UNIVERSITY OF CALGARY

Tumʔi: A Phonetic & Phonological Analysis of a Khoisan Variety

by

Kelly Kilian

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

GRADUATE PROGRAM IN LINGUISTICS

CALGARY, ALBERTA

JANUARY, 2020

Abstract

As part of a linguistic research team I recorded a Khoisan language currently spoken by a linguistic community of three in the northern cape of South Africa. As the location in which this language was discovered is situated geographically close to varieties of both the Khoekhoe and Tuu language families, the question of genetic affiliation and typological similarity within the Khoisan lineages becomes significant. This will be addressed through the analysis of phonetic, phonological and lexical similarities and oppositions between Tumʔi and representative Tuu and Khoekhoe languages (Beach 1938; Bleek 1930; Ladefoged & Traill 1994; Miller et al. 2007).

Overall this project attempts to answer the question, how unique is this undocumented language Tumʔi in comparison to varieties of geographically neighboring Khoisan languages? The analysis is comprised of a detailed description of the vowel and consonant systems, as well as evidence of any contrastive phonetic and phonological features. The clear focus on the analysis of sound contrasts is a consequence of limited data due to speaker competence (Grinevald, 2007). As a result of incomplete acquisition and generational linguistic attrition, the recorded utterances constitute Khoisan content words produced within an Afrikaans framework (Killian 2009). Specific research questions include:

- What is the sound inventory of this language?
- Are there phonation or glottalization contrasts between vowels?
- What click types and accompaniments make up the inventory?
- Are there laryngeal contrasts between consonants?

Results of the analyses indicate the following; Tumʔi shows traces of a phonation contrast, uvular click accompaniment, and evidence of laryngeally marked stops. The phonological typology shares more similarities with the Southern Khoisan varieties of the Tuu family than with

varieties of the central Khoekhoe family. Direct implications for this project include contribution to the current areal typological isoglosses separating the varieties of Khoisan located in southern Africa (Güldemann 2006). The final contribution of this work is the documentation of a moribund Khoisan dialect which has undergone no prior linguistic or anthropological investigation.

Acknowledgments

I have many to thank for the successful completion of this second Masters thesis. My supervisor Dr. William Bennett has provided unending support throughout the processes of data collection, analyses and writing of the thesis. His interest in indigenous languages and language endangerment in South Africa has inspired me to further the research and documentation of endangered languages in South Africa. In terms of the more technical phonetic and statistical aspects of the analyses I would like to thank Dr Stephen Winters. Dr Winters provided extensive guidance in quantitative methods and statistical analyses in his statistics course within the linguistics program. However, even after completing my course work and completing the thesis at a distance, Dr Winters was always available for consultation, and willing to provide advice when needed. I would like to thank Dr Darin Flynn for sharing his knowledge of indigenous language studies and pointing me in the direction of interesting theories and considerations involved in the realm of linguistic fieldwork, indigenous peoples, and the documentation of endangered languages. I would also like to thank the linguistic faculty of the University of Calgary, for all the knowledge and support I was offered throughout my course work and the formulation of my thesis.

Finally, I thank my parents for their immense support throughout my years as an academic. Choosing to produce a second masters and not a doctorate is not the obvious step forward, so having the support of a family encouraging but not pushing me to do more than I was ready for is something I truly appreciate. My parents had so little opportunity awarded to them, so I am especially proud to show my appreciation by grabbing this opportunity to complete a fourth degree.

Dedication

Beginning my research in the documentation of Khoisan languages, I knew that this would be an important part of my progression as both a linguist and an individual. As a coloured person in South Africa my heritage has been stolen from me like so many others. This is a fact based on the revelation that before coloured people were racially labelled by our government we existed as descendants of the original indigenous people. Hence, performing fieldwork and research on Khoisan languages in some sense allowed me access to what is truly my own heritage. What was most difficult to accept about the realities of the endangerment of Khoisan languages, is that the people who had once been nomadic are now simply homeless. It is therefore my honor to dedicate this dissertation to the Tumʔi speaking siblings, my research participants; Elsie, Francina and Robert George, it is my hope that our work together has brought them as much joy and comfort as it has me.

Table of Contents

List of Tables

Table of Figures

Chapter 1

1.1 Introduction

The goal of this research is the documentation and phonological codification of the Khoisan variety Tumʔi, within the fieldwork and constraints of 'salvage linguistics' (Grinevald 2003). Tumʔi is a variety spoken in the small town of Prieska in the Northern cape. The linguistic community consists of a group of three siblings between the ages of 64 and 80. The variety of Khoisan spoken by these individuals has been consensually named Tumʔi, as the speakers had originally referred to their language as *hottentots taal*, which is a derogatory term in South Africa. As outlined in the table of contents, chapter 2 discusses the array of sociolinguistic issues surrounding endangered language fieldwork and documentation. Following this is some general methodological discussion relating to the limitations and constraints of endangered language research. This then leads into discussion of the gaps within the literature concerning the typologies and classification of Khoisan varieties (Du Plessis 2014; Güldemann 2006). Chapter 3 outlines the entire phonological inventory of Tumʔi. This sets the stage for the more in-depth phonetic analyses which focuses on typologically interesting features of the phonological system. The following chapters 4-6 are each dedicated to an in-depth acoustic analysis of a particular feature of interest, including vowel phonation, clicks and click accompaniments, as well as voicing contrasts and other laryngeal features. These topics and structure are designed to answer the following research questions.

1.1.1 Research questions

- What is the sound inventory of this language?
- Are there phonation or glottalization contrasts between vowels?
- What click types and accompaniments make up the inventory?

- Are there laryngeal contrasts between consonants?

Chapter 1 comprises a general outline, accompanied by the main research questions addressed in this paper. Chapter 2 comprises the first content chapter, and specifically discusses the speakers and speaker typologies encountered in this documentation project, as well as issues of linguistic knowledge, loss and insecurity. The chapter includes further discussion of the origin and geographical location of Tumʔi, and introduces discussion of various literature concerning language classification and genealogical relations between Khoisan varieties. General methodological issues are then discussed before the more in-depth acoustic analyses are introduced in the chapters to follow. Chapter 3 constitutes a summary of the entire phonological inventory which answers the initial research question and sets the stage for the deeper investigations into the different sections of the sound system. The sound inventories presented in chapter 3 include brief articulatory descriptions of the segments, as well as discussion of the general and phonological distribution of these segments throughout the data set.

Chapter 4 is dedicated to the analyses of vowels produced with different types of phonation. The phonation analysis includes spectral evidence, which definitively exhibits segments produced with varying degrees of phonation. Statistical analyses provide further evidence of a significant distinction between modal and non-modal groups of segments. Chapter 5 constitutes an acoustic analysis of click types and accompaniments recorded in Tumʔi. The most important discovery of this analysis is the frequent occurrence of the uvular stop accompaniment. A click is generally comprised of one transient burst, followed by irregular oscillations or turbulent noise depending on the click type (Ladefoged & Traill 1994). The lingual-pulmonic segment as defined by Miller et al. (2007) is unique in that the click is comprised of two noise bursts. This click is not only phonetically unique, but also rare in that it is exclusively attested in languages classified within the Tuu family lineage.

The uvular click accompaniment identified in chapter 5 constitutes a unique typological discovery in Tumʔi. Another unique characteristic of the typology is identified in chapter 6, which comprises a selection of acoustic analyses related to voicing contrasts and other laryngeal features. Evidence of the feature [+voice] is typologically significant in that languages of the Khoekhoe family, with the exception of !Ora, do not exhibit voicing distinctions between pulmonic egressive segments (Vossen 2013). However, Khoisan lineages with more complex sound systems such as Tuu encompass multiple languages in which voicing contrasts between egressive segments are attested. Evidence of ejectives constitutes another distinctive phonological characteristic of Tumʔi identified in chapter 6. These segments are produced using the glottalic egressive airstream mechanisms, another typological feature restricted to Southern Khoisan varieties. Chapter 6 also addresses disparities in aspiration between tokens of the voiceless alveolar plosive, which would constitute another feature unattested in the majority of Khoisan varieties classified within the Khoekhoe, and more generally the Khoe-Kwadi lineage. The analyses in chapter 6 considers acoustic evidence of additional aspiration as well as evidence of distinctly unaspirated segments. Based on consistency across multiple tokens of the same lexical item, the aspirated segment [t^h] is included in the consonant system presented in chapter 2, section 2.2.

The content in chapters 4-6 therefore encompass the bulk of the phonetic analyses which reveal the phonological vestiges of the linguistic system of Tumʔi presented in chapter 3. Considering the limitations of the data and related constraints on the analyses, the final description of Tumʔi is relatively extensive in terms of segment transcription and analysis. The structured phonetic analysis of this phonological inventory delivers a reliable sound system for future comparative analyses.

Chapter 2: Sociolinguistic context and methodology

2.1 About the people: Speaker type & knowledge

There are three known speakers of this language; Elsie George, age 80; Francina George, age 70; and Robert George, 64. As siblings raised in the same home they have all been exposed to the same variety of Khoisan, however their levels of exposure and resulting knowledge vary. This variation is mainly attributed to their substantial differences in age, with Elsie noticeably being the authority on the language and Robert being the least knowledgeable speaker. Aside from the age differences there is also the gender distinction, as a non-white male in South Africa Robert was required to leave home, to work on the nearby farms and contribute financially to the household. Hence, Roberts exposure to the language ended when he became a teenager, unlike his sisters who attended school for a longer period and spent most of their time helping in and around the home. Hence, within this linguistic community the speaker knowledge constitutes a continuum of proficiency, though not extending from full fluency (Dorian 1977).

The least amount of data elicitation and recording was conducted with the youngest sibling, Robert George. Robert participated in three recording sessions accumulating to a total of eight and a half hours of elicitation. As one of the youngest of 12 children Robert often emphasizes his lack of knowledge of the language due to his limited interaction with his mother as an adult. However, he attributed the linguistic knowledge he had contributed to the fact that being the baby he spent the most time at his mother’s side as a child. Robert eventually goes on to acknowledge his partial acquisition of what possibly had been an already extremely endangered language. However, he further states that any linguistic abilities he may have acquired at a young age were lost after years of conducting his primary interactions in languages other than Tumʔi. Generally, Robert exhibited the most linguistic insecurity, constantly anticipating his inability to speak or remember anything that might be worthwhile eliciting. At this point Robert will only use a few fixed phrases when communicating with his older sisters in Tumʔi. An example would be Robert saying to Francina he has a weak bladder and needs to leave, which is uttered as "[ǀqu ha̰a̰] (I have a weak bladder), [ek wil |ami] (I want to pee)". The phrase 'I have a weak bladder' is produced in Tumʔi as [ǀqu ha̰a̰], the verb phrase 'I want to' is produced in Afrikaans as [ek wil], the noun 'pee' is then produced in Tumʔi as [|ami], resulting in a mixture of constituents produced in both Tumʔi and Afrikaans, with [ǀqu ha̰a̰] being the fixed phrase in Robert's linguistic repertoire.

Francina, who is the second eldest sibling has substantially more lexical knowledge than Robert, which is evident in her consistency in producing the appropriate lexical items matching previous elicitations. Francina has a charismatic tendency to speak in metaphors. Often when referring to Robert's limited linguistic knowledge she would state that, as the baby, Robert only had access to the crumbs of the language, implying that he could only pick up or acquire whatever he happened to be exposed to. In terms of her own acquisition Francina states that as children they were never explicitly taught the language. Generally, children were not invited to linger around adult conversations and so were kept from acquiring a lot of the more productive linguistic skills. However, being closer to Elsie in age and kinship Francina was able to gain more linguistic knowledge through their interactions. Although not the most knowledgeable of the three speakers the research team relied on Francina most during the elicitation sessions. She often adopted a facilitative role, encouraging participation from the linguistically insecure Robert and the more aged Elsie.

Elsie George is the eldest sibling and most knowledgeable consultant who participated in this project. Elsie spent most of her youth at home performing household duties alongside her mother. She is referred to as a motherlike figure by her siblings, who often talk about how she helped them as children with basic tasks like bathing. Based on the collected data it is clear that Elsie consistently makes phonological distinctions that her younger siblings do not. This is conceivable since she had had the most interaction with their mother as an adult, and therefore made more productive use of the language for a lengthened period of time. However, between the time of the initial data collection undertaking and the final fieldwork effort, Elsie has developed a dementia condition. Three months after discovering this language and establishing contact with these speakers Elsie can no longer make the kind of contribution to the elicitation process that had initially been driven by her extensive linguistic knowledge. This type of deterioration and loss in a linguistic community of only three siblings is substantial and seriously affects the life expectancy of the already endangered language.

The importance of speakers as a source for endangered language research and fieldwork cannot be overstated. Unlike speakers of non-endangered languages, the complete population of an endangered language may consist entirely of marginal speaker types, speakers with limited linguistic competency as a result of several interrelated factors. At the time these speakers were born, Khoisan people had already been subjected to decades of oppression and coercion by the South African government to reidentify themselves as “Afrikaans-speaking Christian ‘coloured’” communities (Killian 2009, pg. 12). This process of reclassification involved not only abandoning their indigenous identities but also their languages. Hence the speakers of Tumʔi acquired an already extremely endangered language--a language used only in the home and rarely directed to them, as the general attitude toward the use of the language was one of trepidation. Since linguistic insecurity is noted as a contributing factor to linguistic competency, I mention a fourth sibling, who could be referred to as a ‘Ghost speaker’ (Grinevald 2007). Despite evidence of exposure based on the other siblings’ knowledge of the language, the fourth sibling refused to participate in the data collection process, denying any knowledge of the language. Even though the fourth sibling had acquired this language to some extent, her own linguistic insecurities stopped her from agreeing to help piece together the remnants of their language. Ghost speakers present an extended challenge for researchers of endangered languages, as the general pool of speakers or linguistic communities we can access is already so small.

Though the ‘semi-speaker’ has been found to be emblematic of the endangered language situation, the consultants involved in the documentation of Tumʔi are better described as ‘terminal speakers’. The negative connotation associated with this profile has caused some debate across the literature on speaker typology, however it most accurately encompasses the distinction between the linguistic abilities of the “semi-speaker” and that of the consultants involved in this project. The speakers of Tumʔi retain a passive knowledge of the language, even the eldest and most knowledgeable speaker has very limited productive skills. Conversation between speakers is conducted primarily in their dominant language Afrikaans, with the insertion of fixed phrases and expressions from their endangered language (Grinevald 2003). The reality of the situation in the case of endangered language study is that the reduced use of the language generally leads to a reduced form of the language (Dorian 1977, pg. 24).

2.1.1 Origin & location

The language as it exists today was discovered in a town called Prieska, in the northern cape of South Africa. According to the distributional mappings of Khoisan languages and migratory routes there is the expectation that speakers of Nama, Griekwa, |Xam, N|uu and possibly |Auni might also be found in the smaller more rural sedentary towns of the northern cape (Kilian 2009). These varieties are classified within the Tuu and Khoekhoe language families which have both been sited around the lower course of the Orange River (Güldemann 2006), including the later analysed N|uusaa which shares most features associated with the Tuu language family (Güldemann 2006). The assortment of Khoisan varieties located in this area are widely attributed to the dichotomy established between the Khoekhoe and Tuu communities approximately 2000 years ago. Hence, the Cape is a unique area in terms of linguistic isoglosses and the distribution of Khoisan varieties (Traill 1995). Most varieties which had previously existed, or currently exist in this area have not been recorded, as a result there remain extensive gaps in the distributional mappings, and missing links in genealogical relations (Güldemann 2006). The consultants were also forthcoming about remaining Khoisan speakers in a town called De Aar, located only 184km south of Prieska. It therefore appears that the speech communities in which Tumʔi and the other possible variety of Khoisan in De Aar exist are extremely small and may only include a few semi- or terminal speakers (Grinevald 2003).

The consultants could not recall much about their mother but refer to her as a "regte hottentots vrou" (a real hottentots woman). They were also certain about her position within the community as a herbalist or herb doctor, hence the speakers' extensive knowledge of herbs and medicinal plants. In view of the fact that the consultants acquired this language from their mother, her place of birth would be better representative of the origins of this language. According to our speakers their mother was born and raised in the Krieberge, formally known as the Kareeberg. This area constitutes an isolated group of hills situated in the area demarcated by the red dividing line in figure 1 below. Also located within this demarcated area are the towns; Vanvyksvlei; Vosburg; and Carnarvon. This entire area is geographically neighboring both Prieska and De Aar, providing motivation for the possible distribution of this particular variety across these areas, see figure 1 below. Generally, however, considering the information above, the area stretching across the northern cape could consist of a range of Khoisan varieties related to either the Khoekhoe or Tuu language families. Two varieties of |Xam were found and recorded in Kenhardt in the late 1800s, which is located only 190km east of Prieska, and the Kareeberge (Traill 1995; Anders 1935).

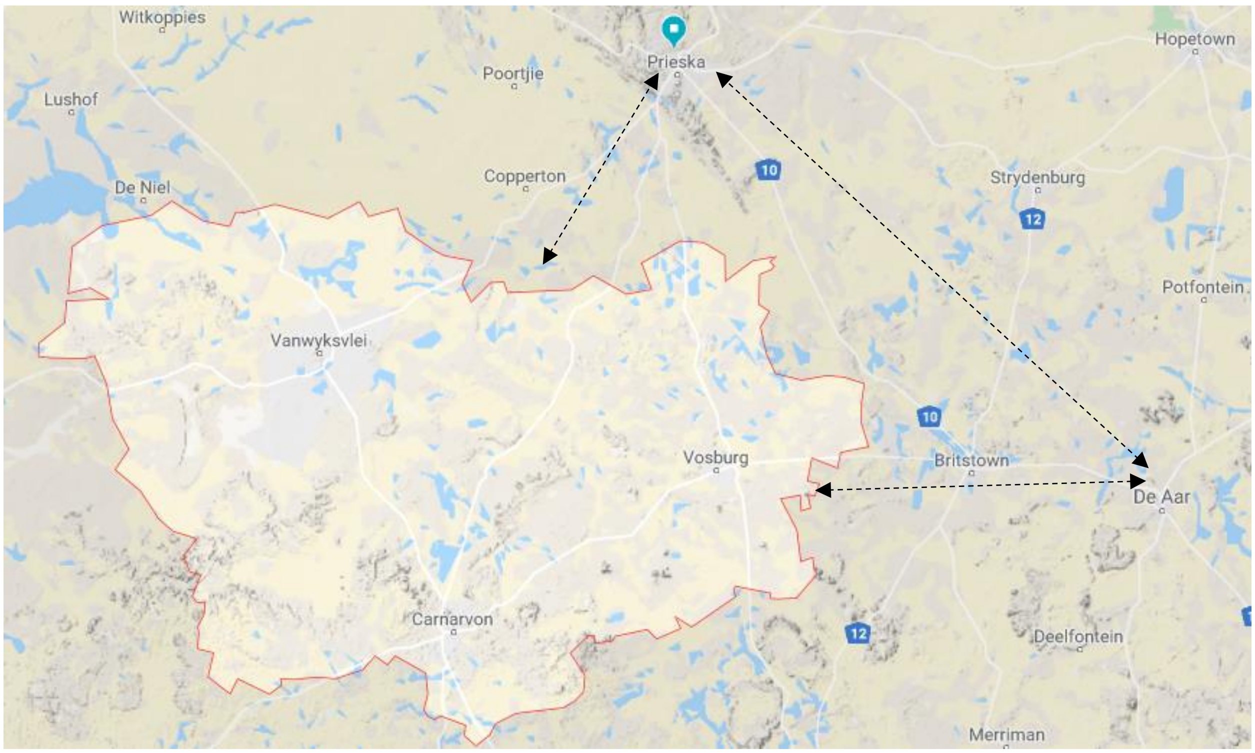


Figure 1: Areas of interest in relation to place of elicitation

Figure 2: Areas of interest and place of elicitation within South Africa

2.2 History & Classification

Established genetic relations between languages are generally based in the comparative-historical framework. Khoisan varieties are therefore classified as belonging to a particular family based on either lexical, syntactic, morphological or phonological similarity (Starostin 2008). Establishing this form of genetic affiliation therefore generally requires a substantial amount of data most commonly constituting a sizeable cognate vocabulary. The ideal situation in terms of data would also include comparable grammatical and morphological information, however, aside from Honken's (1977) analysis, the comparative work on Khoisan has been largely reliant on lexical material (Güldemann 1997). Though many comparative studies include lexical items elicited according to a form of the Swadesh list (Güldemann & Loughnane 2012, pg. 217), the basis for classification is rarely tenable. Comparative research including Greenberg (1963) &

Starostin (2003) are accused of proposing cognate correspondences which do not meet the minimal standards of plausible lexical comparison (Güldemann & Loughnane 2012). At the other end of the spectrum, research performed by Sands (1998) is criticized for adopting exaggerated demands for similarity between possible correspondences, including the requirement for identical click influxes between corresponding cognates. The approach to lexical comparison adopted in this analysis is therefore one which acknowledges the role of language contact in the shared vocabulary observed across the main Khoisan lineages. However, in the same breath, lexical comparisons which constitute essential semantic domains should be heavily weighted in considering genealogical relations (Güldemann & Loughnane 2012), while instances of three-way cognate possibilities as presented in Appendix A, pg. 105-106, indicate a strong affinity between languages rather than instances of typological coincidence (Hastings 2001).

Issues of linguistic genealogical classification are further compounded in the case of endangered languages. Endangered language research is pervaded by uncontrollable factors, including limited available data for particular varieties and even entire language groups (Grinevald 2003). Due to insufficient data, the choice of sample language for comparative analysis is extremely constrained, with only one available option representative of any genetic family (Güldemann 1997). In the case of Tumʔi, a reasonable amount of lexical and phonological similarity is shared with the most extensively documented member of the extinct !Ui family (Du Plessis 2014, pg. 575), ensuring access to data for lexical comparison. The current classification adapted from Güldemann & Loughnane (2012, pg. 216) as presented in figure 3 below, describes a lineage comprised of two isolate languages and three language families. These include the isolated languages Hadza and Sandawe, the Khoe-kwadi family consisting of one isolate language, and one sub-family which is further divided into two sub-families, Kalahari and Khoekhoe. The lineage includes two more families Kx'a and Tuu, which both consist of at least one sub-family and an isolate language. This study is confined to lexical comparisons between Tumʔi and representative languages from the Khoekhoe and the Tuu families. The representative languages considered in the cognate list are of the most accessible of the documented varieties classified within that language family.

Lineages + branches		Languages/dialect clusters	
(1) Hadza		single language	
(2) Sandawe		single language[1]	
(3) Khoe-Kwadi			
Kwadi†		(single language)	
Khoe [= "Central Khoisan"]			
Kalahari			
	East	Shua:	Cara, Deti†, ǀXaise, Danisi, etc.
		Tshwa:	Kua, Cua, Tsua, etc.
	West	Kxoe:	Khwe, ǁAni, Buga, Gǀanda, etc.
		Gǁana:	Gǁana, Gǀui, etc.
		Naro:	Naro, ǂHaba, etc.
Khoekhoe			
	North	(Eini†) DC	
		Nama-Damara DC	
		Haiǁom	
		ǂAakhoe	
	South	(!Ora-Xiri) DC	
		(Cape varieties†) DC	
(4) Kx'a [= Ju-ǂHoan]		Ju DC [= "Northern Khoisan"]	
			North: Angolan !Xuun varieties
			North-central: Ekoka !Xuun, Okongo !Xuun, etc.
			Central: Grootfontein !Xuun, etc.
			Southeast: Tsumkwe Juǀ'hoan, Epukiro Juǀ'hoan, etc.
		ǂHoan	
(5) Tuu [= "Southern Khoisan"]			
Taa-Lower Nossob			
Taa DC		single language complex	
			West: West !Xoon, (Nǀuǁen)
			East: 'Nǀoha, East !Xoon, (Nǀamani), (Kakia), etc.
Lower Nossob		(ǀ'Auni†)	
		(ǀHaasi†)	
!Ui		Nǁng DC:	Nǀuu, (ǂKhomani), (Nǀhuki), etc.
		(ǀXam†) DC	
		(ǂUngkue†)	
		(ǁXegwi†)	

Note: DC = dialect cluster, † = extinct, (older sources), [previously used name]

Figure 3: Khoisan lineage as presented by Güldemann & Loughnane (2012, pg. 216)

Based on the lexical comparison of Khoekhoe languages; Nama and !Ora (Beach 1938; Killian 2009); and Tuu languages, |Xam (Bleek & Lloyd Dic.; Güldemann 2006), N|uu (Brugman, Miller & Sands 2007), N|usaa (Güldemann 2006), and |Auni (Hastings 2001), the lexical items of Tumʔi appear to share more cognates with the languages belonging to the Tuu family. The lexical correspondences between Tumʔi and Tuu languages are evident in the comparison of basic action verbs, masculine and feminine nouns, as well as nouns related to body parts (see appendix A for a comparative table). The strongest cognates include the words for man, child, head, eyes, drink, sleep and leave/walk. Many of these lexical items share direct correspondences with the lexical items reported for |Xam, the most extensively documented and best known of the !Ui languages (Du Plessis 2014, pg. 575), while the general majority of lexical correspondences are observed across the representative Tuu languages. Phonological correspondences are also observed across the lexical items of the Tuu languages and those of Tumʔi and are substantial relative to that observed across the lexical items of the Khoekhoe languages. The phonological correspondences include similarities between click influxes, as well as vowel quality and the use of diphthongs, which are either transcribed as disyllabic vowel combinations in Khoekhoe, or only reported to occur in rapid speech (Beach 1938, pg. 49). Furthermore, there appears to be a high degree of functional weight in the distribution of click sounds within the comparative table as well as the entire collected data; the majority of the lexical items contain clicks. Languages of the Tuu family are generally reported to possess larger click inventories with a heavier functional load than languages of the Khoekhoe family (Güldemann 2006, pg. 106-109; Miller et al. 2009).

Overall, the cognates compared in Appendix A are limited and cannot be presented as a definite indication of a genealogical relation between Tumʔi and the Tuu family lineage. However, these lexical and phonological correspondences provide a basis for the origin of the distinctive phonetic and phonological features identified in the chapters to follow. Furthermore, as discussed above, there are uncontrollable constraints both on the collected data as well as the stock available for cross-linguistic comparison (Güldemann 2006), including issues relating to quality of transcriptions and a general lack of data available for particular varieties, further limiting the available stock (Starostin 2008, pg. 345-346). Due to the vast geographical distribution of Tuu languages, it would be more accurate to refer to any genetic affiliation as a language complex or dialect cluster, hence introducing the possibility that different varieties may be similar or mutually unintelligible depending on their proximity.

2.3 Field work in salvage linguistics

The documentation of endangered languages, also referred to as 'salvage linguistics' is a fast-growing branch of linguistics currently integrating multiple disciplines in projects, extending the scope of linguistics. Linguistic analysis remains the centre of these programs, which by extension are constrained by resources available for documentation (Grinevald 2007). The unpredictable state of resources in the context of endangered language study require that researchers consider the feasibility of documenting a particular language, based on multiple criteria. Therefore, in a less than ideal situation such as in the context of Tumʔi, the project is motivated mainly by the preservation of an obsolescing language and culture with little hope of revitalisation (Ladefoged 1992). Most of the acclaimed work on the documentation of endangered language took place in the early to late 1900s, hence practices in documentation today are considerably more effective in terms of recording methods due to technological advancements and the improvement of elicitation tools. However, what has remained constant are the myriad obstacles and struggles surrounding the methodological process of endangered language documentation (Grinevald 2007).

Many obstacles are solely the consequence of time. In this particular context the speech community is extremely small and had not had any previous experience with anthropological or linguistic documentation efforts. The participants were therefore not jaded in any way by the approach of previous researchers. However, considering the current state of Tumʔi as constituting remnants of a Khoisan language, even defining this as a speech community becomes a challenge. Furthermore, as the first and likely last linguist to enter this community there is an obligation to collect as much material as possible, whether relevant to one’s own research interests or not (Dorian 1977). This had been a particularly difficult task, again due to the effects of time and the indivisible relationship between endangered languages and loss. As mentioned in section 2.1, the consultants who participated in this project are best described as terminal speakers of the language. Therefore, not only is there a constraint on the collective knowledge and data available, but also on access to any natural language use due to the size of the speech community. Loss of speakers within a speech community results in the loss of certain varieties and use in natural settings including traditional ceremonies and cultural practices (Grinevald 2007). An additional effect of loss is triggered by the speaker’s individual loss of linguistic confidence. This particular aspect of loss had a considerable effect on the elicitation process conducted for this project. There were numerous points at which the elicitation process was slowed or interrupted due to confusion or a sense of shame experienced by the speakers.

2.4 Elicitation & data

The elicitation sessions were conducted across two data collection trips, the first only three days long, and the second spanning over six days. The data accumulated on the initial trip is comprised of a total of three hours and five minutes of recorded data. A total of ten hours and six minutes of elicitation was conducted on the second trip, which was significantly more than the initial trip as the second trip was initiated solely to collect data of this specific variety. The method of data collection consisted mainly of single word elicitations using word lists adapted from previous fieldwork expeditions. The word list consists of basic verbs and actions as well as body parts and a few possible nouns relevant to the lifestyle and surroundings of the indigenous community, see appendix G, pg. 119-121 for the original wordlist. Picture prompting was also used as a method of elicitation. During this process participants were shown pictures of indigenous plants and animals taken from Branch et al. 2001, Picker et al. 2002, Iwu 2014, and Van Wyk 2013, for which they were asked to provide a name, generally resulting in the elicitation of a different but semantically related lexical item.

Furthermore, as a consequence of the loss discussed above, the research team was unable to collect any traditional stories or productive interaction solely conducted in Tumʔi. Due to the same limitations the team was also unable to collect any complete sentences uttered entirely in the endangered language. The extent of the syntactic information collected constitutes one or two partial phrases (see Appendix C). The data therefore consists mainly of individual lexical items which do not appear to display any derivable morphological information. Further conditioned by the context of endangered language research there appears to be extensive phonetic, phonological and semantic variation across the lexical items. Within this research framework, variation of this nature is attributed to the lack of any documented norm as many endangered languages such as Khoisan are historically spoken and not written. Ultimately the constraints on elicitation and on the collected data have directly constrained the type of linguistic analysis realizable for the documentation of this language variety.

The elicitation sessions were conducted in the homes of the participants in an attempt to avoid any uncontrollable noise interference, but also to ensure the participants were comfortable.

The equipment used on the initial trip included a zoom h4n recorder as well as a Tosh com DR 40. The microphones were Nady HM-10 mics, with XM-10 phantom power. Two recorders were also used for the second data collection trip, including the zoom h4n recorder and a Hi-MD Walkman MZ-RH1 recorder. In addition to the Nady HM-10 mics we also used the Apex570 Lightweight condenser headset microphone. The analyzable data therefore includes transcriptions and recordings available for phonetic and phonological investigation.

Chapter 3: Phonological inventory

3.1 Vowels

Distinctions in vowel quality are generally determined by three aspects including vowel height, vowel backness, and lip rounding. These aspects are reflected in the acoustic properties of vowels, the most prominent of which are the formants (Ladefoged & Maddieson 1996, pg. 104). Vowel height is proportional to the inverse of the frequency of the first formant (F1), while vowel backness is proportional to the frequency of the second formant (F2), or the difference between the frequencies of the first and second formants. Finally, the degree of lip rounding is generally indicated by the lowering of the second and third formants (F3) (Ladefoged & Johnson 2014, pg. 217). The vowel charts presented in figures 4 & 5 below do not include F3 values, therefore degrees of lip rounding are not a central aspect of this analysis and are based solely on the articulatory features reported in the production of each vowel. The F1 and F2 measures are precisely arranged along the axes to provide the most informative phonetic view of the vowel distributions. The F2 measures are indicated along the horizontal axis, with the value increasing from right to left. The F1 measures are presented along the vertical axis, with the value increasing downward. Plotting the formant measures according to this scale provides a visual representation of the acoustic features of the vowel that correspond roughly to the articulatory dimensions of the vocal tract (Ladefoged & Maddieson 1996, pg. 131).

The vowel systems of Khoekhoe dialects are defined by three distinctive features: the limited number of vowel phonemes, the lack of diphthongs except in instances of very rapid speech, and the phonemic presence of both oral and nasal vowels (Beach 1938, pg. 35). The vowel systems of northern Khoekhoe languages as well as !Ora are reported to include the five basic vowels /i e a (ə) o u/, with the mid-central vowel generally considered neutral and non-phonemic.

These five basic vowels are each reported to have a nasal counterpart, with the exception of the mid-front vowel /e/. Furthermore, laryngeal phonation and related vowel colourings are not reported as distinctive features in any of these varieties (Vossen 2013, pg. 150). Southern Khoisan varieties, though generally reported to exhibit the same basic five vowel system, additionally exhibit attested nasal contrasts, as well as distinctive laryngeal contrasts across multiple varieties (Garellek 2019; Miller et al 2009; Vossen 2013, pg. 208). Based on data collected by Ziervogel & Potgieter, and Dorothea Bleek, a phonetic triangulation of the Tuu language ǁX'egwi, reports: a vowel system with a phonemic oral-nasal distinction, variation in openness across mid-vowels [o] and [e], as well as what are referred to as 'true diphthongs', including [oa] and [əu] (Hastings 2001). Analyses of other reportedly extinct Tuu varieties such as N|uusaa and |Xam report consistent occurrences of diphthongs as VV sequences of unlike vowels (Güldemann 2006, pg. 384, Appendix 1). These transcribed sequences include the combinations [ai] and [ui] observed in the words [tai] 'leave', and [!ui] 'man', which form strong cognates with the same lexical items found in |Xam, |Auni, N|u, and N|uusaa (see Appendix A). Wide-spread analyses of the vowel systems documented for different Tuu varieties therefore report a general presence of diphthongs as well as contrastive phonation, which both appear to be rarely attested in varieties of the Khoekhoe lineage (Vossen 2013).

The vowel charts presented in figures 4 & 5 reflect the vowel system of Tumʔi. The vowels plotted represent the average formant measures of vowels produced by each of the three speakers. Evident across these charts are frequent occurrences of both monophthong and diphthong phonemes, as well as indications of possible openness variation within the mid-vowel range. The distinctions presented in the vowel systems below, as well as the investigation of phonation colouring in Tumʔi, which is discussed in detail in chapter 4, reflect a typology common to Khoisan varieties with more complex sound systems.

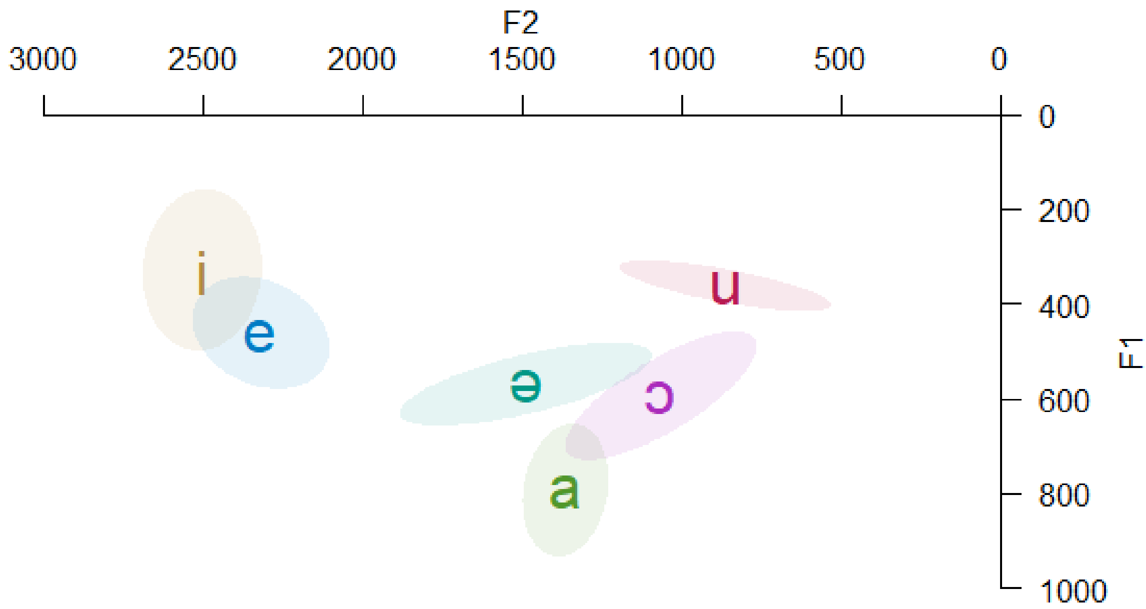


Figure 4: Formant plots of Tumʔi monophthong vowels. The ellipses are drawn according to the covariance calculated for the tokens, and a default confidence interval ellipse.

3.1.1 Notes on monophthongs

The phonemes [i] and [u] constitute the high vowels of Tumʔi. [i] is produced as a close front unrounded oral vowel with little variation across tokens. The phoneme [u] is articulated as a close back rounded vowel and exhibits a wider distribution than its front counter-part [i]. [u] occurs mainly as the first vowel in the initial syllable in a word, indicating that it may be a vowel phoneme designated to particular roots. This distribution may be explained by the fact that lexical roots ending in [u] and [i] are commonly reported in Khoisan languages (Beach 1938, pg. 42). The mid-vowels identified in the inventory include the phonemes [e], [ɔ] and [ə]. The phoneme [e] is articulated as a half-close front unrounded vowel and appears to occur mainly in word-final position, or as the nucleus of the final syllable. The back mid-vowel counterpart to [e] is articulated as the open rounded vowel [ɔ], which appears frequently as the nucleus of the first and second syllable. Based on the vowel chart in figure 4 above, the mid-back vowel is positioned both lower and more centralized within the vowel space than the mid-front vowel [e]. These disparities further indicate that the mid-back vowel is generally positioned closer to the 6th cardinal vowel and is hence transcribed as [ɔ] and not [o] (Ladefoged & Johnson 2014, pg. 218). Finally, the phoneme [ə] is articulated as an unstressed central vowel. The schwa in Tumʔi generally appears in the second or final syllable, only observed in the first syllable of [pəri] 'goat' and Afrikaans loan words such as [məkəs] 'upper thighs' and [vələ] 'wild'. The phoneme [a] is articulated as a low central unrounded vowel, distributed evenly and frequently throughout the data. This unrestricted distribution in position and co-occurrence with other vowel phonemes would align with attested distributions of the low vowel [a] in other Khoisan languages (Beach 1938, pg. 39; Güldemann 2006, Appendix 1).

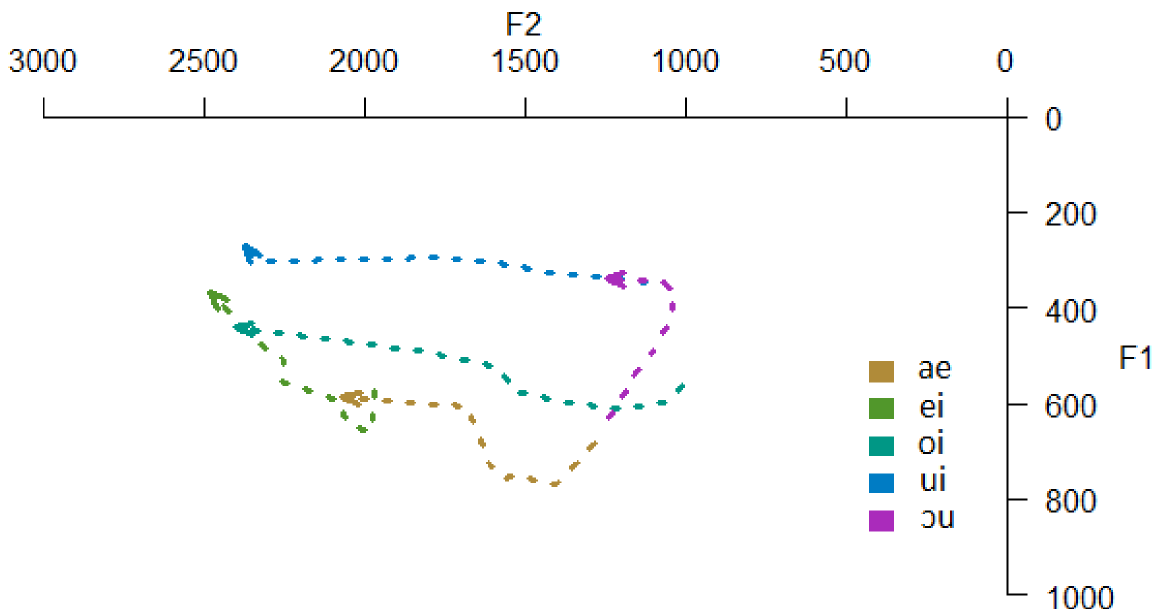


Figure 5: Formant plots of Tumʔi diphthong vowels. The line segments connect the mean values of the formants retrieved at multiple time points throughout the vowel.

3.1.2 Notes on diphthongs

The vowel system of Tumʔi includes a total of five diphthongs, all of which end with a high vowel, either with a preceding low, mid or back vowel. This is a significantly large collection of 'unlike vowel sequences' (Güldemann 2006, pg. 384), which appear to be unrestricted in terms of position within the word as well as in terms of co-occurrence with other vowel phonemes. The phonemes [ei], [oi] and [ui] all reflect a shift to the high vowel [i]. The vowel chart in figure 5 above displays three diphthongs which glide to [i] and end at the same general height and front position. The diphthong [ae] which could be perceived as the higher ending sequence [ai] glides to a distinctly lower front position. Though the phoneme [i] in the diphthong [ai] is generally expected to be articulated as more open and flatter than when produced as a monophthong (Beach 1938, pg. 36), the end point of the low diphthong in the chart above is most accurately transcribed as [ae].

3.2 Consonants

Consonant phonemes are generally described by the characteristics of place of articulation and manner of articulation. As clearly indicated by the columns in table 1 below, the articulators involved in producing egressive consonants in Tumʔi include the lips, the tongue positioned at the alveolar ridge or velum, and the closure of the glottis. The consonant inventory of Tumʔi also consists of different manners of articulation. This includes seven plain plosives produced with complete closure of the vocal tract, followed by a release burst. The aspirated alveolar plosive is the only aspirated segment and is phonetically distinct from the plain voiceless plosive [t]. As discussed later in chapter 6, the voiceless alveolar plosive is the only segment proven to consistently exhibit a contrast in aspiration. The consonant inventory also includes three nasal stops produced with complete closure of the oral cavity at either the bilabial, alveolar or velar places of articulation, with a lowered velum allowing airflow through the nose. Sounds with close approximation of articulators but not complete closure include the six fricatives and single affricate, as well as the labio-velar and lateral approximant presented in the inventory below. Finally, this consonant inventory also includes an alveolar trill which is produced with continuous tapping of the tongue-tip against the alveolar ridge. Generally, a consonant inventory of this size is not commonly observed across Khoisan dialects, with most exhibiting a limited number of egressive consonants. However, there are segments which are commonly observed across the phonological systems of Khoisan languages but are not attested in Tumʔi, including the affricates [ts] and [kx] (Beach 1938, pg. 65-67; Güldemann 2006, pg. 11; Killian 2009, pg. 27).

Table 1: *The Tumʔi non-click consonant inventory*

	Bilabial	Lab-dental	Lab-velar	Alveolar	Velar	Glottal
Plosive	p b			t d tʰ (tʲ)	k g	ʔ
Ejective				(t')		
Fricative	(β)	f v		s	x	h
Affricate				(tʃ)		
Nasal	m			n	ŋ	
Approximant			w			
Lateral approximant				l		
Trill				r		

3.2.1 Notes on Consonants

The [t] phoneme is the most frequently used plosive, occurring both word initially and intervocalically. This is also the only phoneme which may be produced with additional aspiration, or as distinctly unaspirated. Though the degree of aspiration has been exhibited to vary across tokens of the same lexical item, [tumʔi] 'speak', the aspirated segment [tʰ] is produced consistently in tokens of the lexical item [tʰi!qo] 'God', while the unaspirated [t] is consistently produced in tokens of the lexical item [tɔrəŋtɔrəŋ] 'crazy'. [k] exhibits the second highest distribution within the collected data, also occurring both in the word initial and word medial position. The final voiceless stop [p] occurs both word initially and in the word final position, indicating that the voiceless stop consonants are more widely distributed than the voiced stops, [d], [b], and [g], which occur in either the word initial or word medial position and are each transcribed in no more than three lexical items.

The fricatives [s], [f], and [v], are limited to either the word initial position, or the word medial position. The phoneme [h] also occurs both word initially and medially, however it appears to restrictively precede the low vowel [a]. [x] appears to be the most frequently occurring and unrestricted fricative, extending to the accompaniment of clicks in variations of particular lexical items. The bracketed phoneme [β] is transcribed in tokens of a single lexical item [tava] 'handywork'. This bilabial fricative appears to alternate with the phonemes [v] and [w] across tokens of this particular lexical item. The nasal phoneme [n] is frequently transcribed across the data set, however it occurs mainly as an accompaniment to one of the four click types in chapter 5 section 5.5.1. The actual phoneme [n] as well as the phoneme in the form of a click accompaniment appear in both the onset and coda position of the syllable, however in the word initial position the phoneme [n] is only ever present as a click accompaniment. The other pulmonic egressive nasal phonemes [m] and [ŋ] share similar wide-spread distributions, however the phoneme [ŋ] is clearly more restricted in its distribution than [m], occurring only in the coda position.

Particular phonemes recorded infrequently and inconsistently include [l], [tʃ] and [t^j]. The lateral approximant [l] is recorded solely in the Afrikaans borrowing [vələ] 'wild', which is used in reference to different types of medicinal plants. This includes lexical items like [vələals] and [vələ-kier], which both refer to unspecified plants used for medicinal purposes. The phonemes [tʃ] and the [t^j] are among those bracketed in table 1, which means though these segments have been transcribed, their phonemic contrastiveness cannot be verified. The segments [tʃ] and [t^j] are both transcribed in tokens of the lexical item [kut͡ʃaka]/[kut^jaka] 'go out'. Consequently, these phonemes are bracketed and not definitely proposed as phonemes of the consonant inventory, due to alternating production across tokens of a single lexical item, and their absence in any of the other lexical data.

3.3 Clicks

Click segments are a unique class of sounds defined by the rarefaction of air concealed by two articulatory closures (Ladefoged & Maddieson 1996, pg. 246). Numerous descriptions of clicks are presented throughout the literature on Bantu and Khoisan languages, and are generally concerned with the phonological oppositions between different clicks. However, in-depth phonetic analyses of the click systems of particular Khoisan languages have provided evidence for extending these phonological distinctions and identifying more fine-grained differences between the articulation of different clicks (Ladefoged & Traill 1984). The general acoustic analysis of clicks is concerned with the waveform and spectra which align with the release of the anterior closure. The analysis of this closure release provides a phonetic interpretation of the sound produced as a result of the articulators separating and the rapid change in the configuration of the vocal tract (Ladefoged & Maddieson 1996, pg. 257). The release of the posterior closure is generally described as a voiceless velar accompaniment which is not reflected in the acoustics, or a velar nasal accompaniment depending on the position of the velum (Ladefoged & Traill 1994). However, the stance adopted in this paper is one proposed by Miller at al. (2007) which states that the posterior release is uvular, with possible contrasts in voicing and discrepancies relating to the position of the anterior closure. The proposition that the posterior release is positioned at the uvula and not the velum is based on the analyses of ultra sound data collected from speakers of N|uu. The importance of this claim is related to the discovery of the lingual-pulmonic click, which is acoustically remarkable in that it produces two noise bursts (Miller 2011, pg. 420). The latter burst displays the features of a uvular stop. Following the assumption that this is a voiced counterpart to the general posterior release, both accompaniments are described as uvular (Miller 2010; Miller et al. 2007).

As click sounds are indigenous to Khoisan languages, it is typically challenging to distinguish between particular languages and dialects based solely on their click inventories. Comparative studies of representative languages belonging to different Khoisan families discuss the differences and similarities between these click inventories, with the general diagnosis that languages of the Tuu family contain a larger variety of clicks and accompaniments (Childs 2003). However, results of areal typological analyses have reported the effects of language contact on click inventories to be extensive, with Khoekhoe languages exhibiting a similar reliance on click segments as observed in Tuu languages (Güldemann 2006). Therefore, the distribution of click types within a linguistic inventory, and the functional load of click segments across collected data sets may not be sufficient to distinguish between the phonological typologies of Khoisan varieties classified within different lineages. The click inventory of Tumʔi as spoken today is limited to four distinct click types and three accompaniments. Two of the three accompaniments are attested across most click languages, which provides little evidence of exclusivity or typological distinctiveness (Ladefoged & Traill 1984; Ladefoged et al. 1999). However, the third accompaniment discovered in the inventory of Tumʔi is the audible uvular stop accompaniment which is rarely reported in the documentation of click languages (Miller et al. 2007). As discussed above, this accompaniment is particularly unusual in that the resulting click is composed of two noise bursts. Reports of this click accompaniment are limited to a select number of languages classified as belonging to the Tuu family, hence the presence of this accompaniment in Tumʔi is a typologically remarkable find.

Table 2: *The Tumʔi click inventory*

Click Types → Accompaniments ↓	dental	lateral	palatal	alveolar
plain	ǀ	ǁ	ǂ	!
uvular stop	ǀq	ǁq	ǂq	!q
nasal	ŋǀ	ŋǁ	ŋǂ	ŋ!
velar fricative	ǀx			!x

3.3.1 Notes on Clicks

The plain clicks presented above are the most under-utilized of the click segments transcribed across this data set. Generally, the speakers produce clicks with some form of accompaniment, this is most commonly the pulmonic uvular accompaniment which constitutes a second distinct noise burst. The same frequent distribution of this lingual pulmonic segment has been reported for the languages !Xóõ and |Gui, in which the uvular position appears to be exploited more often than predicted (Güldemann 2001). This unexpected distribution may reflect the speaker's intention to retain the most unique features of the language by extending the use of the infrequently attested uvular accompaniment. As discussed in detail later on in chapter 5, section 5.5.2, the nasal and uvular stop accompaniments are the most frequently occurring posterior release types. However, as discussed in section 5.2, and in more detail in section 5.5.2, the audible uvular accompaniment is what sets apart the click inventory of Tumʔi as typologically complex. The velar fricative accompaniment occurs infrequently throughout the data set; however, this accompaniment is clearly audible and phonetically distinct in particular lexical items such as ǀxei 'give birth' and !xara 'female genitals-type' (see Appendix A, Indexes 62&63).

3.4 Discussion

Beginning with the vowel inventory presented in section 3.1 above, Tumʔi appears to consist of six monophthong vowels and a total of five diphthongs. The monophthong inventory is comprised of the general equidistant high and low vowels commonly attested cross-linguistically. The mid-vowels indicate an uneven distribution in terms of openness, where the front mid-vowel resembles a half close vowel phoneme, and the back mid-vowel constitutes an open more centralized vowel phoneme. According to phonological descriptions of Khoisan dialects, the open-close contrast is generally not attested in languages of the Khoekhoe family, but has been observed in languages of the Tuu family, as well as multiple Bantu languages. The substantial inventory of diphthongs recorded in Tumʔi also constitute a feature remnant of Tuu language typologies. Phonological descriptions of Khoekhoe languages report the use of diphthongs only as a consequence of rapid speech (Beach 1938). Though not included as part of the official vowel inventories presented in figures 3 & 4, chapter 4 discusses the possibility of a phonation contrast in Tumʔi. The analyses presented in chapter 4 do not provide a definitive conclusion concerning phonation as a distinctive feature. However, the results do point toward a significant difference between segment groups exhibiting modal and non-modal phonation.

As discussed in section 3.2 above, the consonant inventory of Tumʔi comprises a substantial variety of pulmonic egressive consonants. As reported across the previous literature and comparative studies, the non-click consonant systems of Khoekhoe languages are limited in terms of voicing and other laryngeal contrasts (Güldemann 2006). Therefore, the phonemic voicing contrast observed between bilabial stops in Tumʔi is an exceptional find, as it distinguishes the phonological typology from that of most Khoekhoe varieties. Among the extended phonemic contrasts is the additional aspiration observed across utterances of the dental plosive [t]. As discussed extensively in chapter 6, section 6.3.2, the segment [t^h] is produced consistently across the tokens elicited for the lexical item [t^hi!qo] ‘God’. Based on this observation, the aspirated segment [t^h] is presented as a phoneme of the consonant inventory in section 3.2 above. As discussed in more detail in chapter 6, section 6.1, contrasts in aspiration are reportedly attested in languages of the Tuu family and not Khoekhoe, with the exception of !Ora (Güldemann 2006; Vossen 2013). Another unique phonetic feature observed in the consonant system of Tumʔi is a complex sequence comprised of a nasal followed by a diffuse bilabial ejective, this segment is discussed in more detail in chapter 6, section 6.3.3. The phoneme [t’] constitutes the only ejective segment in Tumʔi and is rarely attested in Khoekhoe languages. These unique phonological features described in the consonant system of Tumʔi are therefore essential to the comparative analysis of these particular language families.

Finally, analyses of the click inventory reveal evidence of four click types as discussed in detail in chapter 4, section 4.4.1, and a total of three click accompaniments. Two of these accompaniments perform a larger functional role and are discussed in more detail in chapter 4, section 4.4.2. The nasal and uvular stop accompaniments enjoy the widest distribution across the data set, and hence form the bulk of the phonetic analyses of click accompaniments in this paper. While nasal accompaniments constitute one of the most commonly attested segments in click languages cross-linguistically, the uvular accompaniment or lingual pulmonic segment which enjoys the same extensive distribution in Tumʔi is seldomly observed cross-linguistically. The phonetically evidenced report of this accompaniment in the click system of Tumʔi does therefore constitute a unique typological phenomenon. The rest of the clicks and final accompaniment presented in the inventory in table 2 above includes the less frequently observed velar fricative accompaniment and plain clicks. Oddly, plain clicks in Tumʔi share the same narrow distribution as the velar fricative accompaniments. This distribution may be an indication of the speakers' retention of more unique or prominent linguistic features or segments in the language, however it may also reflect the inherent limitations of 'salvage linguistic' research (Grinevald 2003).

Chapter 4: Phonation variation

4.1 Theory

As a contrastive feature, phonation exists both cross-linguistically and particularly prominently across the more typologically complex Khoisan varieties. Languages belonging to the Tuu family have been reported to exhibit contrastive phonation, including Southern Khoisan varieties such as !Xóõ (Garellek 2019) and Ju|'hoansi (Miller 2007) among others. The recorded effects of phonation are realised through adjustment of the glottal aperture formed most commonly by the configuration of the arytenoid cartilages (Garellek 2019). Differences in phonation are therefore attributed to the degree of closure between the vocal folds, and ultimately the degree of voicing activated in the articulation of a segment. As there exists a spectrum dichotomizing voiced and voiceless segments, there are consequently numerous contrastive phonation types reported cross-linguistically. While particular language systems may exhibit as many as eight distinctly measurable voice qualities, for the purposes of this project only three basic types will be considered. Ranging from the most open to the most closed glottal state, the relevant phonation types include: breathy, modal, and creaky voice (Gordon & Ladefoged 2001). The majority of the previous phonation analyses have reported the resultant difference in voice quality as contrastive, serving as a distinguishing feature between particular vowel and nasal segments. However, there are instances in which such differences in phonation occur as idiosyncratic features exhibited only by particular speakers (Huffman 1987).

The basic phonation contrasts distinguishing vocalic segments have generally been identified by measuring differences in the amplitude of the first (H1) and second harmonics (H2). Physiological characteristics, such as the medial thickness of the vocal folds and their aperture at the time of vibration can be interpreted by the H1-H2 spectral slope parameter (Garellek 2019, pg.

3; Huffman 1987, pg. 502). The acoustic analysis of phonation often also includes measuring the harmonics-to-noise ratio (HNR). The HNR measure is reported to indicate the presence of any aspiration noise due to the glottis opening, as well as the distribution of irregular voicing (Garellek 2019, pg. 3). Investigations of phonation types related to the aperture of the vocal folds have been successful in distinguishing between breathy, modal and creaky voice segments through the analysis of spectral tilt and noise measures (Garellek 2019; Gordon & Ladefoged 2001; Huffman 1987; Miller 2007). Although acoustic evidence for different phonation types in earlier investigations have relied on contrasts in the distribution of glottal pulses visible in the waveform, relevant aspects of jitter and shimmer will serve only as a secondary analysis in this investigation. Due to the phonetic variability between speakers and across utterances within the already limited data set, phonation distinctions beyond the engagement of the vocal folds will not be considered. Hence, though various languages exhibit contrasts of phonation types engaging pharyngeal narrowing, the aryepiglottic folds and the ventricular folds, an in-depth investigation of the corresponding acoustic characteristics of these articulatory gestures in Tumʔi would be impractical at this time.

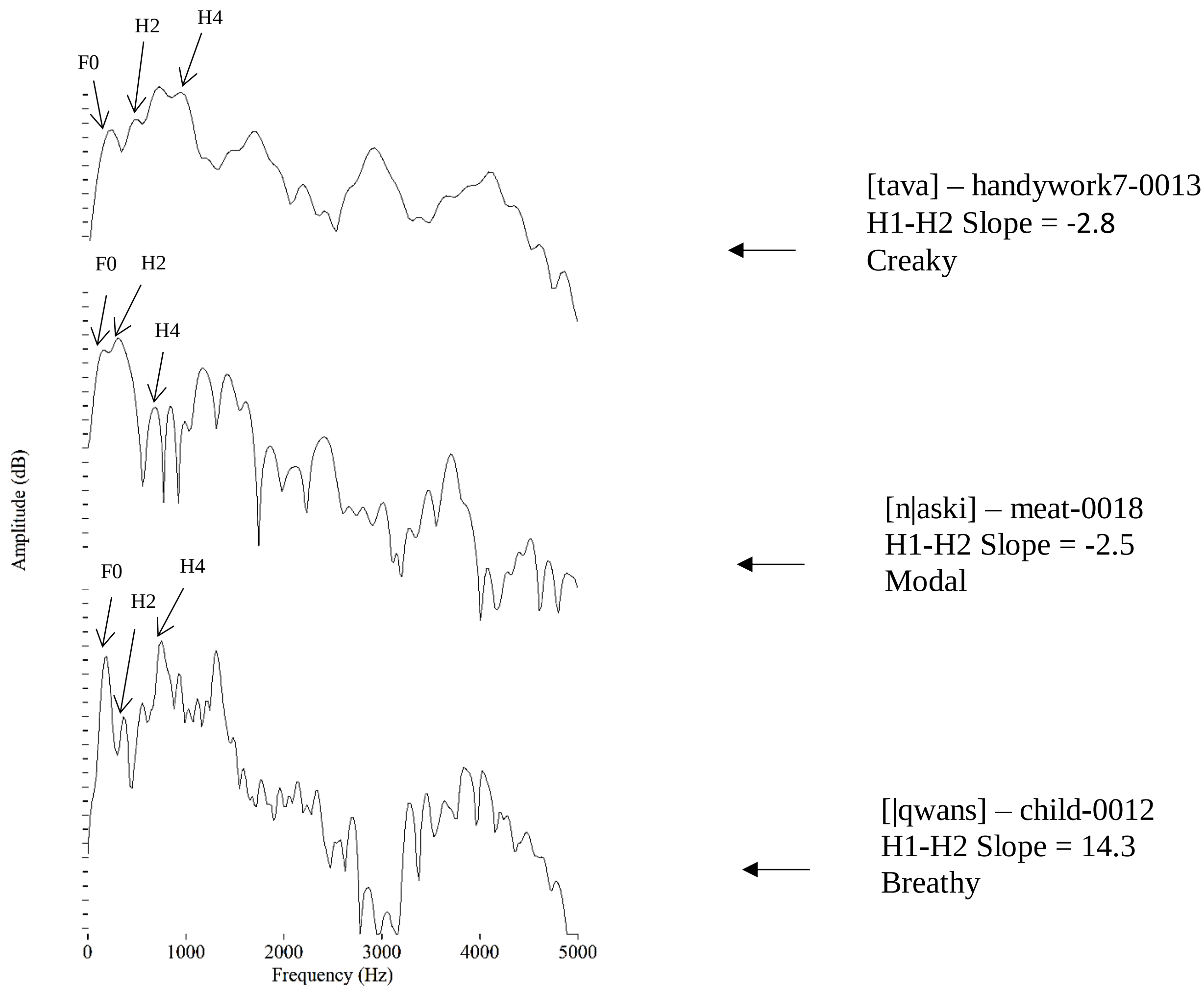


Figure 6: spectra calculated across four glottal pulses for three low vowel segments. tava ‘handywork’; n|aski ‘meat’; |qwans ‘child’.

Individual spectral samples of tokens produced by Elsie and Francina displayed in figure 6 above provide evidence of three possible phonation types based on the H1-H2 slope, as well as the difference in spectral tilt across each sample. There is a clear discrepancy between the H1-H2 slopes across the tokens representing each phonation type. The steepest slope is visible in the breathy spectrum with a value of 14 dB. Both the modal and creaky tokens display a negative H1-H2 spectral tilt. The spectral tilt of the creaky segment is only slightly steeper than that of the modal segment with a value of -2.8. A negative spectral tilt is expected for segments articulated with creaky phonation as well as segments produced with modal phonation (Gordon & Ladefoged 2001). The most distinguishing factor is the spectral slope and the degree at which the intensity drops as the frequency of the harmonics increases. The token representing breathy phonation has a steeper negative slope than either of the other tokens. The modal token exhibits a less steep slope, while the creaky token exhibits only a slightly negative slope. Repeated instances of each spectral tilt are observed across the measured tokens which include all lexical items with the low vowel [a].

Spectral tilt most explicitly reflects the distribution of energy at different frequencies. Breathy vowels are characterized by a high degree of energy in the fundamental frequency, which corresponds to the first harmonic, or H1, while the higher harmonics e.g. (H2 & H4) are characterized by less energy, as the glottal waveform is more sinusoidal due to smoother opening and closing phases. Alternatively, the sharp glottal closure and opening formed in the production of creaky voice corresponds to more energy in the higher frequency harmonics, with relatively less energy in the fundamental frequency, or first harmonic. The expectation is therefore that the H1-H2 spectral tilt should reflect a significantly steeper gradient for breathy vowels in comparison to creaky and modal vowels. The reverse effect is expected for creaky vowels, in which case a greater negative H1-H2 spectral tilt would best distinguish the creaky segment from the modal voice segment. Statistical results from previous phonation analyses provide evidence that the mean HNR for modal voice is expected to be significantly higher than other phonation types (Gordon & Ladefoged 2001; Miller 2007). However, in comparison to creaky voice, breathy segments are expected to have a higher HNR mean due to their limited glottal constriction and resultant noisy energy.

4.2 Method

Table 3: *Tokens included in phonation*

Creaky/Breathy	No. Tokens	Gloss	Modal	No. Tokens	Gloss
ǀqxwa̤ns	4	child	!xamaku	6	thank you
ǀqxwa̤	7	baby	nǀaski	6	meat
ha̤a̤	8	eat	para	4	donkey
tʰa̤va	6	handywork	!ui asa	6	here comes the white man
nǀa̤	6	head			

The contrastive and non-contrastive distinctions between phonation types will be captured primarily using an acoustic analysis of the H1-H2 spectral slope parameter, and information extracted from the contrasted HNR measures, as conducted across previous phonation analyses (Garellek 2019; Gordon & Ladefoged 2001; Miller 2007). Though previous analyses provide evidence of different phonation types exhibiting dynamic changes throughout the vowel duration (Traill 1985), the acoustic measures outlined above are extracted exclusively at the midpoint of each vowel. The tokens analysed for a phonation contrast include a selection of words consistently produced by the eldest speaker with some form of audible non-modal phonation. This phonation set is comprised of 31 tokens which are contrasted with 22 tokens recognised as modal, containing segments exhibiting no additional laryngeal features. All tokens are extracted from utterances produced by both female speakers and are constrained by the amount of elicitation conducted with each speaker as discussed in chapter 3, section 3.1. Presented in table 3 above are all contrasted tokens of the vowel [a] from which phonation measures have been extracted. Though phonation contrasts are observed in tokens of other vowels and diphthongs, these segments are not quantitatively feasible to contrast. Furthermore, it is noted in the analysis of non-modal phonation in the Khoisan variety Ju|'hoansi, that by using only low vowel data such as [a], the H1 and H2 values are more likely to be below the bandwidth of F1, which helps avoid boosting either harmonic due to proximity to F1 (Miller 2007, pg. 134).

Using the computer software Praat, the following measures were extracted from the sample words listed above: the first (H1) and second harmonic (H2), as well as the mean harmonics to noise ratio (HNR). Harmonics measurements were taken from sample spectra which were extracted at the mid-point of each vowel to obtain the most accurate and consistent reading. Each spectrum was derived from a selection of four glottal pulses, beginning and ending at the zero crossings before vocal fold contact. The H1-H2 slope values were calculated in Excel and each segment was coded to indicate the phonation type as either modal or non-modal (creaky/breathy). The choice to compare these measures is based mainly on the results of a linear discriminant analysis (LDA) used to measure phonation in !Xóõ, in which it is stated that distinctions in phonation are most easily perceivable through analyses of the measures discussed above (Garellek 2019, p. 17). Further motivation for the analysis of these particular measures is the phonation analysis conducted by Gordon and Ladefoged (2001) in which spectral tilt is described as a robust and reliable parameter for differentiating phonation. The calculated measures for the modal and non-modal phonation types were then imported to R statistical software, and tested for contrastive significance (Johnson 2011, p. 237-240). The aim of this analysis was to determine whether there is a measurably significant acoustic difference between the modal and non-modal segment groups based on the measures discussed above.

Due to limitations of the sample size, and constraints in terms of categorising the data beyond modal and non-modal groups, the following analyses are comprised of two-sample t-tests, and data visualisation in the form of boxplots. First, Fisher's F-test was used to determine the homoscedasticity of the variables. This tests the homogeneity of the variance, the resultant p-value ($p > 0.05$) indicating that the variances of both samples were not significantly different from each other. A classic two-sample t-test is then run, contrasting the H1-H2 slope and HNR values across the two phonation groups as discussed above. This t-test was used to determine whether the means of the relevant measures captured for the two phonation groups are significantly different. The box plots then provide a visual representation of the disparity between the slope gradients and HNR of the different phonation groups. Though the non-modal phonation group is not specified as representative of a particular phonation type (creaky or breathy), the aim at this point is to determine if there is a modal vs. non-modal voice quality distinction in Tumʔi.

4.3 Results & Discussion

Based on the collected spectral data presented in table 4 below, few of the segments exhibiting non-modal phonation appear to display a negative H1-H2 spectral tilt. The majority of the negative harmonic slopes are observed across the modal segments extracted from tokens of the gloss [ǂxama:ku] 'thank you', see table 4 indexes 1-6. The greatest negative slope -6.2 dB is calculated for a token of the word [ǀqxwa̰] 'baby', which constitutes a non-modal segment, see table 5 index 9. The rest of the tokens for the word [ǀqxwa̰] 'baby' include both substantial and slightly positive H1-H2 slopes within the range of 2-21 dB. This free variation between tokens coded with the same gloss indicate that any significant contrast in phonation is not lexically determined. Though the initial token selection had been decided by the articulation of particular lexical items, the gloss as an independent factor does not appear to play a significant explanatory role in determining the spectral tilt and HNR values of a segment. Therefore, any significant disparity in phonation evidenced by the statistical results presented below are interpreted as instances of free variation. As discussed above, the data is also coded by speaker, however the explanatory value of this factor is not established by the following significance tests.

Table 4: *H1-H2 & HNR spectral measures*: *Modal sample*

Index	H1-H2 Slope	HNR	Gloss	PT	Speaker
1	-0.9	19.047	thankyou5-0021	M	Francina
2	-0.8	18.512	thankyou5.2-0021	M	Francina
3	0.2	8.859	thankyou4-0021	M	Francina
4	-2	18.743	thankyou4.2-0021	M	Francina
5	-5.8	13.516	thankyou3-0021	M	Francina
6	-6.3	20.931	thankyou3.2-0021	M	Francina
7	-0.1	9.901	meat2-0018	M	Francina
8	-2.5	7.694	meat-0018	M	Francina
9	4.9	5.834	meat-0015	M	Francina
10	-3.3	15.374	donkey3-0013	M	Francina
11	17.2	14.322	meat3-0013	M	Francina
12	19.8	8.757	meat4-0013	M	Francina
13	7.4	13.268	herecomesthewhiteman7-0013	M	Francina
14	11.6	9.375	herecomesthewhiteman4-0013	M	Francina
15	9.2	7.642	donkey-0013	M	Elsie
16	14.8	8.881	meat6-0013	M	Elsie
17	17.5	9.659	donkey2-0013	M	Elsie
18	2.9	25.532	donkey5-0013	M	Elsie
19	-1.2	16.681	herecomesthewhiteman-0013	M	Elsie
20	-8.5	22.418	herecomesthewhiteman2-0013	M	Elsie
21	6.7	10.268	herecomesthewhiteman3-0013	M	Elsie
22	1.9	18.69	herecomesthewhiteman5-0013	M	Elsie

Table 5: *H1-H2 & HNR spectral measures*: *Non-modal sample*

Index	H1-H2 Slope	HNR	Gloss	PT	Speaker

1	14.3	9.167	child-0012	C/B	Elsie
2	5.7	6.16	baby2-0013	C/B	Elsie
3	19.8	14.322	eat-0013	C/B	Elsie
4	11.6	14.707	eat4-0013	C/B	Elsie
5	5.5	4.925	handywork1-0013	C/B	Elsie
6	0.8	8.487	handywork2-0013	C/B	Elsie
7	0.8	12.391	handywork6-0013	C/B	Elsie
8	9.8	9.319	handywork7-0013	C/B	Elsie
9	-6.2	5.565	baby-0016	C/B	Elsie
10	21	4.042	baby-181207	C/B	Elsie
11	15	4.821	baby13-181207	C/B	Elsie
12	2.4	1.141	baby7-181207	C/B	Elsie
13	-2	0.852	child2-18207	C/B	Elsie
14	19.8	12.795	eat-18207	C/B	Elsie
15	8.5	13.644	eat5-18207	C/B	Elsie
16	16.3	10.081	eat7-18207	C/B	Elsie
17	16.8	14.71	eat8-18207	C/B	Elsie
18	23.1	7.44	head-0013	C/B	Elsie
19	6.7	12.133	head2-0013	C/B	Elsie
20	8	5.883	head3-0013	C/B	Elsie
21	9.9	9.922	head3-181207	C/B	Elsie
22	10.4	2.358	head4-181207	C/B	Elsie
23	1.5	7.954	head5-181207	C/B	Elsie
24	6.2	16.692	child3-0013	C/B	Francina
25	8.8	14.098	child2-0013	C/B	Francina
26	8.4	14.765	eat2-0013	C/B	Francina
27	17	8.544	eat5-0013	C/B	Francina
28	5.9	11.299	handywork2-0013	C/B	Francina
29	-2.8	11.179	handywork7-0013	C/B	Francina
30	15.3	12.355	baby10-181207	C/B	Francina
31	7.5	12.52	baby3-181207	C/B	Francina

The data presented in table 4 & 5 above was analysed in R (R Core Team, 2012). A two-sample t-test was used to determine whether the spectral tilt and HNR values recorded for the modal and non-modal groups are significantly different. Significance is tested separately for the spectral tilt and HNR values. The first t-test compares the means of the H1-H2 slope measures across the two independent samples, the modal and non-modal phonation groups. The second t-test compares the means of the recorded HNR values across the two independent samples. The slope measures and HNR values therefore serve as the dependent variables representing the two independent populations considered in this analysis. The results of the first t-test comparing the mean H1-H2 slope measures calculated for the two phonation groups is presented in table 5 below.

Table 6: *Results of H1-H2 slope t-test*

t = 2.525, df = 51, p-value = 0.01472
sample estimates:
mean of non-modal = 9.219355
mean of modal = 3.759091

The statistical result presented above yields a significant difference between the H1-H2 slope measure extracted for the two phonation groups. The p-value is less than 0.05 and therefore indicates that the segments included in each population have an inherent difference in spectral tilt. As indicated by the mean values presented above, the non-modal segments are characterised by a steeper slope gradient than the modal segments. This finding aligns with most phonation research which have reported a high fundamental frequency for breathy segments, which results in a steep positive slope between the first and second harmonics. Hence, the result of the initial test would indicate that the undefined non-modal group includes more breathy than creaky phonation.

Table 7: *Results of HNR t-test*

t = -3.1806, df = 51, p-value = 0.002501
sample estimates:
mean of non-modal = 9.492613
mean of modal = 13.813818

The result of the t-test contrasting the HNR values presented in table 7 above also displays a significant difference between the two phonation groups. The p-value is calculated as 0.002 which indicates a high degree of significance. Based on the mean values presented above, the HNR measured for the non-modal group is significantly lower in comparison to that measured for the modal group. The lower HNR observed for the non-modal group provides further evidence that these segments likely form part of the guttural class which includes breathy, glottalized and epiglottalized vowels (Miller 2007). The results of the two-sample t-tests presented above provide evidence of a significant disparity between both the H1-H2 slope measures, as well as the HNR measures of the modal and non-modal phonation groups. Overall, the results of the two-sample t-tests indicate a greater H1-H2 gradient for the non-modal segments, but a smaller HNR measure. Translated, these results indicate that non-modal phonation is attested in Tumʔi, and the non-modal segments most likely consists of breathy segments rather than creaky.

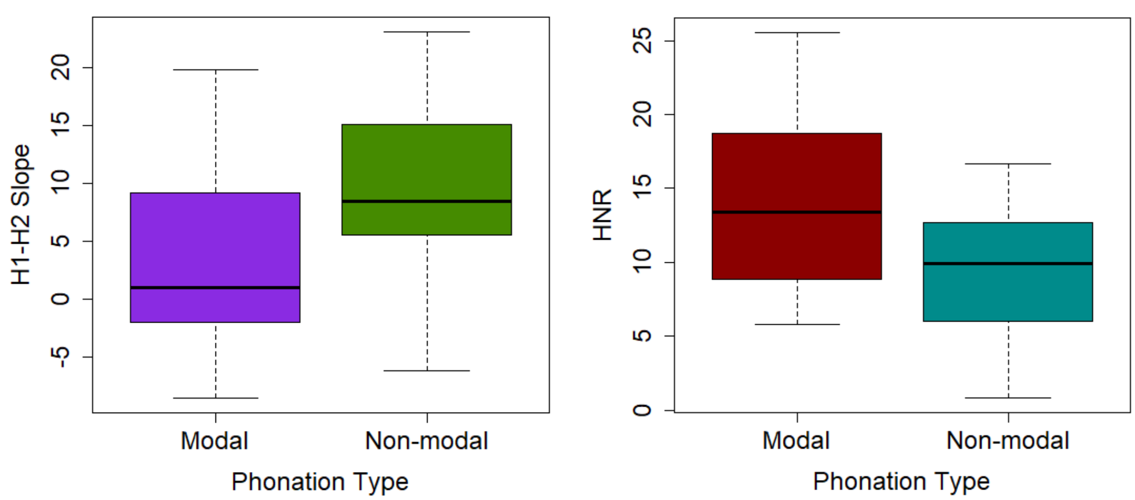


Figure 7: Boxplot diagrams contrasting the means of the H1-H2 slope, and the HNR for the modal and non-modal phonation groups.

The boxplots above provide a visual representation of the discrepancy between the overall means of the H1-H2 slope measures, and the HNR measures of the two phonation groups. The significance of this contrast is interesting because the tokens representative of non-modal phonation have not been distinguished as either breathy or creaky voice vowels, which are predicted to exhibit distinct spectral contrasts. However, despite the indistinct description of the non-modal phonation type, the results of the statistical analyses above show that a form of non-modal phonation is used by speakers of Tumʔi, and that this phonation type is acoustically defined by a steep spectral slope and substantial degree of noise.

We may further speculate on the implications of these results with reference to the results of the linear mixed effects modal performed on !Xóõ (Garellek 2019). The average H1-H2 slope for the non-modal vowels, 9.2 dB, aligns closely with the value reported for pharyngealized breathy vowels found in !Xóõ (Garellek 2019, pg. 24). The average H1-H2 slope value for the modal tokens is then observed to be similar to the value reported for pharyngealized modal segments (Garellek 2019, pg. 24). The results of this analysis indicate an average HNR value calculated for the non-modal tokens that is significantly less than that calculated for the modal tokens. If we are to consider both the modal and non-modal tokens to be pharyngealized segments then we may be able to account for the higher degree of broadband noise measured across the modal tokens. Based on the analysis of !Xóõ (Garellek 2019, pg. 25), the HNR values of the modal pharyngealized segments are considerably higher than that measured for all the other phonation types with the exception of the creaky voice segments. The phonetic characteristics of pharyngealization have been reported to include laryngeal constriction and an epiglottal trill generally produced at the aryepiglottic folds, which would contribute a degree of noise to the voice quality (Garellek 2019, pg. 11).

Furthermore, it must be reported that the eldest speaker who consistently produced particular words with non-modal phonation herself explicates the distinct form of articulation as

"Die verskil kom van agter af, jy moet trill as jy dit se" (the difference comes from the back of the throat, your articulators must trill when you say it). The speaker explains that the sound must be produced at the back of the throat with a consistent trill like mechanism. Not only is there acoustic evidence of a non-modal phonation type, but this special phonation is recognised by the speakers. What is clear based on the data above is that the type of non-modal phonation produced by speakers of Tumʔi may be more phonetically complex than what can be addressed in the scope of this paper.

4.4 Discussion

The analyses above are most certainly constrained by the modicum of data available for the investigation of phonation contrasts. However, the statistical analyses of the spectral data presented in tables 4 & 5 provide evidence of a significant disparity in the phonation of low vowel segments produced in Tumʔi. The results of the t-tests show that the Harmonics and the HNR measures are successful in distinguishing between the modal and non-modal phonation groups. The mean values of these measures do not entirely align with the general expectations reported in previous phonation analyses, however the results are applicable when compared to the measures reported for particular phonation types identified in !Xóõ. Comparison of the slope and HNR measures above with those reported for the pharyngealized breathy and modal segments in !Xóõ indicate that the non-modal segments observed in Tumʔi may be similar to the breathy pharyngeal phonation type observed across multiple Southern Khoisan languages (Vossen 2013). Overall, the results of this analysis definitively tell us that speakers are producing some form of non-modal phonation. Furthermore, the representative spectral analysis in figure 5 above displays a segment reflecting features of a creaky vowel in the form of a negative slope between the fundamental and the second harmonic. Though not apparent in the statistical analyses, this variation cannot be ignored. Hence, the ultimate conclusion is one of idiosyncratic acquisition in the case of the two speakers (Cook 1989). Though both speakers produce non-modal phonation types, there is clear variation in the degree of phonation produced across different tokens as well as different speakers.

Chapter 5: Clicks

5.1 Introduction:

The purpose of this chapter is to determine which phonetic contrasts between clicks recorded for Tumʔi are representative of the phonological structure of the inventory. The scope of this chapter includes the phonetic investigation of all attested plain clicks, as well as the various click accompaniments produced. The analyses of the click accompaniments are centred around the most frequently occurring nasal and uvular accompaniments, with only single spectrograms provided for the description of the less frequent voiced and velar fricative accompaniments. Displayed in table 8 below is the complete inventory of click types and accompaniments attested in Tumʔi, as well as the frequency of each click and accompaniment within the collected data.

Table 8: *The Tumʔi click inventory*

Click Types → Accompaniments ↓	dental	lateral	palatal	alveolar
plain	ǀ (46)	ǁ (4)	ǂ (12)	! (36)
uvular stop	ǀq (19)	ǁq (1)	ǂq (2)	!q (10)
nasal	ŋǀ (16)	ŋǁ (1)	ŋǂ (4)	ŋ! (4)
velar fricative	ǀx (2)			!x (2)

5.2 Literature:

To the English speaker clicks are limited to paralinguistic use, e.g. as a way of indicating disapproval or encouraging horses. To the unfamiliar ear with no phoneme mapping of click sounds, clicks are generally perceived as pops and sizzles introduced in the background. However, in Southern Africa clicks are not uncommon, and the speakers of click languages perceive them in the same manner as all other speech signals. Southern African languages are strikingly marked by the presence of clicks, but as a result of their origin and distribution these sounds constitute an areal phenomenon (Childs 2003, pg. 1-3). Currently clicks are dispersed throughout the sound inventories of numerous African languages including; Khoisan, many Bantu languages, as well as Cushitic languages. Across these languages clicks form only a portion of the consonant inventory, though this may constitute the main portion. While Bantu languages such as Zulu and Xhosa have only incorporated three clicks into their inventories, there are Bantu languages found in Botswana and Namibia which have complete four click systems. However, the click systems in these languages are all derived from the Khoisan language family and hence generally occur at a much lower frequency (Westphal 1971).

Clicks are most common in the consonant systems of Khoisan languages. !Xóõ, !Xũ, and Nama are languages in which they occur frequently. Clicks constitute over 70% of the initial segments in the words recorded for the !Xóõ dictionary (Ladefoged & Traill 1994, pg. 34). Articulatorily, clicks are defined as velaric segments produced with two closures at the posterior and anterior ends of the tongue. These closures cause a suction mechanism; when the anterior closure is released, the rush of air results in a pop, a transient response more intense than that of the general stop consonant (Ladefoged & Traill 1994, pg. 41). The anterior closure is vital as it determines the place of articulation and overall click type, the posterior closure then determines the nature of the following accompaniment. Acoustically, the release of the initial closure is signaled by a transient response, along with some degree of noise depending on the click type (Childs 2003, pg. 1-3). The anterior closure which defines the click type is also distinguishable by the acoustic features of the click release (Ladefoged & Traill 1994; Miller et al. 2009; Sands et al. 2007). Generally, the release of the second closure is not acoustically identifiable, however in certain cases of velar or uvular accompaniment it may be reflected in the acoustics by a second transient response (Miller et al. 2009). This acoustically apparent accompaniment is not ubiquitous across Khoisan languages and has only been verified as an attested phonological contrast in the languages !Xóõ, Nǀuu, ǂHoan, and |Auni (all members of the Tuu family), and in one Khoekhoe language |Gui. Along with evidence of contrast between acoustically apparent and silent posterior closures in Nǀuu, ultrasound analysis of the posterior burst release also indicates uvular articulation (Miller et al 2009, pg. 140-145). Other accompaniments are quite common throughout the Khoisan language family, including nasalization and added pharyngeal features, such as aspiration (Beach 1938).

In terms of sound inventory the Tuu language family holds a reputation for constituting exceptionally complex sound systems, including languages contrasting voiced and voiceless oral and nasal clicks, see table 9 below. According to evidenced reports on Nǀuu, the degree of voicing contrast exhibited in clicks is comparable to voicing contrasts between pulmonic stops (Miller et al. 2009). The nasal accompaniment is generally audible and acoustically represented by visible low frequency energy extending before and throughout the click burst release (Ladefoged & Traill 1994, pg. 47). A voiced click would resemble the acoustic properties of nasal accompaniment, including low frequency noise preceding the burst release, while aspiration is indicated by longer voice onset time (VOT). Nasal accompaniment is reported to be considerably common across Khoisan languages including but not limited to those found in Southern Africa (Beach 1938;

Ladefoged & Traill 1994, pg. 46). Acoustically identifiable velar and uvular accompaniments are however not as ubiquitous and are associated mainly with dialects of the Tuu family (Miller et al. 2007; Vossen 2013). In light of the cognate comparisons discussed in section 2.2 and presented in Appendix A, pg. 99, one might consider evidence of such a unique feature to be an indication of genetic affiliation rather than a simple instance of typological coincidence.

Table 9: *N|uu click inventory (Miller et al. 2009 pg. 132)*

LINGUAL					
	Labial-uvular	Dental-uvular	Alveolar-uvular		Palatal-uvular
			Central	Lateral	
Stop	ʘ	ǀ ǀʰ ᶢǀ	ǃ ǃʰ ᶢǃ	ǁ ǁʰ ᶢǁ	ǂ ǂʰ ᶢǂ
Nasal	ŋ̊ʘˀ ᵑʘ	ŋ̊ǀʰ ŋ̊ǀˀ ᵑǀ	ŋ̊ǃʰ ŋ̊ǃˀ ᵑǃ	ŋ̊ǁʰ ŋ̊ǁˀ ᵑǁ	ŋ̊ǂʰ ŋ̊ǂˀ ᵑǂ
LINGUO-PULMONIC					
Stop	ʘ͡q	ǀ͡q ǀ͡qʰ	ǃ͡q ǃ͡qʰ	ǁ͡q ǁ͡qʰ	ǂ͡q ǂ͡qʰ
Affricate[a]	ʘ͡χ	ǀ͡χ	ǃ͡χ	ǁ͡χ	ǂ͡χ
LINGUO-GLOTTALIC					
Affricate[b]		ǀ͡χ’	ǃ͡χ’	ǁ͡χ’	ǂ͡χ’

According to Ladefoged & Traill 1994, any click will be classified as belonging to one of five possible categories, including bilabial, dental, lateral, palatal, and alveolar, with retroflex clicks reduced to a type of alveolar articulation. These types are easily distinguishable by their acoustic properties. Generally, these clicks can be distinguished by the intensity of the transient response, and any extended noise component. Hence, an analysis of the waveform is considered the most convenient practice for describing clicks (Ladefoged & Traill 1994). The transient response refers to the immediate acoustic effect of the separating articulators; it constitutes an impulse response wave that is shaped by the size and configuration of the vocal tract cavity. The measurable turbulent noise then reflects the aerodynamic properties of the click release. During articulation of the noisy clicks, the primary articulators such as the tongue or lips are released more slowly. The divide between click types is determined by which of these features dominate the anterior release. As reported across multiple analyses, the bilabial, dental and lateral clicks are defined by their noisiness which is generally determined by the rise time and duration of the burst. The waveform of the abrupt clicks are dominated by the transient response and a shorter burst duration (Beach 1938; Ladefoged & Traill 1994, pg. 40-41; Miller et al. 2009; Sands et al. 2007).

The same click descriptions have also been evidenced using spectral data. Analyses of clicks produced in !Xóõ exhibit evidence of a broad divide in the spectral distribution of particular click types. One division is marked at 2500 Hz, which is equivalent to 14 Bark. The expectation is that the dental and palatal clicks display more energy above this frequency, while the lateral and alveolar clicks display more energy below it. Dental and palatal clicks are generally expected to display a peak at about 4000 Hz, equivalent to 18 Bark. However, reports have mentioned a significant peak in the spectra of dental clicks between 900-1500 Hz, similar to the spectra of bilabials. The spectrum of the alveolar click is expected to exhibit two low frequency peaks at 800 Hz and 1200 Hz. Hence, a relatively steep negative spectral tilt should be expected for the alveolar click (Ladefoged & Traill 1994, pg. 42).

5.3 Method

Due to difficulties and constraints on the data collection process as discussed in sections 3.3 and 3.4 above, the description of all contrasts are confined to the analysis of representative tokens. The extracted data include waveforms and spectra which best reflect the discrepancies between the click segments. All extractions include a selection of the waveform beginning at the onset of the burst release. The extracted waveforms are generally limited to a 20ms window. Any extended or limited time frames are due to the inherent differences between burst durations and are later specified where relevant. All spectral samples constitute 10ms spectral slices extracted from the onset of the burst release. The extracted 10ms sound samples are analyzed with Linear Predictive Coding Spectra (LPC (burg)). The prediction order is set to 28, and the time window set to 10ms to include all formant activity and the entire waveform. Due to considerable differences in burst length between click types, the spectral samples are generally restricted to a shorter time window. While various approaches to the spectral analysis of clicks have been applied in the investigation of Khoisan languages, each reports the same general conclusions. Different analyses include spectral frequencies displayed on a Bark scale or LPC spectral analysis, while others even include functions which transform the frequencies and amplitudes to reflect the human ear's perception of disparities in loudness (Ladefoged & Traill 1994, pg. 42). Analysis like the latter auditory spectra may provide a less ambiguous representation of the distinguishing spectral characteristics, however a comprehensive overview of these varying analyses point to the same spectral characteristics needed to distinguish between the different click types.

5.4 Data

A tally of the items in the data set indicates that 63 of the total 125 words begin with a click segment. Therefore, about 50% of the collected data begin with a click segment as opposed to a pulmonic egressive consonant or vowel segment. 19 of the total 125 words contain word medial and intervocalic clicks; 9 of these constitute words which also contain a click segment in the word initial position. Including the additional 10 items with word medial segments, clicks occur across a total of 73 lexical items, and are therefore distributed across 58% of the entire data set. The segments counted as clicks include all those which exhibit a click-like burst and can be acoustically classified as belonging to one of the click types discussed above. After phonetic transcription and thorough acoustic investigation of the click segments recorded across all the individual tokens, it is clear that many of these segments are not produced consistently and may exhibit phonetic variation across different tokens. Phonetic analyses of multiple tokens representing the same lexical items indicate variation in the click type produced. Though this data is passable for proving the use of multiple click types within this dialect, it is not sufficient for defining any predictable phonological environments. Hence, the collected data cannot be analyzed to determine any definite relation between a particular click type and its phonological environment. This is exhibited in the variation between segments extracted from tokens of the same word as presented in figure 8 below. However, along with this unexplained variation, there are particular tokens which exhibit some consistency between corresponding click segments.

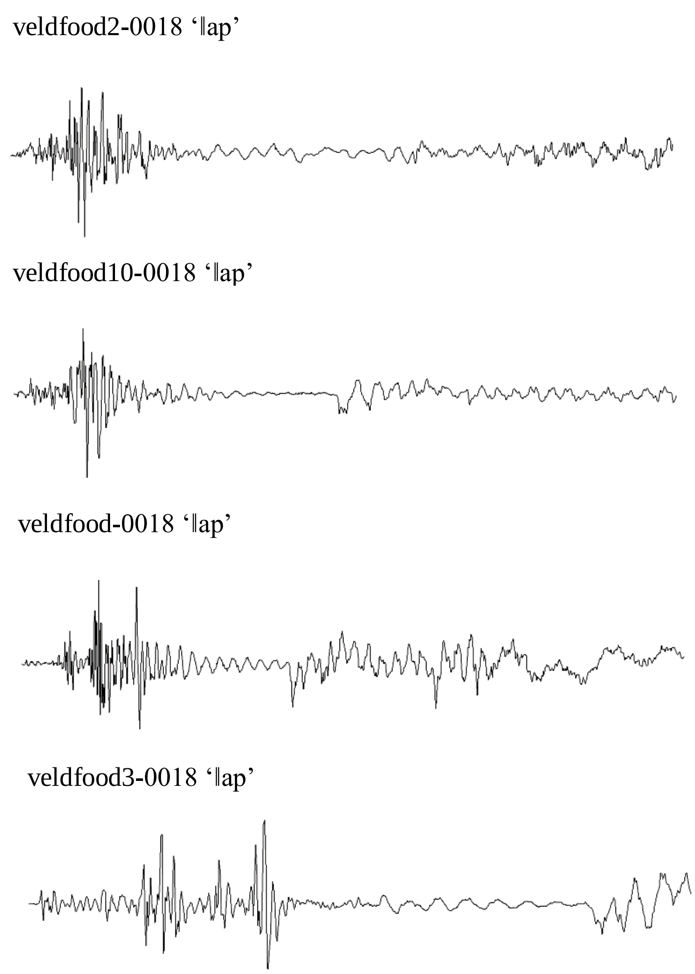
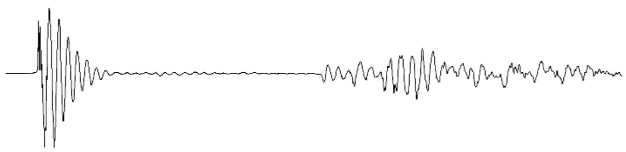


Figure 8: The click burst for each token. ǁap ‘veld food’

Figure 8 above consists of four corresponding segments extracted from four different tokens of the same lexical item. The segments extracted from ‘veldfood2-0018’ and ‘veldfood10-0018’ offer equally good representations of the expected lateral click waveform. The waveforms of segments extracted from ‘veldfood-0018’ and ‘veldfood3-0018’ exhibit more peculiar shapes. The waveform presented for ‘veldfood-0018’ displays the same delayed transient as expected for a lateral click waveform, however this is followed by a few milliseconds of turbulent noise and a second transient. The waveform presented for ‘veldfood3-0018’ exhibits a similar pattern, with a lengthy rise time to the maximum amplitude, and what appears to be two transient responses. Though there is obvious variation between these segments, they also share visible commonalities. Each of the waveforms are characterized by a long rise time and a substantial amount of turbulent noise. Hence, the general shape of these waveforms still resembles the shape of the lateral click more closely than any other click type. Based on these occurrences, as well as instances of consistency as presented in figure 9 below, it is proposed that variation across corresponding click segments may occur to an extent. Therefore, segments for particular lexical items are produced consistently across all tokens, while tokens of other lexical items exhibit variation. A description of this variation is necessary for assessing the generalizability of the following phonological findings.

grasshopper2-181206 ‘!arnǂa’

grasshopper3-181206 ‘!arnǂa’



grasshopper-181206 ‘!arnǂa’

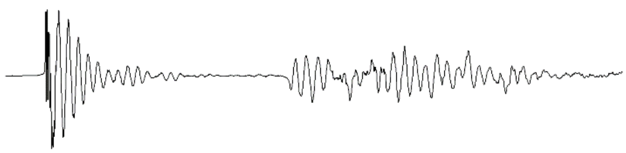


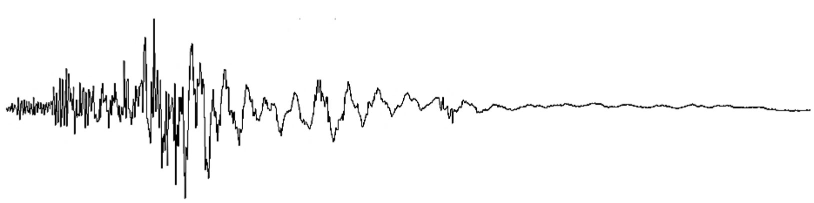
Figure 9: The click burst for each token in !arnǂa ‘grasshopper’

The click segments recorded for a specific lexical item are either produced consistently, as displayed in figure 9, or one may expect variation between corresponding segments extracted from different tokens as displayed in figure 8.

5.5 Analysis

5.5.1 Plain clicks

man2-0015 ‘|ui’



| Dental

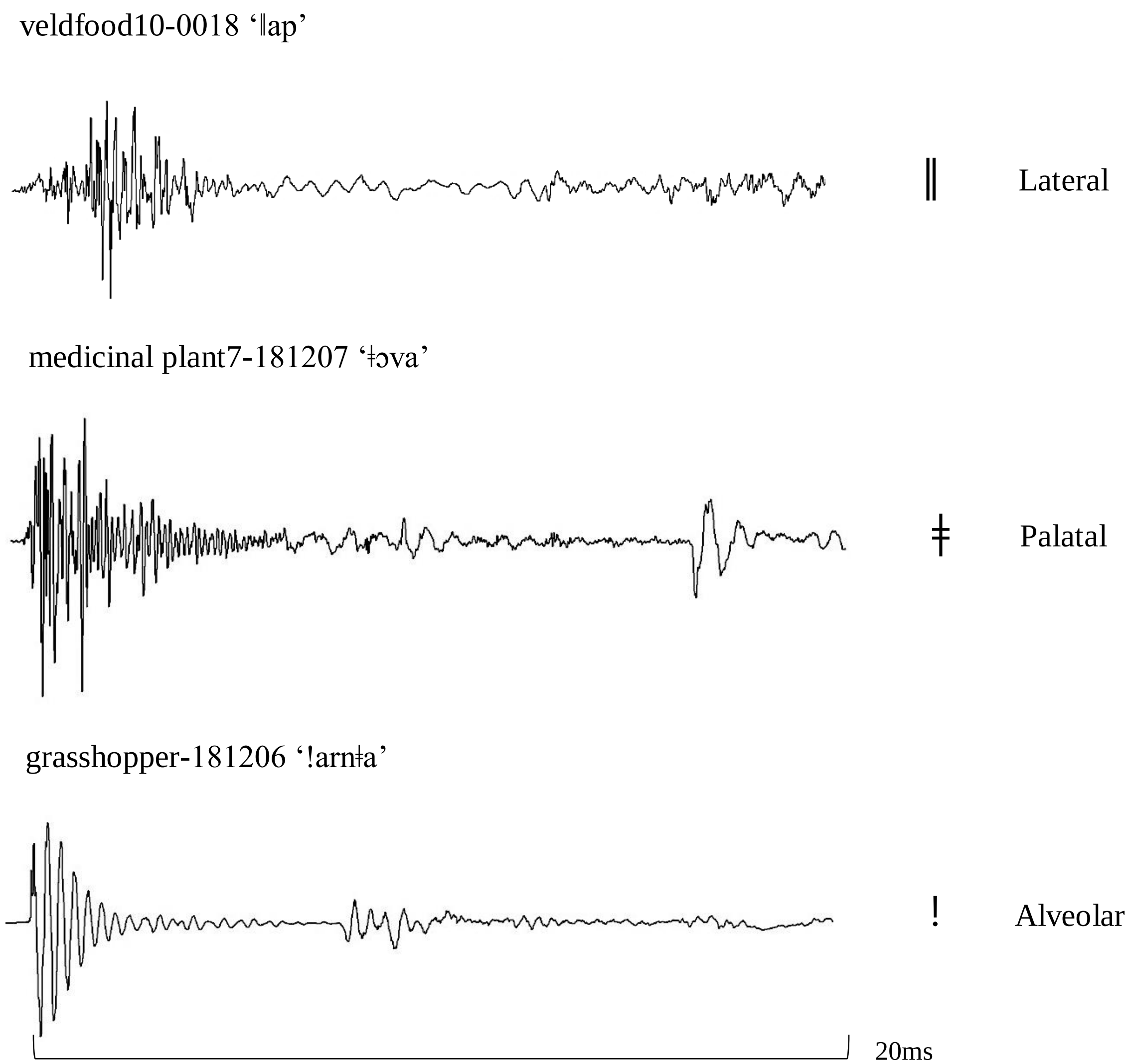
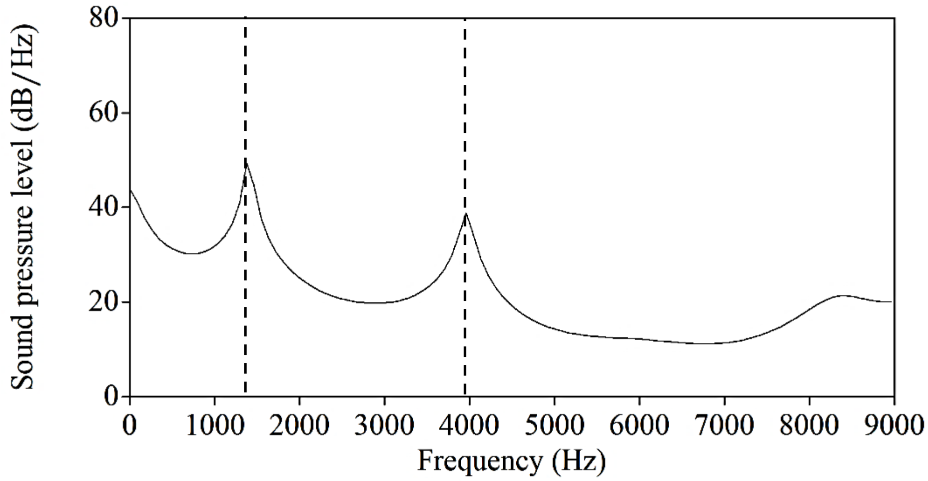


Figure 10: The release burst of four click types, extracted from tokens of four different lexical items. |ui ‘man’, ǁap ‘veld food’, ǂɔva ‘medicinal plant’, !arnǂa ‘grasshopper’.

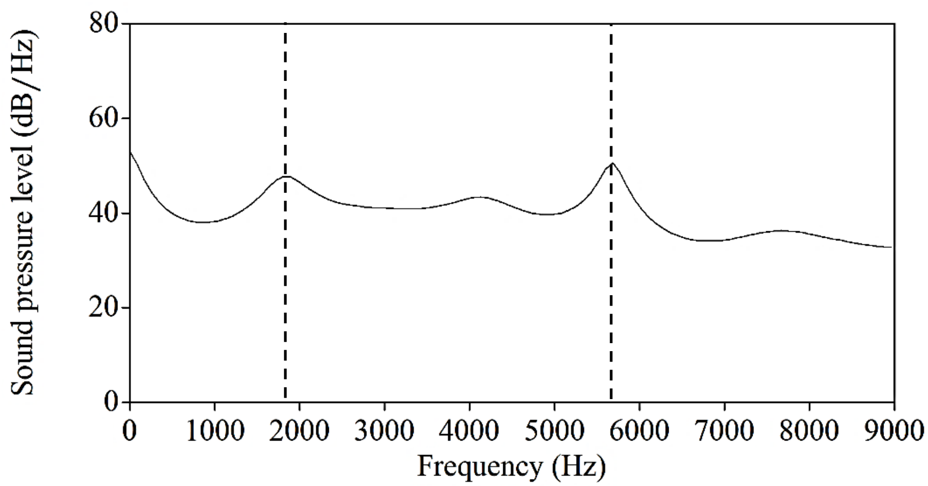
The representative waveforms displayed in figure 10 each constitute a 20ms window of the anterior release burst of one of four click types. As predicted, we see extended turbulent noise following the release of the dental click. Similar to dentals, the lateral click also has a longer burst duration, with a low amplitude transient response. The dental and lateral clicks are both characterized by the lengthy burst duration and low amplitude of the transient response. The palatal click also appears to exhibit a lengthy burst duration, however if one separately considers the difference in time taken for each click to reach maximum amplitude, the difference between the palatal and noisier clicks becomes more apparent. Unlike the dental and lateral clicks which exhibit extensive rise times before reaching maximum amplitude, the palatal click reaches maximum amplitude in under 2ms. Along with a relatively short rise time, the transient response of the palatal click is significantly intense and dominates the waveform. Another waveform with a short rise time and dominated by the transient response is the token representing the alveolar click. The release burst is characterized by an intense transient which progresses into damped oscillations.

While there is clearly visible turbulent noise following the release burst of all three of the initial clicks, the release burst of the alveolar click constitutes eight uniform oscillations. The waveforms of the noisier clicks constitute a range of irregular oscillations following the initial release. Based on the distribution of energy across this selection of waveforms a distinction can be made between the noisy versus abrupt clicks. The most distinguishing features identified in the analysis of these waveforms include the difference in rise time and the intensity of the transient response. Translated to an articulatory account, the palatal and alveolar clicks exhibit an earlier and more intense transient due to the rapid rate at which the articulators separate. The dental and lateral clicks therefore reflect a slower rate of closure release (Ladefoged & Maddieson 1996, pg. 258). Though the waveforms presented in figure 10 are not accompanied by any visible indication of a pulmonic burst following the anterior release it is generally understood that all clicks are produced with a posterior closure. Hence, this acoustically low-intensity event is generally transcribed as a velar [k] or uvular [q] accompaniment (Ladefoged & Maddieson 1996; Miller 2010; Miller et al 2009).

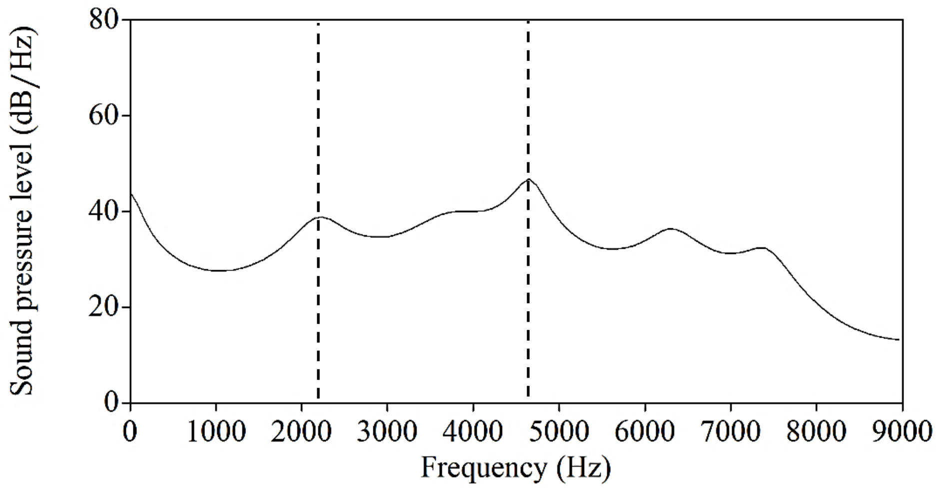
man2-0015 ‘|ui’



veldfood10-0018 ‘ǁap’



medicinal plant7-181207 ‘ǂɔva’



grasshopper-181206 ‘!arnǂa’

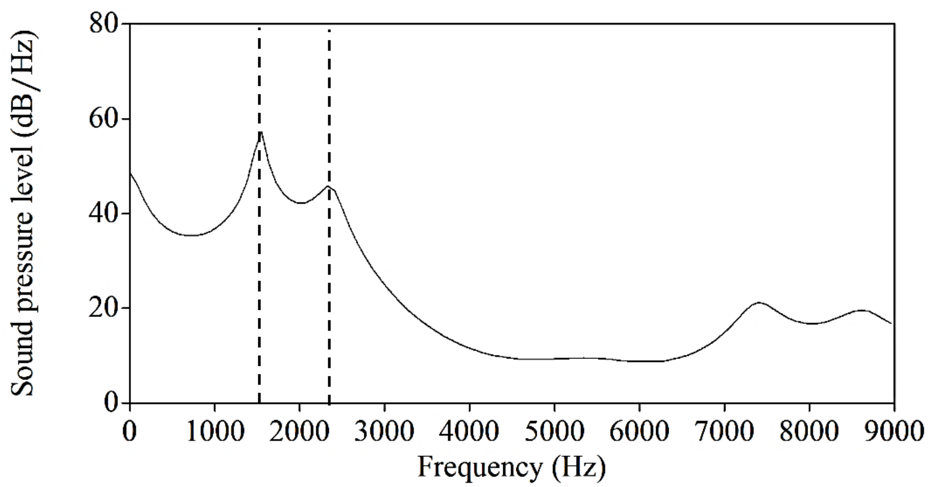


Figure 11: The LPC spectra of four click types, extracted from tokens of four different lexical items. |ui ‘man’, ǁap ‘veld food’, ǂɔva ‘medicinal plant’, !arnǂa ‘grasshopper’.

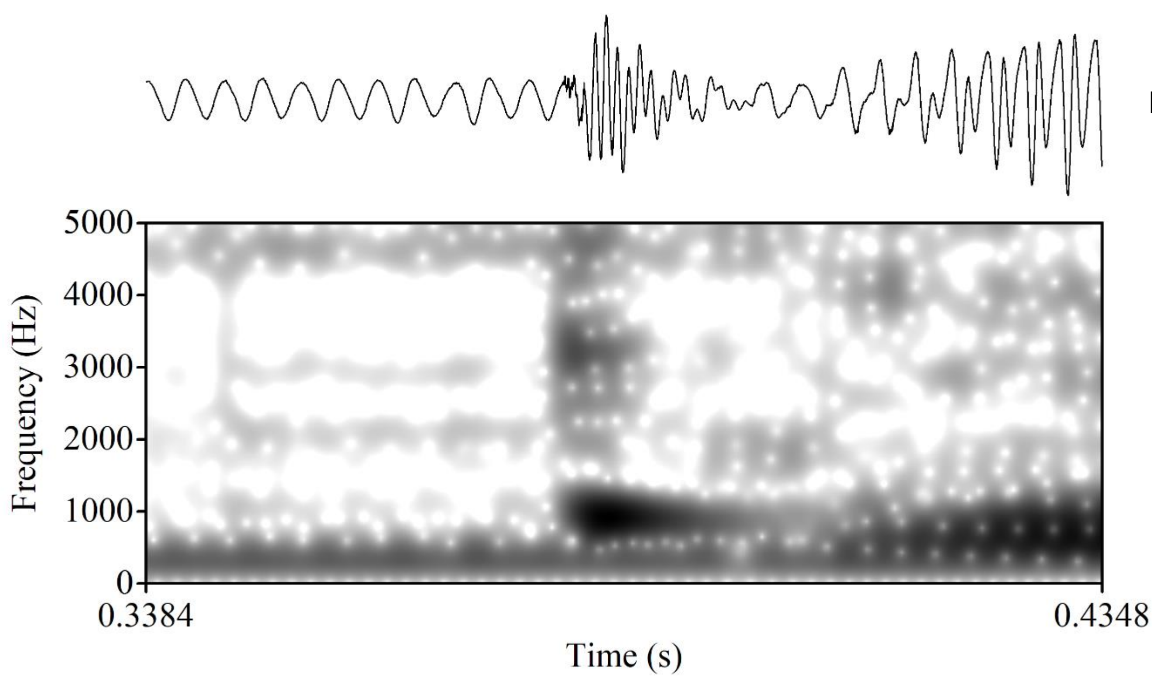
The spectral samples presented in figure 11 exhibit predictable differences in spectral frequency across the four identified click types. Two identifiable peaks are present across all spectra, which provide points of reference for comparison. The spectrum for the dental click exhibits one peak at 1400 Hz, and a second peak at 4000 Hz. The frequency of both peaks fit the expected spectral distribution of dental clicks as reported in Ladefoged & Traill 1994, pg. 42. While the first peak at 1400 Hz reflects the similarity between the distribution of dentals and bilabials, the second peak reflects a common observation across dentals which aligns with that expected for the distribution of palatals. The result of the difference in frequency and amplitude between these peaks is a slightly negative spectral slope. The spectrum for the lateral clicks exhibits two peaks present at a similar amplitude, the first peak at about 2000 Hz, and a second peak at about 6000 Hz. The lateral click therefore appears to display a flat distribution, though previous analyses predict more energy distributed above 2500 Hz (Miller et al. 2009, pg. 140). The two peaks exhibited in the palatal click spectrum are present at just above 2000 Hz and at about 4700 Hz. The second peak is about 10 dB higher than the first, creating a slight positive slope. This observation aligns with previous analysis in that most of the energy is distributed at the higher frequencies. Finally, the spectrum for the alveolar click constitutes an initial peak at 1500 Hz, and a second peak at about 2300 Hz. Although the initial peak resembles the initial peak observed in the dental spectrum, these click types are distinguished by the position of the second peak. The second peak is reported to reflect the volume of the lingual cavity and hence appropriately distinguishes the larger dental cavity from the more compact cavity formed for the articulation of the alveolar (Miller et al. 2009, pg. 140)

5.5.2 Click accompaniments

The main accompaniments discussed as part of this investigation include the nasal accompaniment and the uvular accompaniment. The velar fricative accompaniment which appears less frequently across the data set will only be briefly described. The following analysis therefore focusses on the accompaniments most frequently transcribed. As discussed in section 5.2 above, the nasal accompaniment is one of the most common accompaniments reported to occur across click languages (Maddieson et al. 1999), which also occurs frequently in Tumʔi. The uvular accompaniment takes the form of a pulmonic stop following the initial click burst. Also discussed in section 5.2 above, this particular voiced stop accompaniment is much less common, having only been reported to occur in languages of the Tuu family (Ladefoged & Traill 1994; Miller 2010; Miller et al. 2007).

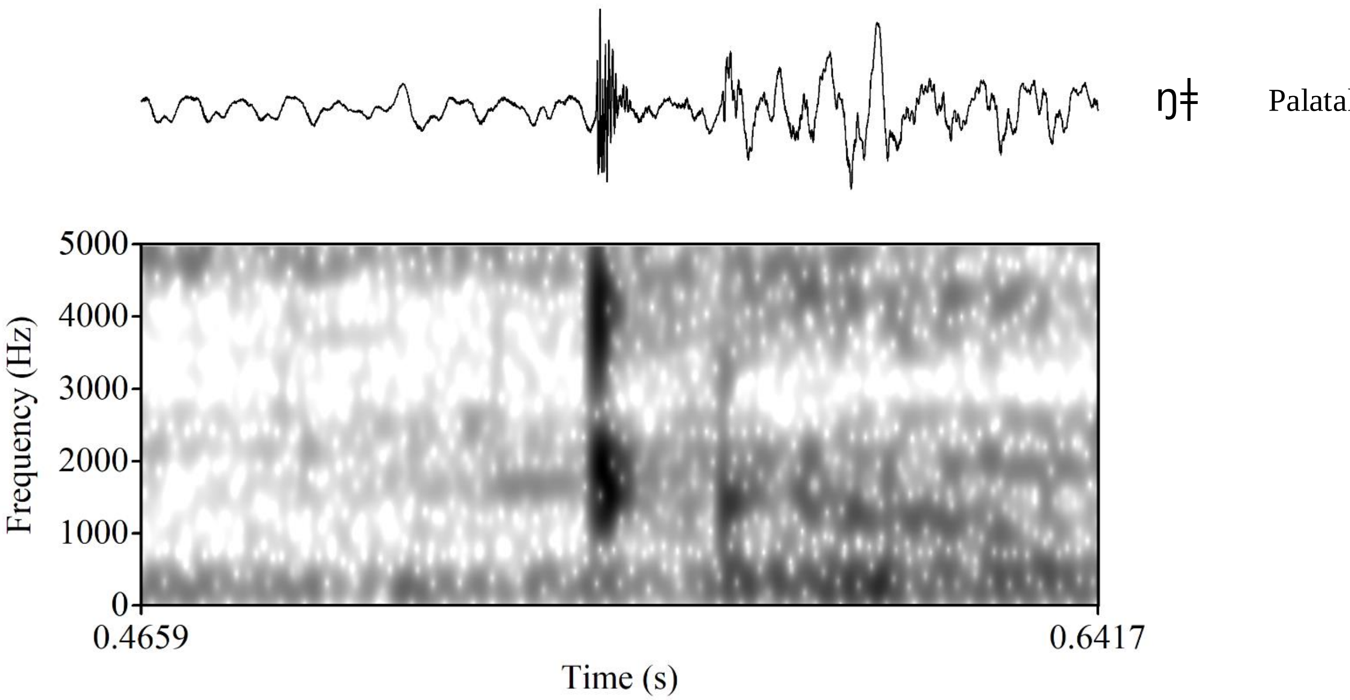
5.5.2.1 Nasal

drunk-0016 ‘n!ɔrɔ’

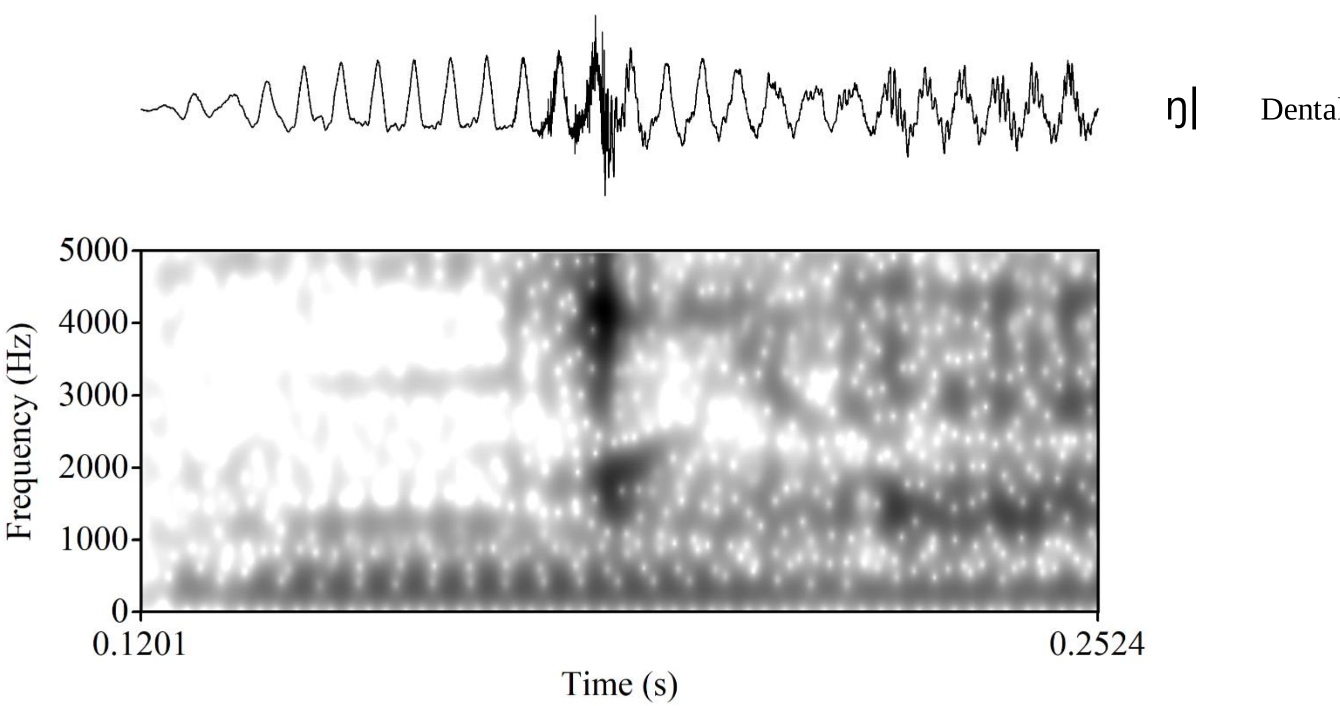


ŋ! Alveolar

crush9-181206 ‘sein‡ama’



head-181207 ‘n|a’



jail-0016 ‘sein|eŋ’

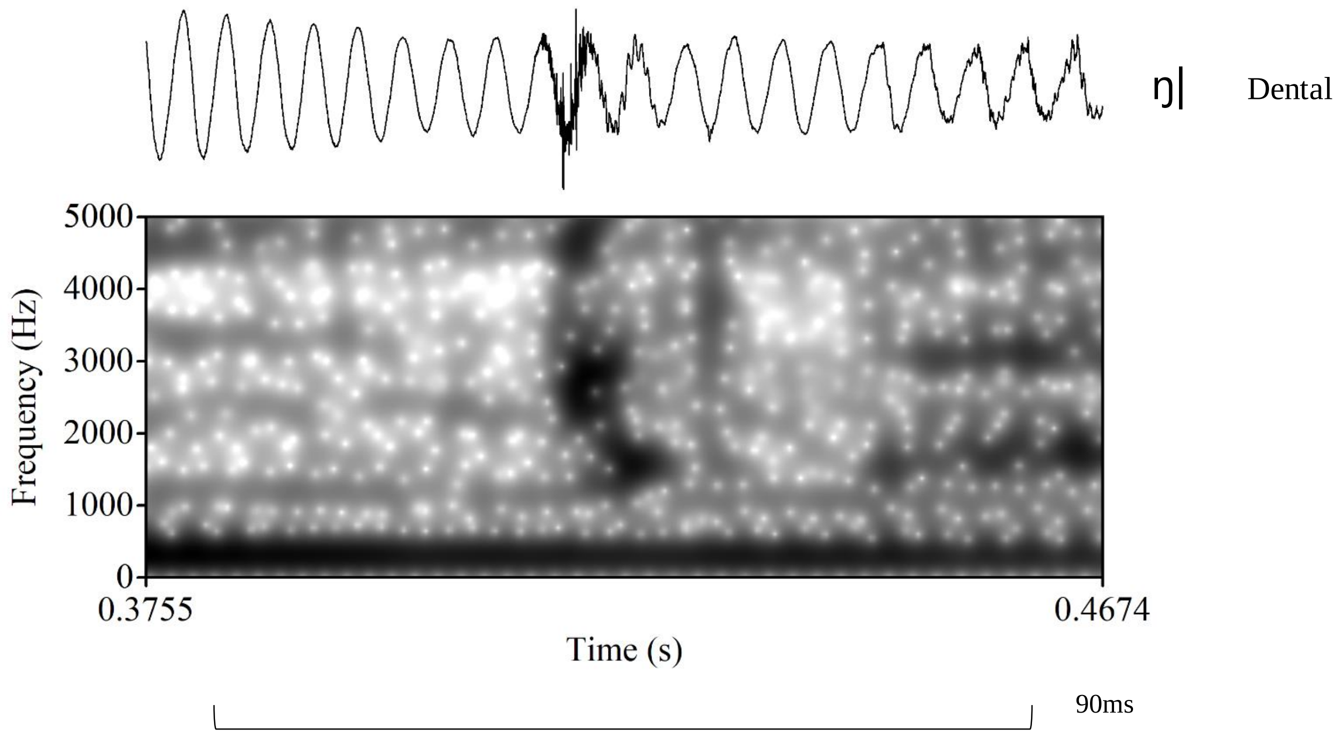


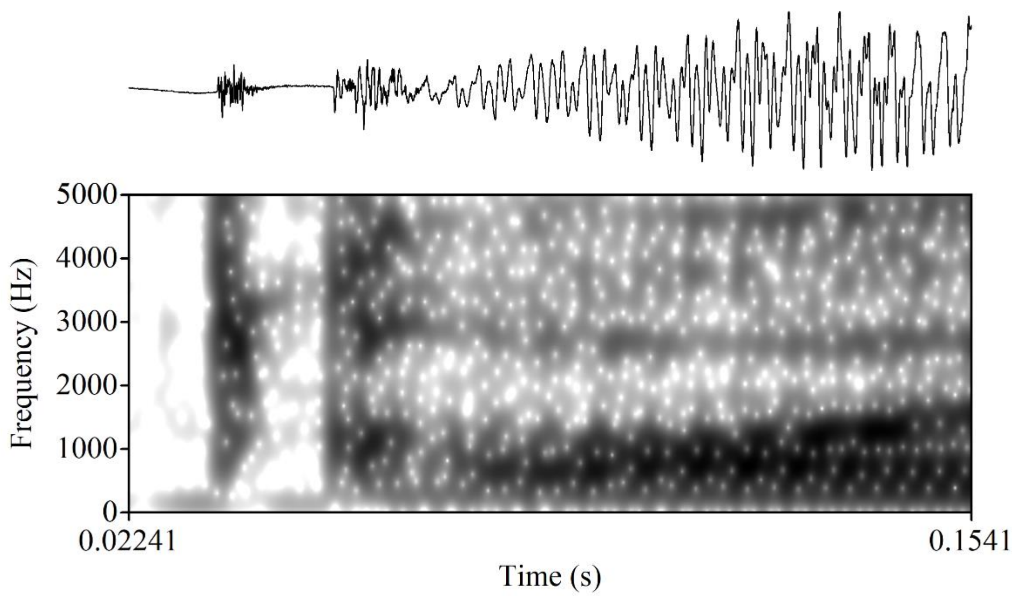
Figure 12: The release burst of four nasal clicks, extracted from tokens of four different lexical items. [ŋ!ɔrɔ] ‘drunk’ & [ŋ|a] ‘head’ represent word initial clicks with nasal accompaniment. [seiŋǂama] ‘crush’ & [seiŋ|eŋ] ‘jail’ represent word medial clicks with nasal accompaniment.

Observation across the entire data set indicates a total of 22 cases of nasal accompaniments with alveolar, palatal and dental clicks. Instances of nasal accompaniment include lexical items such as [ŋ!ɔrɔ] ‘drunk’ and [ŋ|a] ‘head’ presented in figure 12 above, in which the nasal aspect constitutes the efflux of the word initial click segment. However, this accompaniment is also observed to occur in the word medial position, in lexical items such as [seiŋǂama] ‘sugar’ and [seiŋ|eŋ] ‘jail’. Word medial in the scope of this paper refers to intervocalic click segments generally observed in the onset or C_2 position. Overall, nasal accompaniments are observed to occur with three particular click types, in two positions. The waveform presented for the word [ŋ!ɔrɔ] ‘drunk’ below exhibits a burst resembling the transient response expected for an alveolar click produced and released with a lowered velum. The acoustic representation of the stimuli indicates extensive glottal activity preceding the release burst of the click. The waveform representing [seiŋǂama] ‘sugar’ exhibits the burst of a palatal click with nasal accompaniment in the word medial position, displaying irregular glottalization, with fewer periods.

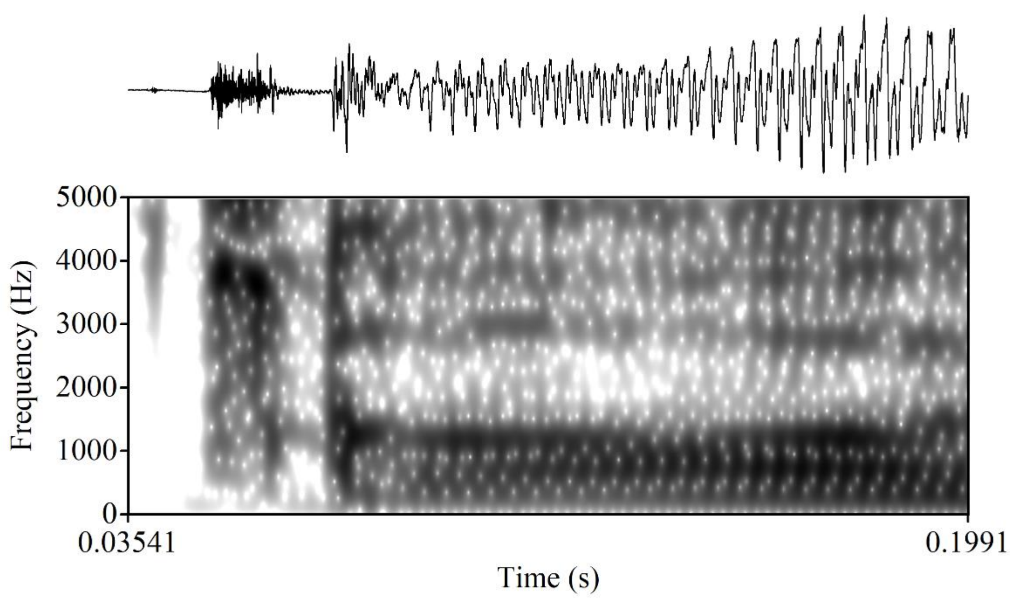
The distribution of the nasal accompaniment within this data set is constrained to three of the click types defined in section 5.5.1 above and does not appear to exhibit any controlled voicing contrast. The token representing the lexical item [ŋ|a] ‘head’ constitutes the waveform of a dental click with nasal accompaniment, hence differing from the more abrupt clicks described above. This waveform is also characteristic of nasal accompaniment produced with a smaller glottal aperture, which is signaled by the extensive glottal activity preceding the anterior release and extending through to the onset of voicing of the following vowel. As variation exists throughout every occurrence of these segments, I propose two possible explanations for the differences between the nasal accompaniment across these tokens. First, in the word initial position, instances of nasal accompaniment are purposefully more voiced due to the phonological prominence of the position. Second, the difference between the absence of glottal activity and delay of voicing onset following the burst, is another attribute of the general variation which is now an idiosyncratic feature of each speaker’s dialect (Cook 1989). However, the final token representing a second word medial segment exhibits nasal voicing consistently preceding, and throughout the release burst of the dental click segment. Therefore, though the first explanation would imply an interesting positional contrast, any disparities in voicing and glottal activity between the segments presented in Figure 12 (Ladefoged & Traill 1994), are likely the result of uncontrolled variation.

5.5.2.2 Uvular

woman-0012 ‘ǁqoita’



woman2-0012 ‘ǁqoita’



woman3-0013 ‘ǁqoita’

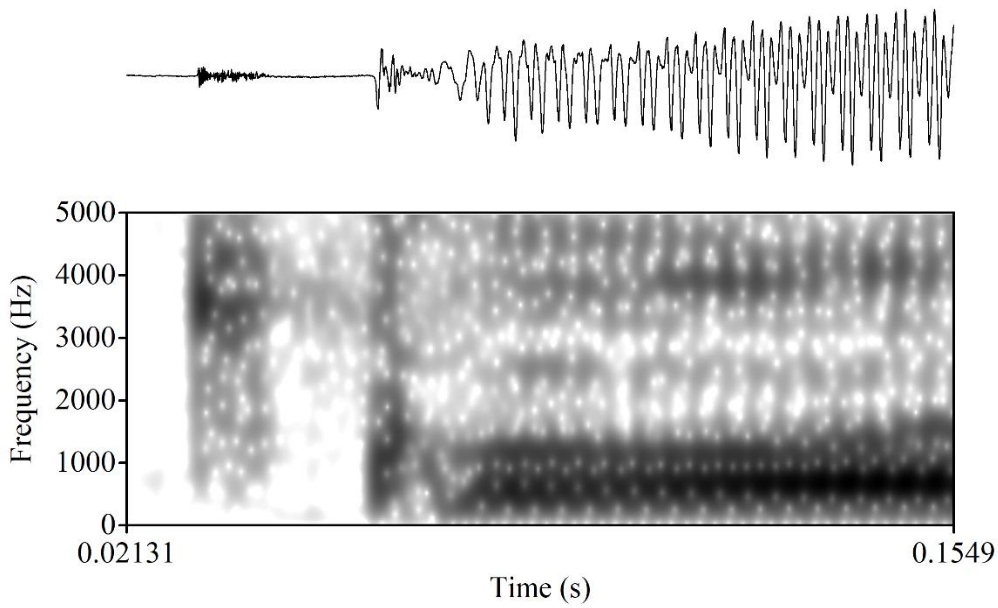


Figure 13: The release burst of three lingual-pulmonic segments, extracted from tokens of the same lexical item. ǁqoita ‘woman’.

The second accompaniment type is presented in figure 13 above. This accompaniment is consistently transcribed in tokens of as many as 29 lexical items but is also observed to occur sporadically among tokens of other lexical items. The choice to transcribe the pulmonic burst following the posterior constriction as the uvular stop [q] is based on extensive investigation of the airstream mechanism involved in producing clicks, and the articulatory nature of the posterior constriction (Miller et al. 2009). The click segments displayed in the waveforms above each exhibit two bursts preceding the onset of the vowel. The initial burst signals the anterior release of the click segment, while the second reflects the release of the posterior closure produced at the upper end of the tongue root. The clicks produced by the release of the anterior closures across all the presented samples all resemble features of both noisy and abrupt click types. The initial waveform for the token ‘woman-0012’ does not exhibit any intense transient but exhibits a long rise time to maximum amplitude. The considerable rise time extends to the waveforms for tokens ‘woman2-

0012' and 'woman3-0013', however differences in burst duration and the intensity of the transient indicate substantial variation. The waveform for token 'woman2-0012' exhibits a longer burst duration comprised of irregular oscillations. The final waveform for token 'woman3-0013' is similar in that it constitutes a longer rise time and burst duration. However, the oscillations following the transient are regular and distinct, similar to those expected for the alveolar click. Hence, as discussed in the initial analysis of click types in section 5.5.1, the speakers exhibit variation between corresponding click segments across tokens of the same lexical item.

The clicks presented in figure 13 each display an additional transient caused by the release of the posterior constriction. Each waveform displays an initial noise burst caused by the release of the anterior constriction, followed by a period of silence, and a second noise burst. The second burst is one caused by a significant pressure build up behind the posterior constriction which produces a pulmonic burst at the point of release (Miller at al. 2009). The pulmonic burst present in each of these tokens appears to resemble the preceding lingual burst in terms of the intensity of the transient and burst duration. The pulmonic burst for token 'woman-0012' exhibits a short transient with a slightly longer burst duration than the preceding lingual burst. The accompaniment in token 'woman2-0012', which follows a more intense lingual burst, appears to exhibit a more intense transient response than the pulmonic burst in the token 'woman-0012'. The pulmonic bursts for each of these tokens appear to extend for the time of the burst duration and immediately transition to the onset of the vowel. In the case of the final token 'woman3-0013', the period of silence preceding the pulmonic burst is relatively longer than observed across the other tokens. The extended pulmonic burst also appears to exhibit a lot of low frequency energy following the release burst which may be the formants of the vowel beginning to emerge through the brief aspiration. Hence, it appears that the nature of the pulmonic burst resembles that of the preceding lingual burst in terms of the rate of constriction and release. The transient of the pulmonic burst for the token ‘woman3-0013’ also appears to correspond to the intensity of the transient response exhibited for the preceding lingual burst, displaying a more abrupt release characterized by more uniform oscillations.

Overall, the use of this audible posterior release is phonetically evident and exhibited across each of the four click types identified in the inventory of Tumʔi. Hence, this accompaniment is not constrained in terms of co-occurrence with any particular click type, having been observed as an extension of both abrupt and noisy clicks. As predicted in previous analysis of Tuu languages such as !Xóõ, N|uu, and ǂHoan, the pulmonic burst associated with this accompaniment tends to vary in accordance with the preceding lingual segment type (Miller et al. 2009 pg. 149). Hence, the results of the analysis above do present the question of a corresponding link between the nature of the uvular burst accompaniment and its preceding lingual burst.

5.5.2.3 Additional accompaniment

The velar fricative accompaniment recorded in Tumʔi is significantly limited in its occurrence. As indicated in section 4.1, table 4, the velar fricative accompaniment occurs only with the dental and alveolar clicks.

Female genitals-0013 ‘ǁxa̰ra’

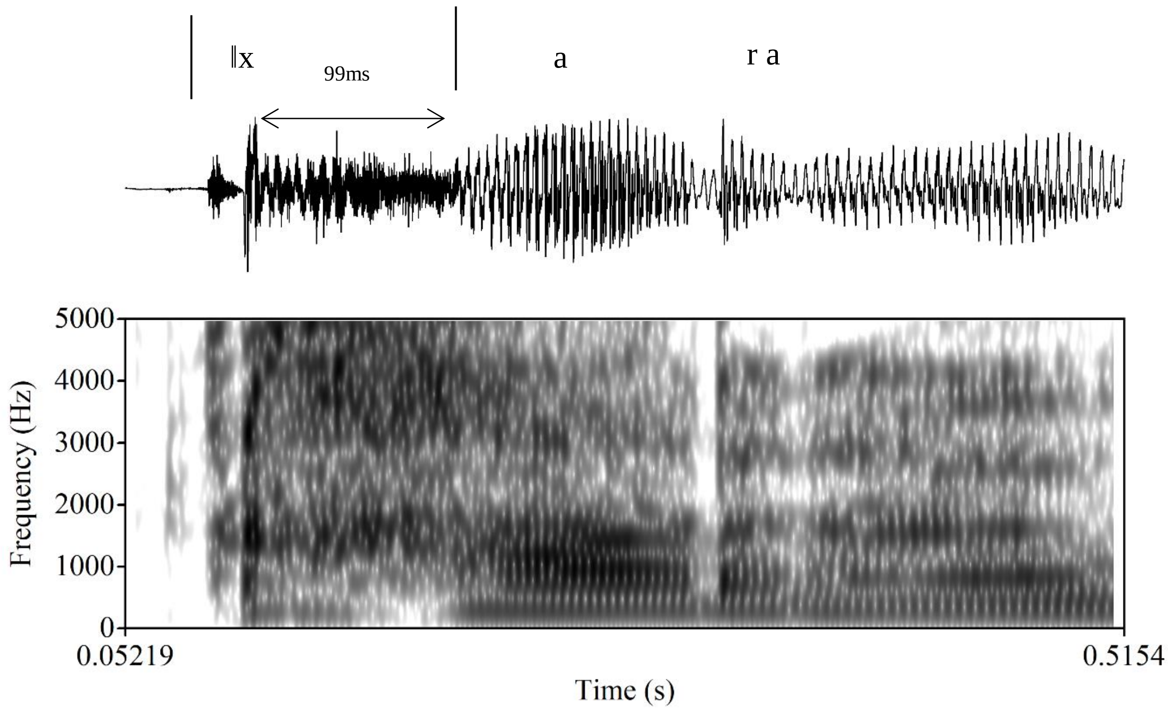
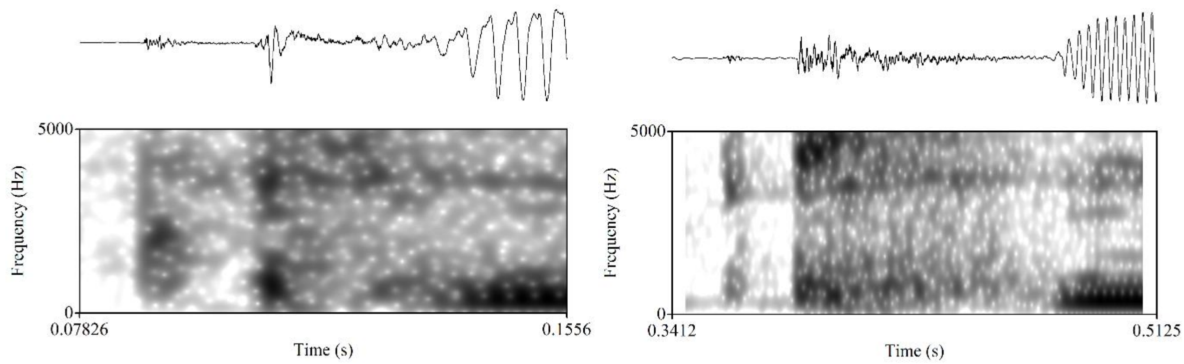


Figure 14: The entire lexical item containing the dental click with velar fricative accompaniment, extracted from the lexical item ǁxara ‘female genitals’.

The initial segment demarcated in figure 14 constitutes a lateral click followed by a velar fricative accompaniment. The initial transient closely resembles that of a noisier click with a low intensity response followed by a series of irregular oscillations. Following the dampened oscillations is a second transient which expands into the fricative portion of the segment extending across 99ms. The frication exhibited in this accompaniment is apparent both in the waveform and spectrogram presented in figure 14. The fricative portion of the waveform is characterized by extensive turbulent noise, while the corresponding portion of the spectrogram displays a significant degree of energy beginning at 1000 Hz and extending to the highest frequency.

5.5.3 Word initial vs. word medial clicks

coffee-0013 ‘|qum|qum’



coffee-0016 ‘!um!qum’

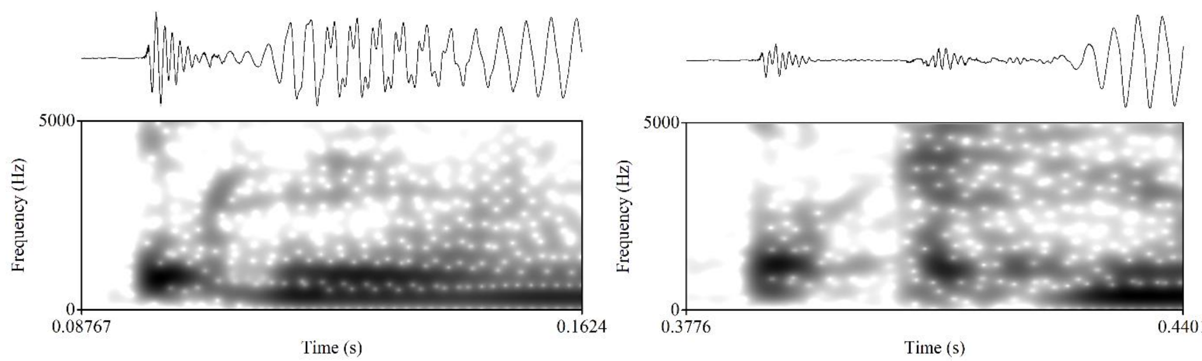


Figure 15: The release bursts of two sets of waveforms, extracted from tokens of the same lexical item. |qum|qum ‘coffee’. Each waveform set includes an initial waveform representing the word initial click segment, and a second waveform representing the word medial click segment.

As specified in section 5.4 above, 19 lexical items are phonetically transcribed with a click segment in word medial position (see Appendix B). The waveforms presented for the tokens ‘coffee-0013’ and ‘coffee-0016’ indicate that the same variation reported in the word initial position extends to the click types recorded in the medial position. This variation is exhibited in the disparities between the click sets extracted from each token. The waveforms representing the first token ‘coffee-0013’ are both comprised of a lingual burst produced with an anterior dental constriction, followed by a pulmonic burst. The initial lingual burst and the second pulmonic burst both constitute relatively noisy segments with irregular oscillations. Hence, the first set of waveforms provide evidence of similar click types occurring in both the word initial and word medial position. The set for token 'coffee-0013' also serves as a representation of the uvular accompaniment occurring in word medial position. Therefore, as observed in the distribution of nasal accompaniment, the pulmonic stop is not positionally restricted.

The second token 'coffee3-0013' reiterates the inconsistency observed across click segments recorded for a single lexical item. The initial waveform presented for this token displays an alveolar click almost immediately followed by the onset of voicing of the following vowel. The medial waveform extracted from this token set indicates an identical release burst for the initial click, however unlike the word initial click, the medial click segment is accompanied by a pulmonic stop. As discussed above, this accompaniment is characterized by an initial lingual burst followed by a period of silence and a second burst caused by the release of the posterior constriction. Based on the characteristics displayed across these waveform sets, there appears to be some correlation between the burst characteristics of the word initial click and that of the medial click. This proposed similarity is further exhibited in the transcription of the click initial and click medial lexical items, in which these segments generally appear to agree in place of articulation throughout the data set. However, the tokens presented above also indicate that the presence of a pulmonic accompaniment following one click would not necessarily determine the same accompaniment for the other. Therefore, the proposed similarity is not expected to apply to the click accompaniment.

5.6 Discussion

Based on the above analyses, clicks in Tumʔi are produced with one of four possible anterior releases, dental, lateral, palatal or alveolar. The distribution of clicks throughout the data set indicate that, as a class of sounds, clicks carry a heavy functional load in the contemporary use of Tumʔi (Güldemann 2006, pg. 108). Furthermore, considering a good proportion of these words such as ‘man’ [|ui], ‘woman’ [|qoita], ‘drink’ [|qxwa̰], and ‘look’ [n|a!qoi], etc., may be described as belonging to the basic vocabulary, there is reason to believe that the functional role of clicks could be considered a true characteristic of the now obsolescing language (Maddieson, Ladefoged & Sands 1999, pg. 62). Though representative tokens provide evidence for the use of each of the click types listed above, the data also displays considerable variation across tokens of the same lexical item. Therefore, the recorded production of these click types does not necessarily reflect a controlled distinction recognized by the speakers. The speaker’s consistency in the articulation of particular clicks varies between lexical items. Hence, the tokens presented for the lexical item ‘grasshopper’ [!arnǂa] may exhibit more consistency across corresponding click segments than the tokens presented for the lexical item ‘veldfood’ [ǁap], as seen in figures 8 & 9 above. Phonetic analysis and transcription of the collected data indicate that this variation between corresponding segments extracted from tokens of the same lexical item extends to the production of click accompaniments.

The tokens representing the nasal accompaniment in the analysis above as well as in the phonetically transcribed lexical inventory indicate a restriction in distribution. Nasal accompaniment occurs with the alveolar, palatal and most commonly the dental click, but does not occur with the lateral click. Based on the analysis in section 5.5.2.1 above, there does not appear to be a controlled voicing contrast between the nasal accompaniments in word initial or word medial position. Instances of variation are also apparent across the tokens representing the uvular accompaniment, however this variation is expected to occur as the consequence of differences between the preceding click type (Miller et al. 2009, pg. 149). The significance of this accompaniment however, lies in its distinct acoustic characteristics, and frequent distribution throughout the data set. The limited occurrence of this accompaniment across Khoisan languages provides some indication that it may be genealogically related to languages exhibiting the same phonological structure. The lingual-uvular segment has been reported to occur in four languages of the Tuu family !Xóõ, |Nuu, ǂHoan, |Auni (Ladefoged & Traill 1994; Miller et al. 2007), and one Khoe language, |Gui, which has had extensive contact with !Xóõ (Nakagawa 2006; Vossen 2013, pg. 181). Hence the use of this accompaniment presents a typological characteristic similar to a number of languages classified as belonging to the Tuu family.

Chapter 6: Voicing, Aspiration & Laryngeal contrasts

6.1 Introduction

The purpose of this chapter is to discuss the possibility of particular phonemic contrasts between consonants recorded in Tumʔi. The analysis below is confined to contrasts in laryngeal activity, including voicing, aspiration and use of the glottalic egressive airstream mechanism. Representative tokens of a single lexical item are used to exhibit the disparities in aspiration across multiple tokens. The phonemic contrast is evidenced by the regularity with which the phoneme /t/ is produced as [t^h] in the lexical item 'God' [t^hi!qɔ]. The analysis of voicing is limited to the comparison of one minimal pair, the comparison of corresponding segments in two different lexical items to determine whether speakers of Tumʔi recognise the distinction, and whether a voicing distinction is a feature of the consonant inventory of this language. Finally, ejective segments are observed in the data set, however this form of laryngeal articulation appears to be limited to the alveolar plosive [t] as is the case with full aspiration. The nasal segment in the word [tumʔi] 'speak' is commonly produced with a period of voicing followed by a diffuse burst and glottal stop. This production is distinct from instances in which the nasal is followed by a plain glottal stop; these segments are analysed as post-nasal glottalized stops.

6.2 Literature

Based on previous typological analyses the consonant systems of Khoekhoe and |Xam are similar in overall organization. The major difference between these systems concerns the size of the inventory, which in turn is a typological consequence of an extended variety of distinctive features attested in Tuu languages (Güldemann 2006). Khoekhoe languages lack laryngeally marked stops including both aspirated egressives and ejectives; observations across the inventories indicate that these segments have been replaced with aspirated and ejective fricatives. The same laryngeally marked stops are however also unattested in the consonant inventory of |Xam which is considered typologically unusual (Vossen 2013, pg. 211). Previous analyses attribute this shared lack of complex stops to the phonological process of affricate lenition. For example, the segments /t^h/ and /k^h/ in both Khoekhoe and |Xam have undergone a phonological process of lenition introducing the segments /ts/ and /kx/ into the inventory (Beach 1938, Vossen 2013, pg. 211). Also lacking from the phonological inventories of Khoekhoe varieties is a voicing distinction. Egressive segments with the feature [+voice] are absent from Khoekhoe varieties with the exception of !Ora and |Gui (Güldemann 2006; Vossen 2013, pg. 153). Therefore, any evidence of the laryngeally marked stops, such as those discussed above, or a phonological voicing distinction would naturally align with the phonological typology of Tuu languages.

6.2.1 Phonetics

Though it is understood that contrasts in voicing are attested cross-linguistically, previous analyses have characterized this distinction in some languages as a difference in aspiration (Ladefoged & Johnson 2014, pg. 57). The presence or absence of additional airflow following the stop burst can determine the classification of a particular consonant as voiced or voiceless. In these cases, the degree of voicing of a segment is therefore determined by the period of aspiration following the plosive burst release. This period of voicelessness following the release of the articulators is recorded as a measure of voice onset time (VOT). The measure of VOT must however be considered against a continuum of varying degrees of aspiration and glottal aperture (Cho & Ladefoged 1999, pg. 226). The VOT value of a segment may be used to distinguish between contrasts in voicing as well as aspiration. A low or negative VOT value provides evidence for the classification of a voiced segment, however these segments may be further distinguished as voiced with a negative VOT and unaspirated with VOT of zero or close to zero. A substantial positive VOT value provides evidence in support of a voiceless classification; the relative VOT values of these voiceless segments are then considered to determine the presence of contrastive aspiration (Ladefoged & Johnson 2014, pg. 152). As noted by Cho & Ladefoged (1999), statistical clumping of distinctive measures like VOT is not very tenable, however they propose a plausible division of VOT values into four phonetic categories. The first includes segments with a VOT of around 30ms, forming a category of unaspirated stops. A VOT of around 50ms would signal the second category, one of slightly aspirated stops. Another would be at around 90ms signaling aspirated stops, and any VOT greater than this would be classified as belonging to a fourth category comprised of highly aspirated segments (Cho & Ladefoged 1999, pg. 223). The benchmark measures noted above are included only as points of comparison and should not be considered the quintessential standards for determining the presence of a voicing distinction.

Unlike the phonetic distinctions discussed above, ejective segments are not distinguished based on their VOT value. Ejectives are produced differently, in that the vocal folds are pulled together and moved up during closure. Air is compressed in the pharyngeal cavity due to the upward shift of the closed glottis and later released. Therefore, unlike the pulmonic egressive plosive and aspirated pulmonic egressive plosive, the ejective is at times characterized by two bursts; making both the waveform and spectrogram useful references. A textbook ejective waveform like this would include the release burst of the plosive, which should be as long as 120ms, followed by a second release burst marking the release of the glottal closure (Ladefoged & Johnson 2014, pg. 138). Generally, however, ejective segments are most easily identifiable by the interval between the release of the glottal closure and the onset of voicing of the following segment, which acoustically resembles a period of silence with some reverberation of the stop closure. Additional information is provided in the analyses of spectrograms, including the presence or absence of energy between the preceding plosive and following vowel, which would reflect instances of additional aspiration. The spectrogram also provides information for distinguishing between different places of articulation through the frequency distribution in the spectrum of the release burst as well as the distribution of the vertical striations across the burst spike.

6.3 Method

VOT is measured by analyses of the waveform, beginning from the release burst of the plosive, to the onset of voicing signaling the beginning of the following vowel. Corresponding wideband spectrograms are provided with each waveform as additional evidence of the voicing characteristics. Spectrograms are set to a 0.005s window length, with the dynamic range set to 45dB, for added clarity between burst spikes and continued voicing. These figures are contrasted against the distinguishing characteristics and the standard measures predicted for each of the different plosive types presented in section 6.2.1 above. VOT values are measured to determine disparities in the extent of aspiration between the different segments under comparison. The number of bursts distinguish the pulmonic egressive plosives from ejective segments produced with a glottalic airstream mechanism. The frequency of the burst noise then provides a general indication of the particular plosive type, in terms of the place of articulation.

6.3.1 Data

The following analysis is confined to the comparison of segments extracted from a select number of lexical items. Aspiration distinctions are investigated across representative tokens of the lexical items [tumʔi] ‘speak’, [tʰi!qɔ] ‘God’ and [tʰava] ‘handywork’. The entire inventory has only one lexical item in which an ejective egressive occurs, hence the analysis of the ejective is limited to a different representation of the alveolar plosive in [tumʔi] ‘speak’, which appears to alternate between a pulmonic and glottalic egressive airstream mechanism. The data used in the investigation of the voicing contrast includes the minimal pair [pəri] ‘goat’ and [bəri] ‘bread’, as well as the lexical items [duːr] ‘expression of distance’ and [torəŋ torəŋ] ‘crazy’. This data is limiting particularly for proving the existence of the second voicing contrast, as any disparity would not reflect a phonemic contrast but could be attributed to an allophonic distinction. The credibility of the second pair comparison is further weakened by the fact that the lexical item [du:r] is borrowed from Afrikaans and therefore does not reflect a native lexeme of Tumʔi. Overall, the data available for these analyses is severely limited, therefore the only definitive purpose of the following analysis is to provide evidence for the attested use of these particular phonetic distinctions in Tumʔi.

6.4 Analysis

6.4.1 Aspiration contrast

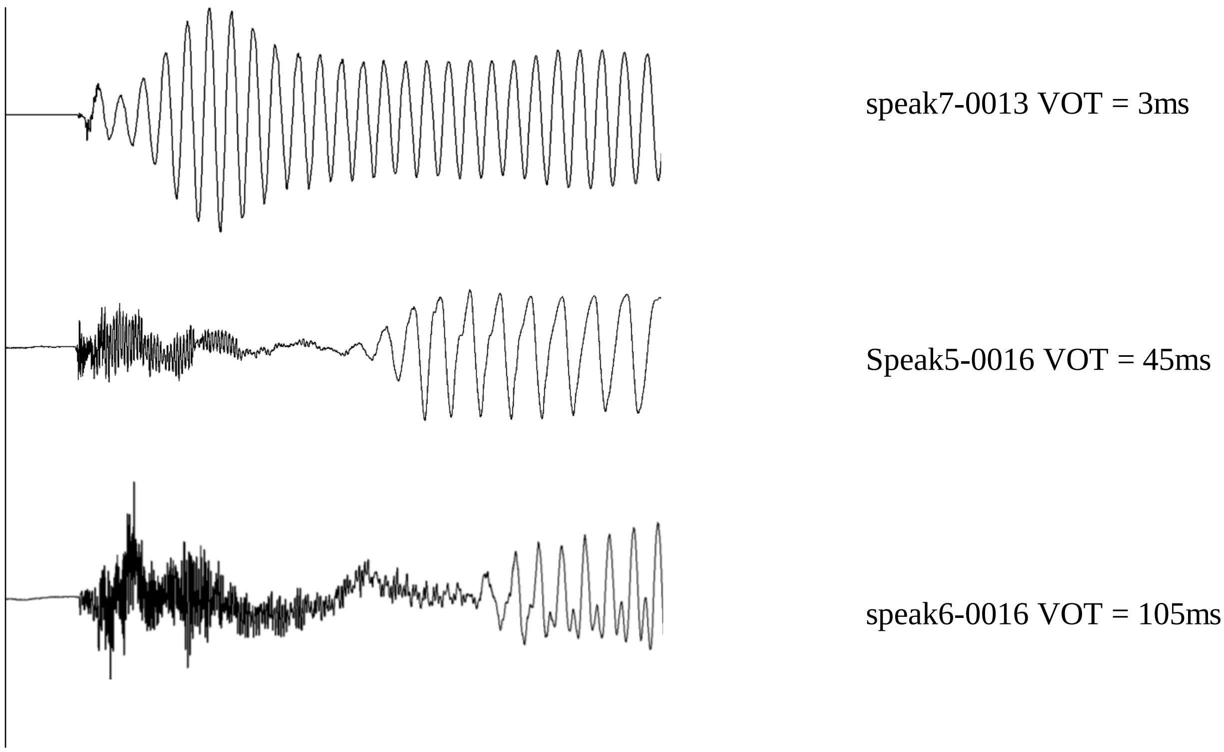


Figure 16: The release burst of the segment [t], extracted from three different tokens of the lexical item tumʔi ‘speak’.

Presented in figure 16 above are instances of substantial disparity in the period of aspiration with which the voiceless alveolar segment [t] is produced. Tokens of a single lexical item [tumʔi] ‘speak’ are produced with varying degrees of aspiration, ranging from entirely unaspirated with a VOT of 3ms to definitively aspirated with a VOT of 105ms. The noise bursts displayed in figure 16 are restricted to the voiceless alveolar plosive, though there is clear variation across the accompanying aspiration; each of the segments are definitively voiceless. This claim is based on the audible difference between the segments produced in the articulation of the lexical item [tumʔi] ‘speak’ versus the segments produced in the articulation of [duːr] ‘expression of distance’. Apart from varying degrees of aspiration, these alveolar segments are identified as voiceless based on the lack of voicing during the closure phase of the stop as well as characteristics of the noise burst observed across each spectrogram as presented in figure 17 below. The formant transitions demarcated by red formant contours calculated in Praat, all display similar features characteristic of a voiceless alveolar plosive. In each case the formants appear to extend through the burst and aspiration period; F2 is generally situated at about 2000 Hz and F3 at 3000 Hz. F2 and F3 take a slight dip during the aspiration period; this is more obviously exhibited in the spectrogram for the token speak6-0016 which is the most aspirated segment. At the onset of the following vowel the third formant generally remains steady around 3000 Hz, while the second formant dips to just above 1000 Hz. Most importantly however each spectrogram appears to exhibit little to no energy distributed at the level of the first formant (for all spectrograms see Appendix E).

Speak5-0016 VOT = 45ms

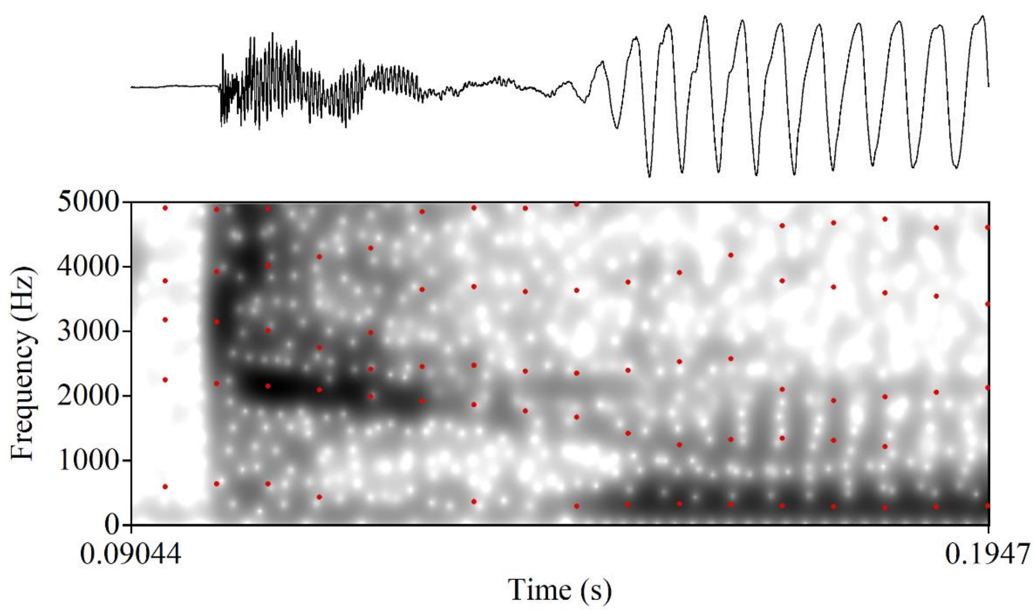


Figure 17: The release burst of the segment [t], extracted from a single token of the lexical item tumʔi ‘speak’.

Though these tokens are all representative of the voiceless segment [t] they exhibit substantial differences in aspiration. The initial token exhibits a 3ms VOT measured from the beginning of the noise burst to the onset of the vowel. The expected VOT for a segment with some form of aspiration is between 40-60ms, therefore this voiceless segment has markedly been produced without aspiration. The second token, speak5-0016, produced for the same lexical item exhibits a lengthier period of aspiration with a VOT of 45ms. The second token presented here might be considered a representative example of the aspiration predicted to accompany the voiceless aspirated segment. Displayed in the final token speak6-0016 is what Cho & Ladefoged (1999) would classify as a highly aspirated segment. Characterised by a VOT value of 105ms, this segment displays the longest VOT of all the recorded tokens of [tumʔi] ‘speak’ (see Appendix F). Considering however, that the tokens presented above as well as the VOT values presented in Appendix F have all been extracted from tokens of the same lexical item, these disparities in aspiration cannot be attributed to any controlled phonological disparity. Hence, the phonetic distinction in this case is likely a reflection of speaker variation or may reflect the general inconsistency with which more complex features are retained and reproduced.

god2-0013 ‘t^hi!qɔ’

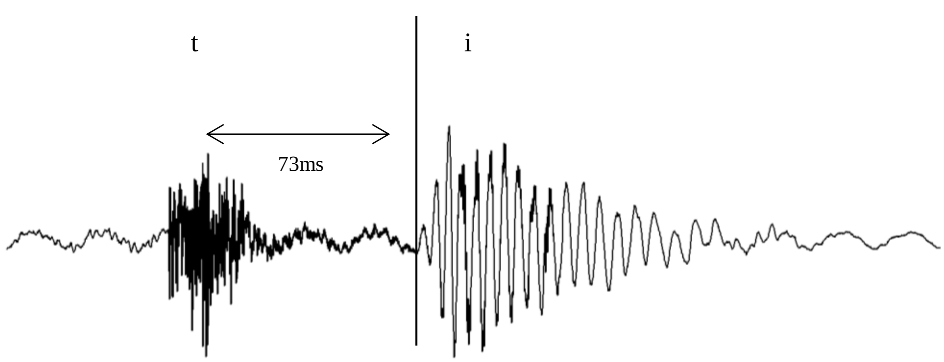


Figure 18: The release burst of the segment [t] and the following vowel [i], extracted from a single token of the lexical item t^hi!qɔ ‘God’.

Though token analyses of the lexical item [tumʔi] 'speak' exhibit considerable variation in aspiration length, the lexical item [tʰi!qɔ] 'God' provides evidence of consistent extended aspiration accompanying the segment [t]. A representative token of the first two segments of this lexical item are provided in figure 18 above. The section of the waveform demarcated as [t] indicates a considerable burst followed by a period of semi-regular variations. The VOT value of this segment is measured at 73ms which is longer than that expected for a regular voiceless segment as predicted by Ladefoged & Cho (1999). While the analysis of the tokens in figure 16 above is only useful for exhibiting the phonetic disparities in aspiration, the lexical item presented in figure 18 provides some evidence of the consistent utilization of extended aspiration. The speakers therefore appear to vary equally in their utilization of this aspiration continuum, yet tokens of the lexical item [tʰi!qɔ] 'God' provide evidence of some level of remaining control in the distribution of aspirated segments (for all tokens see Appendix H).

6.4.2 Ejectives

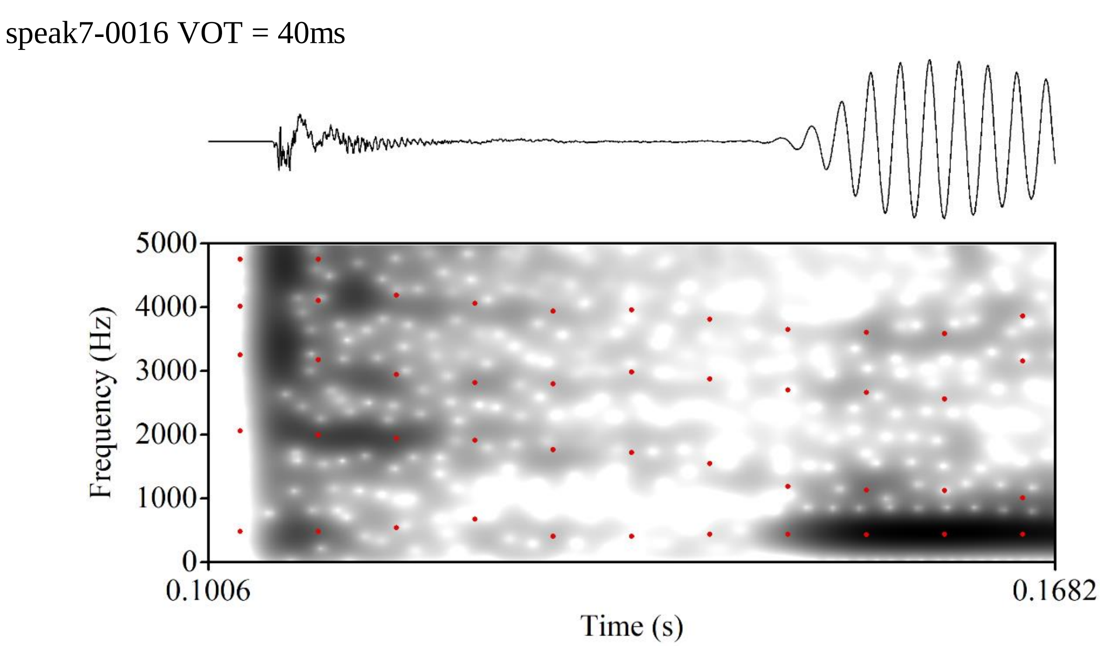


Figure 19: The release burst of the segment [t], extracted from a single token of the lexical item tumʔi 'speak'.

Variations of audible glottal releases accompanying the segments [m] and [t] are observed in Tumʔi. The most convincing acoustic representation of this is observed in tokens of the lexical item [tumʔi] 'speak'. Figure 19 above constitutes a representative token of the alveolar ejective [t']. The characterization of this segment as an ejective is based on the intensity and duration of the noise burst as well as the reduction to silence before the delayed onset of the vowel. The raised larynx in the articulation of the ejective causes a high degree of compression of the air in the oral cavity, which then results in a more intense burst at the release of the oral closure. The oral and glottal closures are produced simultaneously, however the release of the oral closure precedes that of the glottal closure during articulation. The extended period of silence following the release burst represents the interval between the oral and the glottal release. This is further exhibited in the spectrogram by the complete lack of energy between the noise burst and the following vowel as seen in figure 19 above.

The acoustic signal of the ejective is distinct from that of the aspirated plosives presented in figures 16, 17 and 18. Hence the initial alveolar segment of the lexical item [tumʔi] 'speak' alternates between different degrees of aspiration as well as airstream mechanism. The extensive variation observed across tokens of this single segment could be attributed to the chaotic nature of a dying language phonology. The phonological system may have been reduced to the extent that these distinctive features have only been retained for this particular segment and furthermore have lost any contrastive property.

6.4.3 Post-nasal glottal and glottalized stops

Speak2-0013 ‘tumʔi’

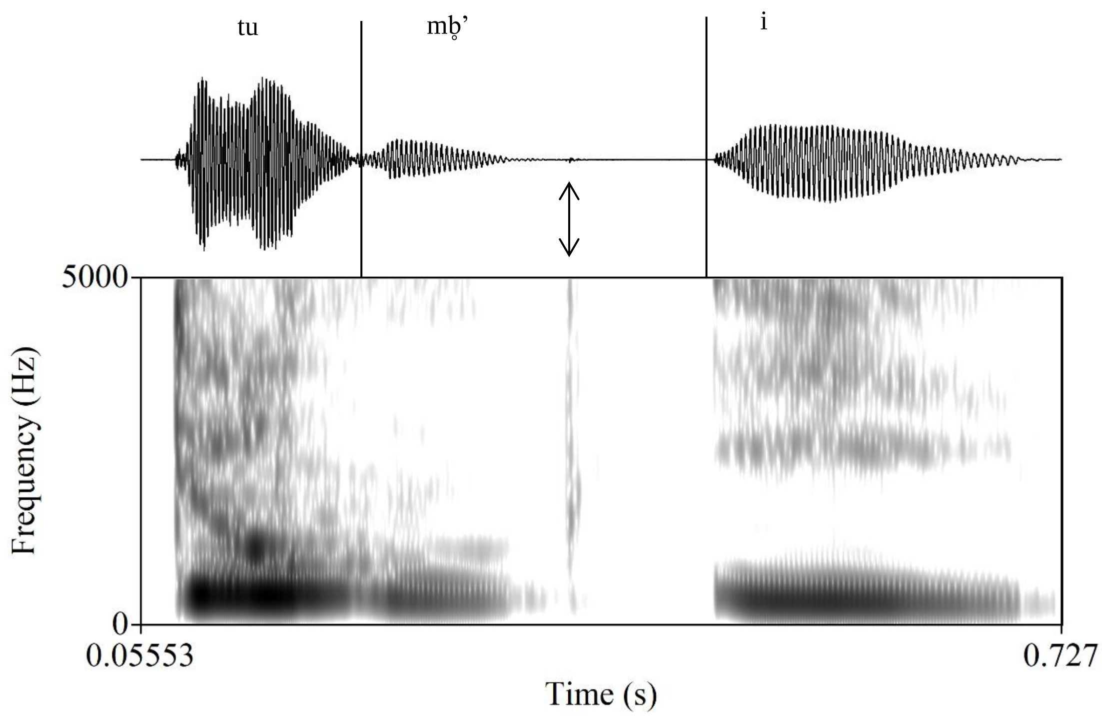


Figure 20: The voicing and release burst of the segment [m], extracted from token 2 Rec 0013, of the lexical item tumʔi ‘speak’.

The section of the waveform marked as the segment [m’] is characterized by a period of nasal voicing which diffuses into extended reverberation followed by a small noise burst and an extended period of silence. As expected, the spectrogram displays energy only across the lower frequencies during the articulation of the nasal. Following the nasal, the energy dissipates until the point at which the labial articulators separate producing a distinct release burst. The release of the oral closure is marked by a double-sided arrow connecting the noise burst in the waveform to the spike of energy displayed in the spectrogram. The bilabial stop produced at the separation of the articulators is comprised of the release burst as well as a period of silence preceding the voicing onset of the vowel.

Speak4-0016

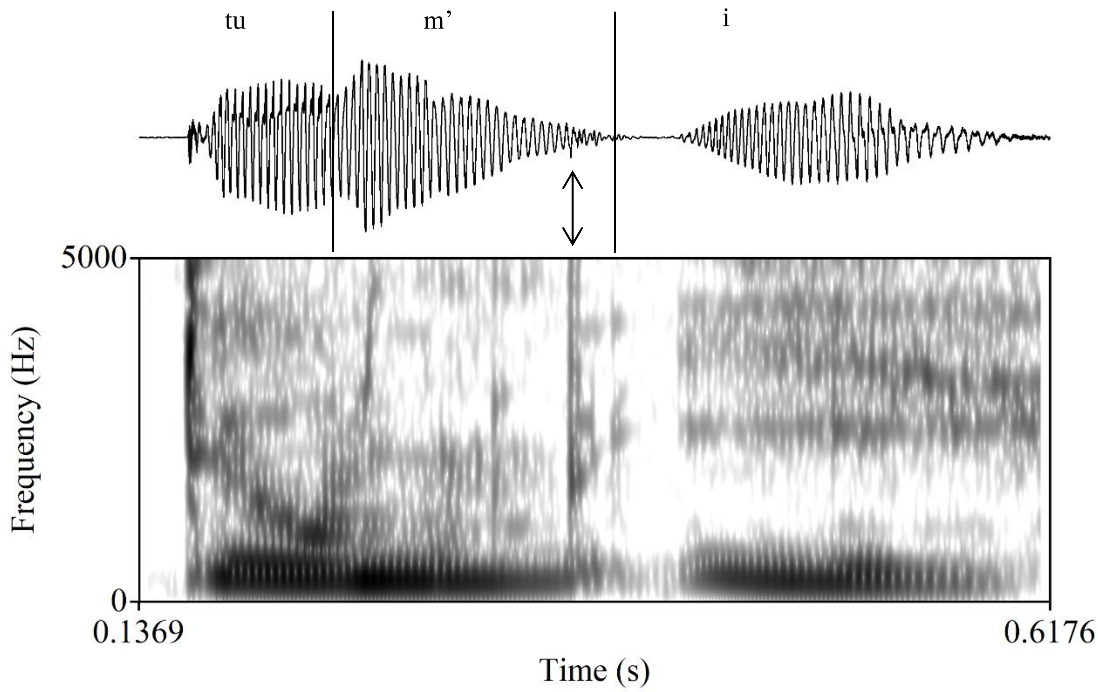


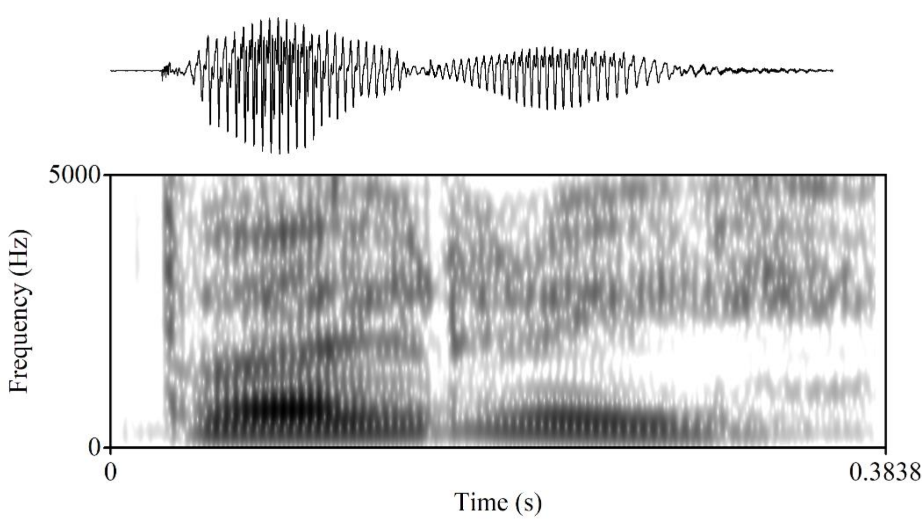
Figure 21: The voicing and release burst of the segment [m], extracted from token 4 Rec-0016, of the lexical item tumʔi ‘speak’.

The release burst in the segment above occurs 51ms after the onset of the nasal voicing, in this token the release occurs before the nasal voicing dissipates. Similar to the waveform in figure 20 the section demarcating the nasal segment presented in figure 21 above is characterized by a distinct noise burst following the nasal voicing. However, unlike the token discussed above, the noise burst occurs during the final damped oscillations of the preceding voicing. The noise burst in the second token is therefore less distinct, however as indicated at the position of the double-sided arrow, there is a discernable spike of energy following the nasal segment. Also observed in the token above is a period of silence following the final reverberation of the preceding stop burst. This period exhibiting no glottal constriction or airflow is characteristic of a glottal stop which may be produced simultaneously with the stop as the ejective segment discussed in section 6.3.2, or the glottal stop may be produced entirely separate from the preceding bilabial stop.

As is the case with most phonetic contrasts discussed in previous sections, there are tokens in which the segment [m] is followed by a period of silence without any identifiable burst in either the waveform or the spectrogram. Though most recorded utterances of this lexical item have been transcribed with a post-nasal stop segment, the issue of variation remains. However, considering that every token is transcribed with the nasal segment consistently followed by a glottal stop constituted by a distinct period of silence, one may argue that the absence of a noise burst in those instances may represent an obsolescing feature. Therefore, in certain instances, at the point which the bilabial stop release would have occurred there is now only a simple glottal stop following the nasal.

6.4.4 Voicing contrast

bread2-0013 ‘bəri’



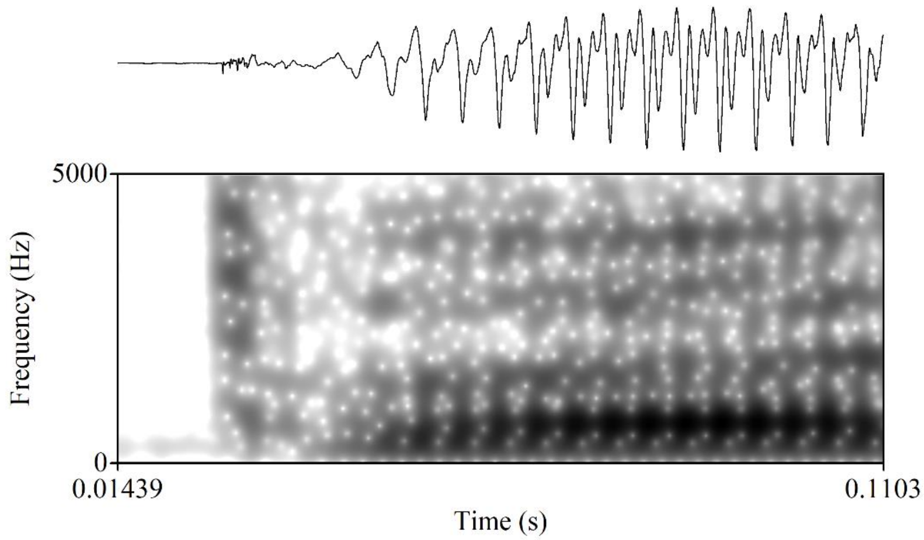


Figure 22: The release burst of the segment [b], extracted from one token of the lexical item bəri 'bread'.

Figures 22 & 23 constitute representative tokens of the segments [b] (figure 22), and [p] (figure 23) recorded for lexical items indexes 80 & 97 (see Appendix B). The closure release of the segment presented in figure 22 above indicates a clear noise burst both in the waveform and the spectrogram. The waveform displays a small transient followed by random variations in noise. The noise burst is indicated clearly in the spectrogram by a corresponding spike of intensive energy distributed up to 5000 Hz, extending from the level of the first formant (F1). The VOT value of this plosive is 13ms (for VOT values of all tokens see Appendix D), this VOT is therefore lengthier than expected for a voiced segment. The second waveform and spectrogram presented in figure 22 constitutes a smaller extract of the segment, including the segment under investigation and the following vowel [bə]. This slightly enhanced display provides a better indication of the energy distribution across the noise burst, extending from the lowest frequency level to the highest frequency. Though the spectrograms in figure 22 indicate a noise burst extending to the lower frequencies, the voice bar does not fully extend across the burst (Ladefoged & Johnson 2014, pg. 199). Hence, though this segment exhibits the audible impression of a voiced obstruent, analysis of the waveform and spectrogram indicate that this segment is not voiced but rather exhibits the characteristics of an unaspirated stop.

goat5-181207 ‘pəri’

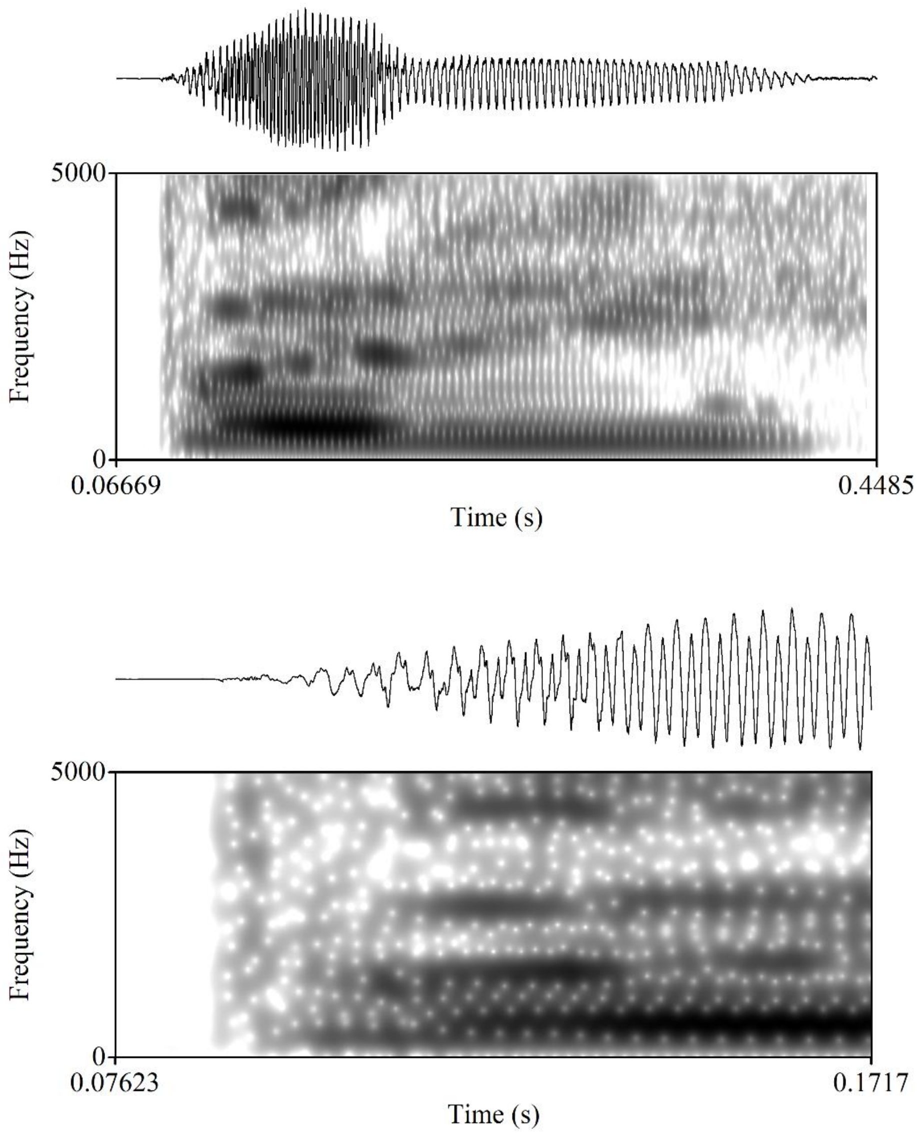


Figure 23: The release burst of the segment [p], extracted from one token of the lexical item pəri ‘goat’.

Presented in figure 23 above is the impressionistically voiceless counterpart of the bilabial plosive presented in figure 22. The waveform presented for the segment [p] in figure 23 reflects a weak transient relative to that presented in figure 22. The minimal visibility of the burst spike in the spectrogram further indicates the faint transient and low-level intensity produced at the release burst. The spectrogram not only indicates a faint spike, but the distribution of energy across this spike appears to be a lot more diffuse than that observed in figure 22. Considering the zoomed-in waveform and spectrogram presented in figure 23 consisting of the segments [pə], there is clearly less energy extending to the low frequencies of the burst spike. Hence, the noise burst produced at this closure release is visibly less intense than that of the segment presented in figure 22. The segment transcribed as voiceless does not however have the positive VOT value expected for aspirated voiceless segments (Ladefoged & Johnson 2014, pg. 152). The VOT value of this segment extracted from token goat5-181207 is 22ms (for all VOT values see Appendix D). Though this VOT value is not substantially positive, collective analyses of VOTs across voiceless segments in multiple languages indicate a VOT range for bilabials beginning as low as 10ms (Cho & Ladefoged 1999, pg. 219).

Overall, these segments exhibit an audible contrast which is perceived as a difference in voicing. However, the acoustic analysis does not indicate contrasts characteristic of a voicing distinction. Though the VOT values presented in Appendix D are generally lower across tokens of the voiced segments, the calculated means are not significantly different. Furthermore, the tableau includes outlying segments which cannot be excluded from the analysis but undeniably skew the overall comparison. The complete analyses of the noise bursts and VOT of the representative tokens generally provide evidence of a contrast in intensity rather than voicing. Furthermore, the fact that these segments are a minimal pair provides the evidence necessary to substantiate a phonemic distinction.

expression of distance3-0015 ‘duːr’

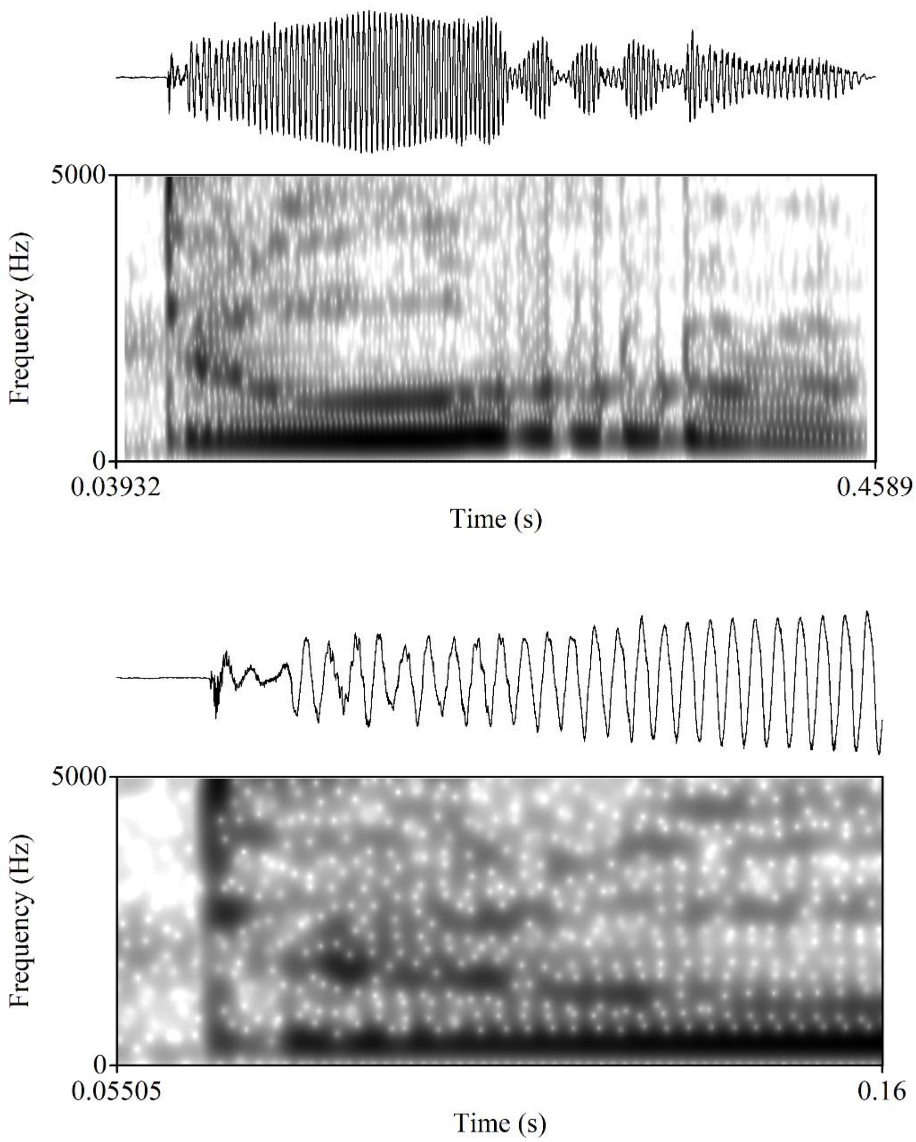


Figure 24: The release burst of the segment [d], extracted from one token of the lexical item duːr ‘expression of distance’.

crazy2-0013 ‘tɔrəŋtɔrəŋ’

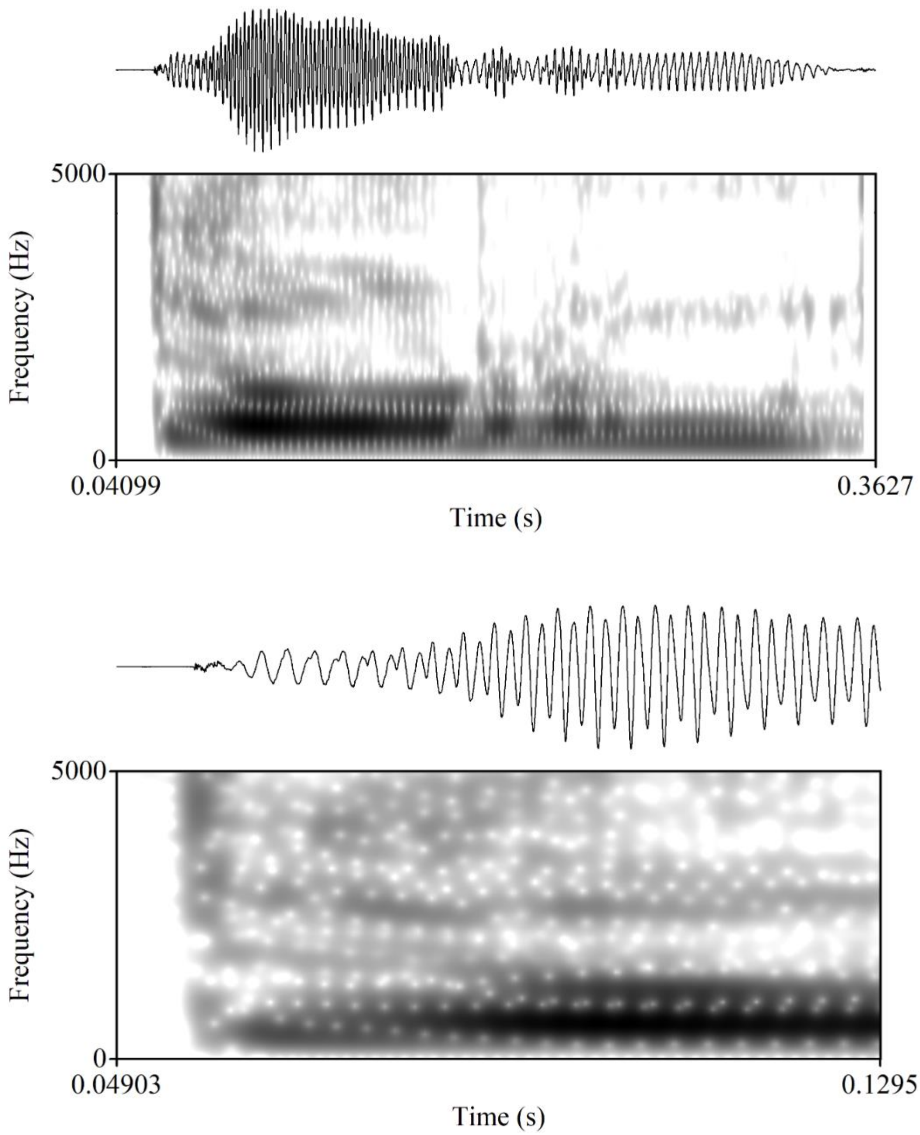


Figure 25: The release burst of the segment [t], extracted from a single token of the lexical item tɔrəŋtɔrəŋ ‘crazy’.

Presented in figures 24 & 25 above is a second contrast which audibly resembles a voicing distinction. Figure 24 constitutes a representative token of the alveolar plosive perceived as [d], the counterpart [t] which is perceived as voiceless is also represented by a single token in figure 25. Similar to the contrast discussed above, there is a disparity in the intensity of the noise burst between the segments presented in figures 24 & 25. The waveform presented in figure 24 indicates a relatively intense transient almost immediately followed by the onset of voicing of the following vowel. The spectrograms presented in figure 24 display a distinct burst spike extending across the lower and higher frequencies, indicating a louder burst. Furthermore, the noise burst displayed in the waveform presented in figure 25 differs from that of the segment in figure 24; the segment in figure 25 displays a substantially shorter transient. The spectrograms in figure 25 indicate less intensity relative to that displayed in the spectrogram of figure 24. The noise bursts in the spectrograms of figure 25 indicate a substantial degree of noise distributed between 2500 kHz-5000 Hz as expected for the alveolar plosive, while the noise burst of the alveolar in figure 24 exhibits substantial energy across the lower frequencies as well as above.

Due to limitations of the data set there are fewer tokens of each of these lexical items. The number of tokens captured for each lexical item is also uneven and therefore cannot be presented comparatively in a table as provided for the tokens of [pəri] ‘goat’ and [bəri] ‘bread’. However, for the purpose of transparency I report the VOT values of each segment; the VOT of the perceived voiced stop presented in figure 24 is measured at 8ms, while the VOT of the perceived voiceless segment presented in figure 25 is measured at 13ms. The difference in VOT value between the voiced and voiceless alveolar segments is greater than the bilabial segments presented in figures 22 & 23 but is not substantial enough to distinguish between either segment set. Hence, again it appears that these segments are not contrastive in their VOTs, and instead exhibit a substantial contrast in intensity. Furthermore, unlike the segments presented in figures 22 & 23 above, these lexical items do not constitute a minimal pair, therefore any proposed distinction cannot be defined as phonemically contrastive. As noted above in section 6.3.1, the lexical item [du:r] ‘expression of distance' is borrowed from Afrikaans. As a result, the intensity distinction observed in figure 24 cannot unambiguously be attributed to the phonetic typology of Tumʔi.

6.5 Discussion

The aspiration contrast recorded in Tumʔi is particularly interesting as it does not appear to be pertinent to distinctions in voicing; this is clear in the similarity of VOT values across multiple tokens of segments perceived to exhibit a voicing contrast. The most substantial differences in aspiration are observed across instances of the voiceless alveolar segment [t]. As displayed in figure 16, section 6.4.1, the period of aspiration following the plosive [t] may range from 3ms – 105ms depending on the token as well as the lexical item. Based on the analyses of these tokens as well as the transcription of the collected data, there appear to be three phonetic contrasts in aspiration. The segment with a measured VOT value of 3ms is consistent with the production of an unaspirated segment, indicating that the intention of the speaker is likely to produce an unaspirated voiceless segment. As voiceless segments with this miniscule VOT value are frequently observed across the data set, it is possible that the unmarked production of the segment [t] in Tumʔi is in fact unaspirated. The production of the aspirated [t^h] however is in some ways more controlled, observed in varying degree across tokens of the lexical item [tumʔi] 'speak', and consistently across tokens of the lexical item [t^hi!qo] 'God' as seen in figure 18 section 6.4.1. Hence, the aspirated plosive [t^h] likely reflects a phonemic contrast within the language.

The analysis in section 6.4.3 displays evidence of a nasal segment followed by what appears to be a voiceless bilabial ejective in certain instances, or a simple glottal stop. The description of this sound is discussed in section 6.4.3 above, however a definitive transcription has not been decided. All tokens of the nasal segment in the lexical item [tumʔi] 'speak' are either produced with a voiceless bilabial ejective, or with a glottal stop. The evidence suggests that the nasal segment precedes the bilabial ejective to form a complex sequence, yet due to a reduced phonological inventory this sequence is sometimes realized as a nasal segment followed by a simple glottal stop. Though minimal, the evidence of an alveolar ejective in section 6.4.2 provides some validation for the existence of the complex sequence described in section 6.4.3. As discussed in section 6.2, sounds produced using the glottalic airstream mechanism are mostly attested in Tuu languages such as N|u (Güldemann 2006, pg. 11), ǀXòõ, and ||Xegwi among others (Vossen 2013). Hence, the attested use of ejective segments in this language would increase the number of typological features that complexify the phonological system of Tumʔi, aligning it more closely to the Tuu lineage. The presence of these complex sounds in the inventory of Tumʔi increase its typological resemblance to Khoisan lineages like Tuu which encompass languages with more complex sound systems.

The perceived voicing contrast between bilabial segments [b] and [p], and alveolar segments [d] and [t], in Tumʔi is demonstrated in the phonetic analyses in section 6.4.4 above. However, only the bilabial segments are considered phonemically contrastive in the language, as these segments are distinguished across a minimal pair comparison. The lenis fortis contrast is further evidenced in the phonetic analysis of the alveolar segments which indicates that speakers of Tumʔi are able to recognize and reproduce the distinction. The fact that the segment transcribed as [d] is extracted from a borrowed lexical item however constrains the applicability of this segmental feature to alveolar segments within the typology of the Tumʔi. The analyses of the representative tokens indicate general variation across the aspiration periods and VOT of the contrasted segments. A comparison of the mean values calculated for two sets of 9 tokens indicate only a 5ms difference in VOT (for VOT table see Appendix D). Though the VOT values reported for the segments transcribed as voiceless are greater, the distinction cannot be defined by the measure of aspiration, and hence may not actually reflect a voicing contrast. Overall the contrasts recorded across both the bilabial and alveolar segments are characterised by the distribution of energy displayed in the spectrogram, specifically the intensity of the energy distributed across the burst spike.

Considering these analyses are based mainly on representative tokens, the only definitive phonemic contrast proposed here is between the bilabial segments which clearly constitute a minimal pair, as seen in section 6.4.4, figures 22 & 23. Though this contrast does not exhibit the acoustic characteristics of a voicing distinction it contributes to the complexity of the consonant system in a manner that is typologically interesting. This contrast in intensity may constitute an entirely unique feature in the same way [+voice] is considered an additional feature in the consonant systems of Khoisan languages, as it is generally attested in more complex sound systems such as those reported for Tuu languages (Güldemann 2006; Vossen 2013, pg. 153).

Chapter 7: Conclusion

The various phonetic analyses presented above come together to form a complex sound inventory consisting of unique features generally unattested in Khoisan varieties with less typologically complex phonological systems. The vowel system consists of six monophthongs and five diphthongs, excluding indications of possible voice quality distinctions discussed in chapter 4. The monophthong system is comprised of the five basic high, low, and mid vowels. However, the disparity in openness between the corresponding mid vowels causes an irregularity, which results in an asymmetric quadrilateral vowel space. This observation introduces the question of whether a phonemic contrast in openness had previously existed as part of the vowel system. That would suggest that the vowel inventory of Tumʔi consisted of seven vowels, which would align with many Bantu varieties and a select number of Khoisan varieties, generally not within the Khoekhoe family. A similar distinction is introduced with the identification of diphthongs, which appear frequently throughout the data set. ‘True diphthongs’ are generally unattested in Khoekhoe languages, except in instances of rapid speech (Beach 1938). The ample observation of monosyllabic unalike VV sequences in Tumʔi might provide further indication that the phonological typology of this language is quite distinct from those classified within the Khoekhoe lineage. Finally, based on the phonetic analysis presented in chapter 4, phonation does not constitute a distinctive feature in Tumʔi at present. However, as discussed in chapter 4, the statistical evidence of a non-modal voice quality distinction might be a remnant of a more complex phonological system.

The uniqueness of the sound system of Tumʔi is further indicated by the observation of phonemic laryngeal contrasts. While phonemic distinctions in aspiration have undergone fricative lenition in most varieties of Khoekhoe and even some of the Tuu family, the speakers of Tumʔi have retained or re-gained a controlled use of at least one aspirated segment. Hence, the sound system of Tumʔi exhibits a feature no longer attested in any of the Khoekhoe languages with the exception of !Ora. Another laryngeal distinction discovered in Tumʔi is the use of the glottalic egressive airstream mechanism. As discussed in chapter 6, the initial alveolar stop of the lexical item tumʔi ‘speak’ may be produced as an ejective, and the bilabial nasal segment may be followed by an ejective, as well. The evidence discussed in chapter 6 suggests that ejectives are phonemically contrastive in the sound system of Tumʔi. As is the case with a few of the other distinctive features identified above, the ejective is only attested in Khoisan varieties with more complex phonological typologies. Overall, the egressive consonant inventory observed in Tumʔi may be described as extensive relative to other Khoisan varieties. The consonant inventory presented in chapter 3, section 3.2, includes a phonemic voicing distinction, and a total of 20 segments.

The same disparity is clear across the different click inventories of the varieties within the different lineages, but to a higher degree. Though certain varieties within the same family have substantially larger click inventories than others, languages of the Tuu family generally encompass larger click inventories. Finally, the click system observed in Tumʔi possibly contributes the most complexity to the sound system, with the identification of the uvular click accompaniment. As discussed in detail in chapter 5, this particular click accompaniment is unique to Tuu varieties, including N|uu and !Xõó (Miller et al 2007). Hence, this unique click efflux provides the most obvious indication of a typological similarity to particular Khoisan varieties, including among others !Xóõ, N|uu, ǂHoan, and |Auni, which all form part of the Tuu lineage. Interestingly this accompaniment constitutes one of the two click accompaniments most frequently observed in Tumʔi, the other being the nasal accompaniment. In terms of attested accompaniments, cross-linguistically the nasal accompaniment is the most common. Though various other analyses of Khoisan varieties have identified a voicing contrast within this efflux type, this is not attested in Tumʔi. Furthermore, unlike the uvular stop accompaniment, the nasal accompaniment is not typologically unusual or remarkable. Overall, the click inventory is comprised of a total of 15 clicks, including the four click types and accompaniments. Considering the size of click inventories in Khoisan languages such as !Xõó and ||Xegwi, an inventory of 15 clicks would appear relatively small. However, the distribution of these clicks across the data set is extensive, which indicate that clicks may have carried a heavy functional load in this Khoisan variety. There are few words in the lexical inventory of Tumʔi which do not include a click segment. Hence, based on observations across the data set, clicks in Tumʔi represent the most important consonant-like segments in the sound system. Setting aside consequences of language contact, this is a feature more commonly associated with the Tuu lineage as opposed to Khoekhoe.

While arguments have been made for the untenability of Khoisan language family distinctions and lineages, comparative research provides evidence of linguistic relations which cannot be ignored. The investigation of the phonological typology of this Khoisan variety was initiated based on evidence of strong lexical correspondences and cognates. The resultant phonological inventory and main findings are useful for filling the typological gaps within the areal isoglosses. In conclusion, the analyses and discussion presented in the paper above indicates that the Khoisan variety now referred to as Tumʔi is one which shares a considerable number of true cognates with languages of the Tuu lineage as exhibited in Appendix A. Furthermore, similar to southern Khoisan varieties including those within the Tuu lineage, Tumʔi exhibits remnant features of what would be described as a typologically complex sound system. The typological similarities and differences are broadly summarised in the table below.

Table 10: *Comparative summary of phonological characteristics*

Phonological characteristics	Tumʔi	Tuu	Khoekhoe
Phonation	Yes	Yes	No
Uvular click accompaniment	Yes	Yes	Attested in one variety
Aspiration	Yes	Yes	Undergone affricate lenition
Ejectives	Yes	Yes	No

References:

Anders, H., 1935. A note on a South Eastern Bushman dialect. Zeitschriftfur Eingeborenen-Sprachen 25 (2): 81-89.

Beach, D., 1938. The Phonetics of the Hottentot Language. Cambridge: University Press.

Bleek, D. F., 1928-30. Bushman grammar: a grammatical sketch of the language of the |xam-ka-!k'e. Zeitschrift für Eingeborenen-Sprachen 19: 81-98/ 20: 161-174.

Branch, W.R., 2001. A photographic guide to snakes and other reptiles of Southern Africa. Cape Town: Struik Publishers.

Chebanne, A., 2012. 'Where Are the Skeletons of Dead Khoisan Languages?' Botswana Notes and Records 44: 81-92.

Cho, T. & Ladefoged, P., 1999. Variation and universals in VOT: evidence from 18 languages. Journal of Phonetics, 27, 207–229.

Childs, G.T., 2003. An introduction to African Languages. Amsterdam, The Netherlands/ Philadelphia, PA: John Benjamins.

Cook, E.D., 1989. "Is phonology going haywire in dying languages? Phonological variations in Chipewyan and Sarcee." Language in Society 18(2): 235–255.

Dorian, N. C., 1977. "The problem of the semi-speaker in language death". Linguistics, 191, 23-32.

Du Plessis, H., 2003. Afrikaans and Khoe: Languages in contact. In K.K. Prah (Ed.). Silenced voices: Studies of minority languages of Southern Africa, (pp. 129-139). Cape Town: Center for Advanced Studies of African Society.

Du Plessis, M., 2014. The damaging effects of romantic mythopoeia on Khoesan linguistics. Critical Arts, 28(3): 569–592.

Elderkin, E.D., 1992. Predictable nasality before East African clicks. Afrikanistische Arbeitspapiere 29: 111-129.

Garellek, M., 2019. Acoustic Discriminability of the Complex Phonation System in !Xóõ. Phonetica, pp.1-30.

Gordon, M. & Ladefoged P., 2001. Phonation types: A cross-linguistic overview. Journal of Phonetics 29: 383–406.

Grinevald, C., 2003. Speakers and documentation of endangered languages. In Peter Austin (ed.), Language documentation and description, volume 1: 52–72. London: Hans Rausing Endangered Languages Project.

Grinevald, C., 2007. Linguistic fieldwork among speakers of endangered languages. In O. Miyaoka, O. Sakiyama and M.E. Krauss (eds.), The Vanishing Languages of the Pacific Rim. Oxford and New York: Oxford University Press, pp. 35-76.

Güldemann, T., 2001. Phonological regularities of consonant systems across Khoisan. Language and Literature Series 16, University of Leipzig Papers on Africa. Universität Leipzig.

Güldemann, T. & Elderkin. E.D., 2003. On external genealogical relationships of the Khoe family. In Brenzinger, Matthias and Christa König (eds.), Khoisan languages and linguistics: the Riezlern symposium 2003. Quellen zur Khoisan-Forschung 17. Köln: Rüdiger Köppe.

Güldemann, T., 2006. The San languages of southern Namibia: linguistic appraisal with special reference to J. G. Krönlein's N|uusaa data. Anthropological Linguistics 48(4): 369-395.

Güldemann, T., 2006. Structural isoglosses between Khoekhoe and Tuu: The Cape as a linguistic area. In Yaron Matras, April McMahon & Nigel Vincent (eds.), Linguistic areas: Convergence and language change (Annual Conference of the North West Centre for Linguistics, November 2002; Manchester, England), 99–134. Basingstoke: Palgrave Macmillan.

Güldemann, T. & Loughnane, R., 2012. Are there "Khoisan" roots in body-part vocabulary? On linguistic inheritance and contact in the Kalahari Basin. Language Dynamics and Change 2: 215–258.

Hastings, R., 2001. Evidence for the genetic unity of Southern Khoesan. Khoisan: Syntax, Phonetics, Phonology, and Contact. ed. A. Bell & P. Washburn. Cornell Working Papers in Linguistics, 18: 225-246.

Huffman, M. K., 1987. Measures of phonation type in Hmong, Journal of the Acoustical Society of America, 81: 495-504

Iwu, M.M., 2014. Handbook of African medicinal plants. CRC press.

Johnson, K., 2008. Quantitative Methods in Linguistics. Oxford: Blackwell.

Keating, P., Garellek, M. & Kreiman, J., 2015. Acoustic properties of different kinds of creaky voice, in 'Proceedings of the 18th International Congress of Phonetic Sciences', Glasgow.

Killian, D., 2009. Khoemana and the Griqua: Identity at the heart of phonological attrition. Helsinki: University of Helsinki dissertation.

Ladefoged, P., 1992. Another view of endangered languages. Language 68: 809-811.

Ladefoged, P. & Maddieson, I., 1996. The sounds of the world's languages. Oxford: Blackwell.

Ladefoged, P. & Traill, A., 1984. Linguistic phonetic description of clicks. Language 60: 1–20.

Ladefoged, P. & Traill, A., 1994. Clicks and their accompaniments. Journal of Phonetics 22: 33–64.

Maddieson, I., Ladefoged, P. & Sands, B., 1999. Clicks in East African languages. In Rosalie Finlayson (ed.), African mosaic: Festschrift for J. A. Louw, 59–91. Pretoria: University of South Africa.

Miller, A., 2007. Guttural vowels and guttural coarticulation in Ju|'hoansi. Journal of Phonetics 35: 56–84.

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

Miller, A., 2011. The representation of clicks. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) The Blackwell companion to phonology. Malden, Mass.: Wiley-Blackwell. 416–439.

Nakagawa, H., 2006. Aspects of the phonetic and phonological structure of the G|ui language. Ph.D. dissertation, University of the Witwatersrand, Johan.

Petrollino, S. & M. Mous., 2010. Recollecting words and expressions in Aasá, a dead language in Tanzania. Anthropological Linguistics 52,2: 206–216.

Picker, M., Griffiths, C. & Weaving, A., 2002. Field Guide to Insects of South Africa. Struik Publishers, Cape Town, South Africa.

Sands, B., Miller, A. & Brugman, J., 2007. The lexicon in language attrition: The case of N|uu. In Doris L. Payne & Jaime Pena (eds.), 37th Annual Conference on African Linguistics, 55– 65.

Somerville, MA: Cascadilla Proceedings Project. http://www.lingref.com/cpp/acal/37/index.html (27 April 2009).

Starostin, G., 2008. From modern Khoisan languages to ProtoKhoisan: The value of intermediate reconstructions. Aspects Comparative Linguistics 3: 337–470.

The Lloyd & Bleek |xam Dictionary - The Digital Bleek and Lloyd lloydbleekcollection.cs.uct.ac.za/dictionaries/services/www/xamdict/

Van Wyk, B. & Van Wyk, P., 1997. Field guide to Trees of Southern Africa. Struik, South Africa.

Vossen, R., 2013. The Khoisan languages. Routledge Language Family Series. London: Routledge.

Westphal, J., 1971. The click languages of Southern and Eastern Africa. In Sebeok, Thomas A. (ed.), Linguistics in Sub-Saharan Africa. Current Trends in Linguistics 7. The Hague/ Paris: Mouton, 367-420.

APPENDICES:

Appendix A

	Tuu					Khoekhoe	
Gloss	|Xam (Bleek & Lloyd Dic.) (Guldemann 2005 & 2006)	|Auni (Hastings 2001)	N|u (Miller et al. 2009)	N|uusaa (Guldemann 2006)	Tumʔi	Nama (Beach 1938, Killian 2009)	!Ora (Beach 1938; Killian 2009)

woman	!kui tait	ǀkẽ/ǀʌn		ǀ'ati	!qoita	Taras	taras
man	!kui		ǂoo	!ui	ǀqui	khoep	khoep
child	!khwãa	ʘpwa			ǀqxwa̰		õa'i
drink	kwa	k'a:a	kx'ai	kx'ũ	ǀqxwa̰	aa	kx'a
speak	tanʔi/ ǂagen		cu	ǂagen	tumʔi	ǂxən	koba
listen	tum-i	tu:ho			tumʔi		komsen
knife	!gwara	gõä	nǂona		gwa̰ra		kõas
rest/sleep	ten	ʘpwa	ʘun	ʘuin	tḛŋ	nǁau!a	ǁ'om
Leave/walk	ta̰ḭ	tãi	!ai	ǁai/ǁ'ai	ta̰ḭ		!ũ
run	!'uuxe			!'uuxai	ǀquxai		!huekx'ãi
bread	bori/bere		peresi		bəri	pere	bereb
eat	ãa	ã	aa	hã	ha̤a̤	ǂu	ǂ'ũ
handywork	tɑba	ǁkari			taβa	tava	
head	nǀa	nǀa:	nǀa	nǀã	nǀã	Tanas	bi !'ap
eyes	taxm/ts'axaiten	ts'a:-xu	ts'əxəm	ts'axen	xaikən	muku	mũb
nose	nǀu-ru	ǀnu:/nǀõ		nǀudu	nǀuɹu		ǂuib
stomach	ǀkoa	ǂke:			ǀgḛḭ		n!aab
beard	nǀum	nǀum			n!uku	nǀumbi	nǀum

Appendix B

Key: Click type

Dental = Den

Alveolar = Alv

Lateral = Lat

Palatal = Pal

Key: Accompaniment

Uvular = Uvu

Nasal = Nas

Velar = Vel

Voiced = Voi

Index	Lexical inventory	Influx & Efflux	Gloss
	Click initial		Total = 63
1	ǀœʊ	Den	get water
2	!oitəs	Alv	Laugh
3	!wa̰ra̰	Alv	Tease
4	ǀɔrise	Den	policeman
5	!ɔrise	Alv	how are you?
6	!uŋka/ !uŋkaha̤	Alv	play dead
7	!ukən	Alv	punish
8	!ɔŋgas	Alv	big container
9	ǁap/ gap	Lat	veld food (recognized by thorns)
10	!uku	Alv	irritable/ argumentative
11	!ukuxãã	Alv	pregnant/ full belly
12	!eipsexat	Alv	medicine for child’s stomach

13	!anirki	Alv	jackals’ knee
14	ǀaiki/ ǁaiki	Den	bat-eared fox
15	ǂamku	Pal	happy heart
16	!uŋka	Alv	medicinal powder
17	!uŋkaha̰a̰	Alv	grind/ mix
18	ǀeina	Den	ouch
19	ǂama:kʰu/ ǂxama:ku	Pal	thank you
20	ǂamaku	Pal	medicine
21	ǁɔmi/ ǁxɔmi	Lat Vel	pregnant/ full figure
22	ǀquxai / ǁquxai	D/L Uvu	leave-death
23	!qoita	Alv Uvu	woman
24	|qui (singular) |quis (plural)	Den Uvu	man
25	ǀqxwa̰/ gwa̰	Den Uvu	baby
26	ǀqəri/ |gəri	Den Uvu	honey beer
27	!qam/ !qami/ kam	Alv Uvu	pee
28	!qɔwa / !qɔva	Alv Uvu	medicinal plant
29	ǀqoi/ |qo̰i/ |qwai/ |xa̰i	Den Uvu	pregnant
30	!qɔra	Alv Uvu	knife/something sharp
31	ǂqwara/ ǀqarɔk/ gwa̰ra	Pal Uvu	knife
32	ǀqwaka	Den Uvu	stink
33	ǀqœʊ	Den Uvu	tell
34	ǀqœʊ/ ǂqœʊ	D/P Uvu	youngest
35	!quruam	Alv Uvu	tasty

36	ǀg̰ḛḭ/ ǀqḛḭ/	Den Voi	sheep
37	ǀqḛḭ	Den Uvu	people
38	ǀg̰ḛḭ	Den Uvu	stomach
39	ǂqam	Pal Uvu	grip
40	ǀqɔm/ !xɔm	D/A Uvu	tenderize
41	ǀqɔmər/ !xɔmər	D/A Uvu	tenderizer
42	ǀquha̰a̰	Den Uvu	weak bladder
43	ǀqara/ tara	Den Uvu	word
44	ǀgams	Den Voi	sickness/ STD
45	ǀgeip	Den Voi	skunk
46	ǀqwa/ ǀqwa̰i/ !qup	D/A Uvu	female genitals
47	!qxu	Alv Uvu	pee-type
48	ǀqxa	Den Uvu	leave it
49	nǀu (singular) nǀuns (plural)	Den Nas	white man
50	nǀu	Den Nas	red stone
51	nǀa/ nǀã	Den Nas	head
52	nǀuɹu	Pal Nas	nose
53	nǂuɹuku	Pal Nas	snout
54	nǀaski	Den Nas	meat
55	n!oɹo/ n!oɹa	Alv Nas	drunk
56	nǀara/ nǀari	Den Nas	comfort
57	nǀoi boom	Den Nas	tree with yellow flowers
58	nǀeitɔ	Den Nas	word used to calm a baby

59	nǀeitœʊ/ nǀeitɔ	Den Nas	nightmare
60	n!ukukwa/ n!ugukwa	Alv Nas	wipe clean
61	n!ɔvɔ	Alv Nas	sugar snack
62	ǀxei	Den Vel	give birth
63	!xara	Den Vel	female genitals-type
	Click initial & click medial		Total = 9
64	ǀuŋǀwa/ !uŋ!a/ !uŋka	D/A Nas	big bum
65	ǂinǀama/ ǀqeinǀqama	P/D Nas	naughty
66	ǂiɹiǁuxa	Pal-Lat	jackal
67	ǀqeiǀkara	Den Uvu	sheep-type
68	!qum!qum/ ǀqumǀqum	A/D Uvu	coffee
69	n!œ̤ʊ̤!a/ n!œʊ!wa	Alv Nas	understand
70	!qara!quru	Alv Uvu	swear word
71	ǀqamǀqu/ ǀqamku	Den Uvu	tasty-type
72	nǀa!qoi/ n!uku/ na!oi	D/A Nas	look
	Consonant initial		Total = 39
73	tumʔi		speak/understand
74	ha̰a̰/ ha̤a̤		eat
75	ha̰a̰/ ha̤a̤		food
76	xaikən (singular) saigəns (plural)		eye
77	saigəns		face
78	tḛŋ		rest/ sleep
79	ta̰ḭ		leave-type

80	bəri/ bə̤ri (singular) brə̤kə̤ti (plural)		bread
81	tʰa̰va/ tʰa̰βa/ tawa		handywork
82	tʰa̤ŋa̤/ tʰaŋa/ taŋa		pain
83	mafuta/ məfuta		oil/ fat
84	kut͡ʃaka/ kutjaka		go out
85	bip/ dip		milk
86	tɔrəŋ tɔrəŋ		crazy
87	tʰaŋa		crazy-type
88	surte		give
89	xumxama/ xumama		meat-type
90	xorkies		smaller type veld food
91	xəmi		dig/ hide
92	xəmi		grind-type
93	xuma		be quiet/ stay silent
94	xɔm		hardened sap
95	tumtum		big toe
96	kuxa		baby jackal
97	pəri		goat/ buck
98	para		donkey
99	vələ als		medicinal ingredient
100	vələ kiər		medicinal ingredient
101	duːɹ		expression of distance
102	xɔra xat		waterhole

103	paka		bury
104	trul		hair
105	bala/ tɹəl		male genitals
106	məkəs		inner thighs
107	xana		a type of weed
108	kama		dagga/ marijuana
109	xɔnjas / tɔŋgas		eye dirt
110	kaiəŋs		fatty parts of the animal
111	skroi		burn
	Consonant initial & click medial		Total = 8
112	sein|ama/ seinkama	Den Nas	sugar water
113	sein‡ama	Pal Nas	crush
114	sumn|um/ sumn!um	D/A Nas	chew
115	tʰi!qo	Alv Uvu	god
116	sein|eŋ	Den Nas	jail
117	nuku|ən	Den	whisper
118	nuku!wa	Alv	get ready
119	sin|e/ si|e	Den Nas	cut
	Vowel initial		Total = 4
120	ixwa		now/truth
121	asa		see
122	œʊtəre		ask
123	eisevarkmag		ingredient for hotnots powder

	Vowel initial & click medial		Total = 2
124	ukuǁən / hukuǁən	Lat	I don't know
125	aritamsumnǀa	Den Nas	big thank you

Appendix C

Index	**Fixed phrase inventory**	**Influx & Efflux**	**Gloss**
1	ǀqui asa	Den Uvu	here comes the man
2	ǂuku !aɹa !ui asa	Pal-Alv	careful here comes the white man
3	n!uku !wara	Alv	look the police
4	!ɔri sese:se / !ɔrise asa / !oɹisa asa	Alv	here comes the police
5	nǀu aːsa	Den Nas	white man coming
6	eit xətɔf		dressed up (male)
7	eit xəwails		dressed up (female)
8	daai !qoita is !oris	Alv Uvu	that woman is pregnant
9	œʊtəre xwa		ask for something to drink
10	œʊrərɛxəha̤a		ask for something to eat

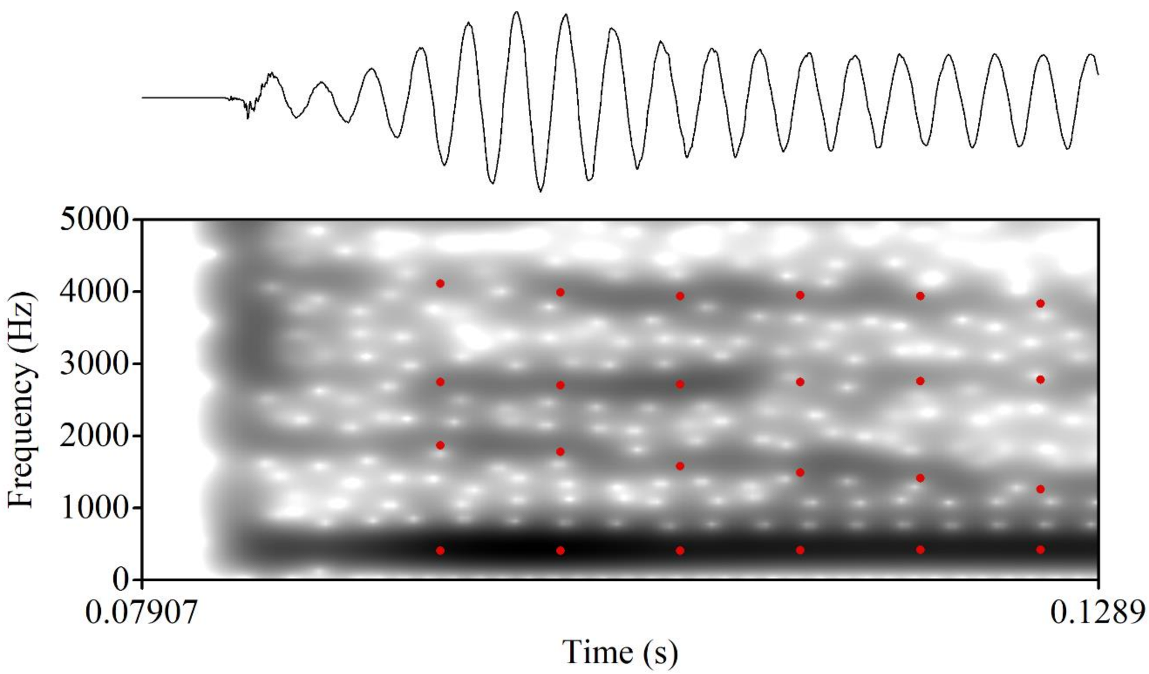
Appendix D

Voiceless token:	VOT	Voiced token:	VOT

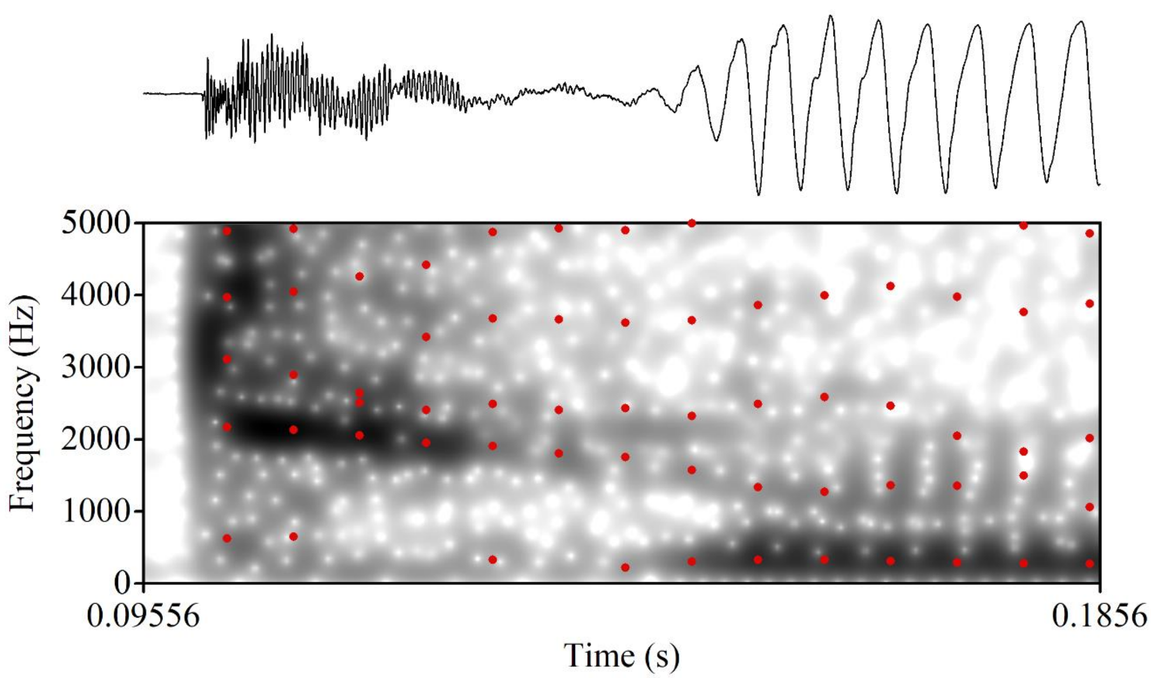
goat-0015	15ms	bread2-0013	12ms
goat-181206	12ms	bread3-0013	13ms
goat-181207	18ms	bread-0013	11ms
goat3-181207	18ms	bread1-0013	11ms
goat4-181207	14ms	bread4-0013	16ms
goat5-181207	22ms	bread2-181209	14ms
goat6-181207	25ms	bread3-181209	30ms
goat7-181207	28ms	bread4-181209	14ms
goat-181207	21ms	bread7-181209	9ms
	Mean = 19.2		Mean = 14.4

Appendix E

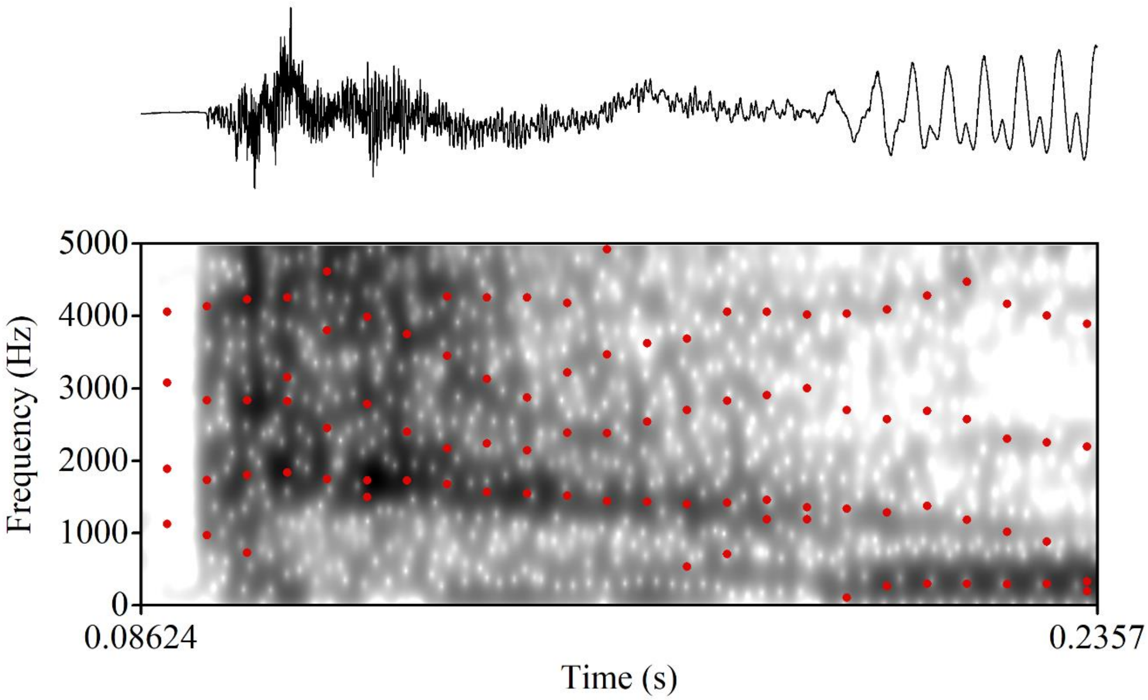
speak7-0013 VOT = 3ms



speak5-0016 VOT = 45ms



speak6-0016 VOT = 105ms



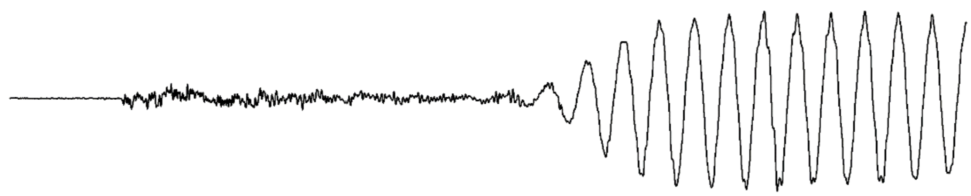
Appendix F

Index	Token:	VOT		Token:	VOT
1	speak-0013	12ms	22	speak-181207	54ms
2	speak2-0013	7ms	23	speak2-181207	68ms
3	speak3-0013	6ms	24	speak3-181207	80ms
4	speak4-0013	15ms	25	speak4-181207	51ms
5	speak5-0013	41ms	26	speak5-181207	69ms
6	speak6-0013	20ms	27	speak6-181207	35ms
7	speak7-0013	3ms	28	speak7-181207	54ms
8	speak8-0013	10ms	29	speak8-181207	62ms

9	speak9-0013	21ms	30	speak9-181207	31ms
10	speak11-0013	35ms	31	speak10-181207	39ms
11	speak12-0013	31ms	32	speak11-181207	82ms
12	speak-0016	63ms	33	speak12-181207	57ms
13	speak2-0016	60ms	34	speak13-181207	46ms
14	speak3-0016	23ms	35	speak14-181207	44ms
15	speak4-0016	10ms	36	speak15-181207	71ms
16	speak5-0016	48ms	37	speak-181209	42ms
17	speak6-0016	105ms	38	speak2-181209	61ms
18	speak7-0016	40ms	39	speak3-181209	51ms
19	speak8-0016	7ms	40	speak4-181209	31ms
20	speak-181206	30ms	41	speak5-181209	40ms
21	speak2-181206	56ms			

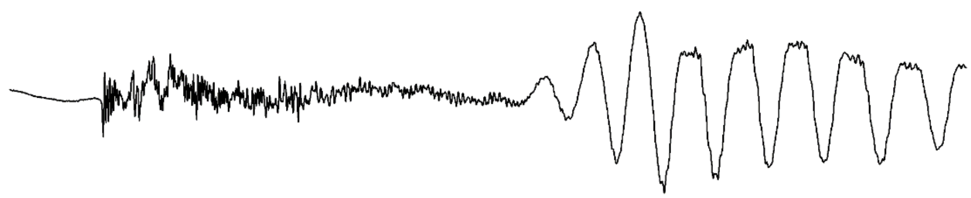
Appendix H

God1-0013 VOT = 35ms

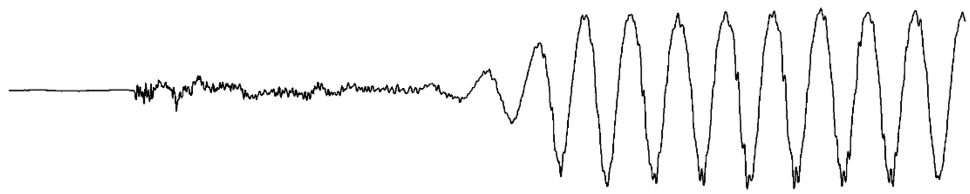


God2-0013 VOT = 74ms

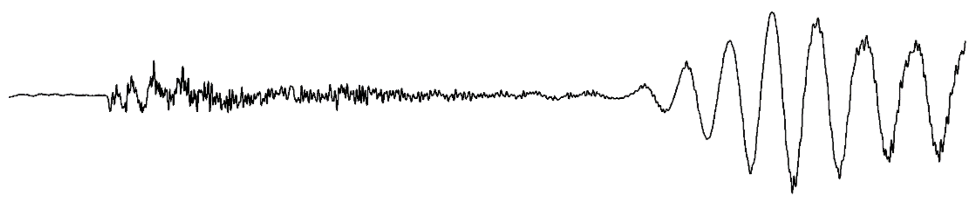
God4-0013 VOT = 35ms



God5-0013 VOT = 26ms



God6-0013 VOT = 46ms



God7-0013 VOT = 64ms

Appendix G

Wordlist:
we (feminine)
woman (nom, singular)
you (masculine, singular)
drink
rest
sharp
answer
one
cat

lion
breathe
five
yellow
child
white man
I
dog
to admonish
great
to call
griqua
death
to run
now
how are you?
how
Le Valliant:
to drink
mouth
bird
liver
alive
god
head
tongue
tiger
wolf
milk
one
two
three
four
you (masculine, singular)
head
eyes
hand
ears
mouth
tooth
why do they laugh?

bread
fire
man
woman
speak
drink
animal
water
house
meat
rain
sun/dry
story
son
daughter
father
mother
bird
grass
ground
sky
moon