as 60-70 percent of the world's languages are tonal, with the majority of African languages distinguishing lexical items by tone (Odden 1995; Yip 2002). In spite of its ubiquity across the African continent, tone has been one of the last phonologically relevant features to be incorporated into descriptions (Elderkin 2004:4). Like most African languages, Tsua incorporates tone for lexical contrast but does not display the mobility of shifting, spreading or floating tones typically found in neighboring Bantu languages. Brugman (2009:117) asserts that shifting, spreading or floating tone phenomena have not been reported in any Khoisan language. Nevertheless, while Tsua lacks the tonal mobility found in many African languages, it does have an extreme tonal depressor effect following voiced obstruents, aspirated obstruents, and the glottal fricative /h/. Detailed evidence for Tsua tonal depression is presented in Chapter 4.

The goal of this chapter is to establish the Tsua tonal melodies in roots with non-depressor consonants and propose their representations. A “tonal melody” is defined here as a permitted tone sequence docked to the tone-bearing units (TBUs) of a bimoraic root, resulting in the form [T_1T_2], with the mora as the TBU. First, an overview of tone in Central Khoisan languages is presented to give the Tsua tonal melodies context. The overview includes descriptions for which accurate and detailed tonal information is available. Recent work on Central Khoisan by Vossen (1997) has demonstrated that it can be treated as a genetic unit (Elderkin 2004:4; see also Honken 2008:187). Thus, what might be expected in Tsua can be illustrated by limiting the overview to Central Khoisan tone.

3.2 Central Khoisan tone

3.2.1 Khoekhoe branch

The pioneering work in Beach (1938) is considered by many to be the first phonetically accurate description of the Khoekhoegowab language's tone inventory (henceforth Khoekhoe).[17] He identifies six major tone melodies and labels them as High-rising, Mid-rising, Low-rising, High-falling, Mid-falling and Low-mid level, considering each melody as a unit. Moreover, Beach suggests that four of the Khoekhoe melodies are derived by a split from two proto-melodies using comparative data from the closely-related language Korana [!Ora in Güldemann 2006]. He also observes Khoekhoe tone sandhi forms, in which citation melodies systematically change in certain morpho-syntactically conditioned environments. These forms are described in later work on Khoekhoe such as Haacke (1999).

Contrary to the unitary treatment of tone melodies in Beach (1938), Hagman (1977) takes a compositional approach to each melody. He posits three register tones, High (H),

[17] Khoekhoegowab is the endonym and Nama/Damara the bipartite exonym.

Middle (M), and Low (L), arguing that each Khoekhoe root consists of two morae. Each mora is associated with one of these three tones. Hagman finds six root melodies of the nine possible in a three tone system: HH, HM, MH, MM, LH and LM. He makes the generalization that the first mora may have any of the three tones but the second mora may only have a High or Middle tone. Hagman notes the tone sandhi forms but does not fully account for their patterns, most likely because his primary concern is syntax.

Haacke (1999) is a detailed analysis of Khoekhoe tone and Haacke and Eiseb (2002) is an extensive Khoekhoe dictionary. Haacke posits four tone levels that he labels using numbers 1 through 4, following the convention in research on Chinese tone (Chao 1930). 1 indicates the lowest tone while 4 indicates the highest. Haacke identifies six major citation melodies and their corresponding sandhi forms. Two of the sandhi forms, /21/ and /22/, each occur with two of the six main citation melodies.

(10) Khoekhoe citation and sandhi melodies as recapitulated in Haacke (2008:158)

Melody	**Citation Form**	**Sandhi Form**
Double-Low	/12/	/21/
Low-Rising	/13/	/13/
Low (level)	/22/	/22/
High (falling)	/32/	/21/
High-Rising	/24/	/22/
Double-High	/43/	/43/

There are sixteen theoretically possible sets of tonal combinations in the four tone system in (10). Haacke addresses the six attested versus sixteen possible issue in part by arguing that there are a number of minor melodies, although he considers them exceptional since they are rare. The combination of the major and exceptional melodies accounts for twelve of the sixteen possible sets.

Brugman (2009) is a dissertation that focuses on Khoekhoe prosody. She addresses the citation melodies and sandhi forms but differs from Haacke (1999) in her analysis. She describes six citation melodies that occur on CVV, CVCV and CVN roots: Superlow (SL), Low (L), Superlow-Low (SL-L), High (H), Superhigh (SH) and High-Superhigh (H-SH). Brugman observes that the "level" SH and H melodies fall significantly throughout their production and the SL melody rises slightly. She attributes these realizations to the phonetic implementation of morae unspecified for tone, writing: "The four monotonal melodies associate with the first mora of the root, leaving the second mora unspecified, while the two tones in contour melodies associate with one mora each" (p. 120). Her analysis of the relationship between citation and sandhi forms is summarized in (11).

(11) Khoekhoe citation and sandhi melodies from Brugman (2009:132)

Melody	Citation Form	Sandhi Form
Superlow-Low	SL-L	SL-L
Superlow	SL	L-SL
High	H	L-SL
Low	L	L
High-Superhigh	H-SH	L
Superhigh	SH	H

3.2.2 Kalahari Khoe West branch

This subsection covers three languages from the Kalahari Khoe West branch: Kxoé [Khwe], Naro and G|ui. The Kxoé tonal system appears to have six major melodies: High, LH Rising, LM Rising, MH Rising, HM Falling and HL Falling (Honken 2008:186, citing Oswin Köhler's collected materials and Killian-Hatz 2003). Honken notes that both scholars use marking conventions suggesting a rise on the second syllable for the LM Rising melody and a rise on the first syllable for the MH Rising melody in disyllables (fn 4:222).

Visser's (2001) Naro dictionary incorporates a three tone notation system of High (H), Mid (M) and Low (L) to account for six major melodies and one minor for a total of seven distinctive bimoraic melodies. The six major Naro melodies are HH, HL, MM, ML, LM, and LL. The seventh melody is LH. The LH melody has few attestations and may require a re-analysis because there is no clear correspondence to any one Khoekhoe melody according to Haacke (2008:161). Tone is not always indicated in Visser's dictionary. He designates lower case letters to mark melodies that are preliminarily guessed, reserving upper case letters for transcriptions that have a high degree of certainty. It is a clear sign that Khoisan tonal transcription can be quite challenging.

The tonal system of G|ui is rigorously studied in Nakagawa (2006), culminating in a G|ui-Japanese-English dictionary in Nakagawa et al. (2014). Nakagawa uses a three tone system of High (H), Mid (M) and Low (L) to account for six major melodies: three level and three contour melodies. These bimoraic melodies occur with CVV, CVN and CVCV syllable structures labeled as tonological Domain 1 (DOM1) and are transcribed as HH, HM, HL, MM, LM and LL. He proposes a second tonological domain, Domain 2 (DOM2), in which non-roots are limited to the /H/ vs. /L/ two-way tonal contrast. The DOM2 non-roots can be monomoraic or bimoraic, having either one or two TBUs, respectively. Thus, the DOM2 non-roots have a single tone associated with either one TBU or multiply linked to two TBUs.

3.2.3 Kalahari Khoe East branch

Vossen (2013d) briefly describes the surface tones in two Shua subgroup languages, Cara and Deti. Vossen analyzes two surface tones, High (H) and Low (L), that combine to form four combinations on CVCV, CVV and CVN lexical roots: H-H, L-L, H-L and L-H.

The four sequences represent two tone classes each that are distinguished on the basis of diverging tonal behavior in finite and imperfect verb forms, leading Vossen to postulate eight Cara tone classes. These eight classes are labeled H-H (I), H-H (II), L-L (I), L-L (II), H-L (I), H-L (II), L-H (I) and L-H (II). The same tonal sequences for disyllabic lexical roots are observed in Deti but with seven classes diverging in perfect and imperfect verb forms. The classes are H-H, L-L, H-L (I), H-L (II), L-H (I), L-H (II) and L-H (III).

Previous research on Tshwa subgroup tone is virtually non-existent aside from the brief description in Vossen (2013c:104). Vossen proposes four surface sequences on disyllabic roots in Kua and Tsua as H-H, L-L, H-L and L-H, identical to Cara and Deti.[18] These sequences constitute five tone classes based on an investigation of Kua verbal tone: High-High (H-H); Low-Low (L-L); High-Low (H-L); Low-High I (L-H I); and Low-High II (L-H II). L-H (I) and L-H (II) are differentiated on the basis of the divergent tone pattern in the imperfect according to Vossen (in bold).

18 It is not clear at this time whether the term 'Kua' is always used by Khoisan speakers to mean a specific language or as a generic term meaning 'speakers of another language' (Chris Collins, Andy Chebanne, personal communication).

(12) Vossen's five tone classes in the Tshwa subgroup (2013c:104)

Tone class	Sequence	Verb		Present	Pattern	Imperfect[19]	Pattern
1	H-H	ǁáó	'shoot'	kùà ǁáó	LL HH	ǁá-ró-hà	H-H-L
2	L-L	ǁgàũ	'spread (a hide)'	kùà ǁgàũ	LL LL	ǁgàũ-á-hà	LL-H-L
3	H-L	pîî	'suck'	kùà pîî	LL HL	píí-á-hà	HH-H-L
4	L-H (I)	tshàó	'dig'	kùà tshàó	LL LH	**tshà**-ró-hà	**L**-H-L
5	L-H (II)	xùń	'grind'	kùà xùń	LL LH	**xúń**-hà	**HH**-L

Postulating tonal classes 4 and 5 in the Tshwa subgroup as rising tone sequences L-H (I) and L-H (II), respectively, overlooks the Tsua tonal depressor effect following aspiration in *tsh̲ao* 'dig', which can be observed when instrumental evidence and Kua comparative data are taken into consideration. Vossen's analysis does not include Mid tones either. In the case of Kua, which does not have tonal depression, *ts^{h}au* 'dig' is transcribed as High-Mid (HM) in Chebanne and Collins (2014). In the current inquiry, Tsua *ts^{h}ao* 'dig' is analyzed as having a root-initial depressed High tone followed by a Mid tone, resulting in a Depressed High-Mid tonal melody (DH-M), as opposed to Vossen's L-H sequence. *Pii* 'milk; breast; suckle' is analyzed as HL in the current data set, which is in agreement with Vossen. An account that captures the Tsua data set is presented in §3.3 and Chapter 4.

---

19 Usually 'hà' is Perfect, not Imperfect (Chris Collins, personal communication).

3.3 Tsua lexical tone

3.3.1 Citation melodies

The next subsections present a qualitative analysis of the phonetic realization of Tsua tone with the goal of establishing the basic tonal melodies as elicited in citation form, i.e., in isolation, which is standard field practice (Hyman 2010:205; Kutsch Lojenga 2010:7). The moraic TBU analysis for Tsua in the forthcoming subsections is consistent with findings from within Central Khoisan (§3.2) and outside of it (for instance, Collins 2012 and Miller 2013 for the Northern Khoisan languages ǂHoan and Ju|'hoan, respectively). Nasal consonants can bear tone. The high vowels [i, u] perturb tone melody realization by increasing F0 overall compared to tones docked to non-high vowels and are therefore excluded.

Six contrastive tonal melodies have been observed as listed in (13).

(13) Tsua contrastive tonal melodies

Tone Category	**Notation**
a. High-level	acute + acute accents; e.g. áá
b. HM-falling	acute + macron accents; e.g. áā
c. HL-falling	acute + grave accents; e.g. áà
d. Mid-double-rise	macron + macron; e.g. āā
e. MH-rising	unmarked + acute; e.g. āá
f. ML-falling	unmarked + grave; e.g. āà

What is particularly notable about the melodies in (13) is that their respective F0 realizations unfold over time in a way that is reminiscent of Asian languages such as Thai, especially with respect to level tones, posing a challenge for the simple tone associations proposed in autosegmental phonology. For example, the Mid-double-rise melody's F0 rises, falls and then rises again towards the end of the second Mid tone. Each melody has its own distinctive F0 curve shape and onset frequency range, and it is the instrumental comparison of these shapes that is used as one criteria for grouping and transcribing tonal melodies in citation form. Additionally, melody grouping is achieved by using comparative data from Kua and G|ui (Chebanne and Collins 2014 for Kua; Nakagawa 2006, Nakagawa et al. 2013 for G|ui). The Tsua tone categories reliably map to the Kua and G|ui tone categories when checking cognates from the three languages (see the Appendix for cognate lists by tone category). These mappings are presented in Table 17. Phonetic labels are used for Tsua until the tonological analysis in §3.3.10.

Table 17 Tonal category mappings between Tsua, Kua and G|ui

Tsua Tone Category	Kua Tone Category[1]	G|ui Tone Category[2]
High-level	HH	HH
HM-falling	HM	HM
HL-falling	HL	HL
Mid-double-rise	MM	MM
MH-rising	LH	LM
ML-falling	LL	LL

[1]Chebanne and Collins (2014) [2]Nakagawa (2006), Nakagawa et al. (2014)

Kua speaker O was highly adept at grouping tonal melodies in the Kua language through same/different melody judgements. His judgements aided the Tsua groupings via the Kua to Tsua category mappings in Table 17. It should be pointed out that while the mappings are quite helpful in identifying the Tsua categories, the Kua and G|ui tonal categories differ from one another as seen in Table 17. Chebanne and Collins (2014) analyzes the Kua rising tone as LH, while Nakagawa (2006), Nakagawa et al. (2014) analyze the G|ui rising tone as LM. The Tsua analysis presented in §3.3.10 differs from the Kua and G|ui analyses in one respect: there is evidence that the root-initial Low tones in Kua and G|ui do not correspond to root-initial Low tones in Tsua.

Given the nature of tone melody realization in Tsua one question is whether a contour tone unit analysis or a compositional analysis is most appropriate. A contour tone unit analysis, in which a contour behaves as a non-decomposable phonological unit as in Beach (1938), predicts that contour tones can occur on monomoraic forms, while a compositional analysis, in which a contour is a sequence of level tones, predicts that only level tones can occur on monomoraic forms. All of the monomoraic Tsua forms are grammatical morphemes in my data set as discussed in Chapter 2, and they only bear level tones, thus supporting the prediction of the compositional analysis and the mora as the TBU.

A near-minimal CVV sextuplet in (14a) provides support for a six-way tonal contrast in Tsua, with disyllabic examples of the form CVCV showing the same tone patterns in (14b). While the data in (14) focus on the tonal patterns for the CVV and CVCV forms, all six tonal patterns occur with the CVN syllable structure, with each pattern being produced identically to the CVV and CVCV tonal patterns.

(14) Tsua tone patterns on monosyllabic and disyllabic roots

a. Tonal near-minimal sextuplet on monosyllabic roots

Tone Category	**Lexical Item**	**Gloss**
High-level	ǁáé	'to teach'
HM-falling	ǁáē	'to chew'
HL-falling	ǁáè	'to brood (eggs)'
Mid-double-rise	ǁāē	'seed necklace'
MH-rising	ǁāé	'to swear at someone'
ML-falling	ǁāò	'wing'

b. Tonal patterns on disyllabic roots

Tone Category	**Lexical Item**	**Gloss**
High-level	sá.rá	'to fall short'
HM-falling	tsó.rō	'shell; husk'
HL-falling	ǁχá.bà	'hunchback'
Mid-double-rise	ǁχā.mā	'red hartebeest'
MH-rising	dzā.rá	'bird (generic)'
ML-falling	kā.rò	'to run'

3.3.2 Tone plotting methodology

The previous section established the six Tsua contrastive tonal categories through speaker judgements, data from related languages and preliminary measurements. The purpose of this section is to determine the phonetic realization of the six tonal categories by presenting the results of a rigorous acoustic analysis. The analysis will justify the category labelings and transcriptions. First, the tone categories are described in turn using F0 pitch traces and spectrograms plotted by speaker. The purpose is to highlight the relevant characteristics of speaker variability. Second, plots are given to illustrate the cross-speaker defining landmarks for each tonal category.

The cross-speaker F0 averaging is done on a logarithmic scale: mean_f0 = exp(sum(ln(f01-n))/n) in ProsodyPro 5.5.2 (Xu 2013).[20] 5 CVV lexical items with root-initial voiceless consonants followed by non-high, non-nasalized and non-pharyngealized vowels are measured for each category to give the typically produced tonal melodies with the least amount of F0 perturbation. Each speaker produced 3 repetitions of each lexical item: 5 CVV lexical items x 3 repetitions x 3 speakers = 45 utterances per F0 averaged tone melody.[21]

ProsodyPro was set to time-normalize the contours at 10 points per tone, resulting in 20 time-normalized measurement points for each bi-tonal lexical item. Time-normalization is done to allow graphical comparisons across melodies. A grand total of 900 measurements (45 utterances x 20 time-normalized points) are taken for each melody.

---

[20] http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/

[21] As a methodological comparison, Xu (1997:67) measures 48 total utterances from 8 speakers for each of 4 Mandarin tones to establish the mean F0 contours.

3.3.3 High-level melody

Fig. 10 shows spectrograms and F0 traces of the High-level noun /|áḿ/ 'sun' spoken in citation form by speakers S, B and M from left to right. The High-level melody is produced with the highest Hertz value and lowest F0 excursion compared to the other 5 melodies. The lower F0 excursion is one characteristic that makes the High-level melody differ from the Mid-double-rise melody. Nevertheless, all three speakers produce it as phonetically falling to some extent over the course of a given bimoraic root. The slope of the fall is more pronounced for speaker B in her tokens as seen in Fig. 10's middle pitch trace. The fact that the Tsua High-level melody falls a certain extent is unsurprising; in perceptual studies of Thai ‘level’ tones, Abramson (1975, 1978) found that Thai speakers were less confused and accepted synthetic speech with some pitch movement as ‘level’

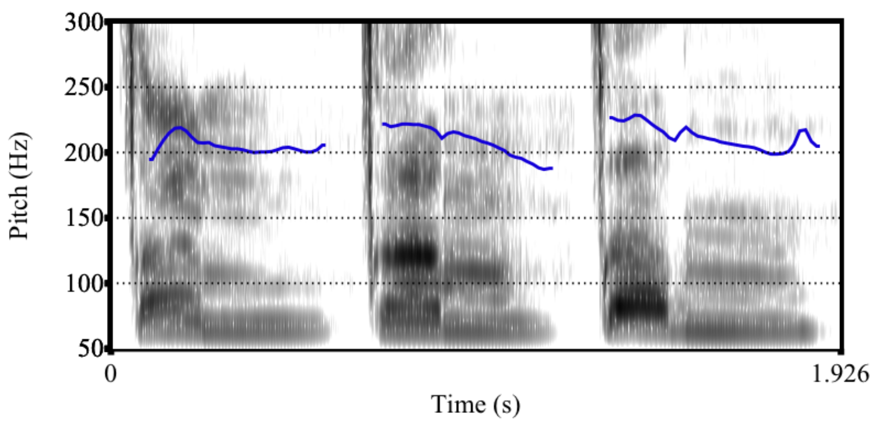


Fig. 10 F0 traces with spectrograms of the High-level noun /|áḿ/ 'sun' spoken in citation form by speakers S, B and M displayed from left to right.

compared to synthetic speech with no pitch movement. Moreover, referring to acoustic studies of Yoruba tones, Laver (1994:467) states: “Acoustic analysis reveals, not surprisingly, that in real speech the pitches of the so-called ‘level’ tones are seldom strictly level.” This is the case for the Tsua High-level melody as well.

The 5 CVV lexical items used to generate the landmarks plot are listed in (15) with Fig. 11 showing the F0 averaged High-level melody across speakers. In terms of cross-speaker landmarks, the High-level contour has a short rise that has an F0 peak at 211 Hz, followed by a gradual fall to 184 Hz, and ending with another short rise to an F0 offset of 188 Hz.

(15) F0 averaged High-level lexical items

Lexical Item	**Gloss**
ǀéé	'blue wildebeest'
ǁʔáá	'bat-eared fox'
ǁóé	'to lie down'
cóó	'heart'
ǀáá	'to cut'

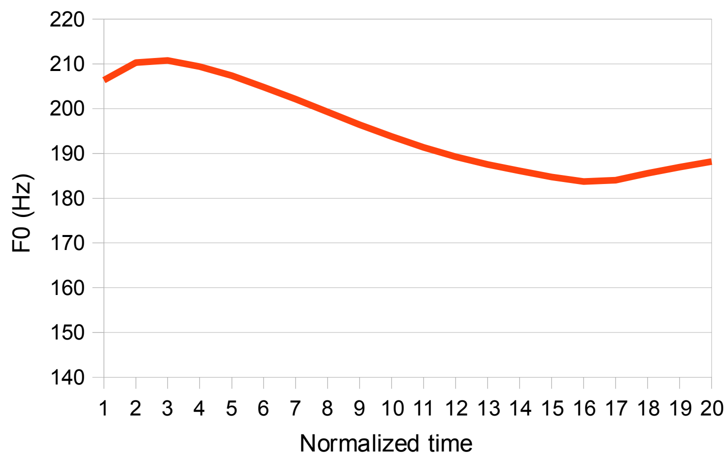


Fig. 11 High-level melody time-normalized and F0 averaged across speakers S, B and M (*n* = 45).

3.3.4 HM-falling melody

Fig. 12 shows F0 traces and spectrograms of HM-falling melodies produced by the three Tsua consultants S, B and M. The HM-falling melody often has a short rise after the F0 onset and then falls to the F0 offset. One property that distinguishes the HM-falling melody from the High-level melody is the lack of a final rise towards the end of the contour for the HM-falling melody. The High-level final rise is more evident in tokens produced by speakers S and M, a rise clearly seen when averaged across all speakers and multiple tokens. All three consultants produce the HM-falling melody with a pitch that falls more dramatically to the F0 offset pitch compared to the High-level melody, but less so when compared to the HL-falling melody shown in the next section.

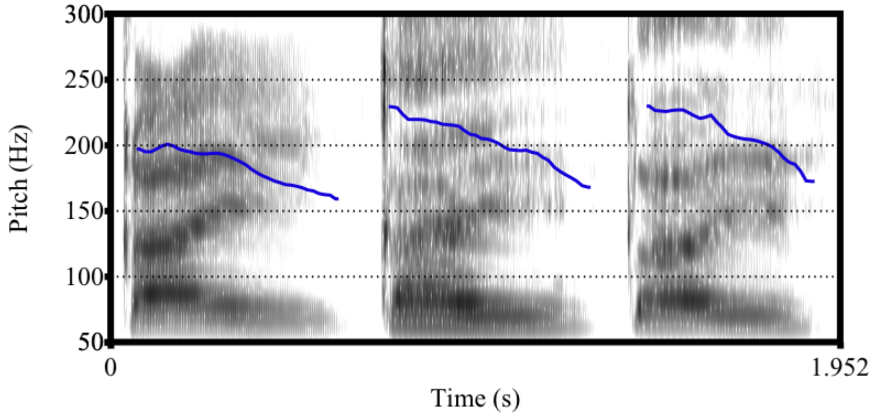


Fig. 12 F0 traces with spectrograms of the HM-falling verb /ǁáē/ 'to chew' spoken in citation form by speakers S, B and M displayed from left to right.

The landmarks plot in Fig. 13 shows that the HM-falling melody's peak F0 value of 207 Hz across speakers is close to the High-level melody's peak F0 of 211 Hz, confirming the root-initial High tone interpretation. Since the 177 Hz offset pitch for HM-falling is lower than the High-level melody's offset pitch of 188 Hz, yet higher than the HL-falling melody's offset pitch of 145 Hz, the second tone is confirmed as a Mid tone.

(16) F0 averaged HM-falling lexical items

Lexical Item	**Gloss**
ǀʔéē	'fire; firewood'
ǀχáā	'meat'
ǁáē	'to chew'
qχ'áē	'to cry'
ǀqχ'áō	'snake'

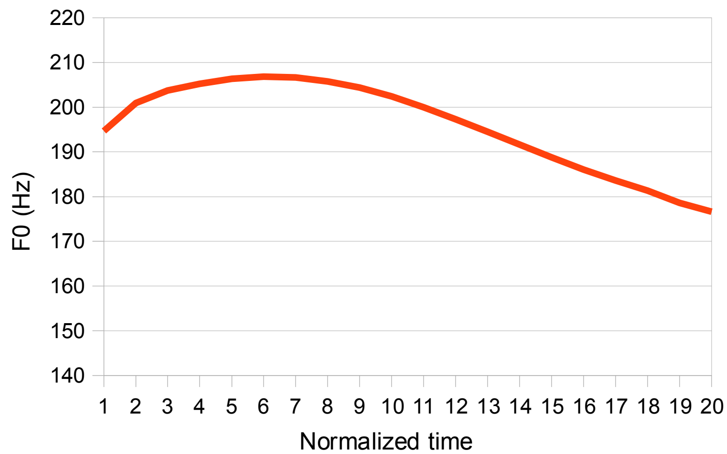


Fig. 13 HM-falling melody time-normalized and F0 averaged across speakers S, B and M (n = 45).

3.3.5 HL-falling melody

Exemplars of the HL-falling melody are shown in Fig. 14 for /|qχ'áǹ/ 'ugly'. There

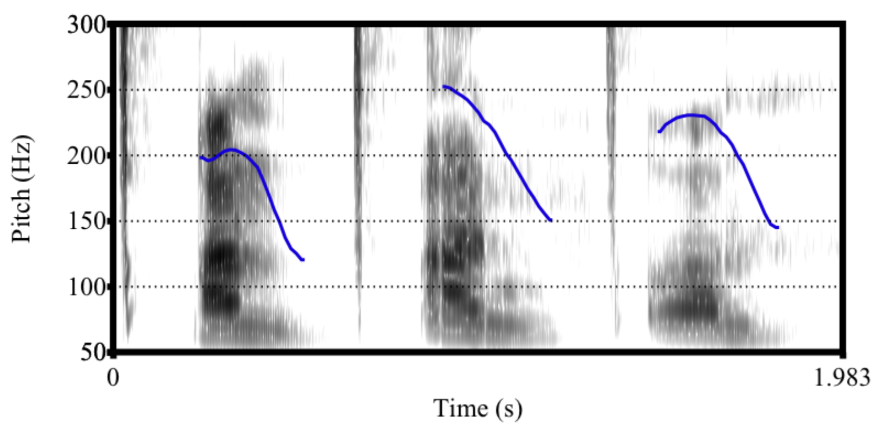


Fig. 14 F0 traces with spectrograms of the HL-falling adjective /|qχ'áǹ/ 'ugly' spoken in citation form by speakers S, B and M displayed from left to right.

is a short rise at the beginning of the contour for speakers S and M before the precipitous fall in many of their tokens. What is important to the identification of this salient melody is the range and speed of the fall. The HL-falling melody has a larger range and decreased duration with respect to the fall compared to the HM-falling melody.

The cross-speaker landmarks plot in Fig. 15 shows that the root-initial tone is confirmed as a High tone since the peak F0 of 209 Hz is close to the peak F0 pitches of the High-level (211 Hz) and HM-falling melodies (207 Hz). The root-final tone is confirmed as a Low tone because it's F0 offset value of 145 Hz is lower than the HM-falling and Mid-double-rise melodies (177 Hz and 174 Hz for HM-falling and Mid-double-rise, respectively).

(17) F0 averaged HL-falling lexical items

Lexical Item	**Gloss**
|ʔáò	'money; blood'
|χóò	'gemsbok'
|ʔóò	'to dry (general)'
qχ’áà	'to drink'
qχ’áò	'man; husband'

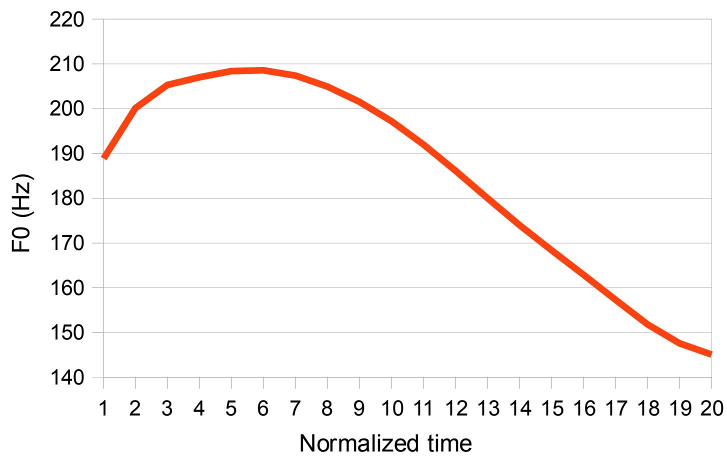


Fig. 15 HL-falling melody time-normalized and F0 averaged across speakers S, B and M (n = 45).

3.3.6 Mid-double-rise melody

The Mid-double-rise melody is shown in Fig. 16 by speaker for the noun /χāē/ 'night'. The Mid-double-rise melody is often characterized by a rise in pitch from the F0 onset, followed by a drop, then a rise towards the end of the melody. The result is a pitch wavering quality. It is possible that this up-and-down wavering is a phonetic effect that helps differentiate the Mid-double-rise melody from the High-level melody, especially in running speech. In many of the consultants' tokens, the initial rise of the Mid-double-rise melody approaches but never quite reaches the High tone pitch levels of the melodies with High tones, i.e., High-level, HM-falling and HL-falling. The Mid-double-rise is consistently higher throughout compared to the Low tone in the HL-falling melody but

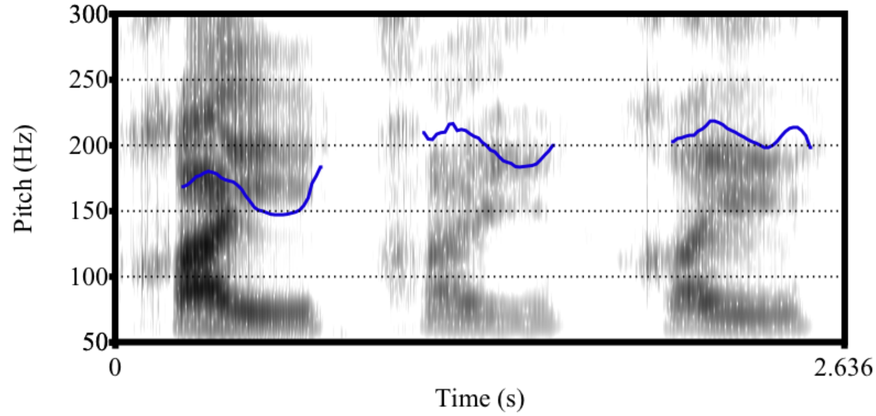


Fig. 16 F0 traces with spectrograms of the Mid-double-rise noun /χāē/ 'night' spoken in citation form by speakers S, B and M displayed from left to right.

lower than the High tones in the High-level, HM-falling and HL-falling melodies.

The cross-speaker landmarks plot in Fig. 17 shows that the peak F0 value of 202 Hz approaches the peak F0 values of High tones (211 Hz, 207 Hz and 209 Hz for High-level, HM-falling and HL-falling, respectively). After the peak F0, the contour falls and stays in the range below High tones but above Low tones, confirming the Mid tone interpretation.

(18) F0 averaged Mid-double-rise lexical items

Lexical Item	**Gloss**
χāē	'night; darkness'
tsqχ'āē	'green; blue'
ǂqāē	'bone marrow'
cχōā	'elephant'
tēē	'to ask'

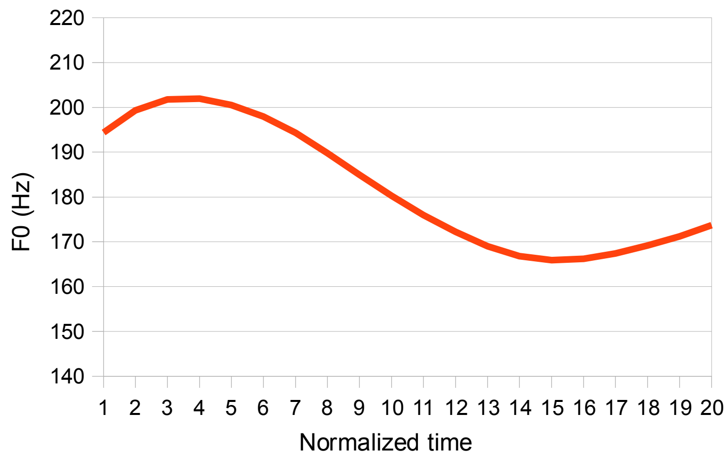


Fig. 17 Mid-level melody time-normalized and F0 averaged across speakers S, B and M (*n* = 45).

3.3.7 MH-rising melody

Typical MH-rising melodies for the Tsua consultants are shown in Fig. 18. The MH-rising melody falls from the F0 onset with a moderate slope until it sharply turns and

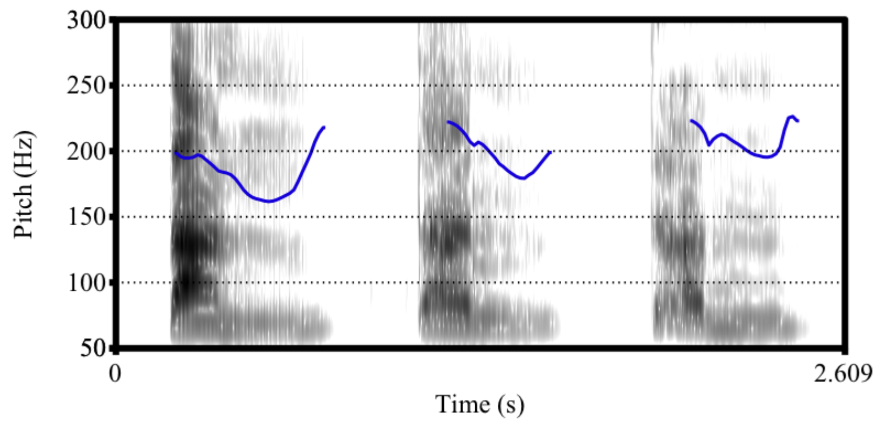


Fig. 18 F0 traces with spectrograms of the MH-rising verb /tā́n/ 'to stand up' spoken in citation form by speakers S, B and M displayed from left to right.

increases quickly to the F0 levels comparable to the High tones in the High-level, HM-falling and HL-falling contours. Speakers S and M consistently produce their rising tone offsets in the High tone range while speaker B is not as consistent as seen in Fig. 18's middle pitch trace.

Fig. 19's landmarks plot shows the initial fall with a moderate slope when multiple tokens are averaged across the three Tsua consultants. The contour takes a turn and increases to the F0 offset value (194 Hz). It is plausible to postulate a root-initial Mid tone interpretation that falls before the turn towards the High tone target as a language-specific phonetic property of this contour.

(19) F0 averaged MH-rising lexical items

Lexical Item	**Gloss**
ǀqχ’āé	'to fall down'
tēé	'bow'
ʔāé	'village'
ʔōó	'to bear fruit'
ǁχāó	'to scrape'

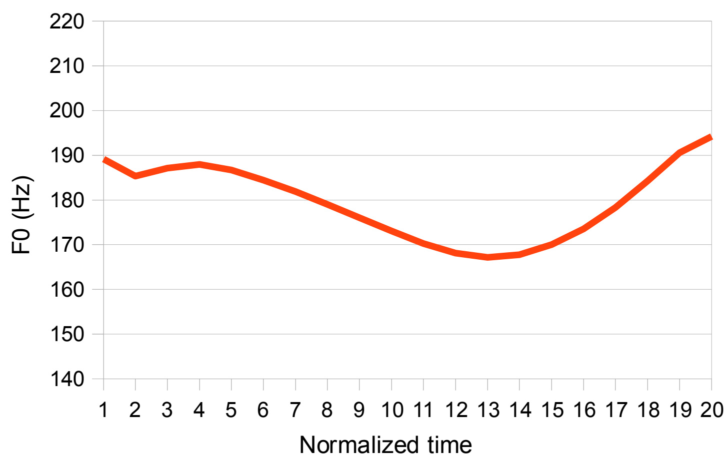


Fig. 19 Fall-rising melody time-normalized and F0 averaged across speakers S, B and M (n = 45).

3.3.8 ML-falling melody

The ML-falling melody is almost a mirror image of the MH-rising melody as the

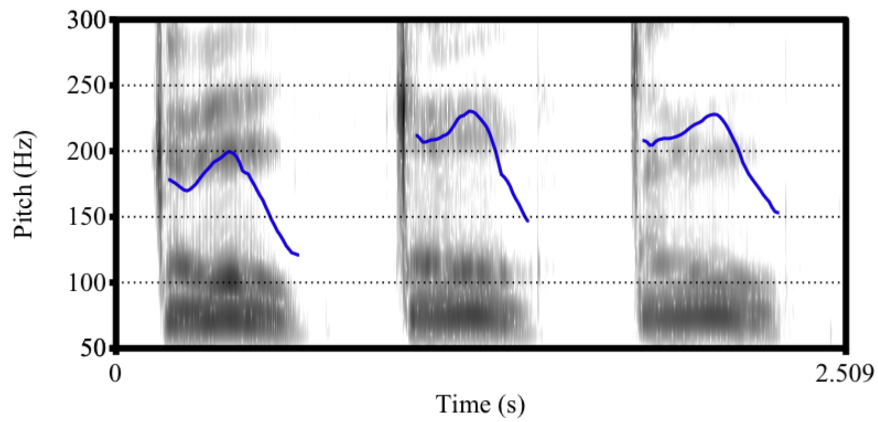


Fig. 20 F0 traces with spectrograms of the ML-falling verb /|ōò/ 'to run after (when hunting)' spoken in citation form by speakers S, B and M displayed from left to right.

nomenclature implies. Typical ML-falling contours are displayed by consultant in Fig. 20 with the verb /|ōò/ 'to run after (when hunting)'. The F0 traces reveal that the ML-falling melody starts with a moderately-sloped rise from the F0 onset for the first half of the contour, followed by a quick fall to the L tone F0 level similar to the HL-falling melody and yet below the M tone F0 range for the HM-falling and Mid-double-rise melodies. The rise before falling could also be interpreted as a language-specific phonetic property. The landmarks plot is shown in Fig. 21.

(20) F0 averaged ML-falling lexical items

Lexical Item	**Gloss**
‖āà	'black korhaan'
ǂqōò	'mud'
‖ōò	'to weed'
kāò	'wet w/ dew'
|ōò	'to run after (when hunting)'

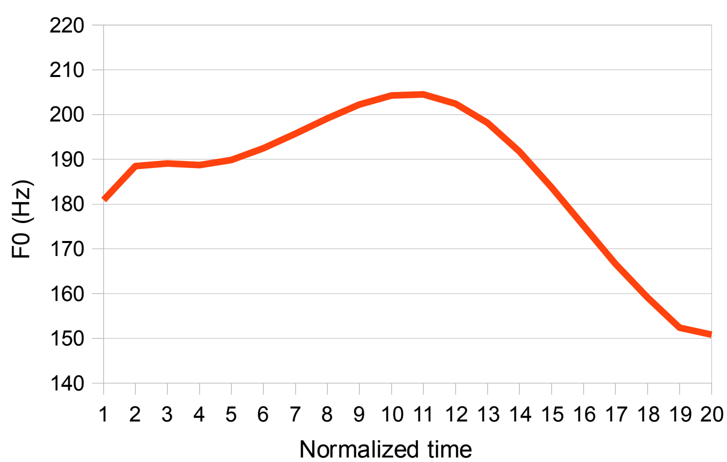


Fig. 21 ML-falling melody time-normalized and F0 averaged across speakers S, B and M ($n = 45$).

3.3.9 Superimposed tone melody traces

The tonal space of the six time-normalized and F0 averaged melodies across the three Tsua speakers is shown in Fig. 22. The High-level, HM-falling and HL-falling shapes have F0 peak values close to or above 210 Hz within normalized time points 1-10, i.e., in the first half of the rime in the CVV syllable. I interpret this as confirmation that the root-initial High tone analysis is appropriate. The Mid-double-rise melody stays below the High-level melody throughout, although the first half of the shape (normalized time points 1-10) is closer in frequency compared to the second half (normalized time points 11-20). The MH-rising and ML-falling contours are close in frequency soon after their respective F0 onsets until they diverge at normalized time point 4. At that point, the MH-rising contour begins to fall and the ML-falling contour starts to rise.

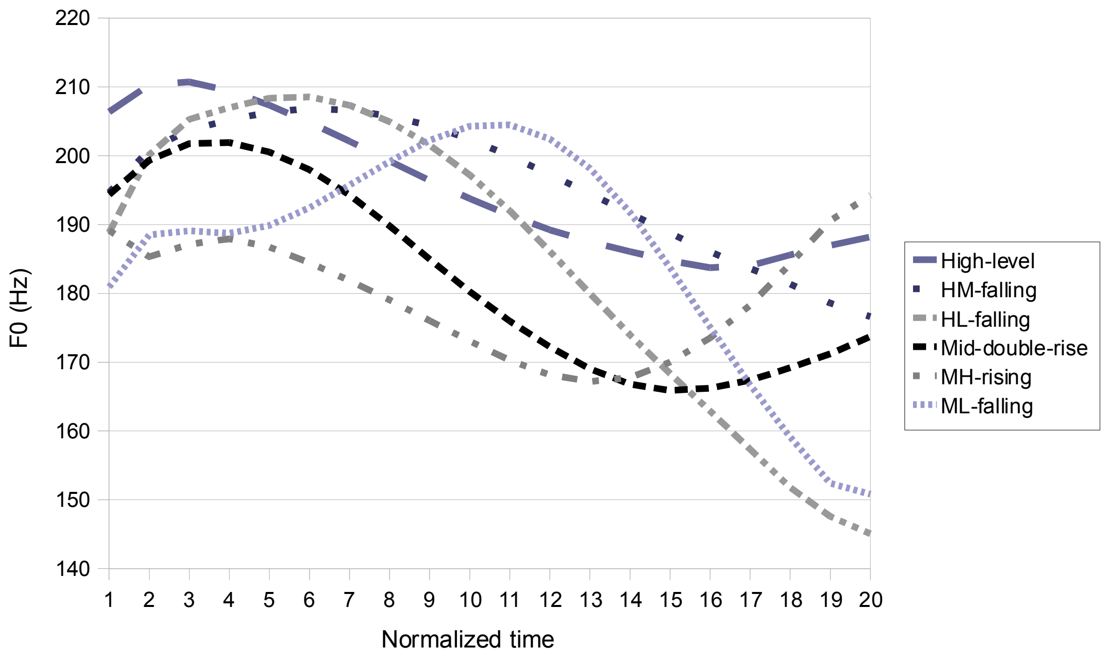


Fig. 22 The tonal space of the six time-normalized and F0 averaged tone melodies across speakers S, B and M (n = 45 per tone melody).

All six of the F0 offsets cluster together in pairs of two at normalized time point 20: the High-level and MH-rising melodies at approximately 190 Hz, supporting a final High tone interpretation; the HM-falling and Mid-double-rise melodies at about 175 Hz, supporting a final Mid tone; and the HL-falling and ML-falling melodies at around 150 Hz, supporting a final Low tone interpretation.[22] Table 18 reports the cross-speaker mean pitch targets in Hertz at salient points in each tonal melody's production from the traces in Fig. 22.

[22] There was broad consensus among the tone researchers at the UMass Amherst Tonal Spaces Workshop, June 2-3, 2014, that tones often hit their targets at the right edge of a domain based on evidence from multiple languages.

Table 18 Mean pitch targets in Hz at salient points in the contours (*n* = 45 per melody)

Tone Category	F0 Onset	F0 Maximum	F0 Minimum	F0 Offset
High-level	206	211	184	188
HM-falling	195	207	177	177
HL-falling	189	209	145	145
Mid-double-rise	194	202	166	174
MH-rising	189	194	167	194
ML-falling	181	205	151	151

3.3.10 Tonological analysis and proposed representations

The tone categories can be straightforwardly analyzed as in Table 19 based on the tone plots and comparative data discussed throughout the previous sections. Nevertheless, the representations should account for some tokens in which the F0 targets are not attained by the consultants for the two monotonal melodies High-level and Mid-double-rise.

Table 19 Tonological analysis of the six contrastive Tsua tone melodies

Tone Category	Bimoraic Sequence	Tonological Analysis
High-level	HH	/H/
HM-falling	HM	/HM/
HL-falling	HL	/HL/
Mid-double-rise	MM	/M/
MH-rising	MH	/MH/
ML-falling	ML	/ML/

I am positing three underlying tones, High /H/, Mid /M/ and Low /L/, to adequately account for the observed melodies. Thus, the two level melodies with the underlying tones /H/ or /M/, reflect the multiple association of one tone to two morae. For instance, the High-level melody is interpreted as having one H tone linked to the first and second morae. The three falling melodies associate sequences of two successive tones, one tone linked to each mora in a root. For instance, the HM-falling melody has an H tone linked to the first mora and an M tone linked to the second. This is consistent with several studies covered in §3.2.

With respect to the two monotonal melodies High-level and Mid-double-rise, there are tokens in which the Tsua speakers produce a phonetically level to slightly falling F0 during the second mora. These tokens do not conform to a particular pattern. This characteristic of the two Tsua monotonal melodies is reminiscent of what Brugman (2009) reports for Khoekhoe as outlined in §3.2.1 for the four monotonal melodies. She writes that while the pitch targets are reached during the first mora, the second mora has a phonetic fall indicating an unspecified root-final tone in her analysis.

A similar phonetic characteristic is reported in Morén and Zsiga (2006) and Zsiga and Nitisaroj (2007) for Thai. The authors state: "During moras with no phonological tone, pitch falls gradually to or within the mid range" (Zsiga and Nitisaroj 2007:347). Given the phonetic properties of the two Tsua monotonal melodies in addition to what has been found in Khoekhoe and Thai, it is plausible that the association of /H/ and /M/ to the second mora is optional in Tsua, leaving the second mora unspecified for tone for some tokens.

The Tsua tone autosegmental representations are proposed in (21). The parentheses for the High-level and Mid-double-rise melodies indicate optionality to account for the tokens where the consultants produce their respective pitches as phonetically level or slightly falling with no specific target during the second mora.

(21) Tsua lexical tone representations

a. Bimoraic sequence in citation forms

High-level	HM-falling	HL-falling	Mid-double-rise	MH-rising	ML-falling
H (H)	H M	H L	M (M)	M H	M L
μ μ	μ μ	μ μ	μ μ	μ μ	μ μ

3.3.11 Differences between Kua and Tsua tone production

A major divergence between Kua and Tsua is the lack of tonal depression in Kua. It is useful at this point to describe other differences between Kua and Tsua tone production, which will set the stage for the analysis of Tsua tonal depression in Chapter 4. The Tsua MH category maps to the LH category in Kua based on numerous cognates (see Table 17 and the Appendix). Fig. 23 shows F0 traces and spectrograms from female Tsua speaker S and male Kua speaker O for the cognate /tan/ 'to stand up' to illustrate the typical F0 shape difference between Tsua MH and Kua LH. In addition, comparative data reveal that the

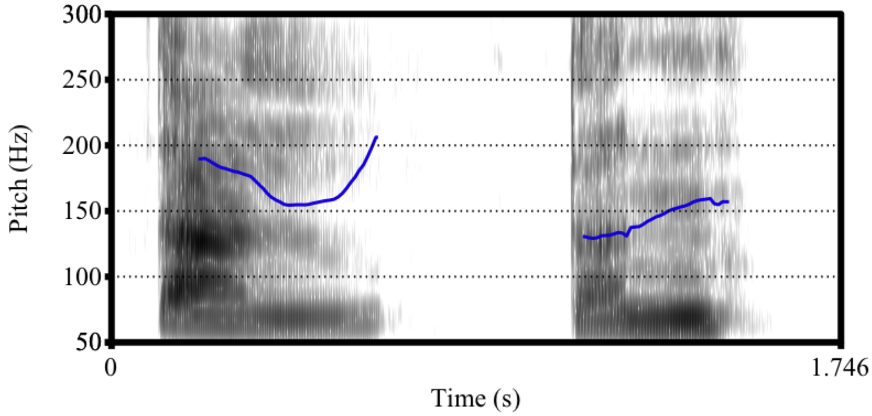


Fig. 23 F0 traces with spectrograms of the cognate verb /tan/ 'to stand up' spoken in citation form by Tsua speaker S (to the left) and Kua speaker O (to the right).

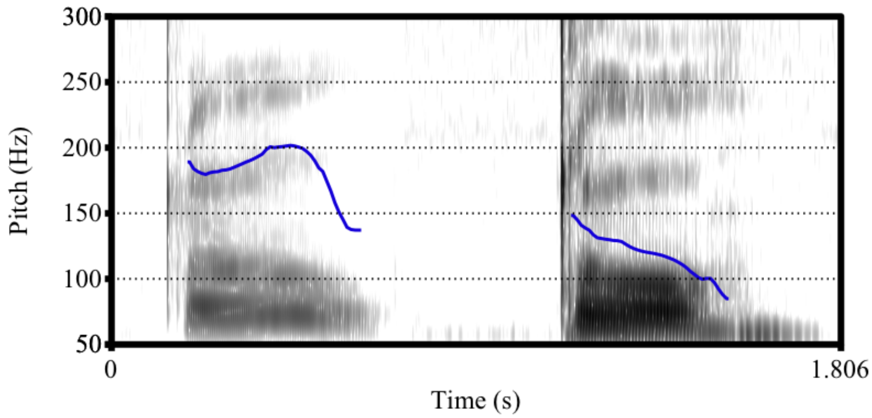


Fig. 24 F0 traces with spectrograms of the cognate noun /ǂqoo/ 'mud; clay' spoken in citation form by Tsua speaker S (to the left) and Kua speaker O (to the right).

Tsua ML category maps to the Kua LL category. Fig. 24 illustrates the difference between the Tsua and Kua productions of their respective tonal categories using the cognate /ǂqoo/ 'mud; clay' as a representative example.

It has been reported that Low tones can be accompanied by breathy voice in G|ui. Nakagawa (2006:97) states: “Words with the /L/ melody or the /LM/ melody are frequently (but not always) pronounced with a distinct phonation type, namely breathy voice. This feature is auditorily clearer in the first half in the case of a word with /LM/. It can therefore be regarded as an optional feature involved in the /L/ tone. Although this feature is not distinctive, it facilitates the /L/ melody to be auditorily more distinct from the /M/ melody.” One question is whether Kua has breathy voice as an optional feature for its LH and LL categories. The next question is whether Tsua has the feature for its corresponding MH and ML categories.

An examination of Kua L tone production for the LL and LH melodies reveals breathy voice for a number of tokens produced by speaker O. For Tsua, the MH and ML melodies, which correspond to LH and LL in Kua, are not produced with breathy voice by any of the consultants. Fig. 25 shows a typical FFT spectrum after the onset of the vowel /a/ in a token of the MH verb /tan/ 'to stand up' produced by Tsua speaker S. The sampling rate is 11025 Hz. The relevant harmonics are labeled: H1 is the first harmonic, aka the fundamental frequency; H2 is the second harmonic; and F1 is the harmonic with the highest intensity in the first formant. At the onset of /a/, H1 is 1.6 dB below H2. Moreover,

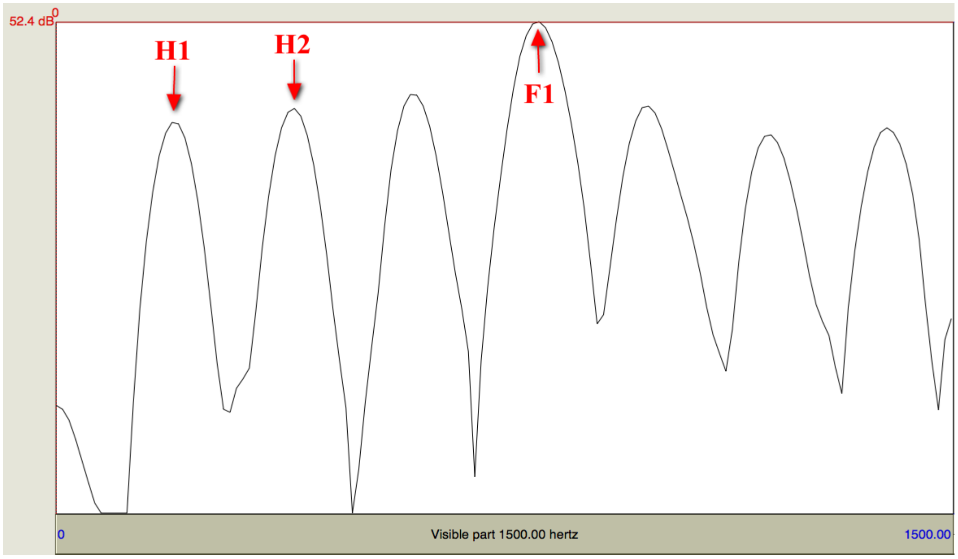


Fig. 25 FFT spectrum of the verb /tan/ 'to stand up' spoken in citation form by Tsua speaker S.

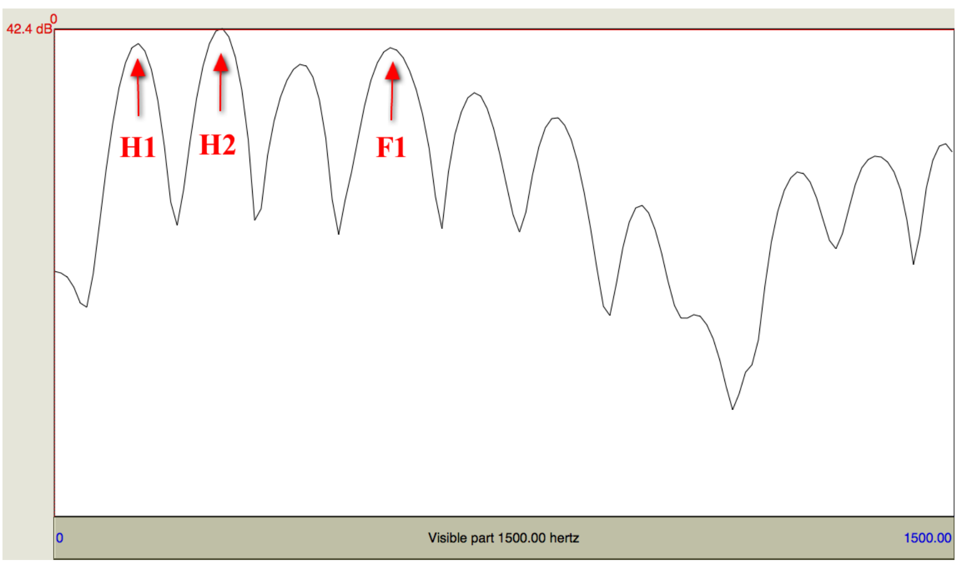


Fig. 26 FFT spectrum of the verb /tan/ 'to stand up' spoken in citation form by Kua speaker O.

H1 is 12.2 dB below F1. The fact that H2 and F1 have more energy compared to H1 demonstrates that the vowel is not breathy voiced, since “vowels with slack or breathy voice have comparatively more energy in the fundamental frequency [first harmonic]” (Ladefoged and Maddieson 1996:317; see also Ladefoged 2003). Contrarily, Kua speaker O produced the LH verb /tan/ 'to stand up' with more breathy voicing: at the onset of /a/, H1 is .5 dB *above* F1 and H1 is 1.9 dB below H2 (Fig. 26). Thus, Kua optionally has Low tones accompanied by breathy voice for LH and LL. The Tsua M tones for MH and ML are not accompanied by breathy voice.

3.4 Summary

The purpose of this chapter was to increase our understanding of the Tsua tonal system through a more detailed examination compared to previous studies of Tshwa subgroup tone. It was argued that Tsua has a six-way tonal contrast with the inclusion of Mid tones in the analysis as opposed to the four-way contrast in Vossen (1997, 2013b). The Tsua Mid tones for these melodies are not produced with breathy voice whereas the root-initial Low tones in Kua and G|ui may be produced with breathy voice. Now that the Tsua tonal categories are established, the next chapter presents a detailed acoustic analysis of tonal depression in Tsua.

CHAPTER 4

4.1 Introduction

The last chapter established six contrastive tonal melodies in Tsua. There are two heretofore unexplored melodies that are the result of a root-initial depressed High tone. The HM-falling and HL-falling melodies become Depressed High-Mid (DH-M) and Depressed High-Low (DH-L) when the root-initial consonant is a voiced obstruent, an aspirated obstruent, or the glottal fricative /h/ (cf. Tables 7 and 10). Tsua is unique with regard to tonal depression. While many languages have depression following voiced or aspirated obstruents, and a few languages by /h/, it is typologically rare for a language to have all three types participate in the interaction of tone and consonant.

Tsua tonal depression is best situated in the context of comparable effects found in the world's languages. Therefore, this chapter is organized in the following fashion. §4.2 gives an overview on consonant-tone interaction in languages of both Africa and Asia. §4.3 presents the tonal depressor effects found in the DH-M and DH-L melodies using F0 instrumental evidence. The chapter concludes with a hypothesis regarding systematic exceptions to tonal depression in §4.4. Predictions made by the hypothesis are tested. The results support the hypothesis. Implications of the results are discussed.

4.2 Consonant-tone interaction in the world's languages

4.2.1 Nguni and Shona group languages

Tonal depression is first described in Beach (1924) for the Southern Bantu Nguni language Xhosa (South Africa; also Botswana and Lesotho). Beach divides syllable-initial consonants into two classes according to their tonal affinities, with one set of consonants associating with the Low tone class, an observation reflecting the interrelationship between consonant types and tone. He writes:

> In Bantu there are numerous examples of the interrelationship of phones and tones, and I hope in another essay to show that all the consonantal initials of Xosa [Xhosa] may be divided into two classes according to their tonetic affinities. For example, in the two Xosa words *into* (meaning "thing") and *imvu* (meaning "sheep"), the syllable *nto* is pronounced slightly higher than the syllable *mvu* because *nt* is a consonant of the first, or high-tone, class, whereas *mv* is a consonant of the second, or low-tone, class...this division of consonantal initials into two classes rests mainly on the presence of [*sic*] absence of *voice* throughout their pronunciation. The fully voiced initials belong mainly to the low-tone class (as *mv* in our example above); whereas the high-tone class contains mainly initial consonants which are at least partially unvoiced (as *nt* above) (Beach 1924:80).

Doke (1926) reports that Zulu, another Southern Bantu Nguni language, has a specific group of segments followed by Low tones while other segments favor High tones.

> Certain Zulu phones (units of speech-sound) naturally favour specific tones. For instance, the voiced explosives, b, d, and ɡ, normally form with vowels low-toned syllables...On the other hand the bi-labial implosive, ɓ, is more often found with high tones. The voiced affricates, too, favour low tones; so also the voiced fricatives, v, z, ɮ, and ɦ...The nasals, m and n, when not used in conjunction with voiced explosives or voiced affricates, favour high tones

> or the middle tone (e.g., n of nɑ), as also does the voiced lateral, l...Voiced clicks favour low tones, nasal clicks usually tones somewhat higher. It seems, therefore, that of consonants, voiced explosives and fricatives, together with sounds in which they play a part, favour low tones; whilst the continuants, which may be used syllabically, favour high tones (Doke 1926:205-206).

It is not until Lanham (1958) when the term "depressor" is introduced to the study of Nguni languages, which specifies a consonant class that triggers a lowering effect on all tones except Low level tones.

> A fairly superficial examination of tonetic data leads to the realization that certain consonants, of which /g/ is an example, have a lowering effect on all except "low level tones"...A conditioning factor associated with consonants of the type which /g/ represents is thus identified from an examination of the tonetic data. Facts associated with this phenomenon of "depressor consonants" which has been identified, are highly pertinent with regard to the identification of tonemes (Lanham 1958:66).

Lanham (1958:70) states that approximately one-third of the total number of consonants in Xhosa act as "depressors" by producing lowered allotones of /H/ and /F/, i.e. phonemic High and Falling tones, respectively. He gives the full set of depressors as: /b, d, dy, g, mh, nh, nyh, v, z, gr, h, dl, dz, j, gc, gq, gx, ngc, ngq, ngx, mb, nd, ng, mv, nz, nj, ndl/ and versions of these consonants that incorporate /w/. Lanham (1960:48) adds that some of the depressor stops are "voiceless up to the breaking of the occlusion...these stops have a fortis release and strong voicing commencing immediately thereafter."

It is later claimed in Schachter (1976) that depressor consonants in the Nguni language Siswati [Swati] form an "unnatural class" because they include phonetically voiced and voiceless consonants: voiceless, unaspirated stops [p, t, k]; voiceless, unaspirated affricates [tf, ts, tʃ]; voiced clicks; voiced fricatives [v, z, ɮ]; the nasal sonorant [ŋ]; and the prenasalized counterparts of depressor consonants (p. 213). Schachter writes: "...the depressor and nondepressor consonant classes of siSwati offer an interesting challenge to one of the fundamental assumptions of modern phonological theory: viz., that the classes of segments that are referred to in phonological rules will in general be natural classes" (p. 215). He appeals to sound change to explain this state of affairs, arguing that depressor consonants formed a natural class of voiced obstruents historically in what he calls "Early" Siswati:

(22) Siswati sound changes (Schachter 1976:214)

	Early	>	**Modern**
a.	b	>	p
b.	d	>	t
c.	g	>	k
d.	dv	>	tf
e.	dʒ	>	tʃ
f.	ŋg	>	ŋ

In addition to reporting tonal depression occurring after voiced and voiceless consonants, Rycroft (1980) claims to have found grammatically conditioned cases of tonal depression in the Nguni languages Swati, Natal Zulu and Ndebele. (23) is an excerpt of the three grammatical conditions.

(23) Grammatically conditioned cases of depression (Rycroft 1980:11)

a. Prefixes of Swati nouns, adjectives and possessive pronouns are affected in this way when undergoing copulative (i.e. predicative) inflection.
b. A noun standing as object in an 'axiomatic negative' construction (i.e. without an object concord in agreement) takes initial depression (plus elision of its first H or F tone, and of its initial vowel if any).
c. With verbs, depression usually occurs as a concomitant of L tone on Subject Concords of the First and Second person, singular and plural, in most Positive Indicative tenses (non-participial). In Remote Past Indicative tenses the tense infix -a- takes initial depression and a rising onset when compounded with a First or Second Person subject concord.

Further research confirms that it is not necessarily the case that all voiced consonants are depressors nor is it necessarily the case that all depressors are voiced. For instance, the phonetic studies in Traill et al. (1987) and Traill (1990) support Schachter's earlier claim by demonstrating instrumentally that the depressor stops are voiceless and without breathy voice at closure or release in Zulu and Swati. What is particularly striking about these studies' results is the depressors often induce *exaggerated* pitch lowering: one female subject in Traill et al.'s (1987) study has an F0 onset difference of ~70 Hz between High and depressed High tones, with the rest of the subjects having a 30-60 Hz differential in Zulu (p. 261). Moreover, Traill (1990:167) reports the onset of a depressed High tone as about 70 Hz lower than a non-depressed High tone in Swati.

Mathangwane (1999) states that the depressors in the Shona group language

Ikalanga consist of voiced obstruents and breathy voice consonants, although some non-depressors are voiced as well. The Ikalanga depressors are transcribed in Mathangwane (1999:175) as: /p^{ɦ}, t^{ɦ}, b, d̤, d, g, v, z, ʒ, ɦ, ts^{ɦ}, tʃ^{ɦ}, bz, dz, dʒ, bg, d^{w}, g^{w}, k^{wɦ}, dz^{w}, z^{w}, w^{ɦ}/. There are four effects in their interaction with tone: (i) block HTS_3 [High Tone Spreading type 3]; (ii) cause H tone delinking, (iii) convert H's to LH rising tones, and (iv) cause tones to be realized lower (Hyman and Mathangwane 1998:208). Mathangwane (1999) also shows that there is a contrast between “non-depressor” aspirated voiceless stops, which do not block High tone spreading and can be followed by a High tone, and “depressor” aspirated stops, which are necessarily followed by a Low tone.

In Nambya, a Shona group language closely-related to Ikalanga, Downing and Gick (2005) argues that there is a contrast between two “f's”: a depressor “f” and a non-depressor “f”, in which High toned verb stems that begin with the non-depressor have a High tone on the first two syllables while High toned verb stems that begin with the depressor pattern with other consonant-initial stems in having a Low tone on the first syllable and a High tone on the second. The paper's phonetic study confirms that the depressor f's are indeed voiceless. Downing (2009:184) summarizes the results of a number of studies by stating: “both the Shona group and the Nguni group have synchronically voiceless depressor consonants, along with more expected voiced depressors. Moreover, there is no evidence that the synchronically voiceless depressor consonants were voiced historically.”

4.2.2 Khoisan languages

Beach (1938:235) describes a characteristic of “Hottentot” [Khoekhoegowab] dialects in which voiced consonants cause tone lowering, arguing that “the parent language of both Nama and Korana had four tones, pronounced very much as in modern Korana...The two lower tonemes...each became subdivided in Nama into two:...the subdivision was produced by voiced consonants.” Beach lists the consonants associated with depressed melodies by stating “of click accompaniments the originally voiced velar click efflux (C*g*), delayed aspiration (C*h*) and the nasal accompaniment (C*n*); of egressive consonants the voiced plosives *b*, *d*, *g*, as well as *ts*, *kx*, *h*, *m* and *n*” (p. 251).

A description of !Xũ was published by Oswin Köhler in 1981, a language considered to be a close relative of Ju|’hoan by König and Heine (2001:2). Köhler notes three basic, underlying tones, High, Mid and Low, but adds two surface tones Extra-high and Extra-low. The Extra-low tone occurs for segmental reasons, as it coincides with pharyngealized or breathy vowels/consonants (Köhler 1981:566-567). König and Heine (2001) has a wordlist of what they term a W2 Northern Khoisan variety, i.e., !Xun of Ekoka. They distinguish four tones: High, Mid, Low and Extra-low (see also König and Heine 2008 for their !Xun concise dictionary with tone markings). Anyanwu and Köhler (2005) reports on an inspection of the wordlist: “Another point of observation is the frequent occurrence of L˩ [Extra-low] tones in the direct neighbourhood of the symbol /h/, which can express either breathiness of a vowel or aspiration of a consonant in the orthography chosen by König and Heine (2001)...there seems to be a correlation between breathiness or aspiration and L˩" (p. 111).

Traill (1985:38-40) shows that certain segments exert a perturbatory effect on tonal contours in !Xóõ, namely, pharyngealized, breathy voiced and glottalized vowels in bimoraic lexical stems. "All the vowel colourings have a depressor effect on the pitch contour, varying from relatively mild...to extreme...it is mostly perfectly clear that, after an initial period of disruption, the pitch contour resumes its normal time course, so that by inspection of the latter part there is never any question about which pitch one is dealing with" (p. 40). He continues by stating "the only forms that have "depressor vowels" are bi-moric lexical stems, and, as noted above, the colourings appear only on V_1, or the first mora. Thus the second mora is always free of any perturbatory influence and always manifests the normal contour."

The goal of a study by Elderkin (1988) is to find evidence for two underlying tonemes in the Northern Khoisan language Žu|hõasi [Ju|'hoan]. His data are taken from the dictionary in Snyman (1975), developing rules for the evolution of the possible tonal patterns from two basic tones. Snyman argues for four tonemes in Ju|'hoan: Hoog [High], Middelhoog [Middle-High], Middellaag [Middle-Low] and Laag [Low]. Elderkin asserts that the High and Low tones occur at very low frequency in the dictionary and can be derived from the Middle-High and Middle-Low tones, a notion that might be extended to Khoisan languages more generally. For instance, by counting the number of occurrences of pharyngealization with different tone patterns, Elderkin writes: “I am prepared to take these figures as indicating that pharyngealization has a lowering effect on a following high tone” (p. 134). He continues: “Throughout the dictionary there are forms which seem to show that it would be pleasant if there was a greater equivalence between certain click accompaniments, pharyngealization, tone lowering and h and x” (p. 142).

Haacke (1999) considers tonogenesis in Khoekhoe through tonal depression and devoicing using a database that became later published as *A Khoekhoegowab Dictionary* (Haacke and Eiseb 2002). The term “tonogenesis” may be concisely defined as the “development of pitch as a contrastive feature” (Schuh 1978a:228). Haacke (1999:57) notes that the co-occurrence of certain consonants with particular melodies is not as consistent as what is maintained in Beach (1938), as Haacke finds counter-examples to some of Beach's co-occurrence claims. Regarding segments that cause depression, “In the case of the glottal voiceless fricative [h], depression is consistent. No manifestations of a co-occurrence of initial [h] with /22/ and /32/ exist in the database, only with the depressed melodies...the bilabial nasal [m]... corroborates Beach: It has been a depressor” (Haacke 1999:59). Haacke goes on to state that [kh/kx^h] and [ts^h] “are without exception attested to be depressors” and “The release consisting of a delayed glottal fricative...very consistently functions as a depressor, similar to the ordinary glottal fricative *h*” (pp. 61-62).

Naro is also affected by tonal depression, based on Haacke's examination of Visser (2001) and his own recordings (Haacke 2008:161). Haacke asserts “...a voiced consonant has triggered the lowering of the tonemes by way of a low-level phonetic process. But where the consonants have not lost voicing as distinctive feature yet, the tonological change is merely allotonic, not tonemic...Thus, the tonogenetic process is not complete in Naro" (pp. 162-163). Moreover, Hai‖om and ǂAkhoe, dialects of Khoekhoe from northern Namibia, "have undergone complete tonogenesis, that is, depression with complete transphonologisation of the voicing distinction to a tonological distinction" but "As depression has not occurred in !Gora [Korana], the extra-low tone /1/ does not occur" (pp. 172-173).

4.2.3 African languages beyond Nguni, Shona and Khoisan

Consonant-tone interaction is a widespread phenomenon across the African continent judging from the number and variety of descriptions over the years. The effects themselves are diverse. Bradshaw (1999:44-45) and Tang (2008:25-26) give comprehensive lists of languages with consonant-tone interaction, the language-specific effects and the sources, which have been partially reproduced here to give the reader a sense of its scope across Africa. Voiced obstruents serve as depressors in all of these cases with the exception of the Kru language Bassa, in which /h/ has an affinity for L tones (Hobley 1964).

Table 20 Consonant-tone interaction in African languages (based on Bradshaw 1999 and Tang 2008)

Afro-Asiatic, Chadic		
Language	**Location**	**Effect**
Ngizim	Nigeria	H spread blocked & L spread conditioned by voiced obstruents (Schuh 1971)
Bade	Nigeria	H spread blocked by voiced obstruents (Schuh 1978b)
Ouldeme	Cameroon	"syllables with a depressor consonant take L" (Swackhamer 1991; de Colombel 1986)
Mulwi	Cameroon, Chad	L inserted after voiced obstruents (Tourneux 1982)
Bolanci	Nigeria	H spread blocked by voiced obstruents (Lukas 1969)
Miya	Nigeria	H docking blocked by voiced obstruents (Schuh 1998)
Niger-Kordofanian, Niger-Congo		
Adamawa-Ubangi		
Suma	Central African Republic	L inserted after voiced obstruents before a H tone (Bradshaw 1995)
Gbaya ɓokota	Central African Republic	Docking of associative H blocked by voiced obstruents (Bradshaw, field notes)
Kwa		
Ewe	Ghana, Togo	L docking conditioned by voiced obstruents in singular imperative (Ansre 1961; Smith 1968; Stahlke 1971)
Ebrié	Ivory Coast	H spread blocked after voiced obstruents (Kutsch Lojenga 1985)
Siya	Ghana	H spread blocked after voiced obstruents (Ford 1986)

Gur		
Dagara-wule	Burkina Faso, Ghana	H spread blocked by voiced obstruents (Somé 1998)
Kru		
Bassa	Liberia	Voiced obstruents and /h/ have an affinity for L (Hobley 1964)
Benue Congo, Narrow Bantu		
Makaa	Cameroon	In associative construction with a final depressor, an L toned vowel is epenthesized or a downstep triggered if other conditions are met (Heath 1991)
Yaka	Central African Republic, Rep. of Congo-Brazzaville	L inserted after root-initial voiced obstruents, L triggers voicing (Kutsch Lojenga 1998)
Kalanga'a	Botswana	Voiced obstruents and aspiration have an affinity for L tone (Downing and Gick 2005)
Benue Congo, Narrow Bantu, Mijikenda		
Digo	Kenya, Tanzania	H spread, H shift and H docking blocked by voiced obstruents (Kisseberth 1984)
Chichonyi	Kenya, Tanzania	Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)
Chiduruma	Kenya, Tanzania	Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)
Chidzihana	Kenya, Tanzania	Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)
Chikambe	Kenya, Tanzania	Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)

Chikauma	Kenya, Tanzania	L inserted and downstep triggered after voiced obstruents; Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)
Chirabai	Kenya, Tanzania	Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)
Chirihe	Kenya, Tanzania	L inserted and downstep triggered after voiced obstruents; Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)
Kigiryama	Kenya, Tanzania	Prefix H fails to shift/spread when verb has initial voiced obstruents (Cassimjee and Kisseberth 1992)

4.2.4 Asian languages

There may be a parallel with findings from Southeast Asian tone languages where earlier consonant voicing became neutralized, being replaced by associated pitch cues (Rycroft 1980:3). Claims regarding the effects of consonants on tone in Asian languages are found in Karlgren (1915) and subsequent work on Chinese. Henderson (1952:151) reports on the effect of phonation type on pitch registers in Cambodian, writing: “The characteristics of the first register are a “normal” or “head” voice quality, usually accompanied by relatively high pitch...The characteristics of the second register are a deep rather breathy or “sepulchral” voice, pronounced with lowering of the larynx...Pitch is usually lower than that of the first register in similar contexts."

Several Indo-Aryan languages such as Standard Hindi have a 4-way contrast among voiceless unaspirated, voiceless aspirated, voiced unaspirated and voiced aspirated plosives (Bhaskararao 1999:337). Of these sounds, voiced unaspirated and voiced aspirated plosives depress the fundamental frequency of a following vowel (ibid:339, citing Ohala 1974 and Kagaya and Hirose 1975). Tonogenesis has occurred via F0 lowering by voiced aspirated plosives in the South-Asian languages Panjabi and Dogri (Bhaskararao 1999:339). Sun (2003:43) reports tone lowering by aspirated onset consonants in two Tibetan dialects: Qiuji and Tiebu. In Qiuji, tone lowering takes place in long rhymes with either breathy or aspirated onset consonants. In Tiebu, tone lowering occurs in open syllables with aspirated onset consonants.

Xu and Xu (2003) studies Mandarin Chinese to clarify the effect of consonant aspiration on the following F0. There are five factors involved in the experiment: consonant, lexical tone, syllable position, tonal context and carrier sentence. The main effects of consonant aspiration and lexical tone, as well as their interaction, are all significant, with a mean difference of about minus 20 Hz for aspirated /t^ha/ versus unaspirated /ta/. The magnitude of the consonant aspiration effect on F0 varies according to the tone itself and the preceding tone. When the preceding tone is High or Rising, the F0 onset is much higher for /ta/ than for /t^ha/. In addition, the effect of aspiration on F0 is greater for the Rising and Low tones than for the High and Falling tones.

In a paper investigating the phonetic realization of the contrastively aspirated affricates /ts^h/ and /dz^h/ in Nepali, Clements and Khatiwada (2007:631) finds the F0 of /ts^ha/ lower than that of /tsa/ by an average of more than 20 Hz for two speakers. The F0 of /dz^hi/ is an average of 13 Hz lower for one speaker and 20 Hz for a second compared to /dzi/ (p. 632). Furthermore, in a preliminary report on the Leng-shui-jiang dialect of Chinese, Caicai (2009) calculates a significantly lower F0 for two out of three tone groups following aspirated obstruents for six speakers (three male, three female). Table 21 summarizes additional findings on Asian languages with consonant-tone interaction.

Table 21 Consonant-tone interaction in Asian languages
(based on Tang 2008)

Sino-Tibetan, Tibeto-Burman		
Language	**Location**	**Effect**
Jingpho	Myanmar	Voiced obstruents and /ɦ/ have an affinity for L tone (Maran 1971)
Manange	Nepal	Aspiration has an affinity for L tone (Hildebrandt 2003)
Sino-Tibetan, Wu Chinese		
Longyou dialect	China	Voiced obstruents have an affinity for L tone (Cao 2002, Zhang 2002)
Shaoxing dialect	China	Voiced obstruents and /h/ have an affinity for L tone (Jixeng Zhang 2006)
Songjiang dialect	China	Voiced obstruents have an affinity for L tone (Bao 1999)
Wenling dialect	China	Voiced obstruents have an affinity for L tone (Bao 1999)
Wenzhou dialect	China	Voiced obstruents have an affinity for L tone (Rose 2002)
Wujiang dialect	China	Voiced obstruents and aspiration have an affinity for L tone (Shen 1994)
Wuyi dialect	China	Voiced obstruents have an affinity for L tone (Bao 1999)
Tai-Kadai		
Kam-Tai		
Kam	China	Aspiration has an affinity for L tone (Long and Zheng 1988)
Mulao	China	Aspiration has an affinity for L tone (Moreton 2006)

Indo-European		
Indo-Aryan		
Kangri	India	/h/ has an affinity for L tone (Eaton 2007)
Austronesian		
Oceanic		
Yabem	Papua New Guinea	Voiced obstruents have an affinity for L tone (Hansson 2004)

4.3 Evidence for DH-L and DH-M in Tsua

We now turn to evidence for tonal depression in Tsua. Recall from §2.2.1 the set of Tsua depressors: /b, d, dz, ɟ, g, ɢ, p^h, t^h, ts^h, c^h, k^h, q^h, h, g|, gǂ, g‖, |ɢ, ‖ɢ, $|^h$, $ǂ^h$, $‖^h$, $|q^h$, $ǂq^h$, $‖q^h$/. There are two melodies with root-initial depressed High tones: DH-L and DH-M. A DH-H melody is not found in the current data set, an absence due to a cooccurrence restriction in which depressors do not occur with HH. This issue is further discussed in §4.4.1. Henceforth, the DH-L and DH-M tonal transcriptions are marked with an 'x' symbol [͓] below the vowel with the depressed H tone, i.e. [á͓à] is DH-L and [á͓ā] is DH-M. Evidence in the form of cognates between G|ui, Kua and Tsua are in the Appendix.

The depressed H tones in Tsua are consistent with the observations in Lanham (1958:74) for Xhosa: "...in passing from a depressor consonant which introduces the syllable, to the immediately succeeding vowel, there is release from the effect of the

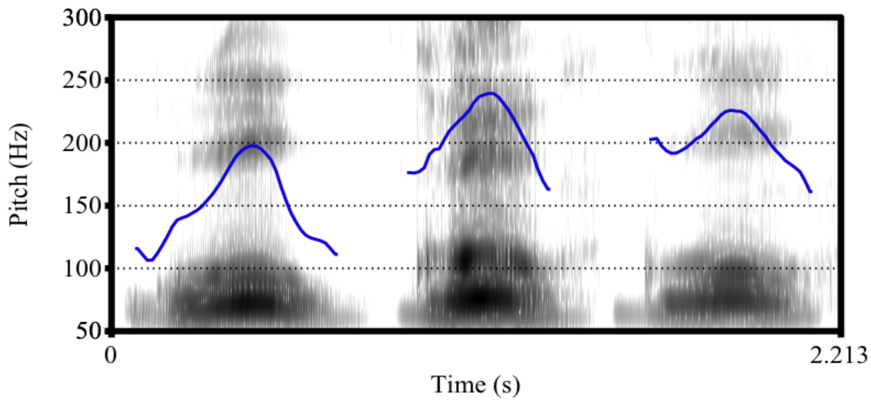


Fig. 27 F0 traces with spectrograms of the DH-L noun: [g͓ǿò] 'aardvark' spoken in citation form by speakers S, B and M displayed from left to right.

consonant and a consequent striving to achieve the required pitch level of /H/...It is clear that under these circumstances a rising glide results and a glide of this kind coinciding with a depressor is a certain indication of /H/. /H/ can thus be realized by an allotone of the shape [/] at various pitch levels."

Typical DH-L contours are seen in Fig. 27 for the three Tsua consultants. These examples are productions of the noun [g͓ǿò] 'aardvark'. For speakers S and B, there is an F0 decrease of 80 Hz or more at the onset of the H tone realization compared to a typical non-depressed H tone. Speaker M has a more modest F0 decrease of about 40 Hz in this example at her pitch contour onset. The contours rise to their respective H tone targets. After the F0 peaks, they fall sharply to the L tone targets.

Fig. 28 is the landmarks plot. (24) is the list of lexical items used to generate Fig. 28. The contour rises from 168 Hz to a peak of 215 Hz at normalized time points 10-11.

After the F0 peak there is a fall to the 145 Hz F0 offset.

(24) F0 averaged DH-L lexical items

Lexical Item	**Gloss**
gó͓ò	'aardvark'
g|á͓à	'Silver tree'
tsʰó͓è	'person'
ɟó͓à	'ash'
tsʰó͓ò	'to pick up'

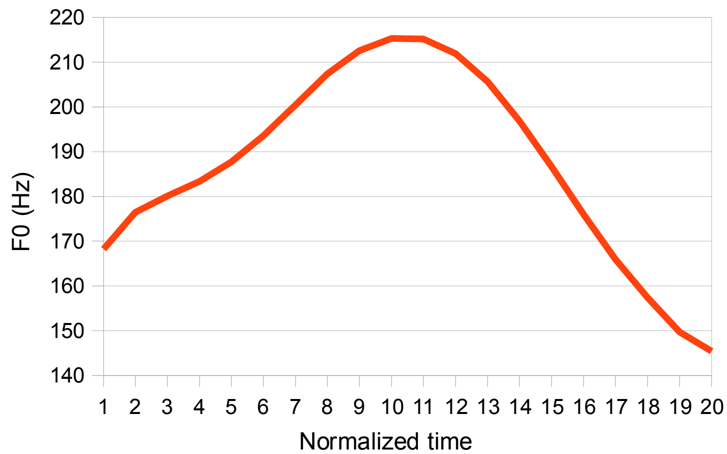


Fig. 28 DH-L melody time-normalized and F0 averaged across speakers S, B and M (n = 45).

DH-M F0 tracings are seen in Fig. 29. In general, the DH-M contour has an F0 decrease in the realization of the first part of the initial tone of about 40-50 Hz compared to the initial tone of the non-depressed HM-falling tonal melody. F0 rises to the H tone target before it falls to the M tone target. There are two characteristics that distinguish DH-L from DH-M: (a) the DH-L F0 onset has a lower hertz value compared to DH-M for some

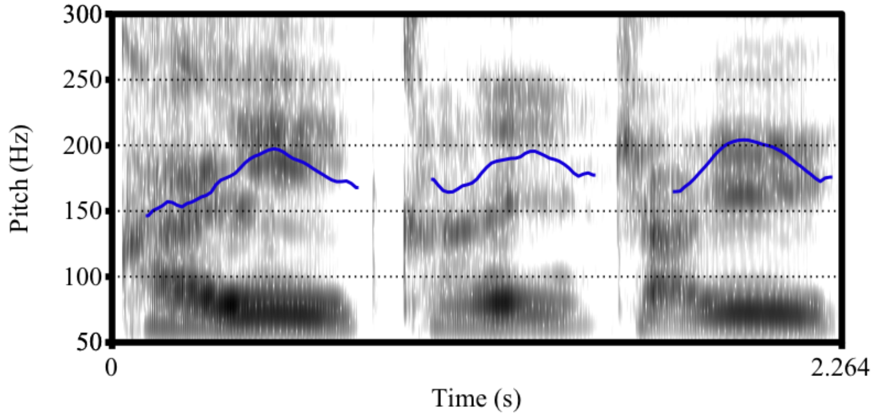


Fig. 29 F0 traces with spectrograms of the DH-M verb: [kʰá̰ē] 'to stab' spoken in citation form by speakers S, B and M displayed from left to right.

tokens and (b) the F0 offset value is higher for DH-M as opposed to DH-L, with the F0 offsets being typical for Mid and Low tones as described in Chapter 3.

The landmarks plot in Fig. 30 has an F0 onset of 171 Hz. There is a short fall to the F0 minimum of 167 Hz at normalized time point 3 before a rise to the F0 peak of 197 Hz at normalized time point 14. A moderate fall to the F0 offset of 176 Hz at normalized time point 20 occurs thereafter.

(25) F0 averaged DH-M lexical items

Lexical Item	**Gloss**
kʰá̰ē	'to stab'
tsʰá̰ā	'water'
tsʰá̰ō	'to dig'
gó̰ā tsúrī	'much later'
g|á̰ā	'to put in smoothly (into sand/fire)'

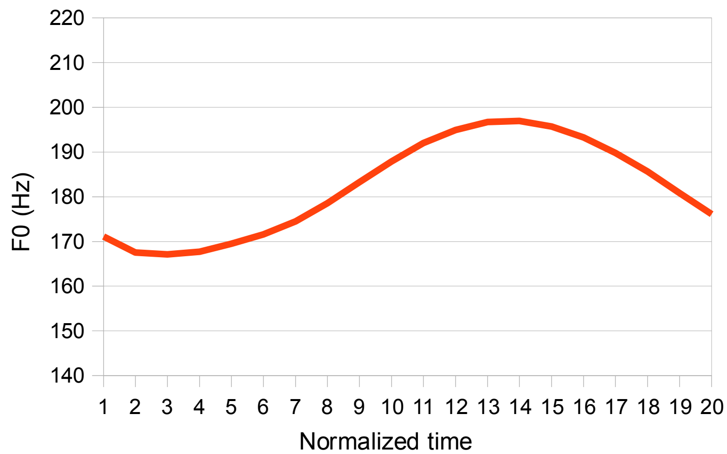


Fig. 30 DH-M melody time-normalized and F0 averaged across speakers S, B and M (*n* = 45).

Fig. 31 shows superimposed F0 traces of HM [ǁáē] 'to chew' with DH-M [gǁá̽m̄] 'thorn'. Fig. 32 compares HL [|χóò] 'gemsbok' with DH-L [gó̽ò] 'aardvark'. The four lexical items are produced by speaker S. The figures are a striking depiction of depressed versus non-depressed contours, thus illustrating how crucial data visualization is for the analysis of Tsua tonal melodies. The F0 lowering for the depressed melodies is apparent at their respective onsets. The curves converge and fall to their F0 offsets. The fact that the curves converge from their midpoints until the F0 offsets confirms the root-final M and L tone interpretations, respectively. It is noteworthy that the F0 lowering is more pronounced for the DH-L melody (100+ Hz) than the DH-M melody (50 Hz) for these particular tokens.

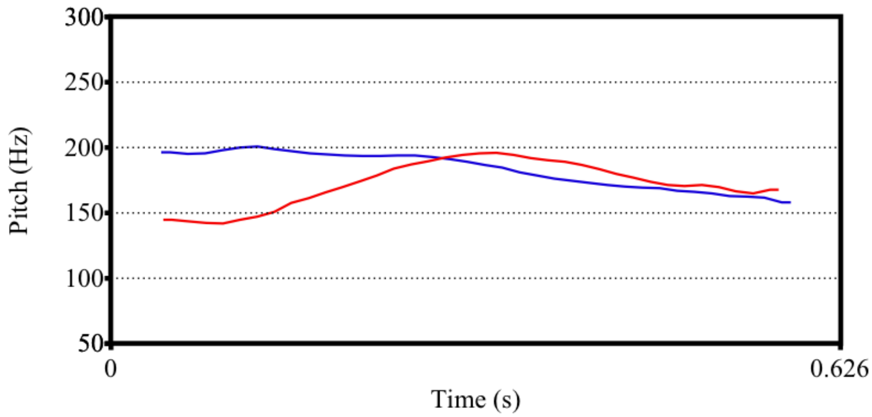


Fig. 31 F0 traces of HM [‖áē] 'to chew' (blue) and DH-M [g‖á̽m̄] 'thorn' (red) to illustrate H tone depression in Tsua as produced by speaker S. F0 lowering is realized at the start of pitch onset. The curves converge and fall to similar F0 offsets.

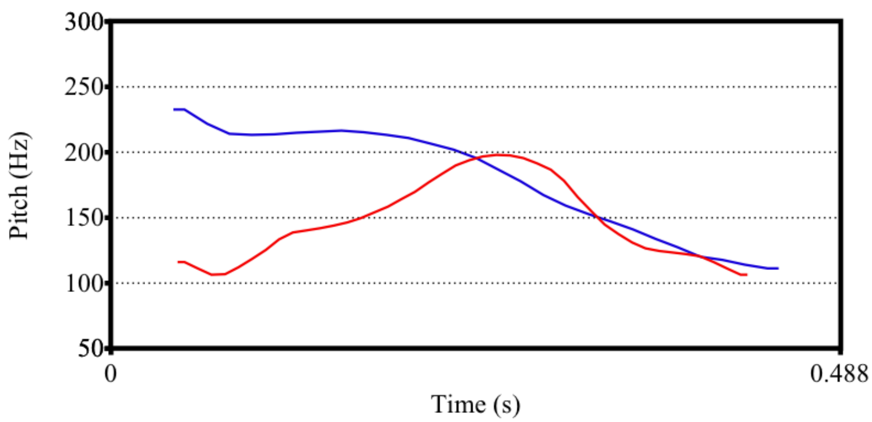


Fig. 32 F0 traces of HL [|χóò] 'gemsbok' (blue) and DH-L [gó̽ò] 'aardvark' (red) to illustrate H tone depression in Tsua as produced by speaker S. F0 lowering is realized at the start of pitch onset. The curves converge and fall to similar F0 offsets.

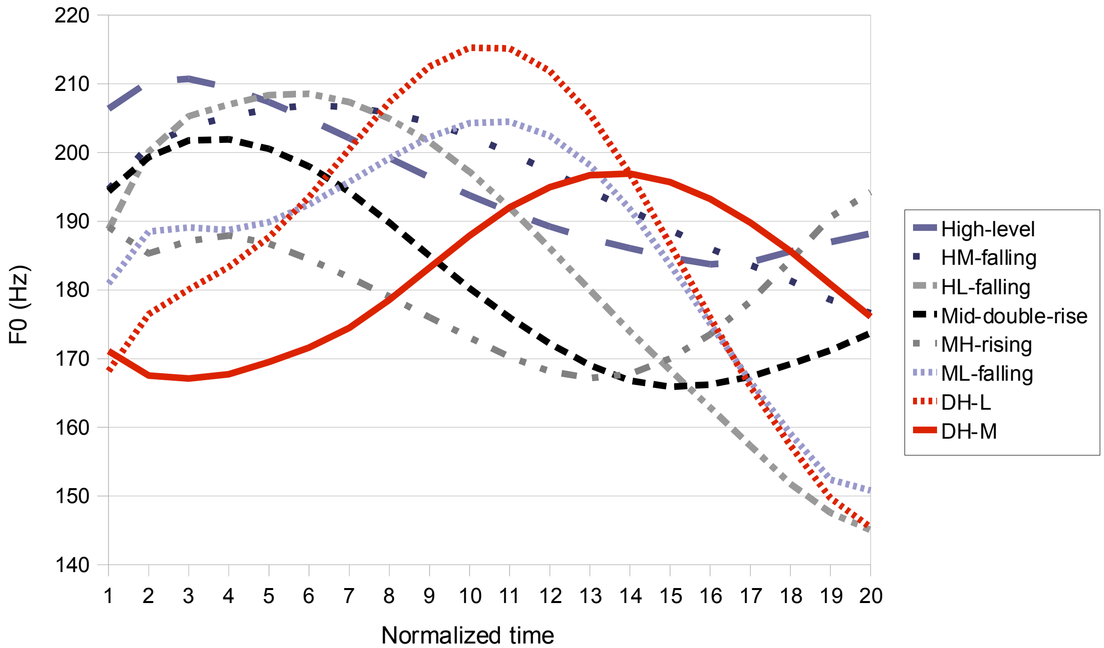


Fig. 33 The tonal space of the six non-depressed and two depressed time-normalized and F0 averaged tone melodies across speakers S, B and M (n = 45 per tone melody). The two depressed melodies have the lowest F0 onset values compared to the six non-depressed melodies. Thus, a root-initial depressed High tone is lower than a root-initial Mid tone.

The tonal space of the six non-depressed plus two depressed time-normalized and F0 averaged melodies across the three Tsua speakers is shown in Fig. 33. The DH-L contour is indicated by the dashed red line. The DH-M contour is shown with the solid red line. Both F0 traces have an onset value of about 170 Hz, an onset value that is the lowest of all the tonal melodies. In other words, a root-initial depressed High tone is lower than a root-initial Mid tone. The DH-L curve rises while the DH-M curve falls from their onsets. The DH-L curve reaches an F0 peak sooner than the DH-M curve, possibly because the DH-L melody has to fall more in an equivalent amount of time to reach its F0 offset. The

DH-L's offset value is quite similar to the HL-falling and ML-falling offset values, thereby confirming the root-final Low tone analysis. The DH-M's offset value is almost exactly the same as the Mid-double-rise and HM-falling offset values. I interpret this as further confirmation that the root-final Mid tone analysis is appropriate.

The tonal space plot in Fig. 33 confirms that the melodies may be distinguished according to the F0 properties: (a) F0 onset, (b) F0 maximum and F0 maximum timing, (c) F0 minimum and F0 minimum timing, and (d) F0 offset. In some cases, a melody's primary F0 properties can have equivalent values. For instance, the F0 minimum and F0 offset for the HL-falling melody are the same value temporally and in terms of frequency (145 Hz). Therefore, the Tsua tone system with its complex curves can be understood in terms of the F0 properties and the concomitant timing relations between them.

4.4 Tonal depression exceptions

4.4.1 The HH and MH tonal split

Tonal depression is not marginal in the grammar because there are numerous examples of DH-L and DH-M. Contrarily, we see a general avoidance of depressor consonants co-occurring with the High-level melody (i.e. HH). There are only 11 HH lexical items with root-initial depressors in the current data set. Interestingly, these are depression exceptions, i.e. the initial High tone is not depressed on the surface, a topic that is addressed in §4.4.3. In other words, there is no evidence of a DH-H melody in the current data set. This is perhaps unsurprising, given that co-occurrence restrictions between tone and onset consonants are often manifested in tone languages where F0 differences indicate lexical contrasts (Chen 2011:612).

For Tsua, the co-occurrence restriction may be the result of a historical tone split. A "tone split" can be defined as a process which results in the multiplication of the number of tones in a language (Maddieson 1974:206). G|ui and Kua HH cognates map to Tsua HH cognates but sometimes map to Tsua MH cognates as well (see the Appendix). A comparable split has been observed in Khwe. Khwe HH and MH often correspond to HH in G|ui (E.D. Elderkin, personal communication). The issue is that there is not an obvious pattern to the tonal splits in Tsua or Khwe. There are root-initial depressors (n = 73) and non-depressors (n = 115) for the Tsua MH lexical items. It may have been the case that depressors caused HH lexical items to surface with an F0 rising contour over the course of Tsua's evolution through lowering. Then, the F0 rising contour merged with the extant MH-rising category. This explanation is plausible. However, it must remain an open question in light of the lack of historical evidence.

4.4.2 /b/ as a non-depressor in C_2 position

There are exceptions in which depression does not occur despite the presence of the conditioning environment. One circumstance is when the depressor /b/ occurs in C_2 position. Recall that §2.3.1 outlines the Tsua phonotactic restrictions on consonants in which /b/ is the only voiced stop able to appear phonologically in C_2 position. It is argued that [d] is an allophone of /r/ in C_2 position when followed by /i/ or /u/. All clicks are restricted to C_1 position. Thus, it is possible that a root-final H tone could be a depression target when following the voiced obstruent /b/. Rycroft (1980:7) provides Swati and Zulu examples of depressed tones that are not root-initial.

One way to test whether a root-final H tone can be depressed is to check the High-level melody, which does occur with /b/ in C_2 position. The MH-rising melody is not a good option because the F0 dip in the middle of the contour happens with or without a depressor in C_2 position, making the results difficult to interpret. The High-level melody, on the other hand, is the most level vis-à-vis F0 compared to the other melodies and may show evidence of a significant decrease in pitch when /b/ is in C_2 position.

The first step is to confirm the depressor effect of /b/ in C_1 position. Fig. 34 presents spectrograms and F0 traces for the noun [bó̰ò χúˤm̀] 'baboon'. Clearly, the initial H tone is depressed (cf. Fig. 27).

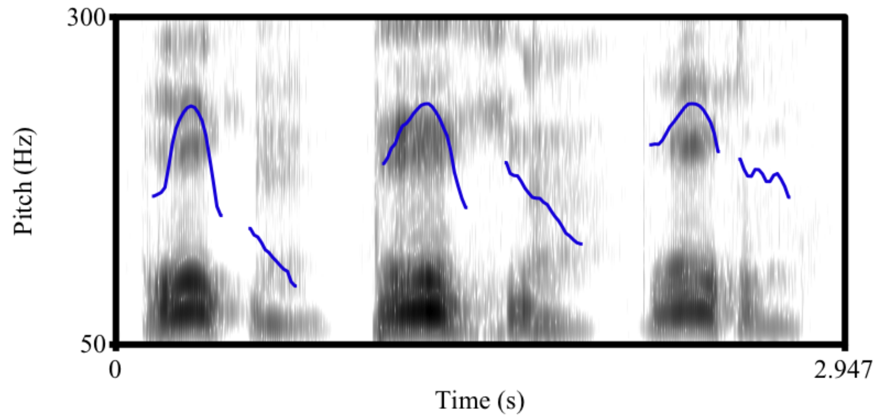


Fig. 34 F0 traces with spectrograms of the DH-L noun [bǫ̀ò χúˤm̀] 'baboon' spoken in citation form by speakers S, B and M displayed from left to right. The F0 trace resembles the DH-L contour in Fig. 27, confirming the DH-L interpretation.

There are not many High-level examples where /b/ is in C_2 position in the current data set. None of the examples in the data set exemplify the typical characteristics of tonal depression — Fig. 35 are F0 traces and spectrograms of the High-level noun /qúbú/ 'blister'. Speakers S and B have phonetic F0 perturbation of 10-15 Hz in the expected direction but speaker M's production shows little if any F0 decrease at the onset of /b/. These results provide evidence that Tsua tonal depression is limited to root-initial position.

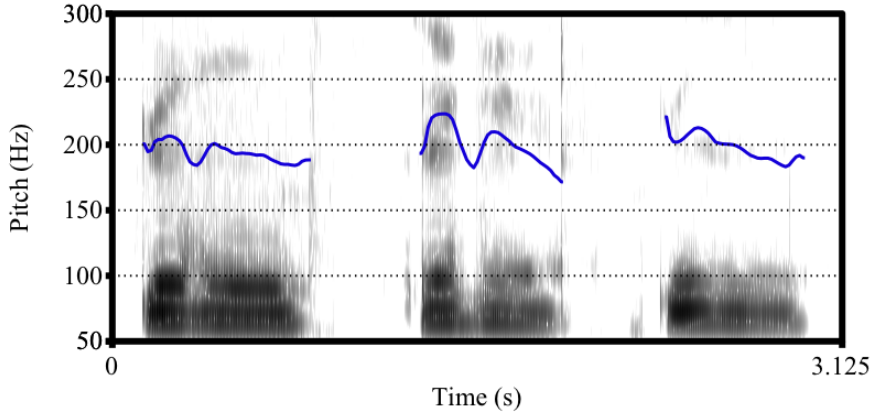


Fig. 35 F0 traces with spectrograms of the High-level noun /qúbú/ 'blister' spoken in citation form by speakers S, B and M displayed from left to right. There is slight F0 perturbation of 10-15 Hz after the production of [b] for speakers S and B. Speaker M has little F0 perturbation in her production. These results provide evidence that Tsua tonal depression is limited to root-initial position.

4.4.3 The relevance of click replacement

An important finding with respect to Tsua tonal depression involves click replacement. Table 22 gives a breakdown of exceptions in the current data set by depressor type. Roots such as /g‖áò/ 'old; old person; to age' were counted once, especially if they were used in compounds, e.g. /g‖úá/ 'stone' versus /ǂʰūū hĩ g‖úá/ 'limestone'. Table 22 lists the depression exceptions count in the first data row. The second data row labeled 'depression expected' is the total number of roots where tonal depression is expected to occur, including the exceptions. Tonal depression exceptions for root-initial H tones constitute 19.3% of the data (Table 22). The 41 total exceptions are broken down by tone melody (Table 23). The 171 roots with tonal depression are broken down by tone melody (Table 24).

Table 22 Root counts of tonal depression exceptions by depressor type

	voiced obstruents	aspirated obstruents	/h/	Total
depression exceptions	33	4	4	41
depression expected	112	88	12	212
				19.3%

Table 23 Root counts of tonal depression exceptions by tone melody

	voiced obstruents	aspirated obstruents	/h/	Total
HH	8	1	2	11
HM	13	3	1	17
HL	12	0	1	13
				41

Table 24 Root counts of tonal depression by tone melody[23]

	voiced obstruents	aspirated obstruents	/h/	Total
DH-M	46	56	4	106
DH-L	33	28	4	65
				171

23 There is no evidence of a DH-H melody as explained in §4.4.1.

The voiced obstruents type has the highest number of depression exceptions with a total of 33. This equals a 29.5% exception rate for that depressor type when divided by the number of roots with root-initial voiced obstruents where depression is expected at 112. It is informative to consider why so many exceptions occur with voiced obstruents by inspecting the cognates between G|ui, Kua and Tsua, which gives an areal approximation of the West to East click replacement patterns as briefly put forth in Traill (1978) and described in the Introduction to Chapter 2. Many of the voiced stops in Tsua have nasal counterparts in G|ui and Kua as observed in Table 25 on the next page. Tsua click replacement for these systematic exceptions comes in three forms: (i) nasalization loss, (ii) click and nasalization loss and (iii) click and nasalization loss plus fronting. Table 25 lists the exceptions for voiced obstruents sorted by click replacement type.

Table 25 G|ui, Kua and Tsua correspondences for the voiced obstruents tonal depressor exceptions by click replacement type where comparative data are available (n = 21).[24]

G|ui[1]	Kua[2]	Tsua	Gloss
ŋǁúá	ŋǁúá	gǁúá	'rock; stone'
ŋǁûì	ŋǁûì	gǁûì	'oil'
ŋǁóē	ŋǁóē	gǁóē	'moon'
ŋǁûì	ŋǁûì	gǁûì	'fat'
ŋǁódì[3]	ŋǁóò rī	gǁóò rì	'nephew; niece'
ŋǁáò	ŋǁáò	gǁáò	'old; old person; to age'
ŋǁárò	-	gǁálò	'to carry something (on the shoulder)'
ŋǁáé	ŋǁáé	gǁáé	'to count; to read'
-	ŋ|éè	g|éè	'there (DEM)'
ŋ!óō	ŋúō	góō	'to be silent; to shut up'
ŋú[4]	ŋgû[4]	gúù	'far; to go far away'
ŋ!árò	ŋárò	gárò	'chameleon'
ŋ!óē	ŋóē	góē	'porcupine'
ŋ!aro	ŋárō	gárō	'to clot; to coagulate; to freeze'
ŋ!úrī	-	gúrī	'to spoil; to damage; to ruin something'
ŋ!áò	ŋáò	gáò	'to hide; to bury'
ŋ!áé	ŋáé	gáé	'to pass by'
ŋ!úū	ŋúū	dzúū	'hut'

24 It is not clear whether Tsua gúù 'far; to go far away' involves click replacement based on the comparative data given here.

ŋǂúú	ŋǂúú	ɟúú	'black'
ŋǂórō	ŋǂórō	ɟórō	'back; backbone'
ŋǂóē	ŋǂéē	ɟúē	'to deny; not to trust; to oppose'

[1]Nakagawa et al. (2014) [2]Chebanne and Collins (2014) [3]Tanaka (1978) [4]Vossen (1997)

Nasalization loss is apparent from G|ui to Kua to Tsua or from Kua to Tsua in the first eleven examples in Table 25. It is striking to observe that the dental [|] and lateral [ǁ] clicks are the most stable out of the four Tsua click types, being preserved in all of the cases. This observation is consistent with the claim that of the four basic clicks in Khoe languages, only the non-affricated (alveolar and palatal) clicks are replaced by clickless consonants while the affricated (dental and lateral) clicks remain stable (Traill 1986). Looking at the first example 'rock; stone', click replacement via nasalization loss is exemplified as:

(26) Tsua click replacement via nasalization loss

G|ui (West) ➞	**Kua ➞**	**Tsua (East)**	
ŋǁúá	ŋǁúá	gǁúá	'rock; stone'

Click and nasalization loss are evident in six of the exceptions. All six examples involve alveolar click replacement. We can compare /ŋ!óē/ 'porcupine' in G|ui with the Kua and Tsua cognates /ŋóē/ and /góē/, respectively.

(27) Tsua click replacement via click and nasalization loss

G|ui (West)	➞	**Kua**	➞	**Tsua (East)**	
ŋ!óē		ŋóē		ɡóē	'porcupine'

The third type of click replacement, click and nasalization loss plus fronting, is apparent in the cognates for 'hut'. We might expect the root-initial velar nasal consonant [ŋ] to correspond to the voiced velar stop [ɡ] in Tsua, following the pattern seen in (27). However, the Tsua correspondent is the voiced alveolar affricate [dz]:

(28) Tsua click replacement via click and nasalization loss plus fronting

G|ui (West)	➞	**Kua**	➞	**Tsua (East)**	
ŋ!úū		ŋúū		dzúū	'hut'

Independent evidence for Tsua nasalization loss was found after a careful review of Vossen (1997), which has a comparative word list from various Kalahari Khoe East languages. Cognates from the Kalahari Khoe East languages |Xaise, Deti, Cara, Ts'ixa and Danisi as compared to Tsua are listed in Table 26. The Tsua examples have root-initial voiced consonants corresponding to root-initial nasal consonants in the rest of the Kalahari Khoe East languages. For instance, Deti, Cara, Ts'ixa and Danisi all have a root-initial, nasalized lateral click for the verb /ŋ||ai/ 'to sing'; on the other hand, the Tsua cognate /ɡ||ai/ 'to sing' has a root-initial voiced click.

Table 26 Root-initial nasal consonants in Kalahari Khoe East languages compared to Tsua root-initial voiced consonants, based on Vossen (1997)[25]

|Xaise	Deti	Cara	Ts'ixa	Danisi	Tsua	Gloss
ŋgabe	ŋabe	ŋgabe	ŋgabe	ŋgabe	ɡabe	'giraffe'
ŋgao	-	ŋgao	-	ŋgao	gao	'Acacia tree'
ŋ|u	ŋ|nu	ŋ|u	ŋ|u	ŋ|u	ɡ|u	'leg'
ŋ‖obu	ŋ‖nubu	ŋ‖obu	-	ŋ‖xobu	ɡ‖ubu	'to shake'
-	ŋ‖ai	ŋ‖ai	ŋ‖ai	ŋ‖ai	ɡ‖ai	'to sing'
ŋgabi	ŋabi	ŋgabi	-	-	gabi	'to turn over'

The data in Tables 25 and 26 suggest that the nasal form is the earlier form if we use the comparative method of linguistic reconstruction (Crowley and Bowern 2010). For instance, the sound change ŋ! > ɡ occurred historically in Tsua. Tsua has been less conservative in terms of the historically nasal consonants compared to the other languages in Tables 25 and 26. Thus, a diachronic sound change in Tsua involved nasalization loss and in some cases the root-initial consonants underwent fronting. These roots did not originally have depressed H tones, since nasalized clicks and nasal non-clicks did not trigger tonal depression historically and do not involve depression synchronically. These particular H tones never became depression targets. The lack of F0 lowering in these cases may be motivated diachronically.

---

25 Vossen (1997) transcribes nasal clicks with a tilde above the click symbol and voiced clicks with a tilde below the click symbol. I have modified Vossen's transcriptions by using [ŋ] to indicate nasal clicks, e.g. [ŋ|], and [ɡ] to indicate voiced clicks, e.g. [ɡ|], to make the root-initial consonant comparisons easier to see. No further modifications were made.

What remains uncertain is why tonal depression did not affect these H tones *after* the sound changes as a synchronic rule. One possibility is the consonant-tone interactions were lexicalized by the time click replacement commenced. Therefore, the roots in question retained their original H tones after click replacement, even though the tones were eventually preceded by voiced obstruents. What might be unusual here is the apparently frozen-in-time nature of tonal depressor effects, since depression occurs following voiced obstruents that did not originate as nasalized clicks or plain nasal consonants. In other words, it may be that depression was a synchronic rule at an earlier stage of the grammar but the depressed tonal contours had been lexicalized by the time click replacement took place. Alternatively, there may be a synchronic depression rule, with the roots in Table 25 constituting lexical exceptions. This topic is addressed in Chapter 5.

One question that arises is whether the co-occurrence of root-initial nasal clicks with respect to the six non-depressed melodies is restricted. Root-initial nasal clicks are attested in 5 out of the 6 non-depressed melodies. Table 27 has examples for each tonal melody. Only the ML melody does not co-occur with a nasal click in the current Tsua data set. It is not clear at this time whether this restriction will hold if more data are collected. While it cannot be definitively stated that root-initial nasal clicks co-occur with the 6 non-depressed melodies, we can say that they do co-occur with root-initial H and M tones.

Table 27 Examples of nasal clicks co-occurring with five of the six non-depressed tonal melodies in Tsua (from Mathes and Chebanne 2013)

Melody	**Tsua Example**	**Gloss**
HH	ŋ‖áḿ	'to scrape'
HM	ŋ‖úm̄	'nighthawk'
HL	ŋ‖áù	'to spray liquid'
MM	ŋ‖ūm̄	'cloud'
MH	ŋ‖āí	'game stick'
ML	-	-

The proposal above can be formulated as the hypothesis:

(29) **Tsua Depression Exceptions Hypothesis:** Tsua voiced obstruents that were nasal sonorants or nasalized clicks historically are not depressors synchronically.

There are three predictions made by the hypothesis: first, root-initial voiced clicks in G|ui and Kua should correspond with depressed Tsua H tones, whether the clicks have been replaced or not; second, depressed H tones in Tsua should never correspond to root-initial nasals in Kua and G|ui; and third, any lexical items with root-initial voiced obstruents that entered the grammar after the Tsua sound changes may not be subject to depression if the depression pattern is not productive.

Evidence supporting the first prediction is displayed in Table 28. The top five cognate sets show root-initial correspondences between G|ui, Kua and Tsua where the voiced alveolar click has been replaced but tonal depression still occurs in Tsua. The tone melodies are in parentheses after each cognate for clarity.

Table 28 Cognates where tonal depression occurs in Tsua but not in G|ui or Kua

G|ui[1]	Kua[2]	Tsua[3]	Gloss
g!ãĩ (HM)	gãĩ (HM)	gãĩ (DH-M)	'steenbok'
g!oo (HL)	guo (HL)	goo (DH-L)	'aardvark'
g!ui (HM)	gui (HM)	gui (DH-M)	'rope; trapping string'
g!am (HL)	gam (HL)	gam (DH-L)	'to throw'
g!ari (HL)	-	gari (DH-L)	'to thatch'
gum (HL)	gum (HL)	gum (DH-L)	'to blow'
ɟua (HL)	ɟua (HL)	dua (DH-L)	'kudu'
ɟuu (HL)	ɟuu (HL)	duu (DH-L)	'eland'
g‖am (HM)	g‖am (HM)	g‖am (DH-M)	'thorn'
g‖uu (HM)	g‖uu (HM)	g‖uu (DH-M)	'chest'

[1]Nakagawa et al. (2014) [2]Chebanne and Collins (2014) [3]Mathes and Chebanne (2013)

With respect to the second prediction, there are no instances of Tsua depressed H tones following voiced obstruents with cognate nasal correspondences in Kua or G|ui in the data set. There is also evidence to support the third prediction: recent Setswana loan words are not subject to tonal depression. For example, the Tsua HML /débēdù/ 'dewlap; chicken jowl' and HM /díā/ 'to delay', which should have root-initial depressed H tones, are likely the Setswana loans *lebedu* 'dewlap' and *diiwa* 'to delay', respectively. At least 2 of the 4 exceptions for the glottal fricative /h/ are loans: the HL /húmà/ 'to become rich' and the HH /hírá/ 'to hire' are from Setswana *huma* 'to become rich' and *hira* 'to hire'. Thus, by appealing to sound change, at least 25 out of 41 exceptions can be plausibly accounted for pending the availability of supplementary comparative data.

4.5 Summary

This chapter has detailed the extreme nature of Tsua tonal depression. The F0 of a depressed H tone is almost 100 Hz lower than a non-depressed H tone in some instances. For example, speaker S produces extremely depressed H tones for the DH-L category in many of her tokens. Moreover, there is a systematic class of exceptions that historically had root-initial nasal consonants. These consonants are not depressors synchronically even though click replacement has caused them to become part of the voiced obstruents depressor type.

Loan words are not subject to depression either. The nature of sound change is such that diachronic segmental changes do not appear to alter the fundamental properties of consonant-tone interaction insofar as Tsua is concerned. The findings in this chapter suggest that the depression pattern is phonological, not phonetic. An interesting question is whether Tsua speakers would apply tonal depression to nonsense words. A nonce word experiment would help understand where depression stands in the grammar by providing evidence about whether a depression rule is still present in Tsua's phonology. It is a question that is further explored in Chapter 5.

The Tsua Depression Exceptions Hypothesis presented in this chapter gives us two crucial pieces of knowledge. First, Traill's observation of the geographical West to East click replacement tendencies has consequences not just for the Tsua clicks themselves but also for understanding the Tsua tonal system. Therefore, future research on Khoisan tone should look at comparative lexical data because click replacement patterns may disguise tonal depression in other Kalahari Khoe East languages. Second, the Tsua evidence suggests that click replacement involves the nasalization accompaniment, which undermines a claim in Traill and Vossen (1997:23-24): “...accompaniments consist of independent articulations involving glottal states (voiceless, voiced, glottal stop), voice onset time, nasalization and ejection...Neither process [click loss and click replacement] involves the accompaniment.”

The F0 lowering effects of voiced and aspirated obstruents have been conflated until this point. It is essential to know if there are any F0 curve differences between depressor types and the theoretical implications for consonant-tone interaction. Chapter 5 teases apart tonal depression by depressor type through the use of Smoothing Spline Analysis of Variance (SS ANOVA), a domain general statistical tool for the rigorous comparison of curves along multiple reference points. The objective is to further understand the fundamental phonetic properties of Tsua consonant-tone interaction.

CHAPTER 5

5.1 Introduction

An understudied aspect of the phonetics of consonant-tone interaction is the effect of different depressor types on F0 within a language. One challenge is finding a language that has more than one tone depressor type. Tsua's rare depression pattern addresses this challenge. In this chapter, pairwise comparisons between tonal depression following voiced obstruents versus aspirated obstruents for the DH-L and DH-M tonal categories are conducted. The glottal fricative depressor /h/ could not be included due to a lack of sufficient data to carefully control the syllable and segmental structures for precise comparisons (cf. Table 22).

The study in this chapter also looks at how a depressed H tone is produced depending on the identity of the following tone (i.e. M or L). Moreover, confirming the absence of an effect of a root-initial depressor on a following Mid tone is undertaken to assess consonant-tone interaction theoretical issues. Examining these research questions will sharpen our understanding of the phonetic implementation of tonal depression in Tsua, which in turn should advance our knowledge of consonant-tone interaction in general.

A rigorous quantitative method is required to compare Tsua's complex, post-depressor F0 curves. Chapters 3 and 4 illustrate why simply taking F0 measurements at vowel onsets and midpoints is not appropriate for Tsua, even for level tones. The dynamicity of the entire curve would be lost. Thus, important acoustic phonetic detail is missed, thereby undermining a complete understanding of Tsua consonant-tone interaction.

The solution employed here is Smoothing Spline Analysis of Variance (SS ANOVA), a domain-general statistical tool for comparing curves along multiple reference points (Gu 2002, 2013). Davidson (2006) utilizes SS ANOVA as a method for comparing tongue curve shapes as captured by ultrasound images. SS ANOVA is used for the analysis of formant contours in Baker (2006), DeDecker and Nycz (2006) and Simonet, Rohena-Madrazo and Paz (2008). Recent research on tone in Asian languages has incorporated SS ANOVA as well (e.g. Moisik et al. 2013; Chuang et al. 2013; Yiu 2014). SS ANOVA provides the optimal vehicle for studying the phonetic implementation of Tsua tonal depression.

The SS ANOVA statistical tool is used to address the research questions in (30). The topic of z-score normalization is discussed in §5.2.1.

(30) Tone implementation research questions

Research Question 1 (RQ_1): Does DH-M differ in terms of the F0 curve shapes for the voiced obstruent versus aspirated obstruent depressor types?

Research Question 2 (RQ_2): Does DH-L differ in terms of the F0 curve shapes for the voiced obstruent versus aspirated obstruent depressor types?

Research Question 3 (RQ_3): Do depressed H tones differ in terms of the F0 curve shapes depending on the identity of the following tone (M or L)?

Research Question 4 (RQ_4): Is there a consistent F0 lowering interaction between root-initial depressors and Mid tones? The answer is yes if the F0 onsets, subsequent rises and F0 maxima show significant and consistent differences in normalized z-score values when root-initial depressors followed by Mid tones are compared with root-initial non-depressors followed by Mid tones.

5.2 Methods

5.2.1 Time and F0 z-score normalization procedures

Quite a few normalization techniques have been proposed since the 1960's when they were first used for vowel classification and comparison (Zhu 1999). Tone normalization aims to reduce between-speaker variance so that curve comparison is possible. The pairwise comparisons in the subsequent sections use measurements taken via a two-step normalization procedure. First, a Praat script is run to measure F0 at 100 equidistant time points for the bimoraic lexical items under investigation (50 per tone).[26] Second, extracted F0 values at the 100 time points are z-score normalized in accordance with Rose (1987, 1991). The standard formula in (31) is used to calculate the within-speaker F0 z-scores.

---

[26] Many thanks to Christian DiCanio for his Pitch Dynamics Praat script.

(31) F0 z-score normalization formula (based on Rose 1987:347)

$F0_{norm} = (F0_i - F0_{mean}) / \sigma$

$F0_{norm}$ is the normalized value as a z-score at time point X (1-100). $F0_i$ is an observed F0 value. $F0_{mean}$ is the set of F0 mean values of all time points and all melodies by speaker. σ is one standard deviation of all time points and all melodies by speaker. Recall in Chapter 3 that lexical items with root-initial voiceless consonants followed by non-high, non-nasal, non-pharyngealized vowels were chosen to establish the typically produced tonal melodies with the least amount of F0 perturbation. 5 CVV lexical items per tonal category x 6 tonal categories x 3 repetitions/speaker for a total of 90 tokens are measured to calculate $F0_{mean}$ and σ for each speaker (Table 29; see (15) - (20) in Chapter 3 for the lexical items). The total number of measurements per speaker *n* that go into the z-score calculation is computed as: 90 tokens x 100 measurements per token = 9000.

Table 29 $F0_{mean}$, σ and total number of measurements *n* by speaker

Speaker	$\mathbf{F0_{mean}}$	**σ**	***n***
S	168.62	26.48	9000
B	194.96	23.30	9000
M	200.13	23.67	9000

5.2.2 Smoothing Spline ANOVA for comparing F0 curves

The z-score normalized values for the 100 points per repetition are loaded into the statistical package R version 2.9.2 (R Core Development Team 2009). The data are fit using smoothing splines, which are piecewise cubic polynomials. Smoothing splines provide curves that smooth discrete, noisy data (Craven and Wahba 1979:377). The optimum amount of smoothing is estimated to find the best fit via a smoothing parameter λ. One approach to the non-parametric smoothing estimation is via the minimization of a penalized least squares score. The least squares function is in (32). The first term discourages the lack of fit to the data, the second term penalizes the roughness of η, and the smoothing parameter λ controls the trade-off between the two conflicting goals (Gu 2002, 2013).

(32) Least squares function (Gu 2002, 2013)

$$\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \eta(x_i)\right)^2 + \lambda \int_0^1 \ddot{\eta}^2 dx$$

The smoothing parameter λ is crucial to the spline estimate. If λ is large, the curve will be smoother whereas a smaller λ attempts to fit each of the individual data points (Davidson 2006). Craven and Wahba (1979) discusses an automatic method for determining the smoothing parameter.

For our purposes, it is helpful to view an example of the individual data points as fitted with smoothing splines for one of the pairwise comparisons. Fig. 36 is an R plot of the individual data points produced by the three Tsua consultants for the DH-L curve. The light blue data points in Fig. 36 are productions of the DH-L curve following voiced obstruents by all three speakers. The light red data points are DH-L productions following aspirated obstruents. The data points are time normalized into 100 equidistant points (x-axis) and F0 normalized into z-scores (y-axis). Outliers caused by Praat F0 measurement errors were removed. The two smoothing splines are estimated to fit the data points for two contexts: DH-L following voiced obstruents (lower spline) versus DH-L following aspirated obstruents (upper spline).

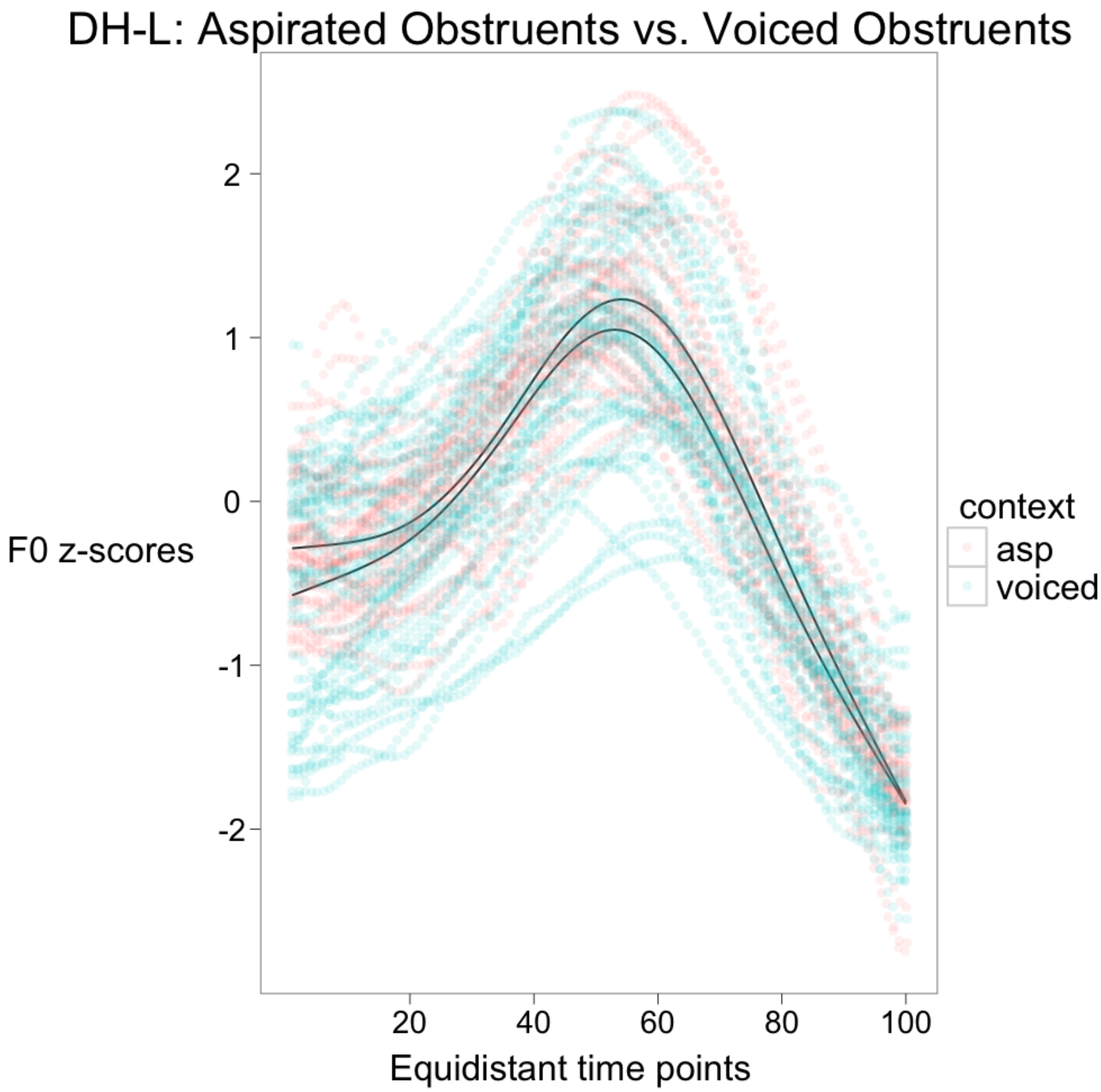


Fig. 36 Example of individual data points fitted with two smoothing splines (the solid black lines) by depressor type in R. The light blue data points are time normalized productions of the DH-L curve following voiced obstruents by all three Tsua consultants (fitted with the lower spline). The light red data points are time normalized productions of the DH-L curve following aspirated obstruents (fitted with the upper spline). Time normalization is achieved by way of 100 equidistant time points (x-axis). F0 z-score normalization values are on the y-axis.

To be sure of a statistically significant difference, 95% Bayesian confidence intervals are constructed around the smoothing splines for the main effects. Thus, similarities and differences between the mean curve shapes for the two contexts can be conclusively determined. The reference points of two curves are significantly different if the 95% confidence intervals do not overlap and are *not* significantly different if they do overlap. The smoothing splines with 95% confidence intervals are plotted in R using the ggplot2 library. The confidence interval differences are described in the following subsections, which segues into an interpretation of the results.

5.3 Results

5.3.1 DH-M aspirated vs. voiced obstruents pairwise comparison

The DH-M curve results are given first. A total of 10 lexical items are measured: 5 with root-initial voiced obstruents versus 5 with root-initial aspirated obstruents. 3 repetitions per consultant are measured for each lexical item. The 10 lexical items are listed in (33). Fig. 37 is the SS ANOVA plot generated in R. The 95% confidence intervals around the smoothing splines are light green for the aspirated obstruents depressor type and light orange for voiced obstruents. Places of confidence interval overlap indicate no statistically significant difference between the two curves.

(33) DH-M pairwise comparison data

a. Aspirated obstruents depressor type

Lexical Item	**Gloss**
tsʰá͓ā	'water'
kʰá͓ē	'to stab'
tsʰá͓ō	'to dig'
tsʰó͓ō	'hide; skin'
qʰá͓n̄	'news'

b. Voiced obstruents depressor type

Lexical Item	**Gloss**
gó͓ā tsúrī	'much later'
dá͓m̄	'tongue'
g|á͓ā	'to put in smoothly (into sand/fire)'
g|á͓ā	'to dry out clothes'
g‖á͓m̄	'thorn'

One area of interest is normalized time points 1-50, i.e. the leftmost section of the plot. This area corresponds to the F0 lowering effect for the DH-M curve by a prevocalic depressor. Wider confidence intervals represent more variability with respect to F0 production by the three consultants. Narrower confidence intervals represent less. The results indicate that there is no significant difference in the production of the DH-M melody whether the root-initial depressor is aspirated or voiced by the three Tsua speakers. This is shown by the overlapping confidence intervals.

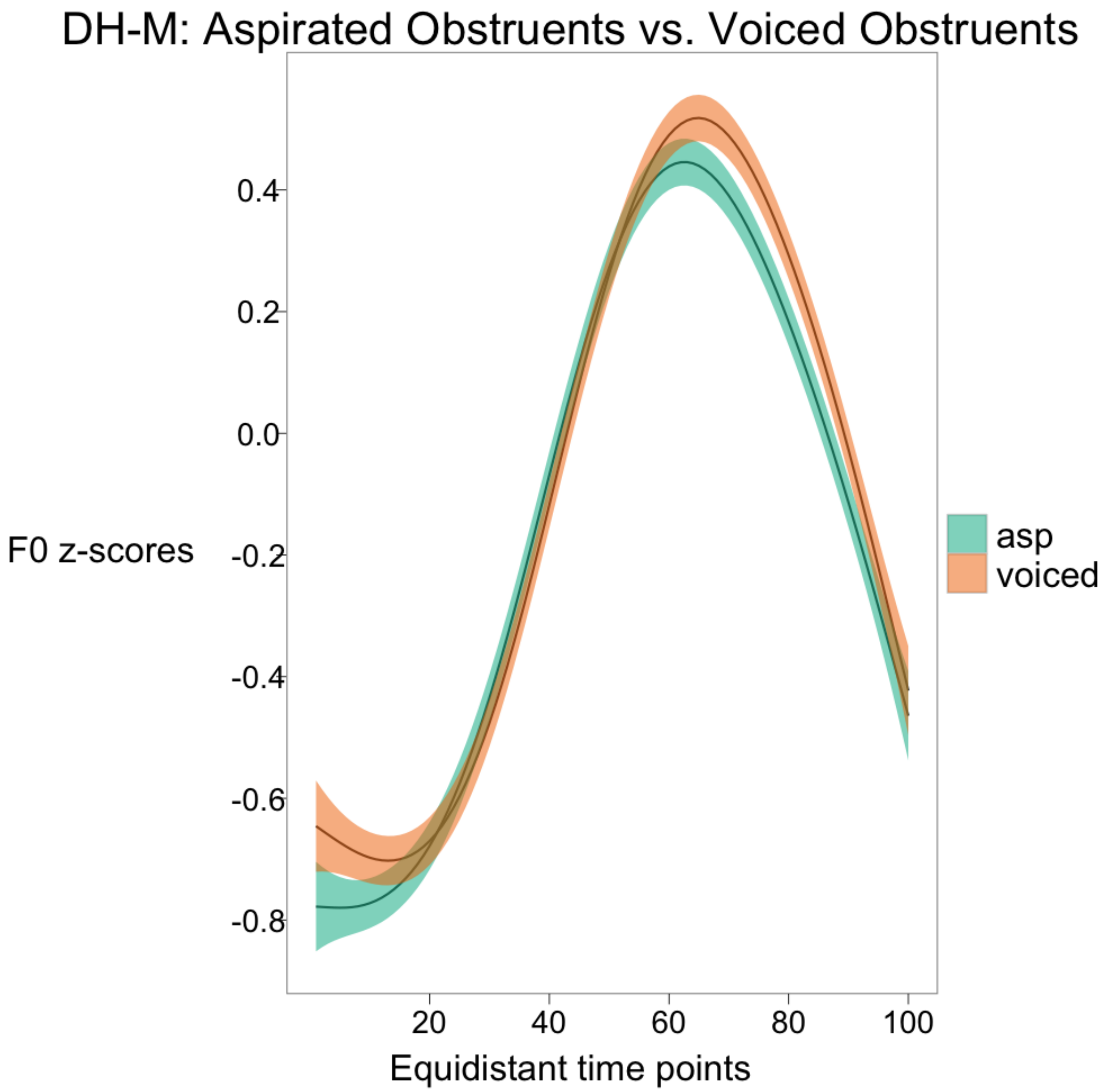


Fig. 37 DH-M pairwise comparison of aspirated obstruents versus voiced obstruents. The x-axis designates the 100 equidistant time points and the y-axis the F0 normalized z-scores. The 95% confidence interval around the smoothing spline for the aspirated obstruents depressor type is light green. The 95% confidence interval around the smoothing spline for the voiced obstruents depressor type is light orange. Places of confidence interval overlap indicate no statistically significant difference between the two curves.

In terms of normalized time points 51-100, i.e. the rightmost section of the plot, there is a significant difference between the two curves after the F0 peak is reached. The difference appears to be marginally significant. Subsequently, the confidence intervals overlap, indicating no significant difference before the F0 offsets. The F0 onsets and peaks display the most variability in comparison to the rest of the curve as demonstrated by the wider confidence intervals. Most importantly, the overall F0 curve shapes are almost identical for DH-M following voiced and aspirated obstruents.

5.3.2 DH-L aspirated vs. voiced obstruents pairwise comparison

The 10 lexical items for the DH-L pairwise comparison are listed in (34). Fig. 38 is the SS ANOVA plot generated in R. As stated earlier, the 95% confidence intervals around the smoothing splines are light green for the aspirated obstruents depressor type and light orange for the voiced obstruents depressor type.

(34) DH-L pairwise comparison data

a. Aspirated obstruents depressor type

Lexical Item	**Gloss**
kʰá͓ǹ	'to crawl'
ǁqʰá͓m̀	'spider web thread'
tsʰó͓è	'person'
tsʰó͓ò	'to pick up'
kʰá͓ǹ	'to creep toward'

b. Voiced obstruents depressor type

Lexical Item	**Gloss**
g|á͓à	'Silver tree'
ɟó͓à	'ash'
gó͓ò	'aardvark'
gá͓m̀	'to toss in the air'
dá͓ò	'to burn (transitive)'

The Fig. 38 plot indicates a significant difference from the F0 onsets until normalized time point 25. From that point, the confidence intervals overlap, showing there is no significant difference until just before the F0 peaks and subsequent falls. The curves are not significantly different prior to the F0 offsets. It can be seen that the F0 curve is higher after aspirated obstruents than voiced obstruents. One major difference between the DH-M and DH-L results is that there are more areas where curve dissimilarities are significant for the DH-L contour. Moreover, the DH-L confidence intervals are narrower than the DH-M's, indicating less variability.

Nonetheless, the overall F0 curve shapes are similar for DH-L following voiced and aspirated obstruents. The DH-M and DH-L plots give us an idea of how much variation is seen within a tonal category depending on the initial consonant. This variation may be implemented by differences in (a) F0 onset, (b) F0 maximum and F0 maximum timing, (c) F0 minimum and F0 minimum timing, and (d) F0 offset. Thus, a tone category's overall curve shape is more important than whether the curve is produced at a slightly higher or lower Hertz value.

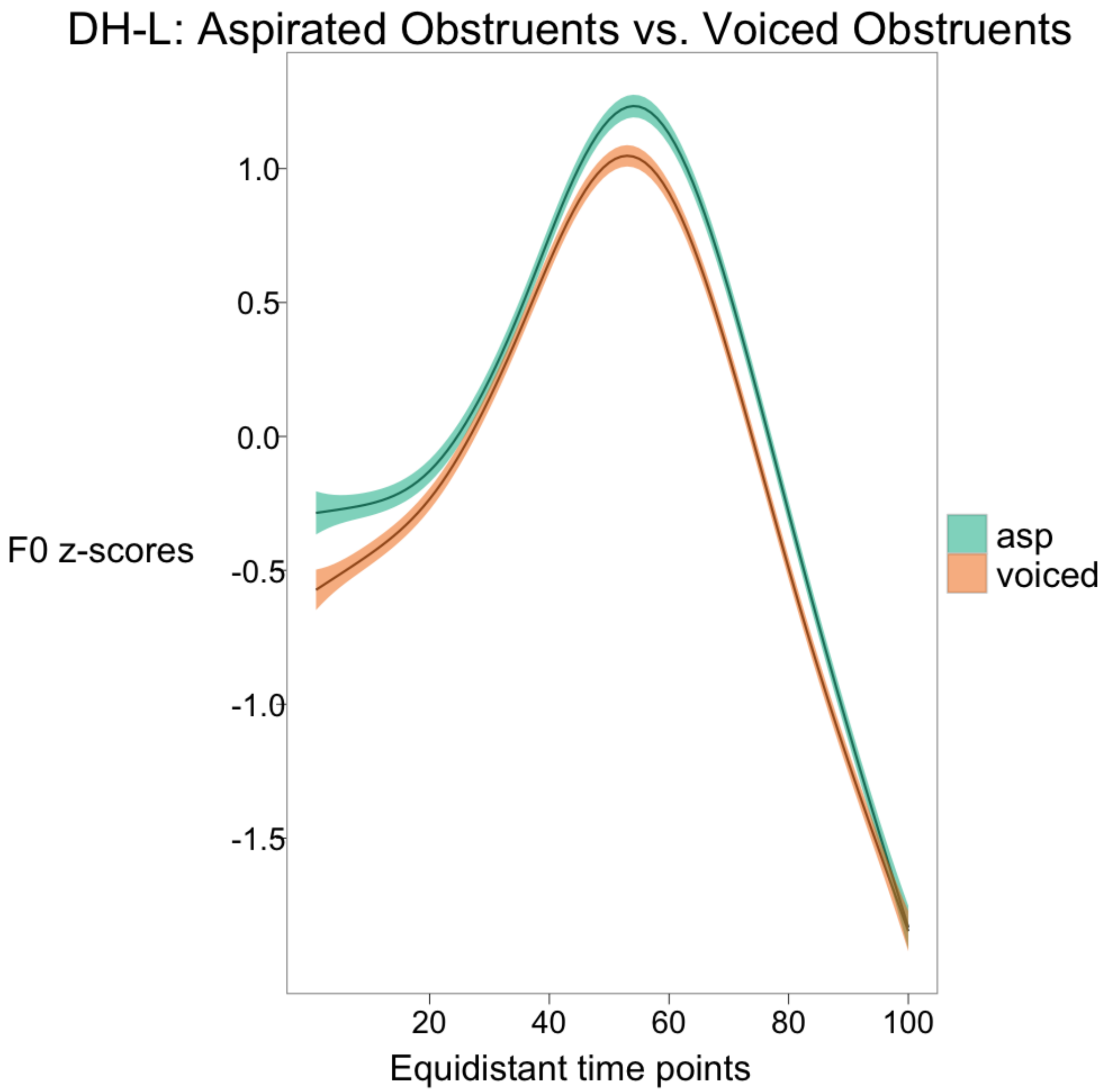


Fig. 38 DH-L pairwise comparison of aspirated obstruents versus voiced obstruents. The x-axis designates the 100 equidistant time points and the y-axis the F0 normalized z-scores. The 95% confidence interval around the smoothing spline for the aspirated obstruents depressor type is light green. The 95% confidence interval around the smoothing spline for the voiced obstruents depressor type is light orange. Places of confidence interval overlap indicate no statistically significant difference between the two curves.

5.3.3 DH-M vs. DH-L pairwise comparisons

There are three pairwise comparisons considered here with respect to DH-M versus DH-L. First, the DH-M aspirated and voiced tokens in (33a-b) are grouped and compared to the DH-L aspirated and voiced tokens in (34a-b) as seen in Fig. 39. Second, the DH-M aspirated tokens (33a) are compared to the DH-L aspirated tokens (34a) in Fig. 40. Third, the DH-M voiced tokens (33b) are compared to the DH-L voiced tokens (34b) in Fig. 41. The purpose is to determine the properties of a depressed H tone by depressor type when followed by a Mid or Low tone.

Fig. 39 reveals that the root-initial depressed H tones are significantly different from one another from the F0 onsets until just before the F0 peaks. The DH-L curve has a higher overall z-score than the DH-M curve until they cross after their F0 peaks. The curves are different throughout. The DH-M F0 offset has a higher value than the DH-L offset, reflecting the root-final M and L tones, respectively.

The two curves are plotted by depressor type in Figs. 40 and 41. Fig. 40 demonstrates that the curves are almost exactly parallel during their initial rises for the aspirated obstruents depressor type. The curves are significantly different except when they cross after their F0 peaks. Interestingly, the plot in Fig. 41 shows that there is not a significant difference between DH-M and DH-L at their F0 onsets for the voiced obstruents depressor type. The DH-M curve falls to its F0 minimum at normalized time point 20. After the short fall, the curves run parallel until their F0 peaks.

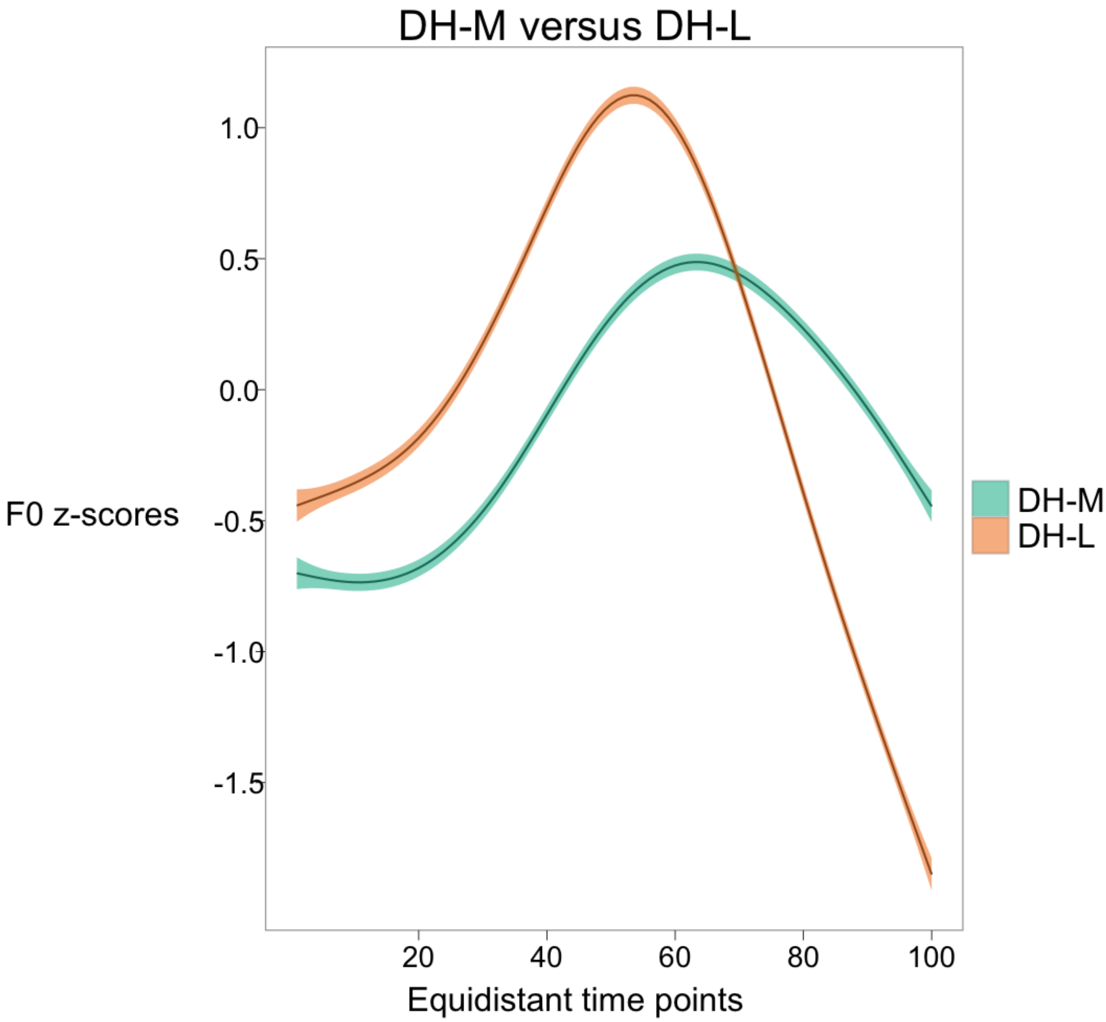


Fig. 39 DH-M versus DH-L pairwise comparison that groups aspirated and voiced obstruents. The x-axis designates the 100 equidistant time points and the y-axis the F0 normalized z-scores. The 95% confidence interval around the smoothing spline for the DH-M tone melody is light green. The 95% confidence interval around the smoothing spline for the DH-L tone melody is light orange. Places of confidence interval overlap indicate no statistically significant difference. The curves are significantly different throughout except where they cross after the F0 peaks. Their respective F0 z-scores are such that the curves run parallel to each other from normalized time point 20 until just before the DH-L F0 peak.

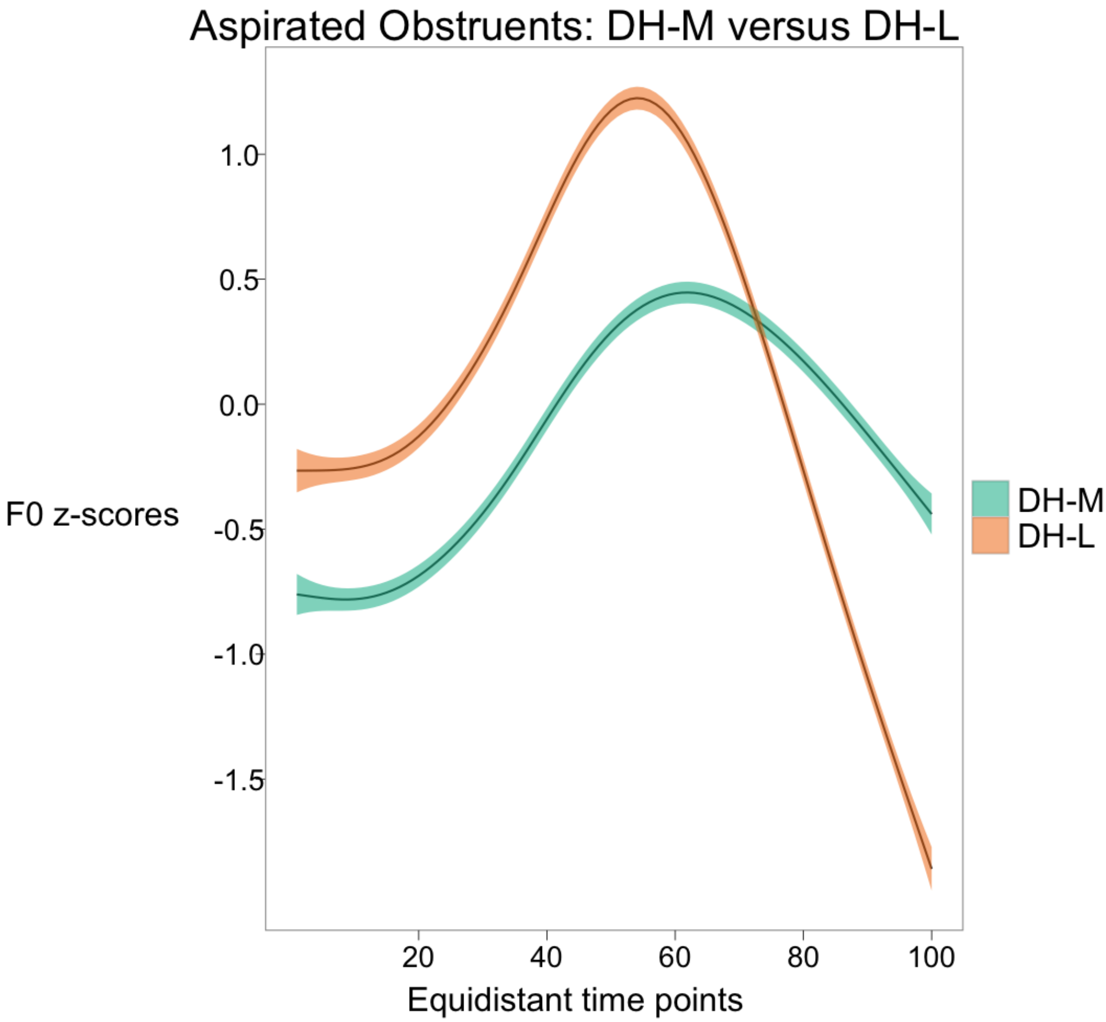


Fig. 40 DH-M versus DH-L pairwise comparison for aspirated obstruents only. The x-axis designates the 100 equidistant time points and the y-axis the F0 normalized z-scores. The 95% confidence interval around the smoothing spline for the DH-M tone melody is light green. The 95% confidence interval around the smoothing spline for the DH-L tone melody is light orange. Places of confidence interval overlap indicate no statistically significant difference. The curves are significantly different throughout except where they cross after the F0 peaks. Their respective F0 z-scores are such that the curves run parallel to each other from normalized time point 1 until just before the DH-L F0 peak.

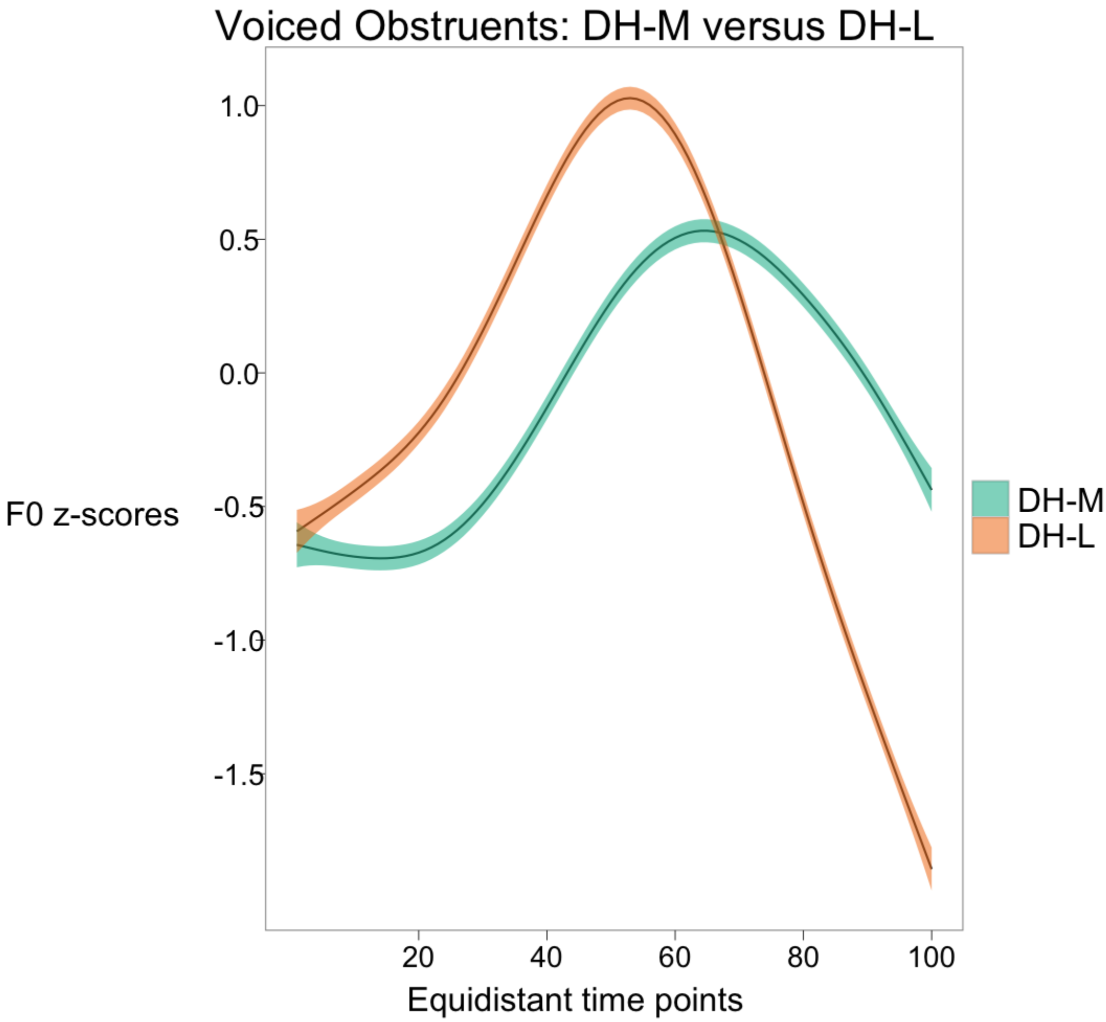


Fig. 41 DH-M versus DH-L pairwise comparison for voiced obstruents only. The x-axis designates the 100 equidistant time points and the y-axis the F0 normalized z-scores. The 95% confidence interval around the smoothing spline for the DH-M tone melody is light green. The 95% confidence interval around the smoothing spline for the DH-L tone melody is light orange. Places of confidence interval overlap indicate no statistically significant difference. The curves are significantly different throughout except at the F0 onsets plus where they cross after the F0 peaks. Their respective F0 z-scores are such that the curves run parallel to each other from the DH-M F0 minimum until just before the DH-L F0 peak.

5.3.4 Root-initial depressors and Mid tones

The potential effect of root-initial depressors on Mid tones compared to root-initial non-depressors constitutes the next comparisons. These involve the three tonal melodies with root-initial Mid tones: MH, ML and MM. There are two comparisons for the MH melody: (i) voiced depressors versus voiceless non-depressors and (ii) aspirated depressors versus voiceless non-depressors. Data set limitations affected the ML and MM categories, as the voiced and aspirated depressors had to be combined versus non-depressors.

(35) MH pairwise comparison data

a. Aspirated obstruents depressor type

Lexical Item	**Gloss**
t^hēé	'to give'
ǁ^{h}āá	'to plane'
c^hēé	'to drop something'
k^hōó	'hot season'
ǂ^{h}ōá	'to cut apart from pain'

b. Voiced obstruents depressor type

Lexical Item	**Gloss**
gǁāń	'melon'
bōó	'ax'
gǁāḿ	'to love'
dzāḿ	'to whip'
bēé	'not'

c. Voiceless non-depressors

Lexical Item	**Gloss**
|qχ’āé	'to fall down'
tēé	'bow'
ʔāé	'village'
ʔōó	'to bear fruit'
ǁχāó	'to scrape'

(36) ML pairwise comparison data

a. Voiced and aspirated depressor types

Lexical Item	**Gloss**
g|āà	'stomach'
ts^hāà	'many'

b. Voiceless non-depressors

Lexical Item	**Gloss**
ǁāà	'black korhaan'
ǂqōò	'mud'
ǁōò	'to weed'
kāò	'wet w/ dew'
|ōò	'to run after (when hunting)'

(37) MM pairwise comparison data

a. Voiced and aspirated depressor types

Lexical Item	**Gloss**
gǁōā	'mountain'
$ǁ^h$ōō	'dried cow dung used for a fire'

b. Voiceless non-depressors

Lexical Item	**Gloss**
χāē	'night; darkness'
tsqχ'āē	'green; blue'
ǂqāē	'bone marrow'
cχōā	'elephant'
tēē	'to ask'

Fig. 42 shows the SS ANOVA plots for the root-initial Mid tone comparisons. For the MH melody, the top left plot demonstrates that Mid tones following non-depressors are produced *lower* than when following aspirated depressors. The top right plot shows that Mid tones following voiced depressors are not produced with a significant difference compared to non-depressors until normalized time point 30. For the most part, the second halves of both plots are quite similar for the root-final tone target.

The reverse is true for the ML and MM contours: F0 is lower following depressors versus non-depressors. The interpretation of this result must be approached with caution. As mentioned earlier, the aspirated and voiced depressors had to be combined, totaling 2 lexical items for each plot due to data limitations. What can be said about the four plots is the F0 curve shapes are similar within each plot. These results confirm that Mid tones do not undergo extreme F0 lowering. Depression specifically targets root-initial H tones followed by non-High tones in Tsua.

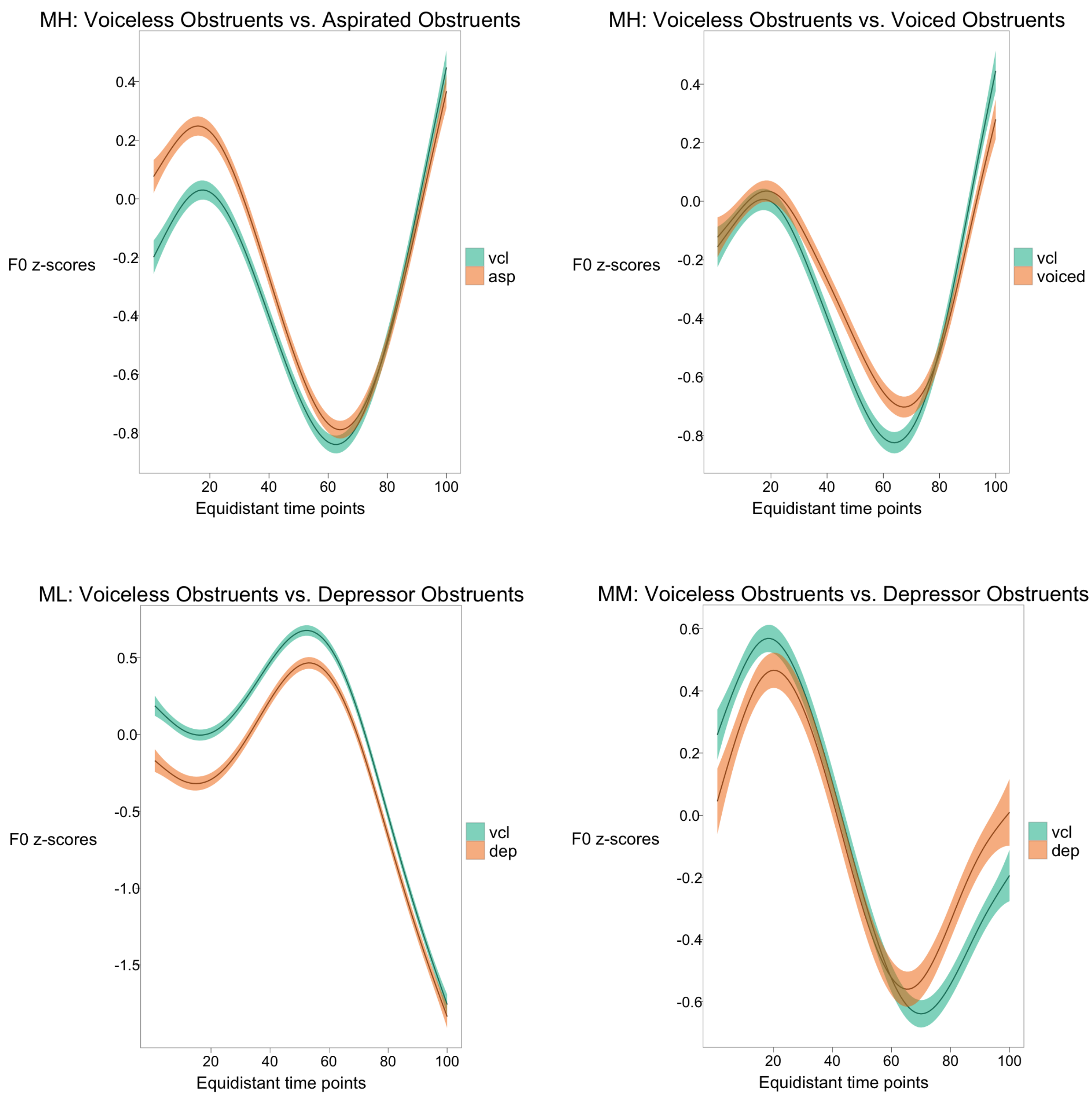


Fig. 42 MH voiceless non-depressors versus aspirated depressors (top left) and voiceless non-depressors versus voiced depressors (top right). ML voiceless non-depressors versus depressors (bottom left) and MM voiceless non-depressors versus depressors (bottom right). The MH contours with root-initial non-depressors are produced lower than with root-initial depressors. The reverse is true for ML and MM: the contours with root-initial depressors are produced slightly lower than with root-initial non-depressors. These results confirm that Mid tones do not undergo extreme F0 lowering.

5.3.5 DH-L vs. ML pairwise comparisons

The final comparisons are between the DH-L and ML tone melodies to confirm that their F0 shapes are distinct with respect to the Depressed H tone as opposed to the M tone in root-initial position. The SS ANOVA results are in Fig. 43. The plot to the left incorporates the DH-L data with aspirated depressors listed in (34a) versus the ML data with root-initial voiceless consonants from (36b). The plot to the right uses the DH-L data with voiced depressors from (34b) versus the ML data in (36b). Both plots confirm that DH and M are distinct. The F0 shapes are not significantly different towards the end of the curves for the root-final L tones.

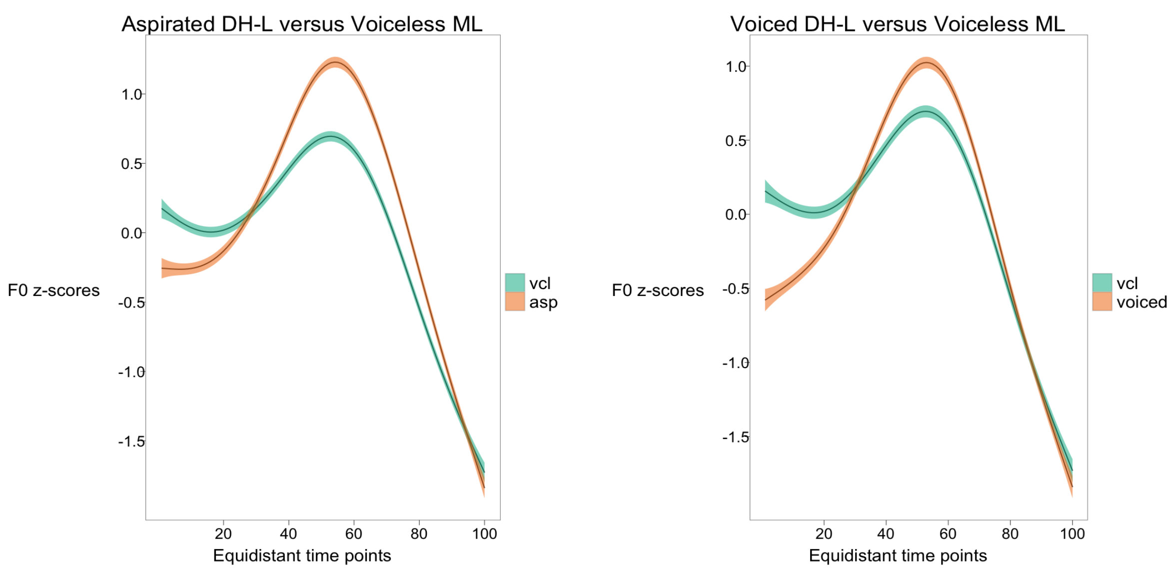


Fig. 43 SS ANOVA results for DH-L versus ML. The plot to the left shows DH-L with aspirated depressors vs. ML with root-initial voiceless consonants. The plot to the right shows DH-L with voiced depressors vs. ML with root-initial voiceless consonants.

5.4 Interim summary

The main question posed in this chapter is how a depressed High tone is implemented with respect to the F0 curve shapes by depressor type for DH-M and DH-L. We observe in Fig. 37 that the DH-M F0 curves are almost identical regardless of whether they are produced with a root-initial aspirated depressor versus a root-initial voiced depressor. The answer to RQ_1 is no. We observe in Fig. 38 that the shapes of the F0 curves for the DH-L melody are quite similar following aspirated obstruents as opposed to voiced obstruents. The answer to RQ_2 is also no.

RQ_3 requires a more nuanced view. Fig. 39 demonstrates that the F0 curves for DH-M versus DH-L are almost parallel from normalized time point 30 until they approach their F0 peaks. The initial F0 rises run parallel to each other when solely considering the aspirated obstruent depressor type for the comparison of DH-M versus DH-L (Fig. 40). The DH-M and DH-L curves following voiced obstruents intersect at their F0 onsets (Fig. 41). However, by normalized time point 30, the curves become parallel until their F0 peaks, after which they fall to their M and L tone targets, respectively. It can be cautiously inferred that the F0 curve shape of a depressed H tone is produced similarly regardless of whether the following tone is M or L, with the caveat that this occurs from normalized time point 30 to the F0 peak.

RQ_4 pertains to whether a root-initial M tone consistently interacts with depressors. Fig. 42 shows that the MH melody is produced slightly *higher* or with no statistically significant difference with respect to F0 following depressors versus non-depressors. The ML and MM melodies are produced slightly lower when following depressors. This F0 lowering is nowhere near 50+ Hz and is not consistent across the Mid tone initial melodies. The answer to RQ_4 is no, confirming that extreme F0 lowering only affects initial H tones.

5.5 Theoretical considerations

5.5.1 Tsua consonant-tone interaction as Low tone insertion

"Depressed High" has been used as a convenient cover term throughout this dissertation, but it could be made more precise now to account for the Tsua data. The structure of [DH] could be analyzed as resulting from Low tone insertion. In this case, the anatomy of [DH] would simply be [LH]. Thus DH-L is [LH-L] and DH-M is [LH-M]. The Tsua tonal depression rule is proposed in (38).

(38) Tsua tonal depression rule with the [+slack] feature

$\varnothing \rightarrow$ L / [-sonorant, +slack] ___ H [-H] #

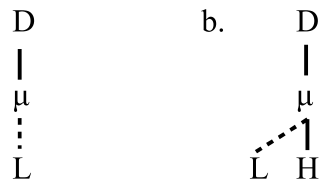
In (38), a Low tone is inserted in the environment following [-sonorant], [+slack] and when followed by a sequence of a High tone plus a non-high tone (i.e. M or L). The rule in (38) appeals to the unifying feature [+slack] as the source of the extreme F0 lowering in Tsua. §5.6 gives a detailed explanation in favor of appealing to this feature. For now, if we assume that voiced obstruents, aspirated obstruents and /h/ can be unified via [+slack], while Mid and Low tones could be represented as [-H], (38) provides a plausible account for the shapes of the depressed contours.

Fig. 1, repeated as Fig. 33 in Chapter 4, clearly demonstrates that the F0 target at the start of the depressed contours is consistent with a Low tone interpretation, because it is clearly distinct from and below a root-initial Mid tone target. Fig. 43 supports the Low tone interpretation as well. Subsequently, the curves rise to their respective High tone targets followed by a fall to the root-final tones. The curves do not show evidence of High tone deletion. The curve shapes in Figs. 37 and 38 corroborate this interpretation. Nevertheless, the SS ANOVA results indicate that the realization of a root-initial Low tone is subject to variability depending on (a) the root-initial depressor type, and (b) the following tone (Figs. 40 and 41). We would need to assume that there is some phonetic leeway in terms of Tsua Low tone production when it is inserted root-initially.

The formalization in (38) is more context dependent compared to analyses of consonant-tone interaction in other languages. Hyman and Mathangwane (1998) adopts the position that a depressor consonant conditions the insertion of a Low tone in Ikalanga. The underlying tonal opposition in Ikalanga is a privative one between H vs. ∅, as in many Bantu languages (Hyman and Mathangwane 1998:197). Simply put, H is specified underlyingly while toneless TBUs receive default L tones after rules of H spreading and/or shifting have applied. Voiced obstruents in Ikalanga serve as depressors and are designated as 'D' in Hyman and Mathangwane's analysis. If a mora following a depressor is toneless, the Low tone links to it (39a). If a mora following a depressor carries a High tone, the linking of the Low tone creates a LH rising contour on the surface (39b).

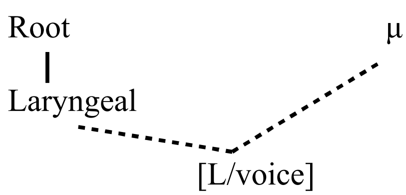
(39) Ikalanga Low tone insertion by a depressor (Hyman and Mathangwane 1998:209)

a. b.



For Bradshaw (1999, 2003), representations such as the ones in (39) are problematic because the relationship between L tone and [voice] is accidental (1999:56). As an alternative, Bradshaw proposes a multiplanar model that incorporates the generalization that tone is segmental as well as suprasegmental with an explicit relationship between L tone and [voice]. The dual nature of tone is expressed by Bradshaw as the [L/voice] feature seen in (40). When the feature attaches to a mora, it is realized as lowered pitch and is suprasegmental in nature. When it attaches to the Laryngeal node of a segment, it is realized as voicing and is segmental in nature. Consonant-tone interaction occurs when the feature attaches to both the Laryngeal node of a segment and to a mora.

(40) Multiplanar model of consonant-tone interaction (Bradshaw 1999:51, 2003:278)



The approaches in (39) and (40) do not adequately account for the Tsua data. Hyman and Mathangwane's formulation predicts that tones other than Low tones should undergo lowering. If a Low tone follows a depressor, that Low tone should not be affected by the Low tone introduced by the depressor because of L + L tonal coalescence, resulting in L. Tonal coalescence is common cross-linguistically and uncontroversial in the literature. On the other hand, we would expect a root-initial Mid tone to become LM rising after the insertion of a root-initial Low tone.

(41) Low tone insertion before a Mid tone

D
|
μ
L M

Post-depressor F0 is expected to be lower than the initial portion of a contour that follows a root-initial non-depressor. Fig. 42 demonstrates the opposite pattern: the MH contour's F0 is *higher* following depressors compared to non-depressors. Any theory of consonant-tone interaction that necessarily ties a depressor to Low tone encounters the same issue when accounting for the Tsua SS ANOVA results.

With respect to Bradshaw's model, it does not address tonal depression following aspirated obstruents or the glottal fricative /h/. Although less common than depression triggered by voiced obstruents, there are languages in which aspirated or /h/ depressors are attested (Tables 20 and 21). The evidence in Figs. 37 and 38 clearly show that the Tsua speakers produce depressed contours almost identically regardless of the depressor type. In other words, the phonetic effect is likely the outcome of the same phonological structure. In sum, Hyman and Mathangwane's model is too general and Bradshaw's model is too specific to adequately capture the Tsua data patterns.

5.5.2 Allotony rules alternative account 1

An alternative to the Low tone insertion account takes Rycroft (1980) as a starting point. Recall from §4.2.1 that Rycroft (1980) is a study of tonal depression in Nguni languages. In Rycroft's analysis, the depressor effect is attributed to the feature [depression]. The resulting post-depressor rising tone is the phonetic realization of /H/, i.e. the rising tone is not phonological. This analysis is problematic, in that the feature [depression] is ad hoc, in part because “it is invoked to apply to a limited phenomenon in a limited number of languages” (Bradshaw 1999:85). However, it may be possible to appeal to features that have been widely attested along with modifications to Rycroft's approach.

The depressed High tones could be considered allotones of the /H/ toneme since (a) the root-initial voicing and aspiration contrasts have not been lost in Tsua, making the surfacing of allotones predictable, and (b) there is an attempt to recover the /H/ on the surface manifested as the F0 peak.[27] Hyman (1972:170) states “a universal phonetic

[27] Thanks to an anonymous reviewer of my submission to the Tonal Aspects of Languages Conference 2014, Nijmegen, The Netherlands, for pointing out (b).

tendency is said to become ‘phonologized’ when language-specific reference must be made to it, as in a phonological rule” (as cited in Hyman 2013:3). Following this definition, Tsua tonal depression has been phonologized because a depressed H tone can be derived by rule from /H/. The depressed H has not been phonologized as an L tone because the root-initial contrasts in voicing and aspiration remain. This is consistent with the literature on tonogenesis (see Haacke 2008 and references therein). If the depressed H had been phonologized as an L tone, then we would expect LL and LM melodies that were historically HL and HM, respectively.

(42) lists potential allotony rule formulations (the glottal fricative /h/ rules have been omitted since statistical tests could not be run on the limited data). (42a) postulates the surfacing of allotone [DH] from a /H/ after voiced obstruents and preceding M or L tones. (42b) posits the same after aspirated obstruents and preceding M or L. Thus, /H/ can be realized allotonically as [H] or [DH] depending on the context.

(42) Tsua allotony rules alternative

a. $$/\mathrm{H}/ \rightarrow [\mathrm{DH}] \;/\; \begin{array}{l}[-\text{sonorant}]\\ [+\text{consonantal}]\\ [+\text{voiced}]\end{array} \;\_\_\_\; \mathrm{M, L}\ \#$$

b. $$/\mathrm{H}/ \rightarrow [\mathrm{DH}] \;/\; \begin{array}{l}[-\text{sonorant}]\\ [+\text{consonantal}]\\ [+\text{spread gl}]\\ [-\text{constr gl}]\end{array} \;\_\_\_\; \mathrm{M, L}\ \#$$

One weakness of this account is that the rules do not tell us much about the anatomy of [DH], especially considering instrumental evidence suggesting a root-initial Low tone. Moreover, (42) does not account for the tonal depression exceptions — [H] surfaces following voiced and aspirated obstruents in a number of cases (see Chapter 4). Therefore, §5.5.3 sketches a second alternative analysis based on the notion that the DH-M and DH-L melodies are best analyzed as two additional contours without depression rules, albeit with restricted distributions.

5.5.3 Additional contours alternative account 2

The Tsua tonal depression rule in (38) may be problematic when we consider the tonal depression exceptions. The exceptions make up 19.3% of the roots with initial depressor consonants. The root counts of tonal depression exceptions by depressor type are repeated in Table 30. Recall from §4.4.3 that 2 of the 4 exceptions for the glottal fricative depressor type are Setswana loan words. It may be that the other 2 are loans for that type as well. It is plausible that the 4 exceptions for the aspirated obstruents depressor type are loans, too. It seems that the crux of the matter has to do with the voiced obstruents depressor type with 33 out of 112 exceptions.

Table 30 Root counts of tonal depression exceptions by depressor type

	voiced obstruents	aspirated obstruents	/h/	Total
depression exceptions	33	4	4	41
depression expected	112	88	12	212
				19.3%

There are two ways to interpret these results. The first interpretation is that a rule such as the one in (38) is present in Tsua's phonology. In turn, the voiced obstruents exceptions, which pattern together as being historically non-depressor nasals, do not undergo the rule via lexical specification. The second interpretation is that the rule is not present in Tsua's phonology. DH-M and DH-L are actually the two additional contours specified underlyingly as /LHM/ and /LHL/, surfacing as [LHM] and [LHL], respectively. Both accounts rely on lexicalization, but in different ways. In the first, there is a productive Low tone insertion rule, but there are lexically marked exceptions that resist the rule. In the second, there is not a rule, as the additional contours are specified in the lexicon.

The first interpretation predicts that novel words with root-initial depressors preceding H tones should undergo the rule. We would expect extreme F0 lowering. The second interpretation predicts that novel words would not undergo the rule. We would not expect extreme F0 lowering in this case. The current data set cannot conclusively distinguish between these two possibilities. There are methods that could in a future experiment.[28]

The probing of speakers' grammars for whether there is a synchronic lowering rule is quite complex. A tone Wug Test is a first step towards checking for the productivity of a tonal depression rule in Tsua. The Wug Test paradigm was pioneered in Berko (1958) to discover how children acquire English morphology. Since then, it has been used by researchers to test rule productivity. The hypothetical tone Wug Test could be designed such that Tsua native speakers listen to monosyllabic words through headphones. They would be asked to give judgements via a forced-choice task in a two-part experiment.

The stimuli would consist of novel CVV and CVN words that have licit segmental sequences in Tsua. Each novel word would belong to one of four Word Types determined by the onset consonant: (i) Voiceless (VCL), (ii) Voiced (VCD), (iii) Aspirated (ASP), or (iv) Glottal Fricative (GF). For each Word Type, all six non-depressed tonal melodies would be used in combinations of two. For example, HH might be paired with ML, MH with HL, etc. The depressed melodies, DH-M and DH-L, would be paired with HM and HL, respectively. These will be called tone pairs.

The test subjects would be prompted by the following instructions: "You will hear a list of word pairs that are not real words in Tsua. If they were real words, how would you pronounce them in your language?" For the first experiment, the test subject would hear a novel word produced with the first melody of the tone pair followed by the second melody

[28] Thanks to Seunghun Lee for discussions related to potential methodologies for rule productivity testing in the field.

of the tone pair. The subject would be asked to pronounce their choice based on what they hear. It is expected that the subjects would choose the depressed melodies over the non-depressed melodies for the VCD, ASP and GF Word Types if a depression rule is present in Tsua's phonology. They may not if a rule is not present, lending support in favor of the two additional contours account.

For the second experiment, the subjects would be asked if novel forms are acceptable when non-depressor consonants are paired with depressed melodies DH-M and DH-L. In addition, the subjects would be asked if stimuli containing depressor consonants paired with the non-depressed contours HM and HL are acceptable. If they are not, then the subjects could be asked to pronounce them in a way they consider acceptable. We would expect HM to be produced as DH-M and HL as DH-L if the depression rule is synchronically active. Further field experiments would need to be devised to confirm the results of these initial tone Wug Tests.

5.6 On the origins of Tsua tonal depression

Some thoughts on the origins of tonal depression in Tsua are in order. If tonal depression has been phonologized in accord with Hyman's definition from earlier, there could be a universal phonetic tendency as the source. Presumably, phonetic F0 lowering took place during Tsua's evolution that was the precursor to the phonological depression observed today. We may hypothesize that the three depressor types (i.e. voiced obstruents, aspirated obstruents, and the glottal fricative /h/) share at least one feature responsible for the F0 lowering, although at first glance it is not clear which feature that would be. There is a promising possibility upon closer scrutiny of the relevant literature.

The effect of prevocalic voiced stops on F0 has been studied for quite some time (e.g., House and Fairbanks 1953; Lehiste and Peterson 1961; Löfqvist 1975; among others). The data from these and other studies make clear that the F0 of a following vowel is higher after voiceless than after voiced stops, and it does not vary as a function of place of articulation (Hombert et al. 1979:39). The common cross-linguistic pattern of F0 lowering by prevocalic voiced stops holds for tonal as well as non-tonal languages. The studies in Hombert (1978) confirm this pattern by presenting the F0 perturbation time course after voiced stops in Yoruba and American English. In sum, the postvocalic F0 lowering caused by voiced obstruents compared to voiceless obstruents has been reported in numerous phonetic studies and can be attributed to slackening of the vocal folds (Halle and Stevens 1971; Kingston and Diehl 1994). The slackening gesture is the articulatory correlate of the [+slack] feature (Halle and Stevens 1971).

The effect of aspirated stops is less consistent with respect to F0 perturbation trends cross-linguistically. §4.2 describes the F0 lowering effect by aspirated obstruents in tone languages spoken in Africa and Asia. However, there are languages in which a *higher* F0 has been reported after aspirated obstruents compared to unaspirated obstruents. For instance, data from Danish (Jeel 1975), Korean (Han and Weitzman 1970), Cantonese (Zee

1980), Thai (Gandour 1974) and Japanese (Ewan 1976) show that an aspirated stop gives rise to a higher F0 at the onset of the following vowel (as cited in Xu and Xu 2003:165). It appears that there is a dichotomy between F0 lowering and F0 raising after aspirated obstruents across languages.

Appealing to the [+slack] feature as the unifying cause of Tsua's synchronic tonal depression is complicated by classificatory systems that consider voiceless, aspirated obstruents as [-slack] (e.g. Halle and Stevens 1971:51). Nonetheless, Chen (2011:622) states: "To the extent that aspirated stops do introduce different perturbation effects as compared to their voiceless unaspirated counterparts, the cross-linguistic variability suggests that speakers may use different strategies to produce aspirated stops in different languages which lead to different perturbation effects." Citing the vocalis muscle effects described in Hoole (2006), Chen continues:

> Although vocal fold activity is a very complex phenomenon and much still remains to be learned, we may hypothesize that languages can differ in the strength of the vocalis muscle effects...It is possible that in Shanghai Chinese at the release of an aspirated stop, such an effect is either weak or not as effective as those after a voiceless unaspirated stop since the glottis opening after an aspirated stop is much greater. This in turn may lead to relative slackness of the vocal folds at the onset of the vowel after an aspirated onset, resulting in a certain amount of extended glottal opening and gradual onset of voicing. Consequently, there is a breathy transition between the aspiration and the vowel voicing with weak glottal buzz, manifested in higher H1 – H2 and *f0* lowering, as compared to voiceless unaspirated stops. In other words, not only the timing of laryngeal adjustments relative to supraglottal gestures (manifested as the VOT values) are important to differentiate stop contrasts, the state of the glottis constriction (e.g., spread vs. constricted and/or stiff vs. slack) can also play

> a role even in cases where VOT seems to be a sufficient cue for phonological distinctions (Chen 2011:622).

If Chen's hypothesis is on the right track, then vocal fold slackness may have caused the F0 lowering following voiced and aspirated obstruents in Tsua, assuming that aspiration is articulated in a similar fashion as described above for Shanghai Chinese.

The voiceless glottal fricative /h/ is an additional complication to the [+slack] possibility, in part because it is classified with the feature [-slack] by Halle and Stevens (1971:51). There are two pieces of evidence to consider in this case. First, we may look at the related language Shua. Westphal's handwritten field notes (n.d.) on Shua verb tense and pronouns have the voiced glottal fricative [ɦ] in the transcriptions. For example, [ɦĩĩ] 'to work' and [ɦiʃe] 'you (fem.)'. The handwritten transcription of [ɦ] does not seem to be Westphal's idiosyncratic way of writing 'h' in general, because he clearly transcribes 'to stab' as [khae] and 'person' as [khwe].

Second, we may consider evidence from Ikalanga, the neighboring Bantu language that has had extensive contact with Tsua, which could have been a source of historical influence on Tsua phonology (Chris Collins, Andy Chebanne, personal communication). Ikalanga has the voiced glottal fricative /ɦ/ as a depressor in its inventory (Mathangwane 1999:175). The voiced glottal fricative /ɦ/ is categorized as [+slack] by Halle and Stevens, which would explain its F0 lowering effect on a following tone. It may be that Tsua's voiceless glottal fricative /h/ was historically voiced and the resulting F0 lowering became phonologized. Recently, /ɦ/ became devoiced in Tsua but remained voiced in the Shua dialects at the time of Westphal's field research.

While a conclusive answer is elusive, this discussion raises intriguing questions about which consonants function as depressors across these languages and what the time course and temporal ordering were for phonologization by depressor type. The results presented in this chapter expand our knowledge of tonal phonetics by showing what is possible in a typologically rare tone system. Furthermore, this chapter highlights the importance of statistical methods in phonetic fieldwork.

CHAPTER 6

6.1 Dissertation summary and conclusions

The primary goal of this dissertation has been to understand the typologically rare and extreme tonal depression pattern in Tsua. I began by describing the complex segmental inventory of clicks, non-clicks and vowels. The click consonant inventory has been asymmetrically reduced due to a process of click replacement often reported in Kalahari Khoe East languages. Click replacement, a process in which certain clicks may be replaced by non-clicks, is thought to occur because of language contact between Bantu and Khoisan languages in Eastern Botswana. Evidence from original fieldwork suggests that Tsua has a Stage 3 click replacement pattern. Furthermore, click replacement has implications for understanding tonal depression in Tsua. Tonal depression occurs after voiced obstruents, aspirated obstruents or the glottal fricative /h/. However, exceptions to tonal depression constitute 19.3% of the data where tonal depression is expected. The exceptions pattern together as either loan words or historically nasal consonants, an apparent casualty of click replacement as far as the nasal consonants are concerned.

SS ANOVAs are employed to further explore the acoustic phonetic properties of Tsua tonal depression by depressor type. The results suggest that the DH-M melody is produced by the Tsua consultants almost identically regardless of whether the onset has a voiced or aspirated depressor. Although the DH-L curve is produced with a slightly higher F0 following root-initial aspirated depressors compared to voiced depressors, the curve shapes are quite similar. It is argued that a tone melody's overall curve shape is more important than whether the curve is produced at a slightly higher or lower Hertz value. Additional SS ANOVAs reveal two vital pieces of information: (i) Mid tones are not targeted for depression; and (ii) DH is produced with comparable F0 shapes when followed by a Mid tone versus a Low tone.

DH may be the result of a Low tone insertion rule. However, the purported rule is not completely isomorphic as the Low tone insertion rule for depression in Ikalanga proposed in Hyman and Mathangwane (1998). Moreover, the Tsua depression pattern is not fully accounted for by the [L/voice] feature in Bradshaw's Multiplanar model of consonant-tone interaction. Formally, I account for the Tsua depression pattern via the Low tone insertion rule in (43).

(43) Tsua tonal depression rule

$\varnothing \rightarrow$ L / [-sonorant, +slack] ___ H [-H] #

The rule proposes that [+slack] is the key feature that unifies the depressor types pending further acoustic and articulatory supporting evidence. What is important is that a Low tone insertion account is plausible given modifications to previous proposals.

Tonal depression exceptions raise the specter that the rule in (43) is no longer present in Tsua's phonology. What is unclear is whether the rule is still productive or the depressed contours are merely remnants of Tsua's phonology from long ago. A tone Wug Test is a first step towards testing for the rule's productivity. If the rule is found to be productive in a future tone Wug Test, this raises the interesting possibility that click replacement may interact with rule application in some way. Such a possibility has not been explored in the Khoisan literature.

Having studied the Tsua tonal system in detail, we are now in a position to ponder the dichotomy between African and East Asian tone systems. The literature is replete with treatments that classify tone languages as belonging to either an ‘African’ or ‘East Asian’ type. Recent studies suggest that there is not such a clear distinction. Some East Asian tone systems have the prototypical African properties, such as tonal reassociation through spreading and shifting, the presence of floating tones, and default Low tone assignment to toneless syllables (e.g., Hyman 2007; Evans 2008). Some African languages have been reported to employ paradigmatic substitutions, such as when a word's default tonal melody is replaced in certain contexts, which is more typically found in East Asian tone languages (e.g., Nakagawa 2006; Brugman 2009). As seen throughout this dissertation, Tsua not only has complex curves that are similar to F0 shapes found in East Asia, but also a tonal depressor effect following voiced and aspirated obstruents, a pattern found in the Wujiang dialect of Chinese. As we learn more about the nature of tone in these two geographically distinct regions, it might be prudent to dispense with this dichotomy altogether.

APPENDIX

G|ui HH, Kua HH and Tsua HH cognates

G|ui (HH)	Kua (HH)	Tsua (HH)	Gloss
!qχ’áó	!qχ’áó	!qχ’áó	'neck'
|áḿ	|áḿ	|áḿ	'sun'
|éé	|éé	|éé	'blue wildebeest'
ǁʔáá	ǁʔáá	ǁʔáá	'bat-eared fox'
tsáó	tsáó	tsáó	'tail'
!áé	káé	káé	'to tie'
ǂáó	ǂáó	cóó	'heart'
ǁóé	ǁóé	ǁóé	'to lie down'
ǁáé	ǁáé	ǁáé	'to teach'
ŋǁáé	ŋǁáé	gǁáé	'to count'

G|ui HM, Kua HM and Tsua HM cognates

G|ui (HM)	Kua (HM)	Tsua (HM)	Gloss
|qχ’áō	|qχ’áō	|qχ’áō	'snake'
|ʔéē	|ʔéē	|ʔéē	'fire/firewood'
|χáā	|χáā	|χáā	'meat'
ǁáē	ǁáē	ǁáē	'to chew'
ǁqórē	ǁqórē	ǁqórē	'claw; fingernail'
tsórō	tsórō	tsórō	'husk; shell'
χám̄	χám̄	χám̄	'lion'
ǁχóbē	ǁχórē	ǁχólē	'to open'
sãã	sãã	sãã	'to heal; to rest'
!χán̄	χán̄	χán̄	'to sew'

G|ui HL, Kua HL and Tsua HL cognates

G|ui (HL)	Kua (HL)	Tsua (HL)	Gloss
|qχ’áǹ	|qχ’áǹ	|qχ’áǹ	'bad; ugly'
|ʔáò	|ʔáò	|ʔáò	'money; blood'
|χóò	|χóò	|χóò	'gemsbok'
ǁqχ’áà	ǁqχ’áà	ǁqχ’áà	'to wash'
ǁχábà	ǁχábà	ǁχábà	'hunchback'
|ʔóò	|ʔóò	|ʔóò	'to dry (general)'
ǁʔãã	ǁʔãã	ǁʔãã	'to fight'
!ʔábò	ʔábò	ʔábò	'to climb up'
|ám̀	|ám̀	|ám̀	'two'
qχ’áà	qχ’áà	qχ’áà	'to drink'

G|ui MM, Kua MM and Tsua MM cognates

G|ui (MM)	Kua (MM)	Tsua (MM)	Gloss
ǁχāmā	ǁχāmā	ǁχāmā	'hartebeest'
ǂqāē	ǂqāē	ǂqāē	'bone marrow'
!χāē	χāē	χāē	'night; darkness'
tsãã	tsãã	tsãã	'meat broth'
ŋ!ābē	ŋābē	gābē	'giraffe'
tsqχ’āē	tsqχ’āē	tsqχ’āē	'blue; green'
|ʔām̄	|ʔām̄	|ʔām̄	'to gossip'
tsχārā	tsχārā	tsχārā	'to hurry'
ǁʔāē	ǁʔān̄	ǁʔān̄	'to dwell'
tsχōrē	tsχōrē	tsχōrē	'to meet a road'

G|ui LM, Kua LH and Tsua MH cognates

G|ui (LM)	**Kua (LH)**	**Tsua (MH)**	**Gloss**
ǁqàbā	ǁqàbá	ǁqābá	'branch'
cèn̄	tàń	tāń	'to stand up'
!qãũ	!qãũ	!qãũ	'cheetah'
gǁàē	gǁàé	gǁāé	'woman'
ɢ|òē	ɢ|òé	ɢ|ōé	'mongoose'
qàm̄	qàḿ	qāḿ	'gallbladder'
bàā	bàá	bāá	'father'
ŋǁàbō	ŋǁàbó	ǁābó	'shoe'
g|èrō	g|àró	g|āró	'ostrich'
gǁàm̄	gǁàḿ	gǁāḿ	'to love'

G|ui HH, Kua HH and Tsua MH cognates

G|ui (HH)	**Kua (HH)**	**Tsua (MH)**	**Gloss**
ǁʔóó	ǁʔóó	ʔōó	'to die'
|qχ’áé	|qχ’áé	|qχ’āé	'to fall'
qχ’áḿ	qχ’áḿ	qχ’āḿ	'mouth; beak'
ǁχóó	ǁχóó	ǁχōó	'to dry up'
ɟéḿ	dáḿ	dāḿ	'tortoise'
dzérá	dzérá	dzārá	'bird (generic)'
ǂχáná	ǂχáná	cχāná	'mucus'
dzérá	dzérá	dzārá	'bird'
gǁáná	gǁáná	gǁāná	'a well'
ǁʔáé	ǁʔáé	ʔāé	'village'

G|ui LL, Kua LL and Tsua ML cognates

Gǀui (LL)	Kua (LL)	Tsua (ML)	Gloss
ǁàà	ǁàà	ǁāà	'black korhaan'
ǁàbà	ǁàbà	ǁābà	'to be hungry'
ǁàmà	ǁàmà	ǁāmà	'dung beetle'
ǂqòò	ǂqòò	ǂqōò	'mud; clay'
!àrò	kàrò	kārò	'to run'
ŋǀàà	ŋǀàà	gǀāà	'stomach'
ŋ!àrè	ŋàrè	gārè	'hoof; foot'
gǀãã	gǀãã	gǀãã	'to try'
ǁòò	ǁòò	ǁōò	'to sweep, weed'
ǀʰãã	ǀʰãã	ǀʰãã	'ant'

G|ui HL, Kua HL and Tsua DH-L cognates

Gǀui (HL)	Kua (HL)	Tsua (DH-L)	Gloss
gǀáà	gǀáà	gǀá͓à	'Silver tree'
g!óò	gúò	gó͓ò	'aardvark'
!ʔáǹ[29]	kʰáǹ	kʰá͓ǹ	'to crawl'
g!úm̀	gúm̀	gú͓m̀	'to breathe on'
g!ám̀	gám̀	gá͓m̀	'to throw s.t.'
ǁʰábà	ǁʰábà	ǁʰá͓bà	'to stumble'
ǁʰóbè	ǁʰóbè	ǁʰó͓bè	'to creep toward'
ǁqʰám̀	ǁqʰám̀	ǁqʰá͓m̀	'spiderweb silk'
ǁʰúmì	ǁʰúmì	ǁʰú͓mì	'to put on apron'
ɟúù	ɟúù	dú͓ù	'eland'

[29] Vossen (1997)

G|ui HM, Kua HM and Tsua DH-M cognates

G|ui (HM)	**Kua (HM)**	**Tsua (DH-M)**	**Gloss**
gǁám̄	gǁám̄	gǁá͓m̄	'thorn'
ɟém̄	dám̄	dá͓m̄	'tongue'
ǁ^{h}ũĩ	ǁ^{h}ũĩ	ǁ^{h}ũ͓ĩ	'to sigh'
górē	gúrē	dzó͓rē	'tree bark'
k^{h}óō	k^{h}úō	ts^{h}ó͓ō	'skin; hide'
ǂq^{h}úrū	ǂq^{h}úrū	ǂq^{h}ú͓rū	'dust cloud'
ts^{h}áā	ts^{h}áā	ts^{h}á͓ā	'water'
!^{h}áē	k^{h}áē	k^{h}á͓ē	'to stab'
ts^{h}áō	ts^{h}áō	ts^{h}á͓ō	'to dig'
ǂ^{h}únī	ǂ^{h}únī	c^{h}ú͓nī	'elbow'

REFERENCES

Abramson, A. S. (1975). The tones of central Thai: some perceptual experiments. In: Harris, J. G., and J. Chamberlain (eds.) *Studies in Tai Linguistics*. Bangkok: Central Institute of English Language, 1-16.

Abramson, A. S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech 21*, 319-325.

Abramson, A. S. (1979). The co-articulation of tones: an acoustic study of Thai. In: Thongkum, T. L., Panupong, V., Kullavanijaya, P., and M. R. K. Tingsabadh (eds.) *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honour of Eugénie Henderson*. Bangkok: Chulalongkorn University Press, 1-9.

Andersson, L-G, and T. Janson. (1997). *Languages in Botswana: Language Ecology in Southern Africa*. Gaborone, Botswana: Longman Botswana.

Ansre, G. (1961). *The Tonal Structure of Ewe*. The Hartford Seminary Foundation.

Anyanwu, R-J, and B. Köhler. (2005). A survey of Northern Khoisan tonology. In: Anyanwu, R-J (ed.) *Tone Systems in African Languages*. Köln: Rüdiger Köppe Verlag, 91-129.

Baker, A. (2006). *Quantifying diphthongs: a statistical technique for distinguishing formant contours*. NWAV 35, Ohio State University, Columbus, Ohio.

Bao, Z. (1999). *The Structure of Tone*. New York: Oxford University Press.

Barnard, A. (1992). *Hunters and Herders of Southern Africa: A Comparative Ethnography of the Khoisan Peoples*. Cambridge: Cambridge University Press.

Batibo, H. M. (2002). The unlamentable loss: The role of attitude in language shift and death. In: Legere, K., Fitchat, S., and F. D. Akindele (eds.) *Talking freedom: Language and democratisation in the SADC Region*. Windhoek, Namibia: Gamsberg MacMillan Publishers Ltd, 193-199.

Batibo, H. M. (2003). Taking the best of both worlds: Integration and identity among the Khoesan speakers of Botswana. In: Brenzinger, M., and C. König (eds.)

Proceedings of the 1st International Symposium on Khoisan Languages and Linguistics. Riezlern/Kleinwalsertal, Rüdiger Köppe Verlag, 284-295.

Batibo, H. M. (2005). *Language Decline and Death in Africa - Causes, Consequences and Challenges*. Clevedon, England: Multilingual Matters.

Beach, D. M. (1924). The science of tonetics and its application to Bantu languages. *Bantu Studies 2*, 75-106.

Beach, D. M. (1938). *The Phonetics of the Hottentot Language*. Cambridge: W. Heffer & Sons LTD.

Berko, J. (1958). The child's learning of English morphology. *Word 14*, 150-177.

Bhaskararao, P. (1999). Voiced aspiration and tonogenesis in some South-Asian languages. In: Kaji, S. (ed.) *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*. Takinogawa City Hall, Tokyo: ILCAA, 337-351.

Bleek, D. F. (1939-40). A short survey of Bushman languages. *Zeitschrift für Eingeborenen-Sprachen 30*, 52-72.

Boersma, P., and D. Weenink. (2010). *Praat: Doing phonetics by computer, version 5.1.25*. URL: http://www.praat.org/, Developed at the Institute of Phonetic Sciences, University of Amsterdam.

Bowern, C. (2008). *Linguistic Fieldwork: A Practical Guide*. New York, NY: Palgrave MacMillan.

Bradshaw, M. M. (1995). Tone on verbs in Suma. In: Akinlabi, A. (ed.) *Theoretical Approaches to African Linguistics*. Trenton: Africa World Press, 255-271.

Bradshaw, M. M. (1999). *A Crosslinguistic Study of Consonant-Tone Interaction*. Ph.D. Dissertation, Ohio State University.

Bradshaw, M. M. (2003). Consonant-tone interaction in Siswati. *Studies in Phonetics, Phonology and Morphology 9*, pp. 277-294.

Brugman, J. C. (2009). *Segments, Tones and Distribution in Khoekhoe Prosody*. Ph.D. Dissertation, Cornell University.

Caicai, Z. (2009). Why would aspiration lower the pitch of the following vowel? Observations from Leng-shui-jiang Chinese. In: *Proceedings of Interspeech 2009*. Brighton, UK, 2299-2302.

Cao, Z. (2002). *Nanbu Wuyu Yuyin Yanjiu [Sound Change in the Southern Wu Dialects]*. Beijing: Shangwu Yinshuguan.

Casali, R. F. (2008). ATR Harmony in African Languages. *Language and Linguistics Compass 2*, 496-549.

Cassimjee, F., and C. W. Kisseberth. (1992). The Tonology of Depressor Consonants: Evidence from Mijikenda and Nguni. *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society: Special Session on the Typology of Tone Languages*, 26-40.

Chao, Y-R. (1930). A system of tone letters. *Le Maître Phonétique 45*, 24-27.

Chebanne, A. M. (2002a). Shifting identities in Eastern Khoe: ethnic and language endangerment. *Botswana Journal of African Studies 16*, 147-157.

Chebanne, A. M. (2002b). Preliminary observations on the sociolinguistic situation of the Cuaa language. In: Legere, K., Fitchat, S., and F. D. Akindele (eds.) *Talking freedom: Language and democratisation in the SADC Region*. Windhoek, Namibia: Gamsberg MacMillan Publishers Ltd, 201-210.

Chebanne, A. M. (2003a). The Khoe and San languages and people of Botswana: The current situation. In: Prah, K. K. (ed.) *Silenced Voices*. Centre for Advanced Studies of African Society, Cape Town, 59-100.

Chebanne, A. M. (2003b). Convergence, identity shifting and language loss in Eastern Khoe. In: Brenzinger, M., and C. König (eds.) *Proceedings of the 1st International Symposium on Khoisan Languages and Linguistics*. Riezlern/Kleinwalsertal, Rüdiger Köppe Verlag, 296-317.

Chebanne, A. M. (2008). A Sociolinguistic Perspective of the Indigenous Communities of Botswana. *African Study Monographs 29*, 93-118.

Chebanne, A. M., and C. Collins. (2012). *Tense and Aspect in Kua: A Preliminary Assessment.* Unpublished manuscript. New York University and the University of Botswana.

Chebanne, A. M., and C. Collins. (2014). *Kua Lexicon*. Unpublished manuscript. University of Botswana and New York University.

Chebanne, A. M., and L. Nyati-Ramahobo. (2003). Language Knowledge and Language Use in Botswana. In: *Proceedings of CSO: 2001 Population and Housing Census Dissemination Seminar*. Gaborone, Botswana, 392-404.

Chen, Y. (2011). How does phonology guide phonetics in segment-*f0* interaction? *Journal of Phonetics 39*, 612-625.

Childs, G. T. (2003). *An Introduction to African Languages*. Amsterdam, The Netherlands and Philadelphia, PA: John Benjamins Publishing Co.

Chuang, C-T, Chang, Y-C, and F-F Hsieh. (2013). Complete and not-so-complete tonal neutralization in Penang Hokkien. In: Lee, W. S. (ed.) *Proceedings of the International Conference on Phonetics of the Languages in China*, 54-57.

Clements, G. N., and R. Khatiwada. (2007). Phonetic realization of contrastively aspirated affricates in Nepali. In: *Proceedings of ICPhS XVI*. Saarbrücken, 629-632.

Collins, C. (2012). *A Preliminary Analysis of the Tonal System of ǂHoã*. Unpublished manuscript. New York University.

Colombel, V. de. (1986). *Phonologie quantitative et synthématique, avec application à l'ouldémé, langue tchadique du Nord-Cameroun*. Paris: Selaf (LCA 7).

Craven, P., and G. Wahba. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik 31*, 377-403.

Crowley, T., and C. Bowern. (2010). *An Introduction to Historical Linguistics*. New York: Oxford University Press.

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America 120*, 407-415.

DeDecker, P., and J. Nycz. (2006). *A new way of analyzing vowels: comparing formant contours using Smoothing Spline ANOVA*. NWAV 35, Ohio State University, Columbus, Ohio.

Doke, C. M. (1926). *The Phonetics of the Zulu Language*. Johannesburg: Witwatersrand University Press.

Doke, C. M. (1937). An outline of ǂKhomani Bushman phonetics. In: Jones, J. D. R., and C. M. Doke (eds.) *Bushmen of the Southern Kalahari*. Johannesburg: Witwatersrand University Press, 61-88.

Dornan, S. S. (1911). The Masarwas and their Language. *South African Journal of Science 8*, 218-225.

Downing, L. J. (2009). On pitch lowering not linked to voicing: Nguni and Shona group depressors. *Language Sciences 31*, 179-198.

Downing, L. J., and B. Gick. (2005). Voiceless tone depressors in Nambya and Botswana Kalang'a. BLS 27, 65-80.

Dressler, W. (1972). The phonology of language death. *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*, 448-457.

Eaton, R. D. (2007). The Phonology of Kangri Tone. *Paper presented at the 13th Himalayan Language Symposium*.

Elderkin, E. D. (1988). Patterns of sound in Northern Khoisan. *African Languages and Cultures 1*, 123-148.

Elderkin, E. D. (2004). The starred tones of Central Khoisan. *Afrika und Übersee Band 87*, 3-77.

Evans, J. P. (2008). 'African' Tone in the Sinosphere. *Language and Linguistics 9*, 463-490.

Ewan, W. G. (1976). *Laryngeal Behavior in Speech*. Ph.D. dissertation, University of California, Berkeley.

Fishman, J. A. (1991). *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon: Multilingual Matters.

Ford, K. C. (1986). The tonal phonology of Siya (a Central-Togo language) and the historical significance of its major functions. In: Bogers, K., van der Hulst, H., and M. Mous (eds.) *The Phonological Representation of Suprasegmentals*. Dordrecht: Foris, 79-94.

Gandour, J. (1974). Consonant types and tone in Siamese. *Journal of Phonetics 2*, 337-350.

Greenberg, J. H. (1963). *The Languages of Africa*. The Hague: Mouton.

Gu, C. (2002). *Smoothing Spline ANOVA Models*. New York: Springer.

Gu, C. (2013). *Smoothing Spline ANOVA Models, Second Edition*. New York: Springer.

Güldemann, T. (2006). *Changing profile when encroaching on hunter-gatherer territory: Towards the history of the Khoe-Kwadi family in southern Africa*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Güldemann, T., and R. Vossen. (2000). Khoisan. In: Heine, B., and D. Nurse (eds.) *African Languages: An Introduction*. Cambridge, United Kingdom: Cambridge University Press, 99-122.

Haacke, W. H. G. (1999). *The Tonology of KhoeKhoe (Nama/Damara)*. Köln: Rüdiger Küppe Verlag.

Haacke, W. H. G. (2008). Tonogenesis *in flagrante*: Tonal depression in Khoekhoegowab, Hai‖om, ǂAkhoe and Naro. In: Ermisch, S. (ed.) *Proceedings of the 2nd International Symposium on Khoisan Languages and Linguistics*. Riezlern/Kleinwalsertal, Austria, Rüdiger Köppe Verlag, 153-184.

Haacke, W. H. G. (2013). Namibian Khoekhoe. In: Vossen, R. (ed.) *The Khoesan Languages*. Taylor & Francis Group, 96-97.

Haacke, W. H. G., and E. Eiseb. (2002). *A Khoekhoegowab dictionary with an English-Khoekhoegowab index*. Windhoek, Namibia: Gamsberg Macmillan Publishers Ltd.

Hagman, R. S. (1977). *Nama Hottentot grammar*. Bloomington: Indiana University.

Halle, M., and K. N. Stevens. (1971). A note on laryngeal features. *MIT Research Laboratory of Electronics Quarterly Progress Report 101*, 198-213.

Han, M. S., and R. S. Weitzman. (1970). Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica 22*, 112-128.

Hansson, G. Ó. (2004). Tone and voicing agreement in Yabem. In: Schmeiser, B., Chand, V., Kelleher, A., and A. Rodriguez (eds.) *WCCFL 23: Proceedings of the 23rd West*

Coast Conference on Formal Linguistics. Somerville, MA: Cascadilla Press, 318-331.

Hasselbring, S., Segatlhe, T., and J. Munch. (2000). *A Sociolinguistic Survey of the Languages of Botswana, vol. 2*. Mogoditshane: Tasalls Publishing and Books.

Hasselbring, S., and T. Segatlhe. (2000). Language Choice in Ten Domains among Khoesan Speakers in Botswana. In: Batibo, H. M. and J. Tsonope (eds.) *The State of Khoesan Languages in Botsana*. Basarwa Languages Project, University of Botswana and the University of Tromsø, 76-99.

Heath, D. P. (1991). Tone in the Makaa Associative Construction. In: Anderson, S. C. (ed.) *Tone in five languages of Cameroon*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington, 3-28.

Henderson, E. J. A. (1952). The Main Features of Cambodian Pronunciation. *Bulletin of the School of Oriental and African Studies 14*, 149-174.

Hildebrandt, K. A. (2003). *Manange Tones: Scenarios of Retention and Loss in Two Communities*. Ph.D. Dissertation, University of California, Santa Barbara.

Hobley, J. (1964). A preliminary tonal analysis of the Bassa language. *Journal of West African Languages 1*, 51-55.

Hombert, J-M. (1978). Consonant Types, Vowel Quality, and Tone. In: Fromkin, V. A. (ed.) *Tone: A Linguistic Survey*. New York, NY: Academic Press, Inc., 77-111.

Hombert, J-M, Ohala, J. J., and W. G. Ewan. (1979). Phonetic explanations for the development of tones. *Language 55*, 37-58.

Honken, H. (2008). The split tones in Central Khoesan. In: Ermisch, S. (ed.) *Proceedings of the 2nd International Symposium on Khoisan Languages and Linguistics*. Riezlern/Kleinwalsertal, Austria, Rüdiger Köppe Verlag, 185-224.

Hoole, P. (2006). *Experimental studies of laryngeal articulation*. Germany: Habilitationschrift, University of Munich.

House, A. S., and G. Fairbanks. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America 25*, 105-113.

Hyman, L. M. (1972). Nasals and nasalization in Kwa. *Studies in African Linguistics 4*, 167-206.

Hyman, L. M. (2007). Kuki-Thaadow: an African tone system in Southeast Asia. *Berkeley: UC Berkeley Phonology Lab Annual Report*, to appear.

Hyman, L. M. (2010). How to Study a Tone Language, with exemplification from Oku (Grassfields Bantu, Cameroon). *UC Berkeley Phonology Lab Annual Report*, 179-209.

Hyman, L. M. (2013). Enlarging the scope of phonologization. In: Yu, A. C. L. (ed.) *Origins of Sound Change: Approaches to Phonologization*. Oxford University Press, 1-29.

Hyman, L. M., and J. T. Mathangwane. (1998). Tonal domains and depressor consonants in Ikalanga. In: Hyman, L. M., and C. W. Kisseberth (eds.) *Theoretical aspects of Bantu tone*. Stanford, CA: CSLI, 195-230.

Jeel, V. (1975). An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants. *Annual Report of the Institute of Phonetics, University of Copenhagen*, 191-211.

Kagaya, R., and H. Hirose. (1975). Fiberoptic electromyographic and acoustic analysis of Hindi stop consonants. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics 9*, 27-46.

Karlgren, B. (1915). *Études sur la phonologie chinoise*. Archives d'études orientales, vol. 15. Leide: E-J Brill; Stockholm: P.A. Norstedt & Søner; Gotembourg: Elanders Boktryckeri A-B.

Killian-Hatz, C. (2003). *Khwe Dictionary*. Köln: Rüdiger Köppe Verlag.

Kingston, J., and R. L. Diehl. (1994). Phonetic Knowledge. *Language 70*, 419-454.

Kisseberth, C. W. (1984). Digo Tonology. In: Clements, G. N., and J. Goldsmith (eds.) *Autosegmental Studies in Bantu Tone*. De Gruyter Mouton, 105-182.

Köhler, O. (1971). Die Khoe-sprachigen Buschmänner der Kalahari. In: Festschrift, K. K. (ed.) *Forschugen zur Allgemeinen und Regionalen Geschichte*. Wiesbaden: F. Steiner, 373-411.

Köhler, O. (1981). La langue !xũ. In: Perrot, J. (ed.) *Les langues dans le monde ancien et moderne*. Paris: Éditions du Centre National de la Recherche Scientifique, 557-615.

König, C., and B. Heine. (2001). *The !Xun of Ekoka. A Demographic and Linguistic Report*. Köln: Institut für Afrikanistik (Universität zu Köln).

König, C., and B. Heine. (2008). *A Concise Dictionary of Northwestern !Xun*. Köln: Rüdiger Köppe Verlag.

Kutsch Lojenga, C. (1985). The tones of the Ebrié associative construction. *Journal of African Languages and Linguistics 7*, 1-22.

Kutsch Lojenga, C. (1998). Depressor consonants and rising tones in Yaka (C10). *Paper presented at the 28th Conference on African Languages and Linguistics*, Leiden University.

Kutsch Lojenga, C. (2010). Procedure for Tone Analysis. *The 3L International Summer School on Language Documentation and Description*, Leiden University Center for Linguistics.

Ladefoged, P. (2003). *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford, UK: Blackwell Publishers Ltd.

Ladefoged, P., and I. Maddieson. (1996). *The Sounds of the World's Languages*. Oxford, UK and Cambridge, MA: Blackwell Publishers Ltd.

Ladefoged, P., and A. Traill. (1994). Clicks and their accompaniments. *Journal of Phonetics 22*, 33-64.

Lanham, L. W. (1958). The tonemes of Xhosa. *African Studies 17(2)*, 65-81.

Lanham, L. W. (1960). *The Comparative Phonology of Nguni*. Department of African Studies, Rhodes University.

Laver, J. (1994). *Principles of Phonetics*. Cambridge University Press.

Lehiste, I., and G. Peterson. (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America 33*, 419-425.

Löfqvist, A. (1975). Intrinsic and extrinsic F_0 variations in Swedish tonal accents.

Phonetica 3, 226-247.

Long, Y., and G. Zheng. (1988). *The Dong Language in Guizhou Province, China*. Dallas: SIL and the University of Texas at Arlington.

Lukas, J. (1969). Tonpermeable und Tonimpermeable Konsonanten im Bolanci (Nordnigerien). *Ethnological and Linguistic Studies in Honour of N. J. van Warmelo, Ethnological Publications no. 52*. Pretoria: Government Printer, 133-138.

Maddieson, I. (1974). A possible new cause of tone-splitting--evidence from Cama, Yoruba and other languages. *Studies in African Linguistics Supplement 5*, 205-221.

Maingard, L. F. (1937). The ǂKhomani dialect, its morphology and other characteristics. In: Jones, J. D. R., and C. M. Doke (eds.) *Bushmen of the Southern Kalahari*. Johannesburg: Witwatersrand University Press, 237-276.

Maingard, L. F. (1961). The Central Group of the Click Languages of the Kalahari. *African Studies 20*, 114-122.

Maran, L. R. (1971). Burmese and Jingpho: a study of tonal linguistics processes. In: Lehmen, F. K. (ed.) *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics, vol. 4*. Urbana: University of Illinois.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Mathangwane, J. T. (1999). *Ikalanga Phonetics and Phonology: A Synchronic and Diachronic Study*. Stanford, CA: CSLI Publications.

Mathes, T. K., and A. M. Chebanne. (2013). *Tsua Lexicon*. Unpublished manuscript. New York University and University of Botswana.

Miller, A. L. (2010). Tongue body and tongue root shape differences in N|uu clicks correlate with phonotactic patterns. In: Fuchs, S., Toda, M., and M. Zygis (eds.) *Turbulent Sounds: An Interdisciplinary Guide*. Berlin: Mouton de Gruyter, 245-279.

Miller, A. L. (2012). *C-V Coarticulation in Consonants with Multiple Lingual Constrictions*. Manuscript submitted for publication.

Miller, A. L. (2013). Northern Khoesan. In: Vossen, R. (ed.) *The Khoesan Languages*.

Taylor & Francis Group, 92-96.

Miller, A. L., Namaseb, L., and K. Iskarous. (2007). Tongue body constriction differences in click types. In: Cole, J., and J. I. Hualde (eds.) *Laboratory Phonology 9*. New York: Mouton de Gruyter, 643-656.

Miller, A. L., Brugman, J., Sands, B., Namaseb, L., Exter, M., and C. Collins. (2009). Differences in airstream and posterior place of articulation among N|uu clicks. *Journal of the International Phonetic Association 39*, 129-161.

Miller-Ockhuizen, A. (1998). Towards a unified decompositional analysis of Khoisan lexical tone. In: Schladt, M. (ed.) *Language, Identity, and Conceptualization among the Khoisan*. Köln: Rüdiger Köppe Verlag, 217-243.

Miller-Ockhuizen, A. (2000). C-V coarticulation and complex consonants: new evidence for the ordering of place gestures in click consonants. In: Fujimura, O., Joseph, B. D., and B. Palek (ed.) *Proceedings of LP '98*. 301-330.

Miller-Ockhuizen, A. (2003). *The Phonetics and Phonology of Gutturals: A Case study from Ju|'hoansi*. New York: Routledge.

Moisik, S., Lin, H., and J. Esling. (2013). Larynx height and constriction in Mandarin tones. In: Peng, G., and F. Shi (eds.) *Eastward Flows the Great River: Festschrift in Honor of Professor William S-Y. WANG on his 80th Birthday*. Hong Kong: City University of Hong Kong Press, 187-205.

Morén, B., and E. Zsiga. (2006). The Lexical and Post-Lexical Phonology of Thai Tones. *Natural Language & Linguistic Theory 24*, 113-178.

Moreton, E. (2006). *Underphonologization and modularity bias*. Manuscript, University of North Carolina, Chapel Hill.

Nakagawa, H. (2006). *Aspects of the Phonetic and Phonological Structure of the G|ui Language*. Ph.D. Dissertation, University of the Witwatersrand.

Nakagawa, H. (2010). Phonotactics of Disyllabic Lexical Morphemes in G|ui. *Working Papers in Corpus-Based Linguistics and Language Education 5*, 23-31.

Nakagawa, H., Sugawara, K., and J. Tanaka. (2014). *G|ui-English-Japanese Dictionary*. (April 2014 preliminary edition). Tokyo University of Foreign Studies and Kyoto

University.

Nienaber, G. S. (1990). Khoekhoen: spelling, vorme, betkenis. *African Studies 49*, 43-50.

Nyati-Ramahobo, L. (2001). Language situation in Botswana. *Current Issues in Language Planning 1*, 243-300.

Odden, D. (1995). Tone: African Languages. In: Goldsmith, J. A. (ed.) *The handbook of phonological theory*. Blackwell Publishing Ltd., 444-475.

Ohala, J. J. (1974). Experimental historical phonology. In: Anderson, J. and C. Jones (eds.) *Historical Linguistics II: Theory and Description in Phonology*. Amsterdam: North Holland, 353-389.

R Development Core Team. (2009). *R: a language and environment for statistical computing.* R Foundation for Statistical Computing. http://www.r-project.org/.

Reetz, H., and A. Jongman. (2009). *Phonetics: transcription, production, acoustics and perception*. West Sussex, UK: John Wiley & Sons Ltd.

Rietveld, A. C. M., and C. Gussenhoven. (1985). On the relation between pitch excursion size and prominence. *Journal of Phonetics 13*, 299-308.

Rose, P. (1987). Considerations in the Normalisation of the Fundamental Frequency of Linguistic Tone. *Speech Communication 6*, 343-351.

Rose, P. (1991). How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency? *Speech Communication 10*, 229-247.

Rose, P. (2002). Independent depressor and register effects in Wu dialect tonology: Evidence from Wenzhou tone sandhi. *Journal of Chinese Linguistics 30*, 39-81.

Rycroft, D. K. (1980). *The Depression Feature in Nguni Languages and its Interaction with Tone*. Grahamstown, Rhodes University. Communication No. 8, Department of African Languages.

Schachter, P. (1976). An unnatural class of consonants in Siswati. In: Hyman, L. M., Jacobson, L. C., and R. G. Schuh (eds.) *Papers in African linguistics in honor of Wm. E. Welmers*. Studies in African Linguistics, Supp. 6, 211-220.

Schapera, I. (1930). *The Khoisan Peoples of South Africa: Bushmen and Hottentots*. London: Routledge & K. Paul.

Schuh, R. G. (1971). Verb forms and verb aspects in Ngizim. *Journal of African Languages 10*, 47-60.

Schuh, R. G. (1978a). Tone Rules. In: Fromkin, V. A. (ed.) *Tone: A Linguistic Survey*. New York: Academic Press, 221-256.

Schuh, R. G. (1978b). Bade/Ngizim vowels and syllable structure. *Studies in African Linguistics 9*, 247-283.

Schuh, R. G. (1998). *A Grammar of Miya*. Berkeley: University of California.

Schultze, L. (1928). Zur Kenntnis des Körpers der Hottentoten und Buschmänner. *Zoologische und Anthropologische Ergebnisse einer Forschungsreise im Westlichen und Zentralen Südafrika 5*, 147-227.

Shen, Z. (1994). The tones in the Wujiang dialect. *Journal of Chinese Linguistics*, 278-314.

Simonet, M., Rohena-Madrazo, M., and M. Paz. (2008). Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish. In: Colantoni, L., and J. Steele (eds.) *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*. Cascadilla Proceedings Project, 72-86.

Smieja, B. (2002). Democratisation vs. language shift and its implications for minority languages in Botswana. In: Legere, K., Fitchat, S., and F. D. Akindele (eds.) *Talking freedom: Language and democratisation in the SADC Region*. Windhoek, Namibia: Gamsberg MacMillan Publishers Ltd, 211-224.

Smith, N. V. (1968). Tone in Ewe. *Quarterly Progess Reports, Research Laboratory of Electronics, MIT 88*, 290-304.

Snyman, J. W. (1975). *Zu|ʔhõansi Fonologie en Woordeboek*. Cape Town and Rotterdam: A. A. Balkema.

Somé, P-A. (1998). L'influence des consonnes sur les tons en Dagara, langue voltaïque du Burkina Faso. *Studies in African Linguistics 27*, 3-47.

Sommer, G., and T. Widlok. (2013). The Sociolinguistic Situation of Khoesan Languages

Today. In: Vossen, R. (ed.) *The Khoesan Languages*. Taylor & Francis Group, 479-481.

Stahlke, H. F. W. (1971). *Topics in Ewe Phonology*. University of California, Los Angeles.

Sun, J. T-S. (2003). Variegated tonal developments in Tibetan. In: Bradley, D., LaPolla, R., Michailovsky, B., and G. Thurgood (eds.) *Language variation: Papers on variation and change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*. Pacific Linguistics, 35-51.

Swackhamer, J. (1991). Tonal Behavior of Podoko Verbs in Independent Clauses. In: Anderson, S. C. (ed.) *Tone in Five Languages of Cameroon*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington, 93-125.

Tanaka, J. (1978). *A San vocabulary of the central Kalahari: G||ana and G|wi dialects*. African languages and ethnography, vol. 7. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.

Tanaka, J. (1980). *The San: Hunter-Gatherers of the Kalahari*. Tokyo University Press.

Tang, K. E. (2008). *The Phonology and Phonetics of Consonant-Tone Interaction*. Ph.D. Dissertation, UCLA.

Thomas-Vilakati, K. D. (2010). *Coproduction and Coarticulation in IsiZulu Clicks*. Berkeley, CA: University of California Press.

Tourneux, H. (1982). Les classes verbales en mulwi (Tchad). In: Jungraithmayr, H. (ed.) *The Chad languages in the Hamito-Semitic-Negritic border area: papers from the Marburg Symposium, 1979*. Berlin: Dietrich Reimer Verlag, 175-182.

Traill, A. (1978). Preliminary report on the linguistic situation amongst the Kua Basarwa living on the cattleposts Bae, Mosetlharobega, Metsimonate in the Central district. *Gaborone, Government Printer, [Appendix 28 in 1979, Kalahari Cattle Posts, vol. 2, ed. Robert K. Hitchcock]*, 252-260.

Traill, A. (1985). *Phonetic and Phonological Studies of !Xóõ Bushman*. Hamburg: Helmut Buske.

Traill, A. (1986). Click replacement in Khoe. In: Vossen, R., and K. Keuthmann (eds.)

Contemporary Studies on Khoisan, vol. 2. Hamburg: Helmut Buske, 301-320.

Traill, A. (1990). Depression without depressors. *South African Journal of African Languages 10*, 166-172.

Traill, A. (1991). Pulmonic control, nasal venting, and aspiration in Khoisan languages. *Journal of the International Phonetic Association 21*, 13-18.

Traill, A., Khumalo, J. S. M., and P. Fridjhon. (1987). Depressing facts about Zulu. *African Studies 46*, 255-274.

Traill, A., and R. Vossen. (1997). Sound change in the Khoisan languages: new data on click loss and click replacement. *Journal of African Languages and Linguistics 18*, 21-56.

UNESCO Ad Hoc Expert Group on Endangered Languages. (2003). *Language Vitality and Endangerment*. Paris: International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages.

Visser, H. (2001). *Naro Dictionary*. D'Kar: Naro Language Project.

Vossen, R. (1984). Studying the linguistic and ethno-history of the Khoe-speaking (central Khoisan) peoples of Botswana, research in progress. *Botswana Notes and Records 16*, 19-35.

Vossen, R. (1991). What Do We Do with Irregular Correspondences? The Case of the Khoe Languages. *History in Africa 18*, 359-379.

Vossen, R. (1997). *Die Khoe Sprachen*. Köln: Rüdiger Köppe Verlag.

Vossen, R. (2013a). Introduction. In: Vossen, R. (ed.) *The Khoesan Languages*. Taylor & Francis Group, 1-12.

Vossen, R. (2013b). Phonetics and Phonology, Shua subgroup, Tshwa subgroup. In: Vossen, R. (ed.) *The Khoesan Languages*. Taylor & Francis Group, 71-75.

Vossen, R. (2013c). Tonology, Tshwa subgroup. In: Vossen, R. (ed.) *The Khoesan Languages*. Taylor & Francis Group, p. 104.

Vossen, R. (2013d). Tonology, Shua subgroup. In: Vossen, R. (ed.) *The Khoesan*

Languages. Taylor & Francis Group, 103-104.

Westphal, E. O. J. (1962). On classifying Bushman and Hottentot languages. *African Language Studies 3*, 30-48.

Westphal, E. O. J. (n.d.). Shua. (BC 1143 (C8), E O J Westphal Papers). Unpublished notes housed at the Rare Documents and Manuscripts Department, University of Cape Town.

Whalen, D. H., DiCanio, C., and P. A. Shaw. (2011). Phonetics of Endangered Languages. *Acoustics Today 7*, 35-42.

Wilmsen, E. N., and R. Vossen. (1990). Labour, Language and Power in the Construction of Ethnicity in Botswana. *Critique of Anthropology 10*, 7-37.

Xu, C. X., and Y. Xu. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association 33*, 165-181.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics 25*, 61-83.

Xu, Y. (2013). *ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis, version 5.5.2*. URL: http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/. In: Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France, 7-10.

Yip, M. (2002). *Tone*. Cambridge and New York: Cambridge University Press.

Yiu, S. S. Y. (2014). Tone spans of Cantonese English. In: Gussenhoven, C., Chen, Y., and D. Dediu (eds.) *Proceedings of the 4th International Symposium on Tonal Aspects of Languages (TAL 2014)*. Nijmegen, the Netherlands, ISCA Archive, http://www.isca-speech.org/archive/tal_2014/, 143-146.

Zee, E. (1980). The effect of aspiration on the F_0 of the following vowel in Cantonese. *UCLA Working Papers in Phonetics 49*, 90-97.

Zhang, J. (2002). *The effects of duration and sonority on contour tone distribution: a typological survey and formal analysis*. New York: Routledge.

Zhang, J. (2006). *The Phonology of Shaoxing Chinese*. Utrecht: LOT.

Zhu, X. S. (1999). *Shanghai Tonetics*. München; Newcastle: Lincom Europa.

Zsiga, E., and R. Nitisaroj. (2007). Tone Features, Tone Perception, and Peak Alignment in Thai. *Language and Speech 50*, 343-383.