The Phonetics and Phonology of Gutturals

A Case Study from Ju|'hoansi

Amanda L. Miller-Ockhuizen

OUTSTANDING DISSERTATIONS IN LINGUISTICS

Edited by

Laurence Horn
Yale University

A ROUTLEDGE SERIES

OUTSTANDING DISSERTATIONS IN LINGUISTICS

LAURENCE HORN, *General Editor*

THE SYNCHRONIC AND DIACHRONIC PHONOLOGY OF EJECTIVES
Paul D.Fallon

GRAMMATICAL FEATURES AND THE ACQUISITION OF REFERENCE
A Comparative Study of Dutch and Spanish
Sergio Baauw

AUDITORY REPRESENTATIONS IN PHONOLOGY
Edward S.Flemming

THE TYPOLOGY OF PARTS OF SPEECH SYSTEMS
The Markedness of Adjectives
David Beck

THE EFFECTS OF PROSODY ON ARTICULATION IN ENGLISH
Taehong Cho

PARALLELISM AND PROSODY IN THE PROCESSING OF ELLIPSIS SENTENCES
Katy Carlson

PRODUCTION, PERCEPTION, AND EMERGENT PHONOTACTIC PATTERNS
A Case of Contrastive Palatalization
Alexei Kochetov

RADDOPPLAMENTO SINTATTICO IN ITALIAN
A Synchronic and Diachronic Cross-Dialectical Study
Doris Borrelli

PRESUPPOSITION AND DISCOURSE FUNCTIONS OF THE JAPANESE PARTICLE
Mo
Sachiko Shudo

THE SYNTAX OF POSSESSION IN JAPANESETakae Tsujioka

COMPENSATORY LENGTHENING
Phonetics, Phonology, Diachrony
Darya Kavitskaya

THE EFFECTS OF DURATION AND SONORITY ON CONTOUR TONE DISTRIBUTION

A Typological Survey and Formal Analysis
Jie Zhang

EXISTENTIAL FAITHFULNESS
A Study of Reduplicative TETU, Feature Movement, and Dissimilation
Caro Struijke

PRONOUNS AND WORD ORDER IN OLD ENGLISH
With Particular Reference to the Indefinite Pronoun Man
Linda van Bergen

ELLIPSIS AND *WA*-MARKING IN JAPANESE CONVERSATION
John Fry

WORKING MEMORY IN SENTENCE COMPREHENSION
Processing Hindi Center Embeddings
Shravan Vasishth

INPUT-BASED PHONOLOGICAL ACQUISITION
Tania S.Zamuner

ORIGINS OF PREDICATES
Evidence from Plains Cree
Tomio Hirose

CAUSES AND CONSEQUENCES OF WORD STRUCTURE
Jennifer Hay

INTERACTIONS BETWEEN MARKEDNESS AND FAITHFULNESS CONSTRAINTS IN VOWEL SYSTEMS
Viola Miglio

MANIFESTATIONS OF GENERICITY
Yael Greenberg

TRUE TO FORM
Rising and Falling Declaratives as Questions in English
Christine Gunlogson

VIETNAMESE TONE
A New Analysis
Andrea Hoa Pham

THE PHONETICS AND PHONOLOGY OF GUTTURALS

A Case Study from Ju|'hoansi

Amanda L.Miller-Ockhuizen

Routledge
New York & London

Published in 2003 by
Routledge
29 West 35th Street
New York, NY 10001
www.routledge-ny.com

Published in Great Britain by
Routledge
11 New Fetter Lane
London EC4P 4EE
www.routledge.co.uk

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2005.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

Library of Congress Cataloging-in-Publication Data
Miller-Ockhuizen, Amanda.
The phonetics and phonology of gutturals: a case study from Ju|'hoansi/by Amanda L.
Miller-Ockhuizen.
p. cm. —(Outstanding dissertations in linguistics)
Includes bibliographical references and index.
ISBN 0-415-96793-7 (hb)
1. Grammar, Comparative and general—Phonology—Case studies. 2. Phonetics—Case
studies. 3. !Xä language—Phonetics. 4. !Xä language—Phonology. I. Title. II. Series.
P217.3.M556 2003
414–dc21 2003046887

ISBN 0-203-50640-5 Master e-book ISBN

ISBN 0-203-57918-6 (Adobe eReader Format)

For Aiya, |asa Dabe,
who always subtly reminded me of my own inner strength,
by simply asking me,
“Aren’t you a woman?”

and

For Xoa||’an |Kunta,
mi !uma te mi are,
who taught me Ju|’hoansi Kokxui

Contents

List of Appendices

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my adviser, Mary Beckman, who was a true mentor, always allowing me to formulate my own ideas about phonetics and about the phonetics—phonology interface, but always ready and eager to provide insights into how they might be grounded in the reality of acoustics and articulation. I would also like to thank Keith Johnson, who listened patiently as I formulated ideas, only softly pointing out where the holes in my ideas lay and allowing me to work them out in my own quiet way. He was also an excellent model for the way in which phoneticians attack problems, and I learned from his example. I also thank Beth Hume who was extremely encouraging and full of enthusiasm for my work. I value most the way in which she sees the creative capacity of phonological theories as equally important as the constraints on those theories. Many of the ideas in this dissertation were developed in Keith Johnson's seminar on *Phonological parsing,* Beth Hume's seminar on *Positional Specification*, as well as Beth and Keith's joint seminar on *The Role of Perception in Phonology.* It is the constant encouragement and excitement about linguistics as an empirical science which unites all three of my committee members, and for which I value all of them more than I can express in words. I would also like to thank David Odden, who helped me to achieve a strong background in phonology, and for his help with the initial set up of the database.

Chuck Kisseberth and Jennifer Cole also had a great impact on my view of phonology. Chuck's view that phonological constraints can be functionally unified was the impetus for my early view of the integration of phonetics and phonology. His belief in the importance of documenting Khoisan phonology, always gave me the encouragement I needed to complete my dissertation, even when I wasn't in close contact with him. Jennifer Cole was also a great source of ideas, and her excitement about the interface between phonetics and phonology was contagious. Simon Donnelly, Khalil Iskarous and Molly Homer also provided friendship and showed interest in my work at UIUC.

I also owe a great deal of gratitude to *the Labbies*, particularly Jennifer Venditti, Kiyoko Yoneyama, Mariapaola D'imperio, Rebecca Herman, Matt Makashay, Pauline Welby and Elizabeth Strand who were always willing to share scripts, their own views about phonetics in general, and quick easy ways to get things done. Allison Blodgett, Laurie Maynell, Janice Fon, Georgios Tserdanelis, Steve Winters, Huang Tsan, Peggy Wong, Craig Hilts, Shelome Gooden, and Satoko Katagiri were always available for helpful advice, and contributed to the warm cooperative atmosphere found at the OSU Linguistics Lab. I would also like to acknowledge the friendship of Marcellino Liphola.

I am grateful to the National Science Foundation for a Dissertation grant, SBR 9705334 (David Odden, Project Director), which supported my stint of research in Namibia from April of 1997 to February of 1998, when I recorded the database used in this dissertation. I am also grateful to the Dept. of Linguistics at OSU for offering additional funding, including, but not limited to a PEGS fellowship. I also am very thankful to Osamu Fujimura, for offering me several quarters of Research Assistantships that allowed me to begin to understand the physiology and neural control of the larynx. In addition to funding, I thank Osamu Fujimura for his patience in explaining difficult acoustic concepts. Seiji Niimi was also a wonderful source of information on laryngeal control when I pursued my research at RILP. I am grateful to the OSU Office of International Studies for a Graduate Student International Dissertation Research Grant, which helped fund the collection of the data for the acoustic case study (chapters 4 & 5), as well as the UIUC Dept. of African Studies for several small research grants.

My undergraduate adviser at UNH, Mary Clark, was also important in molding my career as a linguist. She gave me the perfect answer when we discussed the futility of pursuing a Ph.D. in linguistics. She suggested that I go to the Peace Corps where I could both gain experience as an ESL teacher, as well as learn a rare language, which would help me if I still chose to pursue a Ph.D. in linguistics upon my return. I am glad I listened to her.

The Ju|'hoansi people, who I lived with and taught English to between 1990–1994, helped me to learn their language and nurtured the awe I felt for language discovery. Their deep respect for their language is one of the factors that encouraged me to continue with my efforts, no matter how many obstacles were put in my way. Of course, the people at |Xoan N!huru, Namibia who adopted me into their family in my early days in Namibia, and who agreed to work with me in my more serious quest to describe their language from 1997 onwards, are as much responsible for the creation of this dissertation as I am. Their steadfast hard workmanship, devotion to the task, and inspiration to give me speech that was truly representative of their language made it a real pleasure to pursue my work. Everyone at the village of |Xoan N!huru was involved in this research. I especially thank the headpeople, Kxao Kxami and Old Di||xao, for teaching their children and grand-children to be such responsible workers, and for overseeing the project, and the main consultants whose speech was used for the quantitative part of this dissertation, N!unka |ui (NU), asa Kxao (DK), Kxami Kxao (KK) and Kxao Bo (KX).

In Namibia, I would also like to thank many friends who offered us assistance during the course of many field work trips. I am particularly grateful to Paul and Annetjie Beukes, who provided lodging and assistance in any way they could in the Tsumkwe area. Other dear friends include Gerhard & Charmaine Mouton, Bonny & Julia Kangungu, Johan & Margaret Strauss, and Ivan and Caroline Nell, who were always ready to lend a helping hand. I also thank Wendy Hovelmann, Patti Swarts and others at NIED, as well as the Permanent Secretary of the Dept. of Education, Loini Katoma for their encouragement and the support of the Namibian Government.

I also thank my family on both sides of the Atlantic, for their patience while I strove to understand the Ju|'hoansi language. My husband, Ockie, always supported me and reminded me of the importance of my work for the Ju|'hoansi people, and for Namibia. He made countless sacrifices in order to assure the completion of this dissertation. My parents-in-law, Pieter and Johanna Ockhuizen waited patiently while their children pursued their education, and always encouraged us to pursue our goals and dreams. My own parents, Judy and Oliver Bartlett, and Robert B.Miller, and my sister Alisa and brother Bryan, offered support of many different types.

I also thank the Khoisan academic community. First and foremost, Bonny Sands, whose friendship and scholarly companionship I have found to be inspiring. Bonny was always excited to see and read first drafts of papers and provided helpful comments and suggestions. I have also enjoyed many email conversations with Tony Traill, who always had useful phonetic facts to contribute about !Xóõ, and often about other Khoisan languages as well, on every fact that I ever asked him about. I have also enjoyed stimulating conversation with Hirosi Nakagawa. I have enjoyed friendship and academic support from Willi Haacke, Levi Namaseb, and Ed Elderkin at the University of Namibia. Megan Biesele encouraged me to pursue studies in linguistics when I first met her in Namibia, and has continued to support my efforts in any way that she could ever since.

I would also like to thank Johanna Brugman and Bonny Sands for reading pre-final drafts of this book and assisting me with the editing process.

Preface

This dissertation investigates the phonetics and phonology of guttural consonants and vowels in the Khoisan language Ju|’hoansi. In addition to a set of four contrastive click types, the language contains four phonation type contrasts on vowels, and five phonation type contrasts that are aligned to the release of an initial pulmonic or velaric stop. The release properties of clicks are typically termed click effluxes, but the class of effluxes also includes closure properties such as voicing and nasalizaiton that are not included in the class of gutturals. The phonation type contrasts are produced either through constriction of the glottis, as in glottalized consonants and vowels, spreading of the glottis as in aspirated consonants, uvularized consonants and breathy vowels, or via a constriction of the pharynx produced through a complex of gestures (raising of the larynx, narrowing of the pharynx and retraction of the tongue root and epiglottis). The consonant and phonation type contrasts thus all belong to the natural class of guttural sounds, which are defined acoustically as those that involve either relatively high or relatively low spectral slope values occurring in the C-V transition. Guttural consonants and vowels are also shown to be unified in their effect on F0.

In addition to the rich set of phonation type contrasts, breathy vowels and epiglottalized vowels display timing contrasts. This leads to a three-way timing contrast of the alignment of glottal opening gestures within a syllable, with aspirated consonants aligning the opening phase to the consonant release, breathy vowels and partially breathy vowels aligning a glottal opening gesture with the entire syllable, with the gesture occurring within the domain of a single mora for partially breathy vowels, and over both moras in a fully breathy vowel. A four-way contrast then includes roots that contain no glottal gesture (unaspirated consonants and modal vowels).

The class of Ju|’hoansi gutturals all pattern together in several phonotactic constraints in the language: (1) an OCP constraint that targets both guttural consonants and vowels; and (2) a positional constraint aligning a guttural feature to the first syllable of roots. A subset of these consonants and vowels, which bear the place feature [pharyngeal], are subject to several other constraints: (1) a C-V co-occurrence constraint involving place of articulation; (2) a C-V co-occurrence constraint involving tongue height; and (4) a C-V co-occurrence constraint involving tone. Together, the evidence from these two classes of constraints motivate epiglottal and uvular consonants to be specified for both place features and laryngeal features. Thus, this dissertation resolves the debate between Halle and McCarthy over whether gutturals should be unified through the specification of laryngeal or pharyngeal features. In Ju|’hoansi, they are specified for both place and laryngeal features.

A quantitative acoustic case study reported in part II demonstrates that all guttural vowels involve voice quality cues, consistent with them being specified for a laryngeal feature. Additionally, vowels following guttural clicks display similar spectral slope values and harmonics-to-noise ratios within a similar time domain (C-V transition or tongue dorsum lag phase of clicks) to those found in guttural vowels. These auditorily similarities offer a cogent description of the phonetic grounding of the guttural OCP constraint. The constraint is motivated by parsing ambiguity. The constraint assures that Voice Onset Time (VOT), which is increased for guttural consonants, is available to differentiate roots containing guttural consonants from roots containing guttural vowels.

THE PHONETICS AND PHONOLOGY OF GUTTURALS

PART I

Phonetics and Phonology of Ju|’hoansi Gutturals

CHAPTER 1
Introduction

This dissertation is about the phonetics and phonology of lexical roots in Ju|’hoansi, a Khoisan language with a very rich inventory of consonants, vowels, and tones, and correspondingly rigid constraints on what sequences, segments and combinations of tones and segments can occur at different positions in the root. More specifically, the dissertation has two aims. First, it explores the acoustic properties that unify the class of gutturals in the language. Gutturals are a set of phonemes that include not only consonants with marked pharyngeal and laryngeal articulations, but also vowels with a pharyngeal constriction or one of two non-modal laryngeal stricture properties. Gutturals can be identified as a natural class because of their role in several salient phonotactic constraints. Second, the dissertation describes several phonotactic constraints that operate within the lexical root and that refer in one way or another to guttural features.

All of the phonotactic constraints described in this dissertation enforce acoustic modulation across the root, and help to demarcate the root for lexical access. For example, one of the constraints severely restricts the consonants that can occur in root-medial and root-final positions. This constraint severely limits guttural consonants from occurring anywhere but root beginnings and simultaneously serves to maximize the sonority modulation at the initial CV boundary, minimizing it elsewhere within the root. Another constraint prohibits guttural features from occurring on both the initial consonant and the initial vowel. This ensures a more local acoustic modulation across the initial consonant-vowel sequence and enhances perceptibility of paradigmatic contrasts at the location in the root where the greatest number of segmental distinctions is available. Together, the constraints described here illuminate the functional bases of the Guttural Obligatory Contour Principle and its grounding in more general principles such as sufficiency of contrast (Lindblom and Maddieson, 1988) and ease of parsing of acoustic properties of guttural consonants, which are not intrinsically localized in the acoustic interval defined by such segmental landmarks as the stop burst.

Ju|'hoansi, a Khoisan language spoken in Namibia and Botswana, is an ideal test case of phonotactic constraints, because it has a plethora of sounds, many of which are not found in languages spoken outside of Southern Africa. Khoisan languages in general are known to have the richest phoneme inventories among the worlds' languages (Maddieson, 1984; Traill, 1985). Ju|'hoansi, with 89 consonant phonemes (47 of which are clicks) and 34 vowel phonemes, is second only to !Xóõ in the size of its inventory. It is also one of the best-described Khoisan languages, with a descriptive grammar and dictionary developed by Snyman (1970, 1975), and a more recent expanded dictionary developed by Dickens (1994). Field work involved in this dissertation has extended the recognized lexicon even further. However, little work has concentrated on describing the rich set of phonotactic constraints found in the language. The phonotactic constraints in this language are particularly strong, leaving little doubt that listeners are aware of them.

Many of the constraints described in this dissertation have their basis in perception, as they involve acoustic modulation within the root. The first constraint is a local sequential constraint that rules out the co-occurrence of two guttural consonants and vowels within the same root. This constraint prohibits sequences of sounds that are too acoustically similar, and that have acoustic cues on the same dimension that overlap in the same temporal domain. The constraint serves to rule out contexts where the perceptibility of contrasts would be weak due to masking from similar adjacent sounds. This is an instance of the same universal constraint that causes language inventories to include sounds that are sufficiently perceptually distinct (Lindblom and Maddieson, 1988).

Silverman (1997) shows that inventories containing laryngeals have contrasts that are sufficiently perceptually distinct by assuring sufficient temporal distance between the different types. Ju|'hoansi is, however, a language that contains a three-way timing contrasts in both glottal abduction (aspirated consonants vs. fully breathy vowels vs. partially breathy vowels), and in epiglottal constriction (epiglottalized consonants vs. fully epiglottalized and partially epiglottalized vowels). Therefore, subtle timing contrasts implementing a single acoustic cue are allowed in Ju|'hoansi, but highly similar contrasts involving the same acoustic dimension cannot co-occur, where the acoustic cues involved in contrasts belonging to consonants or vowels are realized over the same temporal domain, namely the C-V transition. Thus acoustic modulation on a given dimension is enforced in order to assure the perceptibility of sequential contrasts.

The second constraint discussed in this dissertation is a positional constraint that makes use of acoustic modulation for a different functional goal, namely to mark morphological root positions. Sounds that have strong acoustic edges, such as high-intensity transients or sustained turbulence with energy concentrated in the upper frequency range are only licensed in initial position. The high intensity of transients and the distribution of energy in the high frequency range compared to neighboring vowels makes for a clear spectral edge at the C-V boundary.

The goal of this second constraint is not to preserve the perceptibility of contrasts as in the guttural co-occurrence constraint, but rather to allow the hearer to parse the incoming speech stream into word-sized chunks through the process of segmentation, by disambiguating root positions through the segmental material allowed in that position. High amplitude consonants and consonants with other rich acoustic landmarks such as stop bursts and guttural release properties (aspiration, uvularization, epiglottalization and glottalization), mark the left edges of roots, while the final position of roots contain a more sonorous vowel or nasal consonant, which is lower in amplitude than the initial consonants found in the language. The first vocalic position is also the richest, bearing the full set of vowel quality and vowel phonation type contrasts. Medial consonants are also very sonorous, and most of these are allophones of initial consonants that don't occur in any other position. The only consonants that occur in all three positions (root-initial, root-medial and root-final) are nasals, but their frequency of occurrence in final position is much higher than their frequency found in initial or medial position, making them good landmarks for root-final position. The uniqueness of sounds found in each position within the root shows that in this language, acoustic modulation assures that each root position is unambiguously marked through the choice of consonants and vowels found in that position. More universal phonetic constraints may also play a role. The fact that only labial nasal consonants occur in coda position of the root may be related to Winters' (2000, 2001) perceptual results, which show that both nasality and labial place of articulation are the most perceptible in post-vocalic position.

Since roots are optimally bimoraic (either monosyllabic or bisyllabic), with at most one extra syllable that arises through epenthesis into non-native clusters in loan-words, the root is also synonymous with a prosodic word in this language (Miller-Ockhuizen, 2001). Deciding precisely whether it is the morphological category root or the prosodic category foot that is being demarcated would require an in-depth investigation of the various types of enclitics found in the language. The phonological and morphological status of enclitics is beyond the scope of this dissertation. However, evidence from loan-word adaptation suggests that many of the phonotactic constraints operate within the domain of the prosodic word. Ju|'hoansi provides evidence that acoustic modulation proposed by Kawasaki (1982) and Ohala (1992) over the syllable must be extended to include other prosodic or morphological categories, in this case the prosodic word.

The third constraint I describe builds on observations made by Elderkin (1988) involving the interaction of guttural consonant release properties and guttural vowels with the tone patterns of roots. All guttural release types cause a lowering of fundamental frequency that is similar to the lowered fundamental frequency values associated with guttural vowels. This interaction is the basis of a phonotactic constraint that rules out the co-occurrence of high tones and guttural consonants and vowels, and results in rising tone patterns over the first half of the root for all guttural consonant initial roots. The natural class of guttural sounds targeted by tonal constraints is identical to the class of sounds targeted by the positional licensing constraint, thus bolstering the existence of the natural class of guttural sounds in the language. The prominent rise in fundamental frequency throughout the root is another form of modulation within the root that may also aid listeners in the process of segmentation, similar to the way stress patterns do in English (Cutler and Butterfield, 1992) and Dutch (Vroomen et. al., 1996).

The acoustic cues involved in gutturals and the temporal position of these cues generates the correct classification of sounds targeted in the constraints, whereas articulatory groupings have proven difficult to motivate. However, the acoustic properties that are targeted are different from those targeted by acoustically based constraints in Aymara, Quechua (MacEachern, 1999) and Arabic (Zawaydeh, forthcoming). I propose a new acoustically-based feature [Spectral Slope] that is specified on all gutturals. Pharyngeals can be either [High Spectral Slope] or [Low Spectral Slope] depending on the degree of constriction found in the pharyngeal region. Epiglottalized sounds are [Low Spectral Slope] given that they are more stop-like (Esling, 1999b), and Ju|’hoansi uvularized consonants are [High Spectral Slope], given that the uvularization is frication noise, articulated with a wide open glottis and concomitantly high degree of airflow. This feature incorporates the fact that constrictions near the glottis have similar acoustic effects to constrictions at the glottis.

The Back Vowel Constraint, which rules out the co-occurrence of central and lateral post-alveolar clicks with following front vowels is unified with a process of pharyngeal assimilation involving uvularized and epiglottalized clicks, and concomitant epiglottalization on vowels. The phonetic basis of this constraint is claimed to be articulatory, grounded in the incompatibilities found with pharyngeal constriction involved in all of the consonants and vowels involved in the pattern, and front vowels. Articulatory grounding of this constraint is left for future research.

This dissertation reports the results of two distinct studies. First, a recorded database was developed and used to identify robust phonotactic patterns in Ju|’hoansi. Second, a well-controlled but smaller acoustic study uses acoustic voice quality measures developed in the literature to investigate spectral properties of guttural consonants and vowels that define the level of acoustic similarity targeted by the Guttural OCP constraint. All guttural consonants and vowels are shown to display relatively high or low spectral slope, as well as increased harmonics-to-noise ratio values compared with modal vowels and modally release consonants, within the C-V transition.

The structure of this dissertation is as follows. In chapter 2, I lay out the inventories of Ju|’honasi consonants and vowels, and describe the acoustic properties that define the class of guttural consonants and vowels in Ju|’hoansi. I also provide acoustic evidence that the class of medial consonants in Ju|’hoansi are all sonorants. Phonetic background regarding clicks is also provided that will allow the reader to understand the phonetic bases of the phonotactic constraints described in Chapter 3. In Chapter 3, I report the results of the database investigation, and describe several novel phonotactic constraints discovered through this investigation. These constraints all target either the class of gutturals, or the class of pharyngeals (including back clicks). The chapter thus provides important evidence that gutturals are marked by both place features and laryngeal features. The database results provided in Chapter 3 also motivate the acoustic study reported in

Chapter 5, which aims to show that the Guttural OCP constraint has its basis in perception. Chapter 4 lays out the methodology and materials used in the acoustic study reported on in Chapter 5. Chapter 6 concludes by pulling together the database results and the acoustic results by explaining the relationship between the phonotactic patterns described in Chapter 3, and the phonetic properties of gutturals defined more broadly in Chapter 2, which are shown to be consistently present in the quantitative study reported on in Chapter 5.

CHAPTER 2
Phonetics of Guttural Consonants and Vowels

2.0.
Introduction

Gutturals form a natural class in Ju|'hoansi. In this chapter, I will show that guttural consonants and vowels are phonetically similar. Pharyngeals and laryngeals are unified phonetically in three different ways. First, they all display noise within the third formant frequency range. Second, they all lower or raise the fundamental frequency. Third, gutturals have an effect on the first formant frequency. In this chapter, I describe broadly the spectral characteristics of Ju|'hoansi guttural consonants and vowels. I provide spectrograms, spectra and fundamental frequency (F0) tracks to illustrate the phonetic properties that define the class of gutturals and the opposing class of non-gutturals. Dynamic properties of the different types of click accompaniments (closure vs. release properties) are shown to be relevant to the definition of gutturals, since only click accompaniments that are realized during the release are gutturals.

Spectral properties of guttural consonants and vowels provide evidence for aspiration, glottalization and epiglottalization on both consonants and vowels as involving the same acoustic properties. The only difference is the timing of the guttural gestures with respect to the click burst.

The acoustic and dynamic aspects of gutturals defined in this dissertation set the stage for the discussion of the acoustic motivation of a new Guttural OCP constraint described in Chapter 3 and for the quantitative acoustic study reported on in Chapters 4 and 5. All of the data provided in this chapter comes from either the recorded 1878 root database (described in section 3.1), or from the more controlled word-list provided in Appendix A.

This chapter is organized as follows. In Section 2.1, I provide the root shape inventory of Ju|'hoansi, and in Section 2.2, I provide the tonal inventory, as background for the rest of the chapter. In Section 2.3, I discuss acoustic properties of the consonant inventory, describing spectral properties that distinguish non-guttural from guttural obstruents. I then turn to a description of the phonetic properties of Ju|'hoansi medial consonants and show that they are all sonorants. In Section 2.4, I describe the vowel inventory, focusing on the spectral properties of guttural vowels. In Section 2.5, I discuss the spectral parallels between guttural consonants and vowels, and show that while the dynamic properties are large enough to support a contrast, the phasing differences are not as large as predicted by Silverman (1997). Rather, consonant and vowel contrasts involving the same articulatory gesture only maintain sufficient contrast, as with consonant contrasts. I argue that the Guttural OCP constraint in Ju|'hoansi allows more similar phasing between consonants and vowels than is found in languages without such a constraint.

2.1. Root Shape Inventory

Since the range of segmental contrasts in the language is determined by the prosodic position in the root, I first briefly outline the inventory of prosodic root shapes found in the language. The inventory of native root shapes is provided in Figure 1. There are a small number of trisyllabic roots, but these roots can all be identified as loan-words. Further, as shown in Section 3.2, trisyllabic loan-words are realized as two or more prosodic words. Therefore, CVCVCV roots are not

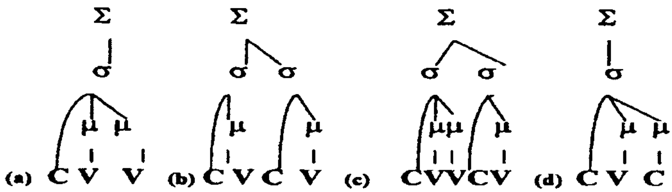


Figure 1. Inventory of Ju|'hoansi prosodic root shapes

considered a native prosodic root shape, and are thus not included in the inventory in Figure 1.

The prosodic generalizations that define Ju|'hoansi roots are that (1) roots are minimally bimoraic (see Miller-Ockhuizen, 2001 and Section 3.2 for evidence from reduplication and duration measurements; and (2) roots are maximally bisyllabic (see Section 3.2 of this dissertation). Phonological evidence is offered in Sections 3.3 and 3.9 that support an analysis whereby a native root equals a prosodic word, which always consists of a single foot. When trisyllabic or longer loan-words are adapted, the root is parsed into two or more prosodic words.

2.2.
Tone Inventory

Ju|'hoansi is a four tone language. Snyman (1975) describes the tones as *Hoog (Super High), Middelhoog (High), Middellaag (Low), and Laag (Super Low).* The tones are transcribed throughout this dissertation with the diacritic symbols offered in Table I, which are adapted from Dickens (1994):

Super-Low	SL	ȁ
Low	L	à
High	H	á
Super-High	SH	a̋

Table I. Tone levels and diacritics

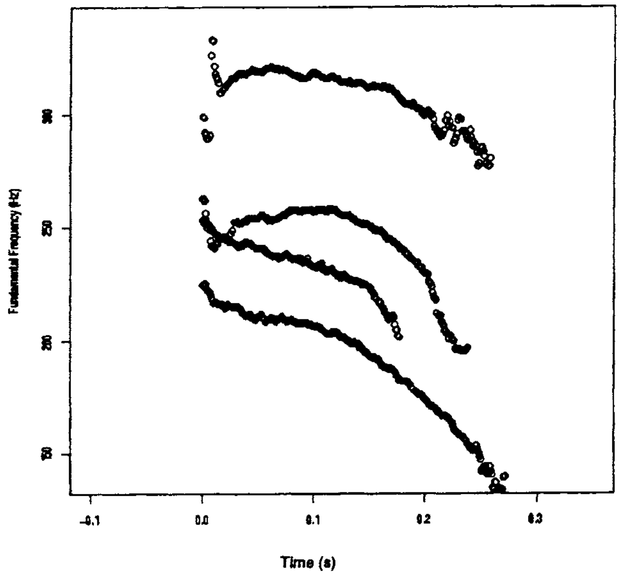


Figure 2. Fundamental frequency tracks associated with the four lexical tone levels in the roots *!ə̋ű* ‘migration’, *ǂáó* ‘heart’, *!əù* ‘bag made from antelope skin’, and *!ə̏ȍ* ‘peregrine falcon’ (single production, Subject DK)

Pitch tracks of CVV roots containing these four tone levels are provided in Figure 2 below. While the SL, L and H tones have more or less equidistant fundamental frequency values, the SH tone is almost twice as far from the H tones as the other three tones are from each other in fundamental frequency. These patterns are consistent with averages provided over multiple repetitions, as displayed in Miller-Ockhuizen (Forthcoming). These pitch tracks were calculated via the inverse of the pitch period for each cycle, with no smoothing. The roots were produced in isolation.

The extreme fall seen at the end of all four tones is a property of the phrasal phonology, and is not present in roots in phrase-medial contexts. The results of the acoustic case study reported in Part II reveals that there is also a tendency to break into breathy voice in phrase-final position.

2.3. Consonant Inventory

Ju|'hoansi is a Khoisan language with a typically large consonant inventory, including velaric plosives (click consonants) and pulmonic plosives (non-click consonants), as well as an extremely large vowel inventory. The particularly salient feature of the consonant and vowel inventories is the large number of contrasting phonation types. In Section 2.3.1, I introduce the consonant inventory, and point out the classes of guttural and non-guttural consonants. I also provide relevant background on different phases of click production, and discuss interesting properties of the Ju|'hoansi inventory. In Section 2.3.2, I focus on some phonetic properties that are unique to click consonants. First, I discuss frequency modulation found with clicks that makes them more similar to sibilants than pulmonic stops. Second, I discuss acoustic cues that signal different click types. In Section 2.3.3, I illustrate spectral and dynamic properties that differentiate non-guttural from guttural obstruents. I show that guttural clicks all involve non-modal voice quality cues following the click burst, while non-gutturals involve other sorts of cues that are located in the closure interval. In Section 2.3.4, I describe the phonetic properties of medial sonorant consonants in the language.

2.3.1. The Obstruent Inventory

The most salient feature of Khoisan consonant inventories is the large number of plosives, and the especially large number of click consonants. Clicks are complex plosives made with a sealed velaric cavity formed between a posterior and an anterior constriction. Clicks resemble non-click plosives in having three phases of production. Abercrombie (1967) describes the three phases of stop production as the shutting phase, the closure phase, and the release phase. Clicks are however, more complex than pulmonic plosives in that there are two constrictions, and thus necessarily two closures and two releases. Thomas-Vilakati (1999) describes click articulation with three phases that parallel the phases of pulmonic stops: (A) the tongue dorsum lead phase, where both anterior and posterior constrictions are made in order to form a cavity (this parallels the shutting phase of pulmonic plosives); (B) the overlap phase, where air is rarefied in order to increase the volume of the velaric cavity (this parallels the closure phase of pulmonic plosives); and (C) the tongue dorsum lag phase, which includes both the release of the anterior constriction and the release of the posterior constriction (this parallels the release phase of pulmonic plosives). Clicks are classified in terms of the anterior place of articulation. Ladefoged and Maddieson (1996:247) refer to this as *click type*. Ju|'hoansi has four basic click types, dental [|], palatal [ǂ], central post-alveolar [!] and lateral post-alveolar [ǁ]. I differ from Ladefoged and Maddieson in calling the two further back click types *post-alveolar*, while they refer to them simply as *alveolar*.

Guttural consonants, which include aspirated, glottalized, epiglottalized and uvularized stops, all contain noise in the C-V transition, which is maintained throughout the following vowel via coarticulation. That is, the distinctive characteristics of guttural consonants are all release properties that follow the stop burst. This is in contrast to the voiced and nasal consonants, which exhibit low frequency energy during the closure phase. Unsurprisingly, voiced consonants in Ju|'hoansi do not pattern phonologically as gutturals. That is, the timing of the guttural cues along with the acoustic properties of the gutturals predicts the phonological patterning.

The acoustic data in this study focuses on the closure and release properties of velaric plosives (click consonants). I focus on click-initial roots in part because this will allow me to compare laryngeal coarticulation found with Ju|'hoansi click consonants to that found in other languages where it has been shown to occur, and in part because the frequency of click-initial roots is extremely high compared with the frequency of pulmonic stop-initial roots (Traill, 1985; and discussed in Section 3.4.1.1 of this dissertation). Moreover, the distributional irregularity between clicks and non-clicks means that it is impossible to find a full set of near-minimal pairs that contrast in phonation type on non-clicks. This focus on roots with initial click consonants allows a more balanced presentation of contrasting sounds with more minimal contrasts. Additionally, the focus on clicks is more relevant to the Khoisan phonetic context, given the high frequency of clicks in root-initial position.

Table II lists click and non-click consonants that are unmarked for extra release properties. The plain uvular fricative [χ] and the glottal fricative [ɦ] listed in Table II are the only unreleased sounds that behave as gutturals with respect to the phonological constraints that target gutturals described in Chapter 3. As can be seen, clicks and non-clicks are parallel in the types of possible closure properties. Nasalization and voicing are included in this table for both clicks and non-clicks, since the acoustic realization of nasality and voicing occurs in the overlap phase of clicks and the closure phase of non-clicks. The glottal stop is present in a very few roots transcribed by Snyman (1975) and Dickens (1994) as vowel initial. Its phonemic status is unclear. Sounds listed in parentheses only occur in a few roots and are marginal in the consonant inventory, occurring in only a few roots, many of which are loan-words.

Nasalized clicks are classified as obstruents, due to the presence of stop bursts (see Ladefoged's 1997 definition of sonorant), and based on their phonological patterning (see Section 3.4). The labial approximant [β], the coronal flap [ɾ], and the dorsal fricative [ɣ] are considered sonorants based both on their phonological patterning with nasals (see Section 3.4), and on the fact that they all have short duration, and maintain voicing and clear formant structure throughout the overlap (closure) phase. The consonants [β], [ɾ], [ɣ] only occur intervocalically, where, given their short duration, voicing is maintained throughout the constriction. They are considered

<table>
<tr><th rowspan="2"></th><th rowspan="2"></th><th rowspan="2">Labial</th><th rowspan="2">Dental</th><th rowspan="2">Palatal</th><th colspan="2">Alveolar / Post-Alveolar</th><th rowspan="2">Velar</th><th rowspan="2">Uvular</th><th rowspan="2">Epiglottal</th><th rowspan="2">Glottal</th></tr>
<tr><th>Central</th><th>Lateral</th></tr>
<tr><td>Glides / Approximants</td><td></td><td>(w), β</td><td></td><td></td><td></td><td></td><td>ɣ</td><td></td><td></td><td></td></tr>
<tr><td>Flaps</td><td></td><td></td><td>ɾ</td><td>(y)</td><td></td><td>(l)</td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Fricatives</td><td>Voiceless</td><td>(f)</td><td>s</td><td>ʃ</td><td></td><td></td><td></td><td>χ</td><td></td><td></td></tr>
<tr><td>Voiced</td><td></td><td>z</td><td>ʒ</td><td></td><td></td><td></td><td></td><td></td><td>ɦ</td></tr>
<tr><td rowspan="3">Pulmonic Plosives</td><td>Voiceless</td><td>p</td><td>t</td><td></td><td></td><td></td><td>k</td><td></td><td></td><td>ʔ</td></tr>
<tr><td>Voiced</td><td>b</td><td>d</td><td></td><td></td><td></td><td>g</td><td></td><td></td><td></td></tr>
<tr><td>Nasal</td><td>m</td><td>n</td><td></td><td></td><td></td><td>ŋ</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Affricates</td><td>Voiceless</td><td></td><td>ts</td><td>tʃ</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Voiced</td><td></td><td>dz</td><td>dʒ</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="3">Velaric Plosives (Clicks)</td><td>Voiceless</td><td></td><td>|</td><td>ǂ</td><td>!</td><td>||</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Voiced</td><td></td><td>g|</td><td>gǂ</td><td>g!</td><td>g||</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Nasal</td><td></td><td>ŋ|</td><td>ŋǂ</td><td>ŋ!</td><td>ŋ||</td><td></td><td></td><td></td><td></td></tr>
</table>

Table II. Inventory of Ju|'hoansi consonants that are unmarked for release type intervocalic allophones of /b/, /d/, and /g/ (Snyman, 1970; Section 3.4). Nasal non-click consonants are sonorants. A sketch of the phonetic properties of Ju|’hoansi sonorants is provided in Section 2.3.4, and the phonotactic constraint that drives this variation is described in Section 3.4.

Ju|’hoansi also contains velaric and pulmonic plosives marked by a rich set of release properties that involve a constriction in the larynx or the pharynx. I adopt the term *guttural* to unify these sounds, expanding the term used by Hayward and Hayward (1989), McCarthy (1991, 1994) and Halle (1995). I will use *guttural* to describe consonants with laryngeal and pharyngeal release properties, as well as plain uvular and epiglottal consonants, and vocalic phonation type contrasts involving laryngeal or pharyngeal articulation. The complete inventory of clicks and plosives bearing guttural release properties is provided in Table III. Tables II and III together list all of the consonants in the Ju|’hoansi inventory, with the exception of the voiced nasal aspirate [mʱ], which only occurs in the diminutive plural enclitic *mʱi*, and never in roots. [mʱ] is therefore irrelevant to the discussion of root-level phonotactic constraints that are the focus of this dissertation.

The voiceless epiglottalized clicks listed in Table III are those that Snyman transcribed as [!x’], and that Dickens (1994) rendered in Ju|’hoansi orthography as *!k* (using the post-alveolar click type to symbolize all click types). Ladefoged and Maddieson (1996:275–77) transcribe them as [k!ˣ’], and refer to them as clicks followed by voiceless velar affricates, assuming a cluster analysis. Miller-Ockhuizen (2000) showed that the fricated portion of the consonant must be more uvular given the higher F1 following these consonants than after analogous plain unaspirated clicks. Voiced epiglottalized clicks are transcribed by Ladefoged and Maddieson (1996:277) as [gk!ˣ’] and by Snyman (1975:33) as [g!ˣ’]. Previous transcriptions imply that these sounds are voiced ejectives, which have not been attested in any other human language. While these sounds do involve larynx raising similar to ejectives, there is no constriction at the glottis, and thus they are not ejectives. That is, larynx raising is one of the movements that participates in epiglottal stricture, along with tongue root retraction, epiglottal retraction, and pharyngeal narrowing. Thus, interpretation of these sounds as voiced epiglottalized clicks allows us to maintain the very strong universal that there are no voiced ejectives found cross-linguistically. The sequential nature of voicing and epiglottalization is shown in Section 2.3.3.2. However, I claim that the sequential nature of voicing is due to the incompatibility of larynx raising and voicing, and these sounds are phonemically voiced throughout.

Several phonetically motivated gaps in the consonant inventory are observed. First, there are no labial ejectives, which are relatively rare among languages of the world (Ladefoged and Maddieson, 1996), and there are no uvularized labial sounds. Second, those ejectives that do occur in Ju|’hoansi are all affricated. That is to say, the coronal and dorsal non-affricated ejectives [t’] and [k’], which are the least marked among languages of the world, do not occur. Third, it is also interesting to note the existence of the pre-voiced epiglottalized click and non-click consonants. As already mentioned, epiglottalization does involve raising of the larynx in production, but does not involve glottal closure. Thus, they are not ejectives.

As can be seen in Figures 3 and 4, the set of guttural release types found on clicks and non-clicks is for the most part parallel. One major difference between pulmonic and velaric plosives is the way they pattern with respect to co-occurrence with voicing and glottalization. Non-click voiced ejectives without an additional uvular release property occur in the language (e.g., [ds’] and [dʃ’]), but pre-voiced epiglottalized clicks without uvularization do not exist. The opposite pattern is found with consonants containing uvularization and ejection. Both voiceless pulmonic and velaric epiglottalized consonants occur, but pre-voiced pulmonic epiglottalized consonants do not occur.

The other major difference in the guttural release properties of click and non-click inventories is that both voiceless and voiced aspirated nasal clicks occur (e.g., [ŋ!ʱ] and [ŋ̥!ʰ]), but

RELEASE	PULMONIC PLOSIVE (NON-CLICK)			VELARIC PLOSIVE (CLICK)			
	ORAL		NASAL	ORAL		NASAL	
	VOICELESS	VOICED	VOICED	VOICELESS	VOICED	VOICELESS	VOICED
ASPIRATED	p^{h} t^{h} k^{h} ts^{h} tʃ^{h}	b^{ɦ} d^{ɦ} g^{ɦ} ds^{ɦ} dʃ^{ɦ}		ǀ^{h} ǂ^{h} ǃ^{h} ǁ^{h}	gǀ^{ɦ} gǂ^{ɦ} gǃ^{ɦ}	ŋǀ^{h} ŋǂ^{h} ŋǃ^{h} ŋǁ^{h}	ŋǀ^{ɦ} ŋǂ^{ɦ} ŋǃ^{ɦ} ŋǁ^{ɦ}
GLOTTALIZED / EJECTED	ts' tʃ'	ds' dʃ' dz' dʒ'		ǀ^{ʔ} ǂ^{ʔ} ǃ^{ʔ} ǁ^{ʔ}			
UVULARIZED	t^{χ} tʃ^{χ}	d^{χ} dz^{χ} dʒ^{χ}		ǀ^{χ} ǂ^{χ} ǃ^{χ} ǁ^{χ}	gǀ^{ʁ} gǂ^{ʁ} gǃ^{ʁ} gǁ^{ʁ}		
EPIGLOTTALIZED	t^{ʜ} k^{ʜ}			ǀ^{ʜ} ǂ^{ʜ} ǃ^{ʜ} ǁ^{ʜ}	gǀ^{ʕ} gǂ^{ʕ} gǃ^{ʕ} gǁ^{ʕ}		

Table III. Inventory of Ju|'hoansi consonants with guttural release types

[1] The glottalized clicks in Ju|'hoansi may involve voiceless nasal airflow, given the fact the strong tendency for these click types to co-occur in roots with nasalized vowels. However, the consonants themselves clearly pattern as gutturals with respect to the Guttural OCP constraint described in Chapter 3. The acoustic and aerodynamic properties of these sounds will be investigated in future research.

there are no voiceless aspirated nasal non-click consonants in the language. In addition, voiced aspirated nasal non-click consonants only occur in a single clitic (e.g., the diminutive plural clitic, [mʱii]), never in roots. The relegation of certain segmental types to clitics is not surprising, as the lexical frequency of other properties is also skewed in clitics. For example, the Super-High (SH) tone, which is very rare on roots, is the most frequent tonal pattern found on clitics, with 6 out of the 7 known clitics bearing SH tone. This means that it is likely that these properties serve to mark clitic boundaries uniquely, just as the phonotactic constraints described in Chapter 3 serve to uniquely mark root positions. The phonotactic properties of Ju|'hoansi clitics, however, is beyond the scope of this dissertation, which focuses on root phonotactic patterns. Further investigation of the distinctive phonological properties of clitics will be left for future research.

An interesting property of the Ju|'hoansi consonant inventory is the predominance of obstruents. Lindblom and Maddieson (1988) show from their review of the UPSID database that languages tend to have a ratio of about 70% obstruents to 30% sonorants. Ju|'hoansi, however, contains approximately 95% obstruents, and 5% sonorants, if we count the weak medial fricatives and flaps as sonorants along with the nasal consonants, and if we treat all clicks (including nasal clicks) as obstruents. Lindblom and Maddieson claim that the 70:30 proportion found almost universally in consonant inventories of diverse languages reflects the size of the phonetic subspaces of the two classes of consonants. Viewed this way, the overwhelmingly high proportion of obstruents to sonorants found in Khoisan languages can be viewed as a reflection of the increase in the obstruent subspace when clicks are part of an inventory. That is, use of the velaric cavity in consonant production increases the phonetic subspace for obstruents, because clicks are in and of themselves more perceptible. The increased salience of click bursts makes formant transitions less necessary for perception, and guttural release properties which tend to somewhat obscure these transitions, more ideal. This idea is pursued further in section 3.4.1.1. Clicks also increase the articulatory subspace through the use of double constrictions to form a velaric airstream.

2.3.2. Some Phonetic Properties of Ju|’hoansi Click Consonants

In this section, I describe some of the phonetic properties of Ju|’hoansi clicks. I provide spectra showing that Ju|’hoansi clicks contain higher or equal intensity energy than vowels at most frequencies over 3000 Hz, while vowels contain higher intensity energy than clicks in the 0–3000 Hz range. I also summarize the literature describing acoustic differences in click noise-burst spectra, and provide acoustic evidence that underlying front vowels following post-alveolar clicks surface as cross-place diphthongs. This demonstrates the co-occurrence restrictions involving post-alveolar central and lateral clicks with front vowels (Miller-Ockhuizen, 2000 and section 3.5.3).

2.3.2.1. Modulation

I first discuss the intensity modulation, which is greater with clicks than with pulmonic plosives. Traill (1994a:171) notes that !Xóõ clicks are “as intense as, or, more often, more intense than the following vowel.” Ladefoged and Maddieson (1996: 259) note that clicks “often have a peak to peak voltage ratio that is more than twice that of the following vowel, meaning that they have at least 6 dB greater intensity.” They claim that this means that clicks sound twice as loud as the following vowel. Traill (1994a) notes that all clicks (except labial clicks) have more intense noise-bursts than pulmonic stops, which are often less intense than the following vowel, and ejectives which are typically equal in intensity to the following vowel. Traill (1997: 110) offers the scale provided in Table IV for intensity of

!, ǁ	>>	ǂ	>>	ǀ, ʘ	>>	t, k, q

Table IV. Scale of intensity of noise-bursts for !Xóõ consonants (from Traill, 1997)

clicks and non-clicks in !Xóõ, based on the overall intensity of their noise-bursts. Traill doesn't include labial pulmonic plosives on this salience scale, because he does not consider the voiceless labial plosive [p] to be a native consonant of !Xóõ. While labial plosives are extremely low frequency in initial position across the Khoisan family, labial sonorants are extremely frequent in medial position. I attribute the low frequency of initial labial obstruents in Ju|'hoansi to the lack of labials with guttural release properties in the consonant inventory (see Section 3.4.1.1).

The data in this study show that Ju|'hoansi clicks are much more intense than non-click consonants. There is, however, substantial token-to-token variability in the intensity difference between the click burst and the following vowel, as has also been show in !Xóõ by Traill (1994a). There is also substantial inter-speaker variability, with female speakers tending to produce extremely intense clicks, and very low intensity vowels. That is, the intensity of click bursts seems to be divorced from the intensity of other consonants in the phrase.

Figure 3 provides waveforms of roots containing all four click types ([|], [ǂ], [!] and [ǁ]), as well as the pulmonic plosives [t] and [k]. In Figure 3, the beginning of the noise-burst is marked with the *b* label, and the beginning of the vowel is marked with the label *v*. Therefore, the entire consonant burst is contained between the *b* and *v* labels. Notice that the bursts in all four initial clicks have a much higher overall amplitude than the bursts associated with initial [t] and [k]. The dental click [|] burst seen in the upper left panel of Figure 3 displays the lowest amplitude among the different click types relative to the following vowel, but even it displays a much higher amplitude than the bursts associated with the pulmonic obstruents [t] and [k], which have much lower overall amplitude than the vowels following them.

Stevens (1998) notes that for clicks with abrupt releases (e.g., [ǂ] and [!]), the high frequency release transient is expected to be about 11 dB greater than that found in pulmonic stops, and the low frequency transient is expected to be 12 dB weaker than that found in pulmonic stops. In non-abrupt noisy clicks [e.g., [|] and [ǁ]), the source consists of a random series of similar pulses that are weaker in intensity, although they are still greater in intensity than the pulmonic stop bursts. The high amplitude bursts result in clicks having particularly salient noise bursts, and particularly weak formant transitions compared with pulmonic stops. Stevens (1998:225) notes that an increase in SPL (sound pressure level) of 10 dB results in a doubling of loudness. Since click noise-bursts are approximately 11 dB greater than pulmonic stop noise bursts, clicks should be slightly more than twice as loud as pulmonic stops.

The higher intensity click bursts mean that Ju|’hoansi clicks are as loud as the following vowel. In contrast, non-click obstruents are quieter than the following vowel. Thus, while the intensity of click bursts compared with pulmonic obstruent bursts makes them overall more salient, it in fact results in less amplitude modulation between the consonant and vowel. That is, the click noise-burst is not much louder than the following vowel. Ju|’hoansi and many Khoisan languages do not have labial clicks. This is likely due to the lower salience of the noise-burst of that click type.

Click consonants exhibit pronounced frequency modulation with the following vowel, compared to pulmonic plosives. Click noise bursts contain energy within a higher frequency range than is employed in pulmonic stop-bursts. That is, clicks are more similar to sibilant fricatives in that both clicks and sibilant fricatives display frequency modulation with the following vowel, while pulmonic stop consonants display amplitude modulation, and little frequency modulation.

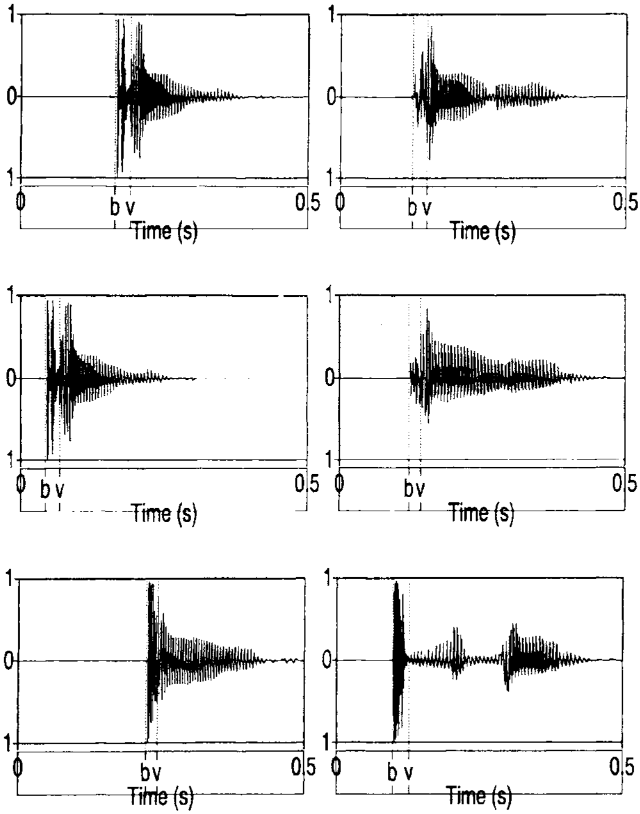


Figure 3. Waveforms of the roots *|àò* ‘buffalo’ and *tàrì* ‘thudding noise’ (above), *ǂàù* ‘giraffe’ and *kàrà* ‘cart’ (middle), *!àù* ‘skin bag’ and *ǁàβà* ‘palm’ (below, Subject KX)

Spectra computed over the entire click burst, and at the onset of the following /a/ vowel for different initial consonant types produced by Subject KX are provided in Figure 5. First, compare the solid black spectrum in panel A, which contains the [|] click noise-burst spectrum, to the solid black spectrum in panel B, which contains the [t] noise-burst spectrum. First, notice the high amount of noise in the spectrum taken at the burst of [|], which is consistent with it being dubbed a noisy click (Sands, 1991; Johnson, 1993; Ladefoged and Traill, 1994). For this noisy click, the energy at all frequencies between 2000 and 8000 Hz is higher in the click burst, which is the solid black line in panel (a), than at the onset of the following /a/ vowel, which is represented by the dotted black line in panel (a), and exhibits energy over the first 4000 Hz. The [t] burst spectrum, represented by the solid line in panel (b), on the other hand, shows energy in the 0–4000 Hz range, which is the same range where the energy is dominant in the vowel spectrum, represented by the dotted line in panel (b). Panel (c) contains the noise-burst spectra for the non-noisy [ǂ] click type, and the following /a/ vowel. For the non-noisy click, the energy is more concentrated above 2500 Hz. This is still much higher than the energy in the following /a/ vowel, which is most prominent in the 1000–2500 Hz range, and is indicated by the dotted line. On the other hand, the spectrum calculated in the [k] noise-burst, represented by the solid line in panel (d), displays energy concentration at all frequencies below 5000 Hz, the same frequency range that is prominent in the following /a/ vowel, represented by the dotted line in panel (d). Panels (e) and (f) provide spectra in the click noise-bursts for [!] and [ǁ]. Since [!] is a non-noisy click type, it again displays a prominent spectral peak in the 3000–4000 Hz range, while the noisy click [ǁ] (Sands, 1991; Johnson, 1993; Ladefoged and Traill, 1994), displays equally high amplitude energy at all frequencies above 2000 Hz. Thus, while we see that the noisy clicks, [|] and [ǁ], display equally high amplitude energy over a wider range than the non-noisy clicks [!] and [ǂ], which display prominent spectral peaks, all of the Ju|'hoansi click types display energy in a higher range than is employed in Ju|'hoansi vowels. Thus, noisy and non-noisy clicks are unified in that they all display increased frequency modulation with the following vowel that is not found between pulmonic stops and following vowels. In this respect, noisy clicks are more like pulmonic sibilant fricatives than pulmonic stops.

Kagaya (1978), Sands (1991), Ladefoged and Traill (1994) and Ladefoged and Maddieson (1996) note that click burst spectra typically show a concentration of energy between 1000–2500 Hz for the more back apical clicks, and above 2500 Hz for the more front laminal clicks, and suggest that these frequencies cue place of articulation. As can be seen in the spectra, the energy at these levels for the appropriate click bursts is higher in amplitude than that found in the following vowel. However, all click types discussed here contain energy in the upper frequency range as well, above 5000 Hz, which contributes to the frequency modulation found between clicks and vowels.

As I argue in Chapter 3 (Section 3.4), the perception of click types (e.g., [|], [ǂ], [!] and [ǁ]) might be improved compared with the perception of place of articulation in pulmonic stops (e.g., [p], [t], [k]), in segments with guttural release properties (e.g. aspiration, glottalization and epiglottalization), given listeners' low reliance on formant transitions in click perception. Sands (1991) shows that formant transitions into the following vowel following Xhosa clicks do not differ across click types, and Traill (1993, cited in Traill, 1994a) claims that there are no significant differences in the formant trajectories following the three Zulu clicks. However, there have been no studies of C-V place coarticulation in a Khoisan language to date.

Traill (1997) claims that the feature <click> enhances properties of stops in various ways, for example through short burst duration, especially in [ǂ] and [!]. He claims that although [!] and [ǁ] are successfully identified by subjects in his perception experiments, even when they contain short bursts of 5.3 ms, their extreme intensity enables the auditory system to undertake frequency resolution, where such resolution would not be possible with equally short weak pulmonic stop bursts. Since clicks can be uniquely identified solely on the bases of their noise-bursts, but were never identified from C-V transitions alone in his experiments, while pulmonic stops within the same language (!Xóõ) can be identified from transitions alone, this is a unique characteristic of clicks (Traill, 1994a). The weakness of formant transition cues may cause listeners to rely more on the stronger noise burst cues

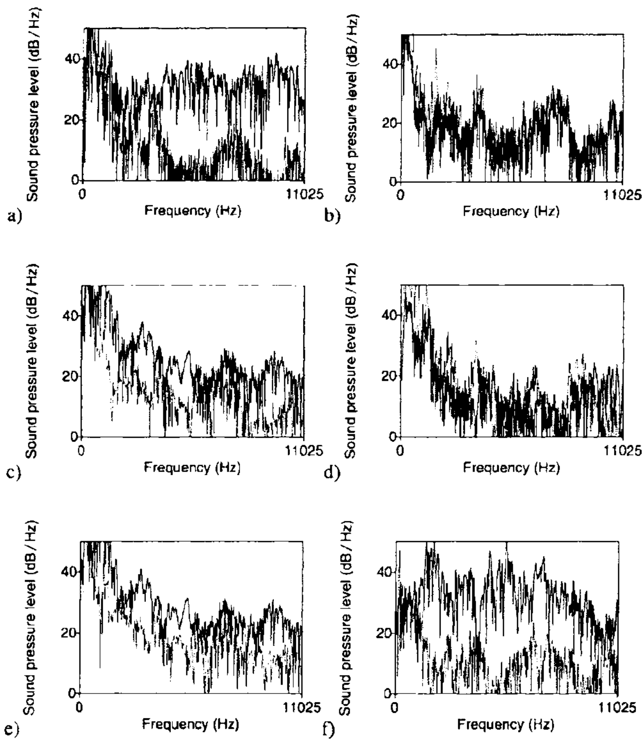


Figure 4. Spectra of consonant bursts and vowel onsets in the roots: (a) *|ə̀ò* 'buffalo', (b) *tə̀ɾì* 'thudding noise', (c) *ǂə̀ù* 'giraffe', (d) *kàɾà* 'cart', (e) *!ə̀ù* 'skin bag' and (f) *ǁàβà* 'palm' (Subject KX, 46.6 ms Hanning window with 14 coefficients, solid line is consonant burst spectrum, and dotted line is vowel spectrum)

available in the signal. Thus, attention to noise bursts aids in identification of click place of articulation, as well as in the process of segmentation. The strong noise bursts might also help explain the high lexical frequency of clicks in a language with a rich set of guttural release properties, since the formant transitions are likely masked by the noise associated with the guttural release properties (see Section 3.4.1.1).

2.3.2.2. Click Type Differences

Differences in the spectra of the more laminal front clicks [|] and [ǂ], and the apical back clicks [!] and [ǁ] noted by Kagaya (1978), Sands (1991), Ladefoged and Traill (1994), and Ladefoged and Maddieson (1996) are somewhat indicative of their different anterior places of articulation. However, place of articulation alone cannot capture the differences. In this section, I provide an overview of the known differences found among different click types, which are possible bases of the co-occurrence restrictions found between clicks with more retracted posterior constrictions and coronal vowel place, known as the Back Vowel Constraint (Miller-Ockhuizen, 2000) and Traill (1985, 1997).

Traill (1994a) shows that there is categorical perception for clicks based on the frequency range emphasized in the noise-burst spectra by !Xóõ listeners. These perception results provide evidence that the energy in the noise-bursts is an important cue signaling click type. Traill developed two sets of stimuli, one set that was synthesized from a 30 ms noise which was spectrally shaped to have a single peak ranging from 4750 Hz to 2200 Hz in 250 Hz steps, and another set that was synthesized from an 8 ms long single pulse with a bandwidth of 100 Hz, spectrally shaped to have a single peak ranging in frequency from 4200 Hz to 1100 Hz in 250 Hz steps. All non-noisy spectra with spectral emphasis within the frequency range from 1100–1600 Hz were perceived as [!], and all non-noisy spectra within the range from 1800–3300 Hz were perceived as [ǂ]. In the noisy spectra, sounds with spectral emphasis in the range of 2000–3250 Hz were always perceived as [ǁ], while spectra with emphasis in the range of 3250–5000 Hz were always perceived as [|]. Thus, click type can be signaled purely from properties of the noise-bursts of the different types.

Further spectral differences have been shown to be important by perceptual studies. Traill (1997) showed that perception of [|] was improved when the natural stimuli that included the vowel context, rather than stimuli that had a neutral synthesized vowel. Traill (1997:115) states that this "…suggests that the anterior closure of the click sometimes does coarticulate with the following vowel." He goes on further to explain that "…this means that the tongue must often move in such a leisurely fashion after release of the anterior closure for this click that the tip must still be raised after the secondary closure has been released and indeed that it must 'override' the coarticulatory effect of this secondary closure." This is in accordance with Miller-Ockhuizen's (2000) finding that the front clicks [ǀ] and [ǂ] exhibit coarticulation with a following [a] or [ə] vowel, as seen through a higher first formant frequency and a lower second formant frequency compared with the same vowels following the back clicks [!] and [ǁ]. However, perceptual improvement due to natural vs. synthetic vowel contexts could also be due to the naturalness of the stimuli.

Miller-Ockhuizen (2000) attributes the higher F2 in the front clicks to the fronting of the vowel from coarticulation with the anterior constriction of the click, and the raised F1 following back clicks to their further back posterior constriction. This could be due either to the fact that only the back constriction is acoustically relevant, or more likely, due to the further back anterior and posterior constrictions found in the back clicks ([!] and [ǁ]).

Thomas-Vilakati (1999:159) shows that the posterior click constriction in Zulu [ǀ] is the farthest forward, followed by [!], then [ǁ], which is the farthest back; and the posterior constriction in [!] and [ǁ] is further back than in the simple velar plosive. IsiZulu doesn't have the [ǂ] click in its inventory. Additionally, Thomas-Vilakati (1999:213) shows that in the production of [!], rarefaction is produced in part by retraction of the tongue dorsum just prior to release, in addition to the tongue lowering found in all click types. She notes that this tongue dorsum retraction may be in preparation for the retraction of the tongue tip constriction present in the anterior constriction, which temporally follows it (p. 132). Preliminary ultrasound investigations of my own articulation of [!] and [ǁ]

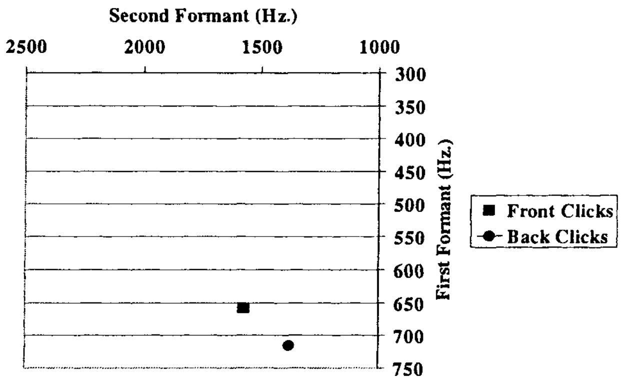


Figure 5. First formant frequency differences associated with the vowel [a] following front ([|] and [ǂ]), and back ([!] and [ǁ]) clicks

in Ju|’hoansi show tongue root retraction in the production of these clicks, which would indicate a more uvular place of articulation than velar. No such tongue root retraction is present in my productions of [|] and [ǂ].

In Miller-Ockhuizen (2000), I show that the vowel /a/ following back clicks, [!] and [ǁ], has higher first formant frequency values than the vowel /a/ following front clicks, [|] and [ǂ]. Figure 5, copied from that paper, illustrates this difference. This figure represents 40 tokens of front click initial roots (4 each of [|] and [ǂ] produced by 5 subjects), and 40 tokens of back clicks (4 each of [!] and [ǁ] produced by 5 subjects).

Spectrograms are provided in Figure 6, which show the realization of underlying front /i/ vowels following the four different click types. As noted in the transcriptions of these roots, and as has been noted by Traill (1985) for !Xóõ and Miller-Ockhuizen (2000) for Ju|’hoansi, the surface realization of /i/ following back clicks [!] and [ǁ] is [əi]. This can be seen in the spectrograms in Figure 6. The two upper panels of Figure 6 show voiced front click-initial roots. As can be seen, there is a short F1 transition at around 1000 Hz, that last about 10 msec. The lower two panels show spectrograms of back

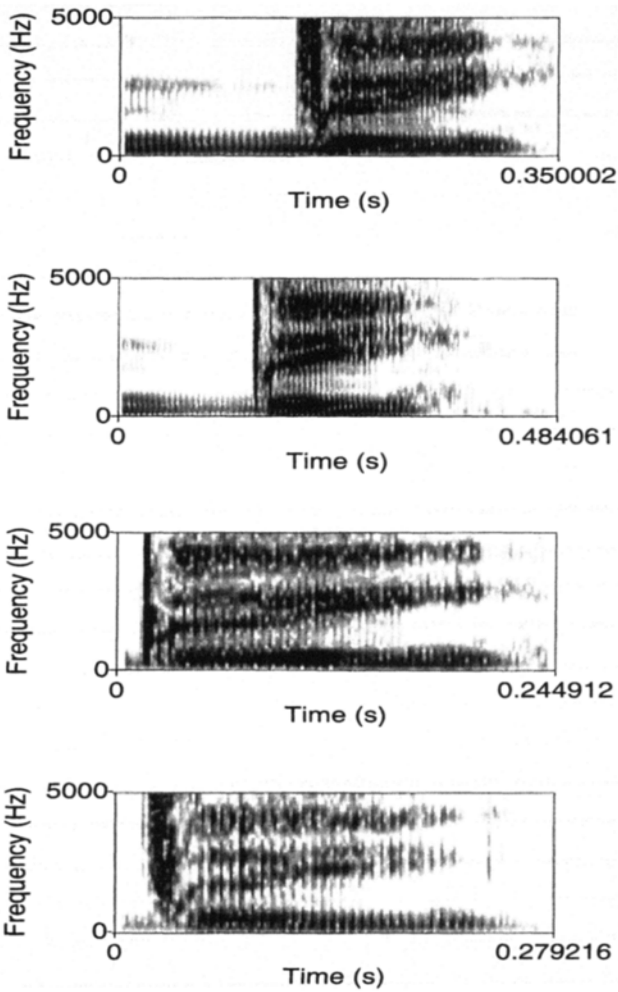


Figure 6. Spectrograms of roots containing the four Ju|'hoansi click types followed by a phonemically front vowel in the roots *g|íí* 'to exit', *g‡ìì* 'wrist', *!àìⁿ* 'neck' and *ǁáíⁿ* 'diaphragm' (cut out of the frame *a*___ 'you(r)___ .', Subject KX)

click-initial roots. As can be seen in these spectrograms, there is a more gradual F1 transition, lasting over the first half of the vowel. As I will claim in Chapter 3 (Section 3.5.3), this is due to retraction through coarticulation with the more uvular back closure found in the two post-alveolar (also known as "back") clicks. That is, in this dissertation, I adopt a phonological analysis whereby all clicks that participate in the BVC are marked with a [pharyngeal] feature, reflecting their further back posterior constrictions. However, future research will investigate the contribution of several articulatory factors that contribute to the lower acoustic noise burst frequencies found in the more back click types. These include place of articulation of the posterior constriction, place of articulation of the anterior constriction, tongue tip constriction shape and tongue body shape.

Cavity volume differences are also found for different click types. In Thomas-Vilakati's study of Zulu clicks, [!] displays the largest cavity volume, followed by [ǁ], and [|], which has the smallest cavity volume. Anterior place of articulation seems to be the least well-controlled by Zulu speakers in Thomas-Vilakati's study, with the placement and posture of the tongue tip and blade in the production of [ǁ] being variable.

2.3.3. Phonetic Properties of Guttural vs. Non-guttural Clicks

In this section, I illustrate the phonetic properties of non-guttural and guttural click consonants. I show that both the acoustic cues associated with gutturals, and the temporal locations of these cues are important in defining the class of guttural consonants. My primary claim is that it is both the similarity of voice quality cues present in guttural consonants and vowels, and the temporal location of these cues, that determines their participation in the guttural natural class targeted by several phonotactic constraints. Guttural release properties only occur on initial obstruents, as will be shown in Section 3.4.2. All guttural obstruents have longer voice onset times than non-guttural obstruents. They also all have spectral noise present in the release portion that is carried out into the following vowel through coarticulation. That is, vowels following guttural consonants display either high or low spectral slope values relative to modal vowels. These properties will be illustrated briefly in this section via spectrograms, pitch tracks and spectra. I label the click consonants using Thomas-Vilakati's phases of click production, with *b* marking the overlap phase of clicks, which corresponds to the closure phase of plosives, and with *c* marking the tongue dorsum lag phase, which includes both the releases of the anterior and posterior constrictions.

2.3.3.1.
Closure Properties of Non-guttural Clicks

In this section, I demonstrate that non-guttural clicks are either unmarked for any laryngeal feature, as is the case with voiceless clicks; or they display acoustic cues located in the overlap phase (e.g., voicing, nasality). In addition, vowels following non-guttural clicks display gently falling spectral slope values, and noise in the C-V transition (the tongue dorsum lag phase of clicks). I assume that the claims made here are relevant to non-click obstruents as well, but non-clicks have not been investigated in this dissertation.

In Figure 7, I provide spectrograms of voiceless and voiced unaspirated click-initial roots. As can be seen, the voiceless clicks display slightly positive voice onset time values. They display little or no energy in the overlap phase (interval labeled B in the spectrogram), and low amplitude short interval of noise between the click burst and the onset of voicing in the following vowel. In the voiced click, the overlap phase (labeled B in the spectrogram) corresponds roughly to the interval of voicing, marked by the low amplitude, low frequency energy seen prior to the click burst. The high intensity click burst is the only acoustic landmark present in the tongue dorsum lag phase of both voiceless and voiced clicks. The absence of noise in the lag phase, following the click noise-burst, explains their patterning as non-guttural consonants, despite the laryngeal articulation involved in voicing. The period of low frequency vibration during the overlap phase sufficiently distinguishes voiced clicks from voiceless clicks, leaving the slight degree of voicing coarticulation in the following vowel to be a more slightly weighted cue. If voicing coarticulation were masked by the presence of similar cues in the following vowel associated to a phonemic vowel phonation type, the voiced consonant would still be perceptible.

There are several articulatory strategies for maintaining voicing in clicks. Snyman (1978:151) provides a spectrogram of a voiced click produced by a single male Ju|'hoansi speaker that exhibits continuous voicing. The voiced click token seen here also displays continuous voicing, as do all of the tokens recorded in this study produced by this speaker. However, clear formant structure can be seen during the overlap phase (interval labeled B) of the voiced click in the lower spectrogram in Figure 7, indicating that the velum is raised. Formant structure during the closure, indicating velum raising, can also be seen in the spectrogram of the voiced unaspirated clicks provided in the upper two panels of Figure 6 produced by Subject KX. Raising the velum allows airflow through the glottis which makes it easier to maintain voicing, and this is a common way of maintaining voicing in stops cross-linguistically. This is the first study to document velum raising in voiced click production.

Other subjects in this study, who maintain a lowered velum, produce discontinuous closure voicing. Female subject NU's productions exhibit discontinuous closure voicing for all fifteen tokens produced in this study. In such subjects' productions, there is a period of voicelessness prior to the click noise-burst, leading to discontinuous voicing in the click. Thus, the degree of continuity of voicing in clicks is subject to a great degree of inter-speaker variation. In this study, it appears that female subjects, like subject NU, are more likely to exhibit discontinuous closure voicing, while male subjects, like subject KK, are more likely to exhibit continuous voicing, using other articulatory strategies to maintain voicing throughout the entire click.

My auditory impression is that many voiced clicks involve some implosion. Implosion, achieved through larynx lowering, is another articulatory strategy that is known cross-linguistically to make it articulatorily easier to maintain voicing during a stop closure. This strategy is another way that some Ju|'hoansi subjects maintain continuous voicing throughout both the overlap phase (interval labeled B in Figure 7) and the tongue dorsum lag phase (interval labeled C in Figure 7) in voiced clicks. The use of larynx lowering, like velum raising, seems to be more common among male subjects. However, future research will investigate whether the variation in continuity of closure voicing is more due to individual speaker variability, or gender.

Figure 8 provides spectra taken at the ¼ and ¾ points in the vowel following voiceless and voiced unaspirated clicks. Here and throughout this chapter, FFT spectra were computed over a 50 ms Hanning window centered over the ¼ point of the vowel, or the ¾ point of the vowel. As can be seen, vowels following both the voiceless and voiced unaspirated clicks display a slightly downward spectral roll-off, which is indicated by the difference in the amplitude of the first harmonic (H1) and the second harmonic (H2) (Bickley, 1982; Ladefoged, Maddieson and Jackson, 1988; Kirk, Ladefoged and Ladefoged, 1993). As noted by Stevens (1998:90), in modal voice there is an approximately −6 dB/per octave roll off. A tone that is an octave higher than a lower tone, is supposed to be about twice that of the original tone, making the second harmonic which is the second multiple of the first harmonic, be about an octave higher than the first harmonic. In Figure 8, the roll off is slightly smaller in the vowel following the voiced consonant (about 1 dB/Hz), than in the vowel following the voiceless consonant (about 3 dB/Hz). For roots containing both initial voiceless and voiced unaspirated clicks, there is an increase in spectral roll-off at the ¾ point of the vowel. Since these roots were produced in isolation, they are in utterance-final position. The increased spectral roll-off at the ¾ point of the vowel is then interpreted as a lapse into breathiness at the end of the utterance. In the quantitative

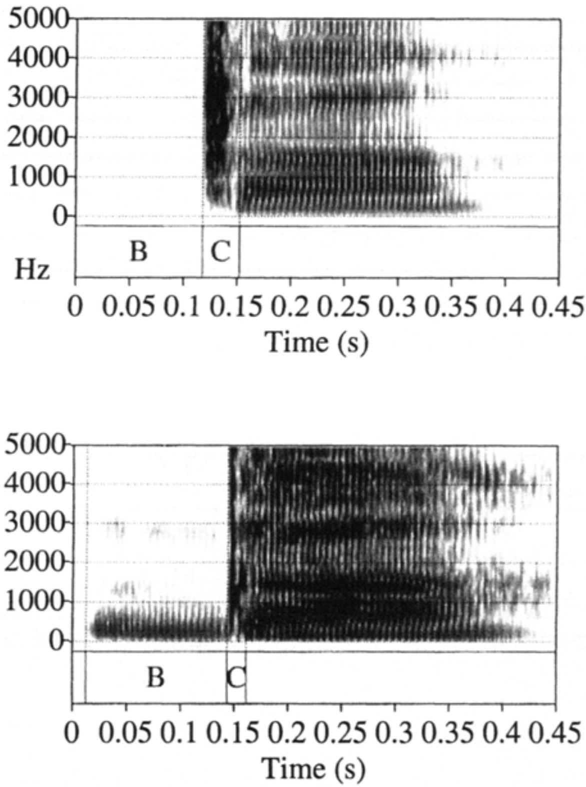


Figure 7. Spectrograms of the voiceless unaspirated click-initial root *ǁáá* ‘to warm hands by fire’ (above) and the voiced unaspirated click-initial root *g!àà* ‘rain’ (below, Subject KK)

results reported in Part II, this lapse into breathiness at the end of the root is shown to be a general property of all roots produced in this context.

Nasal clicks belong to the class of non-guttural consonants, based on their phonological patterning. That is, as will be seen in Section 3.5.3, nasal unaspirated clicks co-occur freely with all guttural vowel types. These patterns persist in spite of the fact that their coarticulatory effect makes following vowels acoustically similar to phonemic guttural vowels in several

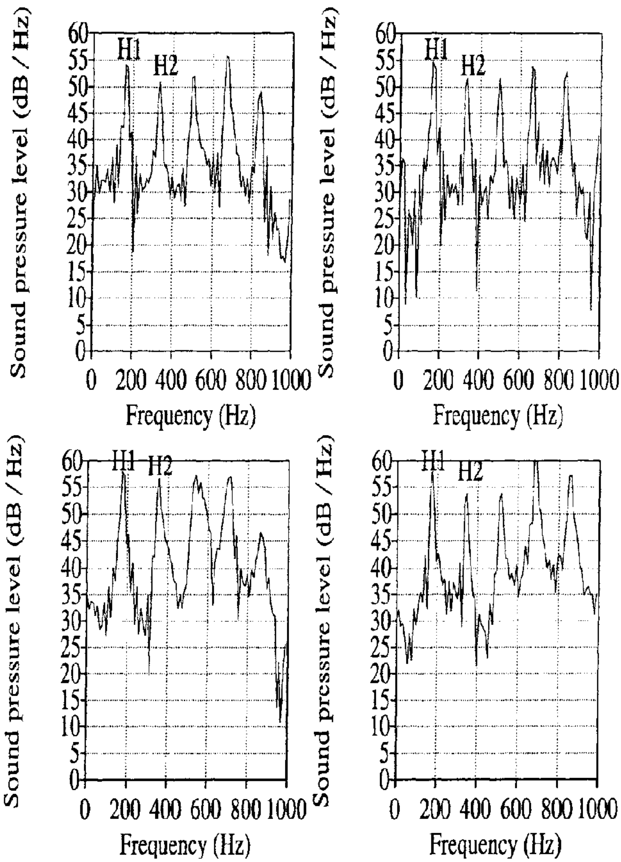


Figure 8. Spectra at the ¼ and ¾ points of the vowel in the voiceless unaspirated click-initial root *ǁáá* ‘to warm hands by fire’ (above) and in the voiced click-initial root *g!àà* ‘rain’ (below, Subject KK, 50 ms Hanning window)

respects (Fujimura and Lindqvist, 1971; Ohala and Ohala, 1993). Cross-linguistically, vowels following nasal consonants display increased F1 bandwidths and overall damping of the signal (Beddor, 1993) due to nasal coarticulation. In Ju|’hoansi this is also expected to be the case, given that nasalization is continuous based on spectrograms reported in Snyman (1978), and the acoustics of nasal-initial roots in the recorded database discussed in Chapter 3 of this dissertation. That is, nasalization in clicks is realized throughout the entire closure, and continues into the following vowel.

A detailed investigation of the acoustics of nasality in Ju|'hoansi is beyond the scope of this dissertation. However, it is not surprising that nasals do not behave as gutturals, given that nasality is cued with a nasal formant during the overlap phase of clicks. That is, a strong cue to nasal consonants is present in the overlap phase. Additionally, it is expected that vowels following nasal clicks do not exhibit higher or lower spectral slope values relative to modal vowels. Similarly, I expect that vowels following nasal clicks do not exhibit higher or lower spectral slope values relative to vowels following oral unaspirated consonants. It is clear that nasal unaspirated clicks do not exhibit noise in the C-V transition. Conversely, both voiceless and voiced nasal aspirated clicks, which pattern as guttural consonants phonologically, do exhibit noise in the lag phase, and vowels following them do display relatively high spectral slope values compared with modal vowels.

F0 patterns also nicely separate guttural from non-guttural consonants. As can be seen by the F0 tracks provided in Figure 9, F0 curves in vowels following voiceless and voiced unaspirated clicks both display a fall in F0 immediately following the click, followed by a relatively flat F0 pattern over most of the vowel. In Figure 9 and throughout, the pitch tracks provided here are the same tokens provided in the spectrograms, and the click bursts are aligned for each of the pairs being explicitly compared. As will be seen in the next section, all guttural clicks perturb the F0 patterns in following vowels, and cause a rise in F0 over the initial portion of the vowel. While there are co-occurrence restrictions concerning consonants and root tone patterns, as shown in Section 3.5.2.3, it is still possible to find minimal pairs for tone and vowel quality. Thus, all of the roots displaying consonantal contrasts in this section contain minimal or near-minimal contrasts in tone, taken from the data used in the quantitative study

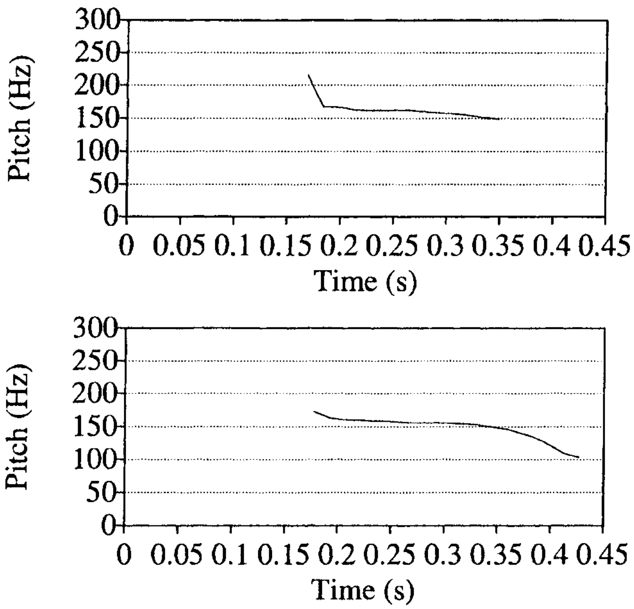


Figure 9. Fundamental frequency tracks associated with the voiceless unaspirated click-initial root *ǂòà* 'sitting mat' and the voiced unaspirated click-initial root *g!àà* 'rain' (Subject KK)

reported on in Part II. In Figure 9, the low-toned root with the diphthong is used in order to have the same lexical tones for the voiced and voiced unaspirated roots compared. In Figures 7 and 8, the high-toned monophthongal root was used as vowel quality control is more important for spectral comparison than tonal control.

2.3.3.2.
Release Properties of Guttural Clicks

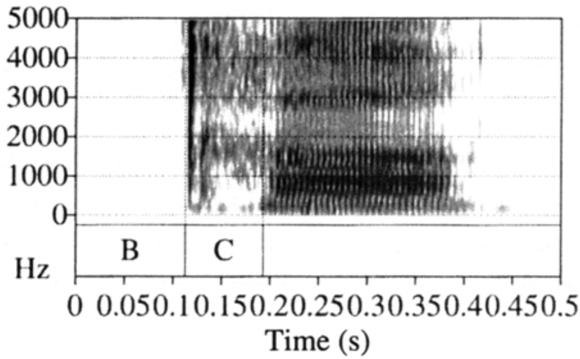
Having shown that non-guttural consonants are all unified in displaying gently sloping spectral slope values in the vowel following them, a lack of noise present in the lag phase of clicks, and falling F0 patterns, I now turn to the phonetic attributes of guttural consonants. In this section, I show that guttural click-initial roots display increased voice onset time, relatively high or low spectral slope values compared to the non-guttural consonants described in the previous section, and steep rises or falls in fundamental frequency compared with the more stable F0 patterns found in non-guttural click-initial roots.

Figure 10 provides spectrograms of roots containing initial voiceless aspirated and voiced aspirated clicks. Voicing in stops, typical of languages that have a four-way VOT contrast, is closure voicing. In some cases the closure voicing in both voiced unaspirated and voiced aspirated clicks is lengthened through the use of pre-nasalization, although this is usually followed by a period of oral voicing, which distinguishes these clicks from nasalized clicks. That is, this is clearly prenasalization, while voicing in nasal clicks is more similar to nasalization found in fully nasal non-click consonants. In plain voiced unaspirated clicks, the period of closure voicing is longer than is found with the voiced aspirated clicks. Comparing the lower spectrogram in Figure 7, with the lower spectrogram in Figure 10, we can see that the voiced unaspirated click has a period of voicing that lasts about 130 ms, while the voiced aspirated click has a shorter interval of voicing that lasts about 100 ms. Shorter voicing is probably part of the contrast, in conjunction with the VOT differences found in the release phase between voiceless aspirates and voiced aspirates. That is, the aspiration itself differs in the two click types, as is indicated in my transcriptions.

In the voiceless aspirated click production seen in the upper panel of Figure 10, the aspiration noise is located in the tongue dorsum lag phase of the click, and is longer than the aspiration noise found in the corresponding voiced aspirated click seen in the lower panel of Figure 10. This illustrates the VOT differences between the two click releases. Both types of aspirated clicks also exhibit increased noise in the third formant frequency range and higher, resulting in the third formant frequency being less intense than seen in the vowel following voiceless unaspirated and voiced unaspirated clicks.

Snyman (1975) transcribes sounds like those shown in the lower panel of Figure 10 as voiced aspirates, but Snyman (1978, 1999), Dickens (1994), and Ladefoged and Maddieson (1996:63) note that there is a period of devoicing in voiced aspirated click and non-click consonants in Ju|'hoansi, and Ladefoged and Maddieson therefore refer to the pulmonic obstruents of these types as clusters with mixed voicing. In fact, as seen in the spectrogram in Figure 11, there is only a slight interval of voicelessness after the click burst, which is followed by a period of voiced aspiration prior to the onset of formants in the following vowel. I therefore analyze these sounds as simple voiced aspirated clicks and attribute the period of devoicing just after the click burst to a wider glottal aperture than is found in analogous Hindi voiced aspirates. Snyman claims that the slight period of devoicing justifies calling these clicks clusters, as this period of devoicing is exhibited in the spectrograms of voiced aspirated clicks in his study, but not in voiced unaspirated clicks. As was noted above, there is inter-speaker variation in the continuity of voicing in voiced unaspirated clicks. Thus, the voicing gap seen in voiced aspirates is not that different from the gap often found in voiced unaspirated clicks in this study.

Vowels following the voiceless and voiced aspirated clicks display a larger positive difference between H1 and H2 relative to the differences found in vowels following the voiceless and voiced unaspirated clicks shown above in Figure 9. Spectra are provided in Figure 11, which show that the H1-H2 values at the ¼ point in the vowel following both voiceless and voiced aspirated clicks are about 6 db/Hz. I assume that this is due to the larger open quotient found in these sounds, following Stevens (1998:90). Interestingly, there is a decrease in the H1-H2 values at the ¾ points of the vowels for these root types, with both the voiceless aspirated and voiced aspirated click-initial roots displaying an approximately 4 dB/Hz spectral roll-off. This is opposite the increase in spectral roll-off found at the ¾ points of the vowels following the voiceless unaspirated and voiced unaspirated clicks shown in Figure 8 above. Similar interactions between lexically specified aspiration and



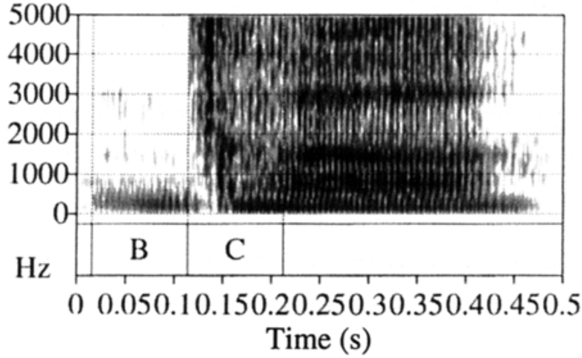


Figure 10. Spectrograms of the voiceless aspirated click-initial root *ǂʰàà* ‘path’ (above) and the voiced aspirated click-initial root *g|ʱàà* ‘place for drying meat’ (below, Subject KK)

prosodically conditioned aspiration (or breathiness) are found in the quantitative study reported in Part II.

Figure 12 provides F0 traces associated with the voiceless and voiced aspirated click-initial roots in Ju|’hoansi. First, notice that the pitch track starts earlier following the voiced aspirated click, because the voicing starts earlier. That is, given that the click-bursts are aligned in these two tokens, the earlier F0 values seen in the trace associated with the voiced-aspirated

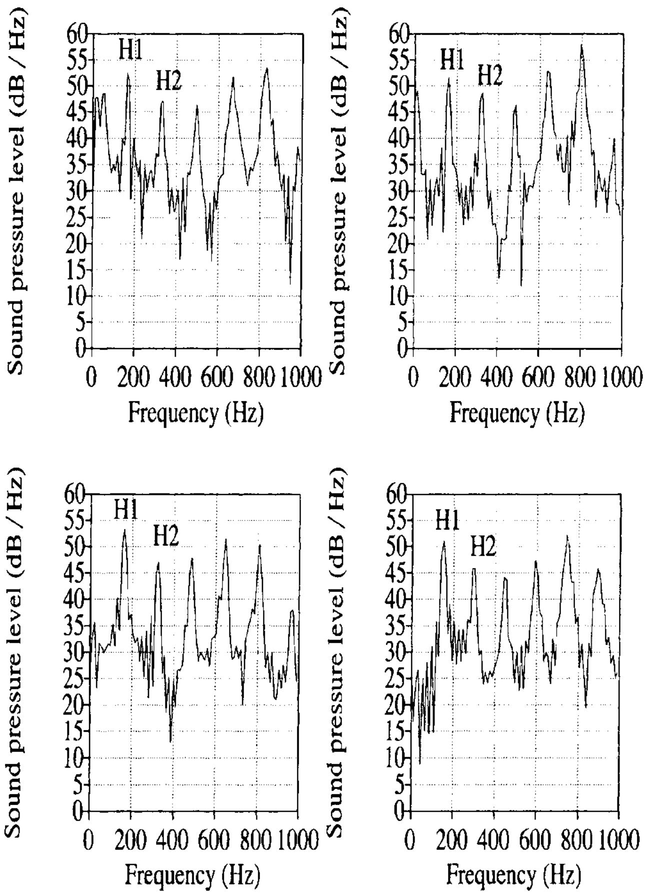


Figure 11. Spectra at 14 and ¾ points of the vowel in the voiceless aspirated click-initial root *ǂ^{h}àà* ‘path’ (above) and the voiced aspirated click-initial root *g|^{ɦ}àà* ‘place for drying meat’ (below, Subject KK, 50 ms Hanning window)

click-initial root is indicative of a difference in voice onset time (VOT), given the dependence of pitch tracking on voicing. The F0 values also start lower following the voiced aspirated click (approximately 120 Hz), than following the voiceless

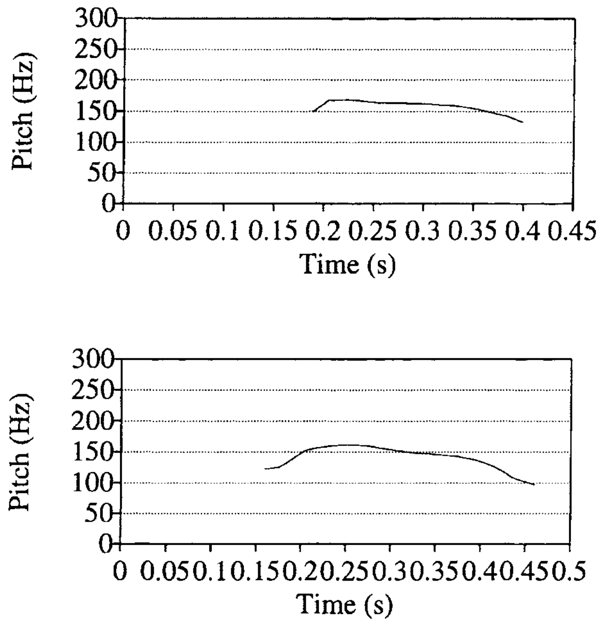


Figure 12. Fundamental frequency tracks associated with the voiceless aspirated click-initial root *ǂʰàà* 'path' (above), and the voiced aspirated click-initial root *g|ʱàà* 'place for drying meat' (below, Subject KK)

aspirated click (approximately 150 Hz), as is typical in languages containing a four-way Voice Onset Time contrast. The pitch of the voiceless aspirated click-initial root is in fact higher over more than half of the vowel, than the voiceless aspirated click-initial root. Both root types exhibit a fall in F0 over the root, but the fall is more gradual in the voiceless aspirated click-initial root than in the voiced-aspirated click initial root. The fall is also slightly greater in the voiced aspirated click-initial roots (about 50 Hz) than in the voiceless aspirated click-initial root (about 40 Hz). All of these attributes seem to be fairly consistent in the data recorded for the quantitative study reported on in Part II.

The uvularization noise found in both voiceless and voiced uvularized clicks is located during the tongue dorsum lag phase (interval labeled C), as shown by the spectrograms provided in Figure 13. The short interval of voicing found in the voiced uvularized click is similar to the short interval found in the voiced aspirated click. Note that in this voiced uvularized token, low amplitude formant structure is again present, indicating raising of the velum as a strategy to maintain voicing during the overlap phase (interval labeled B). In both voiced aspirated clicks and voiced uvularized clicks, the voicing during the tongue dorsum lag phase (the interval labeled C in Figure 13) also distinguishes these sounds from their voiceless counterparts. The voicing present in the aspiration and uvularization in voiced aspirated and voiced uvularized clicks is captured in the transcriptions offered for the voiced aspirated and voiced uvularized clicks throughout this dissertation.

The noise found in the tongue dorsum lag phase (interval labeled C) present in uvularized clicks is higher in amplitude (shown through the darkness of the noise in the spectrogram) than the noise found in aspirated clicks. That is, while formants are somewhat visible during aspiration noise of aspirated clicks seen in Figure 10, they are more difficult to identify in uvularization noise as seen in Figure 13.

Lag Voice Onset Time (VOT) (Lisker and Abramson, 1964) values differentiate uvularized clicks from aspirated clicks. The VOT values for the voiceless uvularized clicks and the voiced uvularized clicks seen in Figure 13 are 62 ms and 67 ms respectively. Compare this to the 73 ms and 45 ms VOT values found for voiceless aspirated and voiced aspirated clicks respectively in the spectrograms in Figure 10 above. These values two points. First, lag VOT values for voiceless aspirated clicks are greater than lag VOT values associated with voiceless uvularized clicks. Second, lag VOT values associated with the release phase of voiceless uvularized and voiced uvularized clicks are different. The lag VOT values for voiced aspirated clicks are shorter than the lag VOT values associated with the voiced uvularized clicks. The voiced uvularized click does not seem to exhibit any voicing during the release phase as is found in the voiced aspirated click. Quantitative results provided in Part II show similar differences. The quantitative results also provide more objective measures of the onset of voicing for the different guttural release properties.

The lead VOT values associated with the voiced aspirated click in the lower panel of Figure 10 is about 100 msec., while the lead VOT value associated with the voiced uvularized click in the lower panel of Figure 13 is about 120 msec, with an approximately 15 msec. voiceless period preceding the click noise-burst.

Vowels following voiceless and voiced uvularized clicks display an increase in noise in the upper frequency range that is similar to that found in vowels following aspirated consonants. This is the result of pharyngeal coarticulation with the following vowel and is similar to the coarticulatory effect found with aspirated clicks. Thus, the presence of noise in the following vowel, and the location of the noise in the C-V transition of roots containing uvularized click-initials unify uvularized clicks with aspirated clicks, and separate them from plain voiceless and voiced unaspirated clicks. Recall that the main cue to voicing is low frequency noise in the closure phase of voiced clicks.

Just as we saw in vowels following aspirated clicks, vowels following voiceless and voiced uvularized clicks exhibit higher spectral slope values than vowels following unaspirated clicks. That is, the difference between H1 and H2 is greater in vowels following these clicks than it is in vowels following voiceless and voiced unaspirated clicks. The higher spectral slope can be seen by comparing the differences between the amplitudes of H1 and H2 in the spectra taken at the ¼ and ¾ points of vowels following voiceless and voiced uvularized clicks in Figure 14, to the spectra in the vowel following voiceless and voiced unaspirated clicks in Figure 9. The higher spectral slope values found in vowels following uvularized consonants is expected to be due to the wide open glottal posture in their articulation, which results in higher airflow.[1] Uvularization and aspiration are thus acoustically very similar in that they both contain

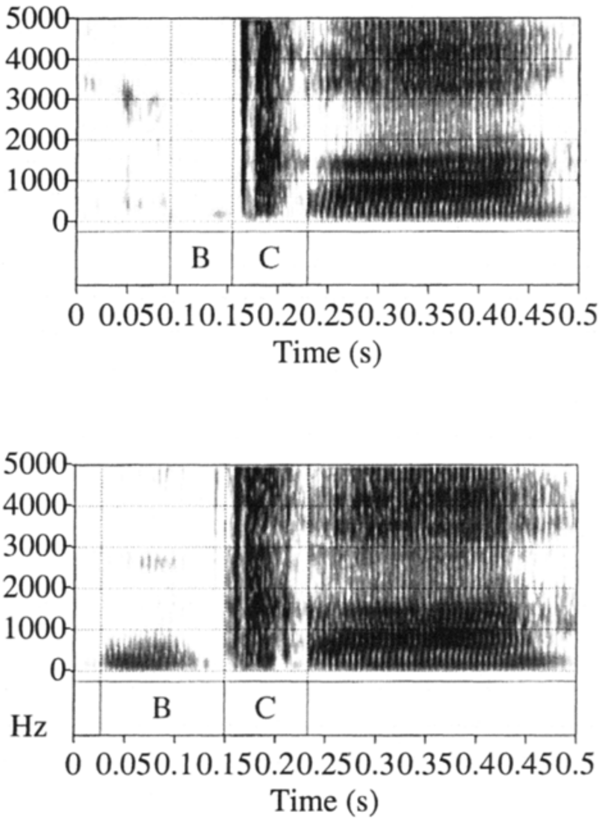


Figure 13. Spectrograms of the voiceless uvularized click-initial root *ǂᵡà̱à̱* ‘moist sand’ (above) and the voiced uvularized click-initial root *g!ʶàà* ‘to take out’ (below, Subject KK)

noise in the tongue dorsum lag phase, which carries over into the following vowel. They are also articulatorily similar in that they each involve a spread glottal posture, resulting in concomitant higher airflow.

Pitch tracks associated with a voiceless uvularized click-initial root, and voiced uvularized click-initial root are provided in Figure 15. In both root types, there is a very slight fall in F0 at the beginning of the root followed by a relatively

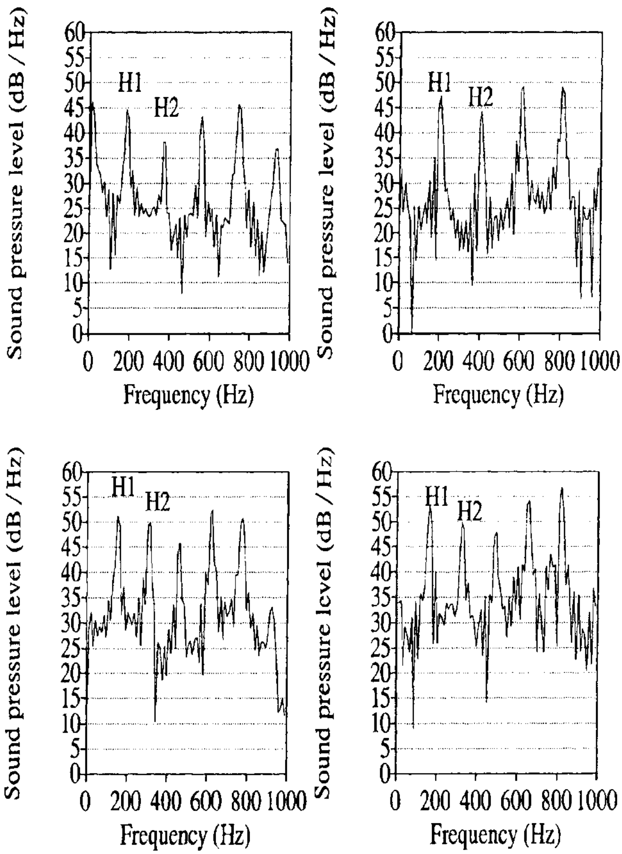


Figure 14. Spectra at the ¼ and ¾ points of the vowel in the voiceless uvularized click-initial root *ǂ$^{\chi}$àà* ‘moist sand’ (above) and in the voiced uvularized click-initial root *g!$^{\text{ʁ}}$àà* ‘to take out’ (below, Subject KK, 50 ms Hanning window)

stable F0 pattern over the rest of the vowel, until the fall in F0

[1] In preliminary fiberscopic experiments with myself as the subject, the glottal posture was even more wide open during the articulation of uvularized consonants than it was during the articulation of aspirated consonants.

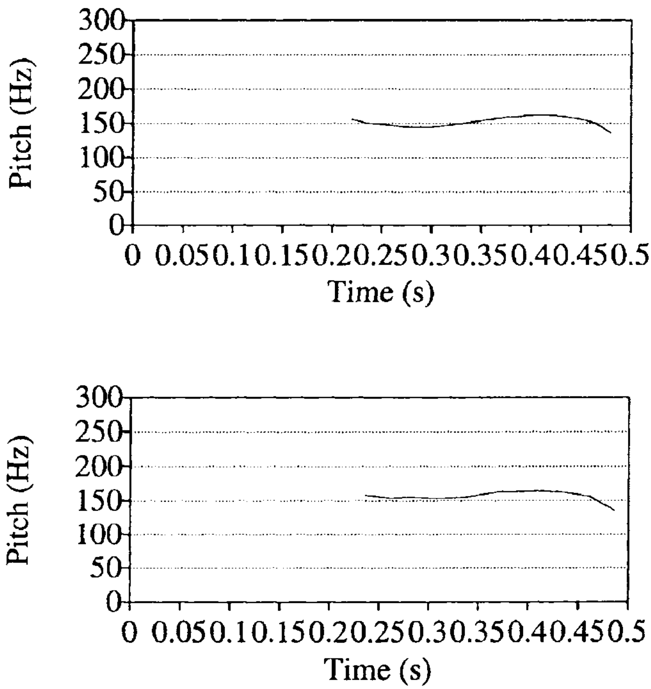


Figure 15. Fundamental frequency tracks associated with the voiceless uvularized click-initial root ǂ$^{\chi}$*àà* 'moist sand' (above) and the voiced uvularized click-initial root *g!*^{ʁ}*àà* 'to take out' (Subject KK)

associated with the utterance final position occurs over the last 50 msec. There do not seem to be F0 differences found between voiceless and voiced uvularized clicks similar to the ones found in voiceless and voiced aspirated clicks. This is expected given the lack of voice lag differences seen in the spectrograms in Figure 13.

In the glottalized clicks, the noise associated with the stop burst is clearly visible before it fades off into silence, as can be seen in Figure 16. The glottalized click-initial roots used in this study contain nasalized vowels, because nasalized vowels always follow glottalized clicks in Ju|'hoansi. The nasalization

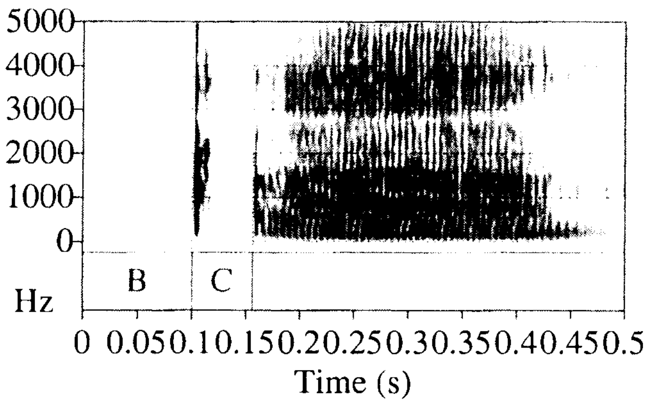


Figure 16. Spectrogram of the glottalized click-initial root *!*^{ʔ}*àà*^{n} ‘to catch up’ (Subject KK)

can clearly be seen through the widening of the first formant bandwidth in the vowel. Perhaps nasal venting is used to offset some of the pressure, resulting in a softer attack of the following vowel, similar to that found in voiceless nasal aspirated clicks in Ju|’hoansi (Traill, 1992).

The epiglottalized clicks involve a period of uvular frication which is followed temporally by a period of glottal abduction. This results in a silent interval, as shown in the spectrogram of the word *!*^{ʔ}*àà* ‘to dry out’ in Figure 17. The uvularization follows the click burst. The uvularization noise is cut off much more abruptly than it is in the uvularized clicks, which is a result of the epiglottal sphincteric action leading to a harder attack. The attack is much harder than that found in the plain glottalized click-initial roots shown in Figure 15. The strong attack can be seen by the fact that the onset of energy in all three formants of the vowels occurs at once, and the formants have high amplitude, as seen by the darkness of the formants at their onsets. Compare this to the onset of the formants following the click bursts in unaspirated clicks in Figure 7, aspirated clicks in Figure 10, and uvularized clicks in Figure 13, where the first formant appears first, followed by

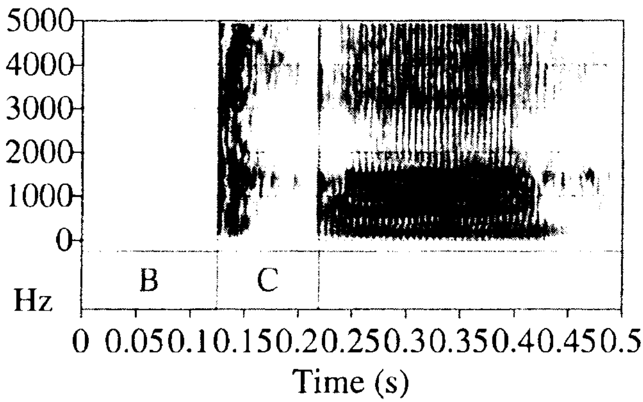


Figure 17. Spectrogram of the voiceless epiglottalized click-initial root *!ᴴàà* 'to dry out' (Subject KK)

the second and third formants, with all formants gradually increasing in amplitude (e.g. gradually getting darker in the spectrogram).

The period of uvular frication is shorter in epiglottalized clicks than it is in the uvularized clicks, which can be attributed to the epiglottal stricture. Notice that the frication interval of the click also has more glottal pulses than do clicks with plain uvular frication, showing that the constriction which temporally follows the interval of frication already has an effect during the frication interval.[2] This may also be attributable to epiglottal trilling as Traill (1986) suggests occurs in the production of pharyngealized vowels in !Xóõ. The vowel following this click is quite short, with the result being similar syllable duration to voiceless uvularized click-initial roots.

Spectra associated with a glottalized click-initial root and an epiglottalized click-initial root are provided in Figure 18. As can be seen, the spectral slope values found in vowels following these clicks (approximately −1 dB/Hz and −3 dB/Hz respectively), defined via the amplitude of H1 minus the amplitude of H2, is much lower than the H1-H2 values found in vowels following voiceless and voiced unaspirated consonants, seen above in Figure 8. Recall that the spectral slope values following the voiceless and voiced unaspirated clicks were about 1 dB/Hz and 3 dB/Hz respectively. The negative H1-H2 values show that the spectral effects of glottalization and epiglottalization are observed on the following vowel for each of these clicks. In glottalized clicks, the smaller spectral slope is presumed to be due to the constriction at the glottis. However, the epiglottalized clicks are produced with a constriction in the pharyngeal region in my productions, formed with pharyngeal narrowing and false vocal fold contraction (Miller-Ockhuizen, In progress). The production of these sounds is similar to Esling's (1999b) productions of pharyngeal stops, although in Ju|'hoansi the constriction is only in the release portion of the consonant. A constriction near the glottis would have the same acoustic effect as a constriction at the glottis. Further research on the production of these sounds by native speakers is necessary to confirm the articulatory mechanisms used in the production of clicks with epiglottalized consonants. The less pronounced low spectral slope values found following glottalized clicks when compared with epiglottalized clicks might be explained by the nasal venting presumed present in glottalized clicks. Nasal venting would lead to a softer attack than is found in vowels following epiglottalized clicks, as seen in the spectrograms.

In glottalized click-initial roots and epiglottalized click-initial roots, there is only a slightly positive spectral roll-off found in the ¾ points of the vowels. The spectral roll-off is about 3 dB/Hz for the glottalized click-initial root, and about 1 dB/Hz for the epiglottalized click-initial root. These values are similar to the 1 dB/Hz and 3 dB/Hz roll-offs that were found at the ¼ point of the vowel following voiceless and voiced

[2] In preliminary fiberscopic experiments with me as the subject, production of these sounds did not involve a constriction at the glottis, but rather a constriction in the pharynx using the false vocal folds.

unaspirated clicks. This suggests that the long range coarticulatory effect of the initial click, which decreases the spectral slope value, interacts with the prosodically conditioned breathiness, which would increase the spectral slope at the end of the utterance. The end result is a fairly modal voice quality. Put into articulatory terms, there is a conflict between the long-range coarticulatory effect which results in glottal or epiglottal stricture, and prosodically-conditional glottal opening. The result is similar to the spectral slope values found in modal vowels at the ¼ point of the vowel.

Figure 19 provides pitch tracks associated with a glottalized and voiceless epiglottalized click-initial roots. The two pitch tracks have a very similar shape, with slightly raised F0 values immediately following the consonant attributed to the consonant effect, followed by a relatively stable F0 pattern over the vowel, and a steep fall at the end of the root, which is due to the roots being produced in utterance final position.

The voiced ejective clicks described by Snyman (1970, 1975), and transcribed by Dickens (1994) as [g!kx’] using the post-alveolar central click to stand for all click types. Snyman (1978, 1999) transcribes this click as a prevoiced click with an ejected velar fricative release. There is clear voicing in this token.

A spectrogram of a voiced epiglottalized click is provided in Figure 20. Note that this token was produced by a different subject than the rest of the data provided in this section. As with modally voiced clicks, the closure voicing in voiced epiglottalized clicks is often somewhat nasal, reflecting opening of the velum in order to ease the articulation of voicing with a double constriction in the oral cavity. The closure voicing seen in Figure 20 is somewhat nasalized.

As can be seen, the voicing and epiglottalization are phonetically ordered, with the voicing being in the closure just as with modally voiced clicks, and the epiglottalization occurring in the release portion of the click. As noted by Snyman (1978), unlike with voiced aspirates and voiced uvularized clicks, there is no voicing in the release part of the click. The lack of voicing in the release portion can be

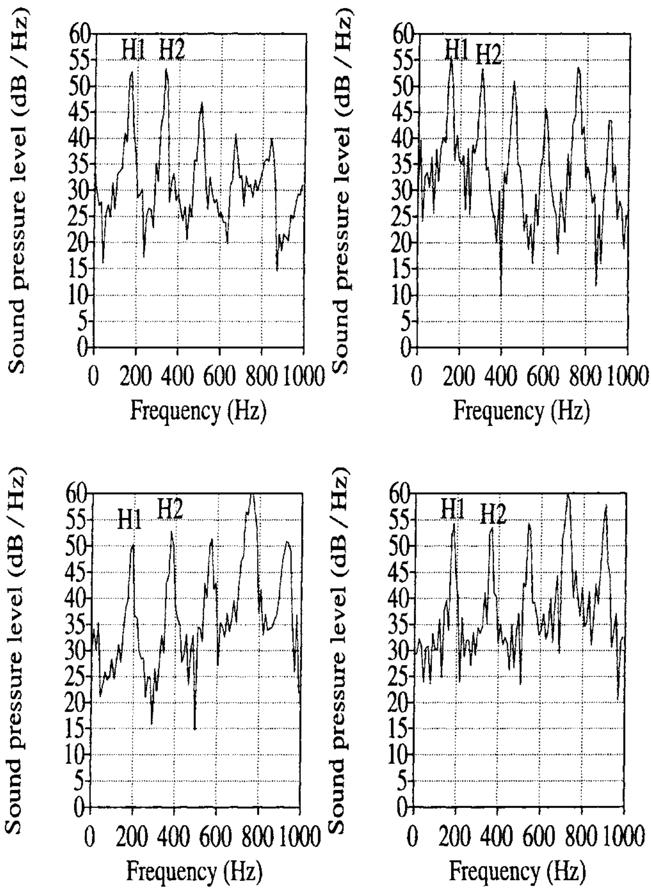


Figure 18. Spectra at ¼ and ¾ points of the vowel in the glottalized click-initial root *!^{ʔ}àà^{n}* ‘to catch up’ (above) and in the voiceless epiglottalized click-initial root *!^{H}àà* ‘to dry out’ (Subject KK)

attributed to the incompatibility of voicing with larynx raising. These segments can then be analyzed as being phonologically voiced throughout. Voiced epiglottalized clicks were left out of the quantitative study presented in Part II because of their marginal status, which made it difficult to get enough minimal pairs for the quantitative investigation.

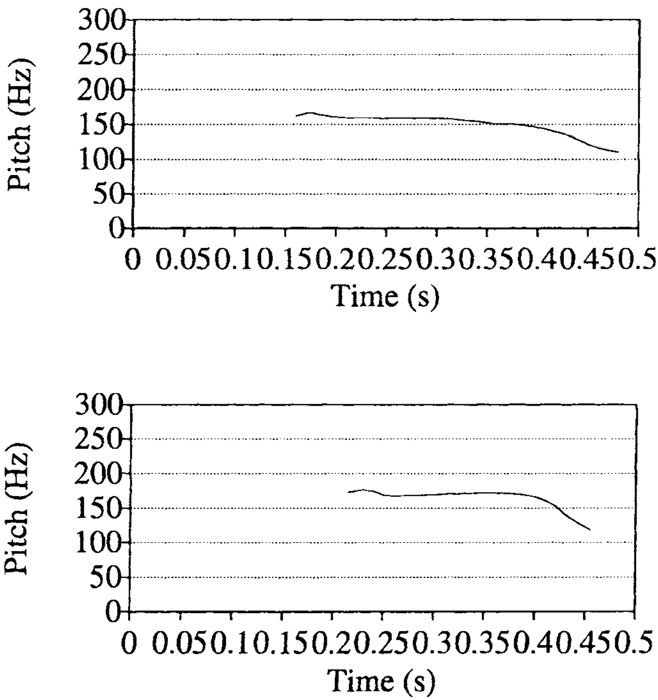


Figure 19. Fundamental frequency tracks associated with the glottalized click-initial root *!ˀàà*ⁿ ‘to catch up’ (above) and the voiceless epiglottalized click-initial root *!^{H}àà* ‘to dry out’ (Subject KK)

The F0 trace associated with the same root is also included in Figure 20. As can be seen, the F0 pattern falls rather drastically at the beginning of the vowel. This is a much steeper fall (about 100 Hz) than the one that occurs after the voiceless epiglottalized click and the glottalized click, shown above in Figure 19.

The natural class of guttural consonants that is targeted by several phonotactactic constraints described in Chapter 3 includes voiceless and voiced aspirates, voiceless and voiced uvularized consonants, glottalized consonants, and epiglottalized consonants. These guttural consonants are all

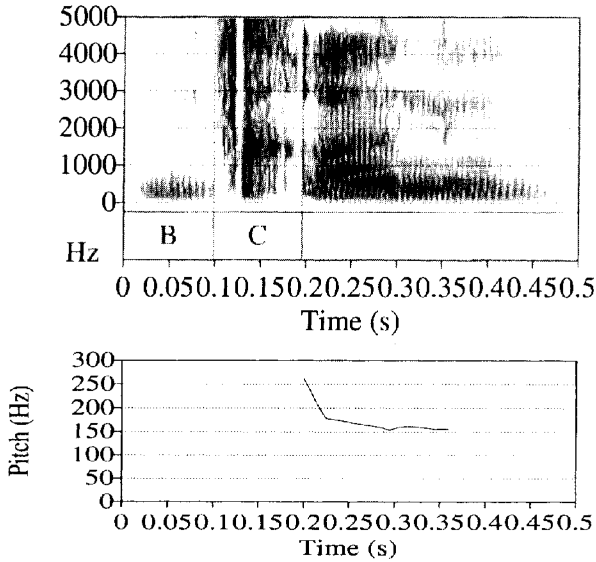


Figure 20. Spectrogram and pitch track of the voiced epiglottalized click-initial root *g|ˤànì* ‘tick’ (Subject KX)

characterized by the presence of noise within the third formant range that occurs in the C-V transition following the noise-burst of the click, and that remains after the onset of voicing in the following vowel. They also all exhibit either high or low spectral slope values (measured by H1-H2) relative to values found on vowels following voiceless and voiced unaspirated consonants. In contrast, the voiced unaspirated clicks and nasalized clicks display low amplitude voicing during the closure preceding the burst, and no noise following the burst (e.g. in the tongue dorsum lag phase of clicks and the release phase of pulmonic stops). Thus, vowels following voiced consonants also do not show any relatively high or low spectral slope values. In these non-guttural clicks, the signal becomes periodic very quickly after the click noise-burst. The voiceless unaspirated clicks show no acoustic properties in the closure (overlap phase) or release (lag phase) relating to phonation, which is why they are considered unmarked on phonetic grounds. This probably also explains their phonological behavior as unmarked sounds (Lombardi, 1994; Bradshaw, 1999). Thus, the acoustic properties of clicks and the temporal position of the acoustic cues present in their acoustic signatures are capable of predicting on phonetic grounds what counts as a guttural consonant, and what counts as a non-guttural consonant in Ju|'hoansi phonology.

2.3.4. Phonetic Properties of Sonorant Consonants

As is clear from the consonant chart in Table II, there are only 6 sonorant non-nasal consonants in Ju|'hoansi. All of these occur in root-medial position, and can thus be analyzed as allophones of initial pulmonic plosives that occur at the same place of articulation. In Chapter 3, I will also discuss evidence from loan-word adaptation that supports this analysis.

Nasals occur in all three consonant positions within the root, and do not differ as much in initial vs. medial position phonetically. In root-final post-vocalic position, only the labial nasal occurs. In this section, I will describe the phonetics of medial vs. initial and final consonants. I will point out evidence in favor of viewing all of the medial and final consonants as sonorants. While nasals occur in all root-positions, they are much more frequent in medial and final positions, as shown in Chapter 3.

The spectrograms in Figure 21 show both a root-initial labial nasal, and a root-medial labial nasal. As can be seen in the spectrograms, the medial nasals are shorter than the corresponding initial nasal consonants. Otherwise, the two labial nasals are similar. Ladefoged (1997) claims that the opposition of [obstruent] vs. [sonorant] is an auditory feature, and points out that sonorants display periodic energy with well-defined formants. Ju|'hoansi labial nasals in both initial and medial positions are clearly sonorants by this definition, although the formants are more clearly visible in this token of an initial nasal, than in this token of a medial nasal. This is a result of the dynamic range. In click-initial roots, a dynamic range of 50 dB often leads to some of the energy in the rest of the root being canceled out.

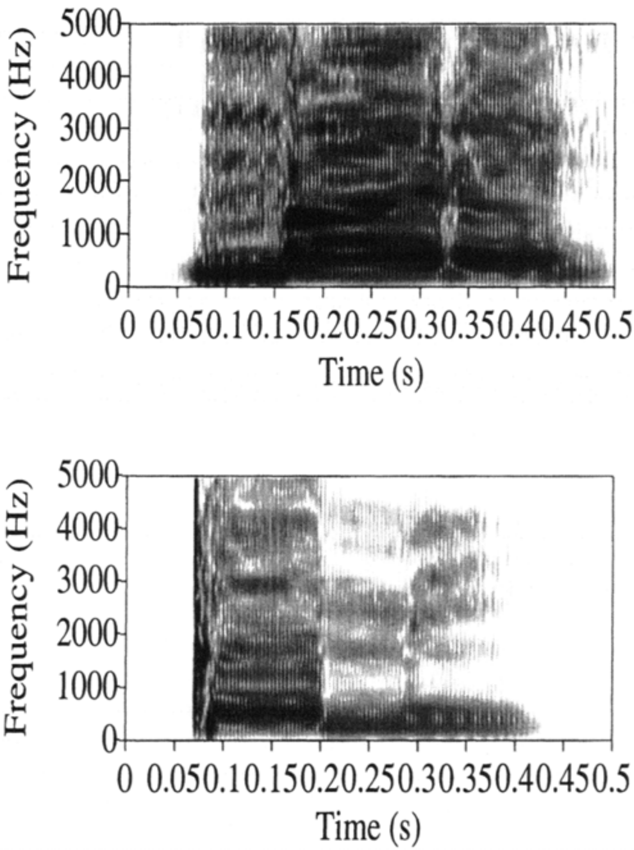


Figure 21. Spectrograms of root-initial [m] in *mə̀rè* ‘bread’ (above) and root-medial [m] in *!ə́mí* ‘tree trunk’ (below, Subject KK)

Spectrograms of roots containing initial and medial coronal nasals [n] are provided in Figure 22. In both of these nasals, the formants are very clear throughout. In this case, the formants are clearer in the medial nasal than in the initial nasal. I take

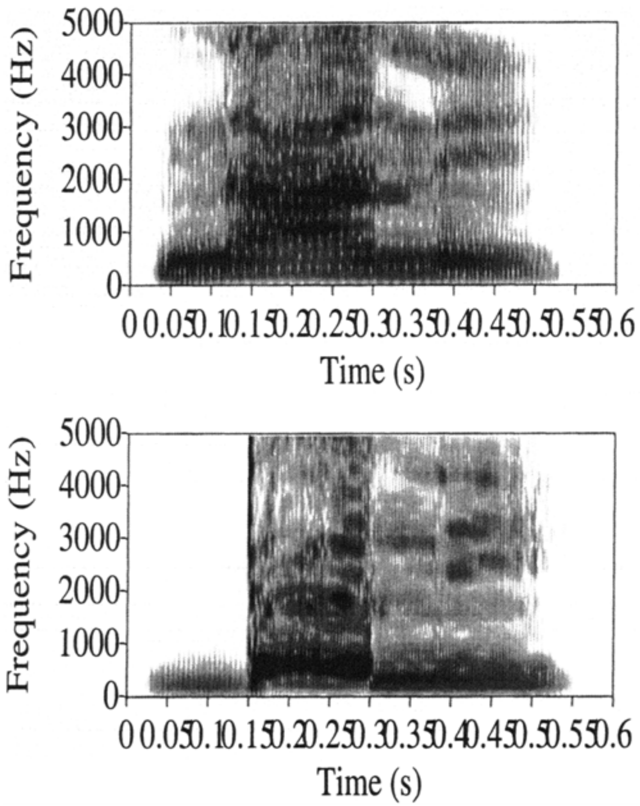


Figure 22. Spectrograms of root-initial [n] in *nȁˤànì* ‘needle’ (above) and root-medial [n] in *dȁˤànì* ‘container’ (below, Subject KK)

the positional differences found between [m] and [n] in the spectrograms provided here as artifacts of the particular productions shown.

The medial coronal flap [ɾ] is visible in the root *mə̀ɾè* ‘bread’ in Figure 23. Voicing is maintained throughout the closure associated with the flap, and all three formant frequencies are visible throughout. A flap is also found in the production of *bə̀ɾè* ‘goat’ in Figure 23. The spectrograms in Figure 23 also show

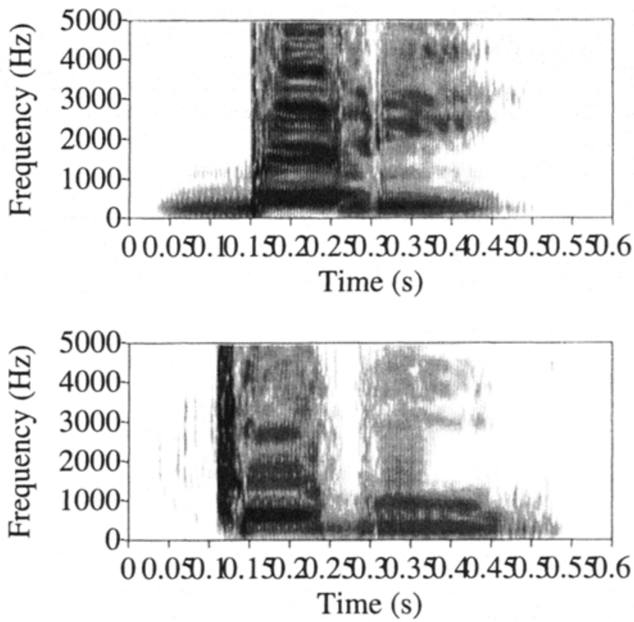


Figure 23. Spectrograms of root-initial /b/ [b] in *bə̀ɾì* 'goat' (above) and root-medial /b/ [β] in *ǁə́βú* 'blood letting horn' (below, Subject KK)

roots exhibiting initial [b] and medial [β]. The medial [β] in the token in the lower panel is a bit longer than the medial coronal flap and nasal. It is similar in length to root-medial [m]. Therefore, the increased length seems to be due to the labial place. The voicing is clearly maintained during the entire closure, and there is frication noise throughout the duration.

In Figure 24, spectrograms of roots containing an initial voiced coronal plosive [d] and a medial coronal flap [ɾ] are provided. The first and third formant frequencies are clearly visible throughout the [ɾ]. There is also no clear burst present in the medial [ɾ]. Compare this to the voiced coronal plosive found in initial position. In the initial coronal obstruent, no formants are visible, and the only energy seen is the low frequency voice bar. The obstruent is released into a clear burst, as is typical of initial plosives in Ju|'hoansi.

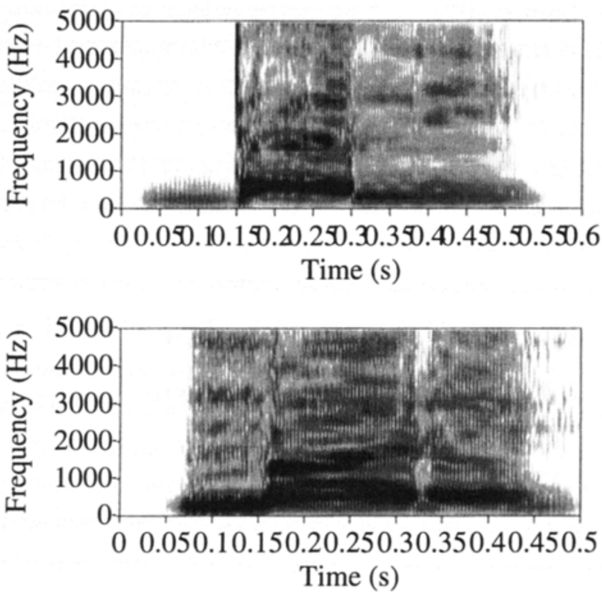


Figure 24. Spectrograms of root-initial /d/ [d] in *dȁˤànì* ‘container’ (above) and root-medial /d/ [ɾ] in *màɾè* ‘bread’ (below, Subject KK)

Spectrograms provided in Figure 25 illustrate the obstruent nature of initial [g] and the sonorant nature of medial [ʎ]. Notice that the initial stop releases into a clear burst, while the medial consonant displays periodic energy throughout its duration, without any noise burst.

Medial [m], [β], [n], [ɾ], [ŋ] and [ɣ] are all clearly sonorants by Ladefoged’s (1997) definition. Ladefoged notes that the feature [sonorant] could also be defined articulatorily as marking consonants that show vocal fold vibration, and no pressure build-up. However, he also notes that it would be difficult to imagine an articulatory connection between these two properties, and claims that the feature [sonorant] is an auditory feature. It is Ladefoged’s (1997) auditory definition of the feature sonorant that correctly groups Ju|’hoansi medial consonants. The classification of obstruents and sonorants discussed here is motivated independently by the phonological patterning of these consonants as described in Section 3.4.

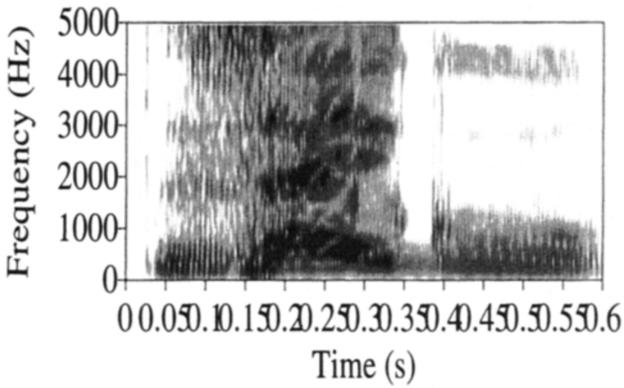


Figure 25. Spectrograms of root-initial /g/ [g] in *gö̀ʱòʱ* ‘grass species’ (above) and root-medial /g/ [ɣ] in *zä̀ˤìɣù* ‘black mamba’ (below, Subject KK)

2.4. The Vowel Inventory

The vowel inventory, which is often left unremarked in the face of the extremely complex consonant inventory, is very rich. Table V provides the inventory of monophthongs. The language has a basic 5-vowel system with just three levels of height, and a two-way contrast on the front/back (and rounded) dimension. While there is no contrast between front and back low vowels, the vowel /a/ clearly behaves as a back vowel with regards to the C-V place co-occurrence constraints described in Miller-Ockhuizen (2000), and in Section 3.5.3. The number of vowels is doubled by an oral-nasal contrast, and multiplied further by an independent four-way contrast in phonation type, with vowels being either modal, breathy, glottalized or epiglottalized.

In Table V, epiglottalization is marked with a superscript [ˤ], following the IPA conventions for marking epiglottalization on consonants. [ʢ] is the symbol for the voiced epiglottal fricative.

Nasalized vowels are transcribed with a superscripted ‘n’, breathy vowels are transcribed with a superscripted [ʱ], the symbol used for the voiced aspirated consonant, and glottalized vowels are transcribed with a superscripted [ˀ], the symbol for a glottal stop consonant, in between the two vowels, which are assumed to associate to the first mora. These superscripted symbols are preferred to standard IPA diacritics for nasalization, breathy and creaky voice, because tones are also marked with diacritic symbols. That is, using standard diacritics for the two vowel phonation types that the IPA marks, as well as for nasalization and tone, makes it difficult to decipher representations of many of the attested combinations of tone, nasality and phonation types found in Ju|’hoansi.

The three non-modal phonation type contrasts in Ju|’hoansi vowels are referred to in this dissertation as guttural vowels based on their patterning with guttural consonants in co-occurrence restrictions found between guttural consonants and vowels described in Section 3.5.1, and on their unified voice quality cues. That is, guttural vowels are similar in that they all contain noise in the third formant frequency range (which results in a decreased harmonics-to-noise ratio shown to exist in Part II), and they all exhibit relatively high or low spectral slopes. The retracted vowel qualities [a] and [ɔ] included in the IPA transcriptions of the epiglottalized vowels are allophonic, and these vowels are phonemically /a/ and /o/ respectively, allowing a parallel phonemic vocalic inventory for vowel backness across the different phonation types.

Figure 26 contains spectrograms illustrating each of the four contrastive vowel phonation types in Ju|’hoansi. It is difficult to see phonation type differences in spectrograms, but there is clearly less energy in the 2000–4000 Hz range in all of the non-modal phonation types compared with the modal vowels in the upper-most panel in Figure 26. The glottalized vowels look more like diphthongs, given the clear diminution of energy over the first half of the vowel, and the higher amplitude energy in the third, fourth and fifth formant frequencies over the second half of the vowel. In addition to the high-intensity, high-frequency noise, they note that the harmonic energy in the

	Front Vowels							Back vowels							
	Non-guttural Vowels		Guttural Vowels					Non-guttural Vowels		Guttural Vowels					
	Modal		Breathy		Glottalized		Epiglottalized	Modal		Breathy		Glottalized		Epiglottalized	
	Oral	Nasal	Oral	Nasal	Oral	Nasal		Oral	Nasal	Oral	Nasal	Oral	Nasal	Oral	Nasal
High	i	i^{n}	$i^{ɦ}$	$i^{ɦn}$	$i^{ʔ}i$	$i^{ʔ}i^{n}$		u	u^{n}	$u^{ɦ}$	$u^{ɦn}$	$u^{ʔ}u$	$u^{ʔ}u^{n}$		
Mid	e	e^{n}	$e^{ɦ}$	$e^{ɦn}$	$e^{ʔ}e$	$e^{ʔ}e^{n}$		o	o^{n}	$o^{ɦ}$	$o^{ɦn}$	$o^{ʔ}o$	$o^{ʔ}o^{n}$	$ɔ^{ʕ}$	$ɔ^{ʕn}$
Low								a	a^{n}	$a^{ɦ}$	$a^{ɦn}$	$a^{ʔ}a$	$a^{ʔ}a^{n}$	$ɑ^{ʕ}$	$ɑ^{ʕn}$

Table V. Ju|'hoansi vowel inventory[3]

[3] Based on phonological evidence discussed in Chapter 3 and Miller-Ockhuizen (2000), vowels in monosyllabic roots are all long. Vowels in each syllable of bisyllabic roots are usually short, although there are some bisyllabic roots that contain initial diphthongs.

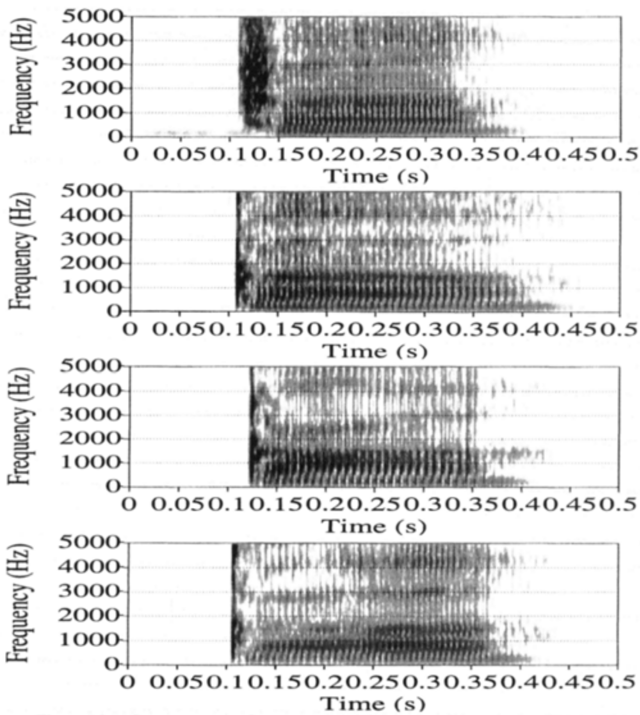


Figure 26. Spectrograms of roots exemplifying the four-way vocalic phonation type contrast: [a] in the root *ǁáá* ‘warm hands by fire’, [aʱ] in the root *!ä̋ʱä̋ʱ* ‘red crested korhaan’, [ɑˀ] in the root *!ä̋ˀä̋ˀ* ‘iron’ and [aˀa] in the root *!ä̋ˀà* ‘dry season’ (from top to bottom, Subject KK)

range between 3000–4000 Hz (in the range of the third formant) is shockingly low in [ɑˤ], making it difficult to identify the third formant.

In addition to the noise present in the third formant frequency range, all of the guttural vowels are similar in that they display high or low spectral slope values throughout the duration of the vowel relative to modal vowels following unaspirated initial clicks. The spectral slope values are seen clearly in the amplitude difference between the first and second harmonics (H1-H2) in the spectra in Figure 27. The breathy vowels display high spectral slope values relative to modal vowels, parallel to the relative differences found between vowels following aspirated consonants and vowels following voiceless unaspirated clicks shown in the previous section. Epiglottalized and glottalized vowels exhibit similarly low spectral slopes relative to modal vowels, parallel to vowels following glottalized and epiglottalized clicks. It is the similarity in the presence of noise, and the relatively high or low spectral slope values found in guttural vowels and vowels following guttural consonants that forms the perceptual basis of the Guttural OCP constraint in Ju|'hoansi (described in Section 3.5.1).

An articulatory motivation for the Guttural OCP constraint in Ju|'hoansi is not tenable. While articulatory incompatibility would account for the lack of co-occurrence between epiglottalized and glottalized vowels and aspirated or uvularized consonants, or likewise the lack of co-occurrence of epiglottalized and glottalized consonants with breathy vowels, which have opposing glottal gestures; it could not explain the blockage of co-occurrence of sounds that are articulated using the same gestures. That is, an articulatory grounding of the OCP constraint described in Section 3.5.1 would not be able to explain why breathy vowels do not co-occur with aspirated and uvularized consonants that involve the same spread glottal posture in their articulation. The co-occurrence of these sounds would be predicted to occur through gestural overlap.

The first formant frequency (F1) values can be seen clearly in the spectra, taken at the ¼ point of the vowel, in Figure 27. As can be seen, the F1 values of guttural vowels are all higher than F1 values found in modal vowels. The F1 value is about 650 Hz in the modal vowel, 750 Hz in the breathy vowel (100 Hz higher than the modal vowel), 800 Hz in the epiglottalized vowel and 730 Hz in the glottalized vowel. Thus, the F1 values in epiglottalized vowels are the highest, followed by breathy and glottalized vowels. That is, the F1 increase associated with epiglottalization is greater than that associated with differences in glottal aperture.

The F1 amplitude values associated with guttural vowels are much lower than F1 amplitude values found in modal vowels. The amplitude value of F1 in the epiglottalized vowel is 48 dB, in the glottalized vowel is 47 dB, and in the breathy vowel is 48 dB. These values are all much lower than the F1 amplitude of the modal vowel, which is 56 dB. Ladefoged and Maddieson (1996) also note the diminution of energy in the 400–700 Hz range (e.g. the first formant range) associated with pharyngealized vowels in !Xóõ. This property is present in all guttural vowels in Ju|'hoansi.

Another difference between the epiglottalized vowels and the other three phonation types, seen in the spectrograms in Figure 26, is the close proximity of the first and second formant frequencies in the epiglottalized vowel, which is a result of the raised first formant frequency. This is due to the extreme retraction of the tongue root and commensurate retraction of the epiglottis, which results in lower, more back vowels than are found in their non-epiglottalized counterparts. This is indicated by the different IPA symbols used for these vowels in the vowel chart. What cannot be indicated in the IPA symbols is the lowering in fundamental frequency that accompanies these vowels. Distributional patterns show that there are twice as many mid epiglottalized vowels as expected, which is consistent with some of these vowels being phonemic high vowels, or at least deriving historically from phonemic high vowels.

Figure 28 contains a scatter plot of the first formant frequency versus the second formant frequency for modal and epiglottalized vowels for two male and two female subjects (the same subjects used in the quantitative study presented in Part II). The [aq] transcription used in the legend is equivalent to [ɑˤɑˤ]. Notice that the first formant frequency associated with the epiglottalized vowels is much higher than that found in the modal vowels for all speakers, and there is little overlap between the two vowel types, indicating that this is a consistent difference among the vowel types. This F1 difference shows that these vowels are lower than their modal counterparts. The second formant frequency is also lower,

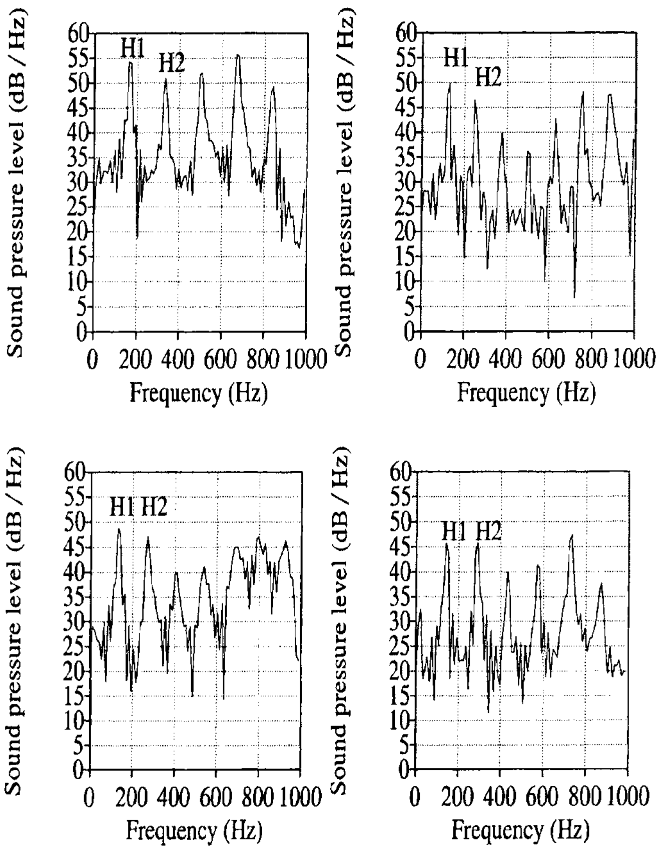


Figure 27. Spectra at the ¼ point of the vowel in the four-way vocalic phonation type contras: [a] in the root ǁ*áá* ‘warm hands by fire’ (upper left), **[aʱ]** in the root *!ȁʱȁʱ* ‘red crested korhaan’ (upper right), **[ɑˤ]** in the root *!ȁˤȁˤ* ‘iron’ (lower left) and **[aˀa]** in the root *!ȁˀà* ‘dry season’ (lower right, Subject KK, 50 ms Hanning window)

indicating a more back vowel. These values are consistent with the IPA transcriptions adopted for them, using the more retracted vowel symbols.

Modal and epiglottalized vowels also display large differences in their third formant frequency values, as indicated by the mean and standard deviation values provided in

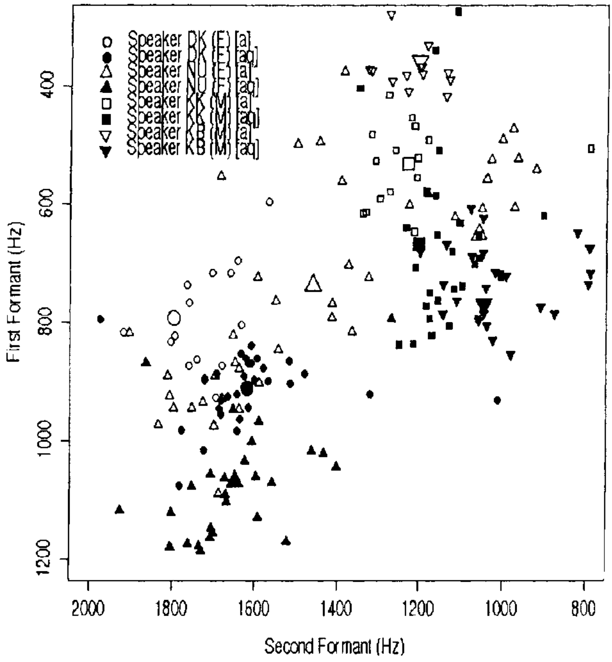


Figure 28. F1 and F2 values associated with modal and epiglottalized low back vowels (modal vowel symbols are unfilled and epiglottalized symbols are filled, means are plotted in larger symbols of the same type used for individual tokens)

Table VI. Notice that Subject KK, who has the smallest F2 difference, also has the largest F3 difference. Given the low F3 values, auditory F2' values (Ladefoged, 1997 and references therein), which take into account the effect of higher formant frequencies in the calculation of F2, would likely be affected for all speakers, including Subject KK. The formant differences associated with epiglottalized vowels are the basis for the constraints between front and high vowels with epiglottal voice quality described in Chapter 3 (Section 3.5.3 and 3.5.4).

In addition to the four-way guttural vowel type contrast, there is also a dynamic contrast in the timing of the epiglottal

	Modal [a]	Epiglottalized [ɑˤ]
Subject DK	3291 (245)	2752 (233)
Subject NU	3507 (308)	2965 (491)
Subject KK	3449 (386)	2827 (370)
Subject KB	3388 (108)	2912 (238)

Table VI. Mean third formant frequency values for modal and epiglottalized low back vowels (standard deviations in parentheses)

	First Mora	*Both Moras*
Breathy	aʱa	aʱaʱ
Glottalized	aˀa	
Epiglottalized	ɑˤa	ɑˤɑˤ

Table VII. Monophthongs vs. diphthongs in voice quality

and breathy voice quality cues, resulting in a contrast between fully and partially breathy long vowels, and fully and partially epiglottalized long vowels, as shown in Table VII. The contours in epiglottalization and breathiness are considered diphthongs in voice quality, parallel to diphthongs in vowel quality such as [əi] and [əe], which also occur in the language.[4] Spectrograms illustrating the timing of four-way breathiness and four-way epiglottalization contrasts in the language are provided in Figure 34 and Figure 37 respectively, for male Subject KK and in Figure 35 and Figure 38 respectively for female Subject NU.

A spectrogram of a root containing a diphthong in breathiness in the upper panel of Figure 29 displays dynamic change in the amount of spectral noise in all of the three formant frequencies, as well as in the noise in the third formant range. Notice that the third and fourth formant frequencies are

[4] The status of the diphthongs [ui], [oe], and [oa] is unclear, since preliminary F1 measurements show that the initial onglide does not differ in height before high and non-high vowels as suggested in the transcriptions.

hardly visible over the first half of the vowel, but become clear in the second half of the vowel, which is an indication of the amount of noise present. The diphthongs in epiglottal quality show dynamic change from a more noisy quality in the V1 position, to a clear quality in V2 position, as well as in the movement of F1 from very high to lower, as can be seen in the spectrogram in the lower panel of Figure 29.

Ladefoged and Maddieson (1996) note the pharyngealization present in the first half of the vowel that is not present in the second half of the vowel for a similar token in the related language !Xóõ. Given that !Xóõ has two vowel phonation types, called pharyngealized and strident, it is not clear whether epiglottalized vowels in Ju|'hoansi correspond to the pharyngealized or the strident vowels in that language. Traill describes strident vowels as simultaneously pharyngealized and breathy, which predicts that they should have a high spectral slope value). However, results of the quantitative study reported on in Part II of this dissertation suggest that the Ju|'hoansi vowels are more like the pharyngealized !Xóõ vowels.

Figure 30 provides spectrograms of the two diphthongs in voice quality produced by female subject NU. The timing of breathiness for her production of the partially breathy vowel is similar to Subject KK's productions. The quantitative results reported in Part II also show that there is complete overlap of the spectral slope and HNR patterns found in voiced aspirated click-initial roots and roots containing diphthongs in breathiness. However, as can be seen in the spectrogram in the lower panel of Figure 30, Subject NU's partially epiglottalized vowel shows a period of somewhat modal phonation following the click burst before the diminution of the second and third formant frequencies over the second quarter of the vowel.

Diphthongs in breathiness and epiglottal quality all rise in periodicity, being more aperiodic in the first part of the vowel. This means that the noise is closer to the release of the consonant, and potentially more confusable with the coarticulatory effect on voice quality of a following vowel from a proceeding aspirated or epiglottalized consonant. However,

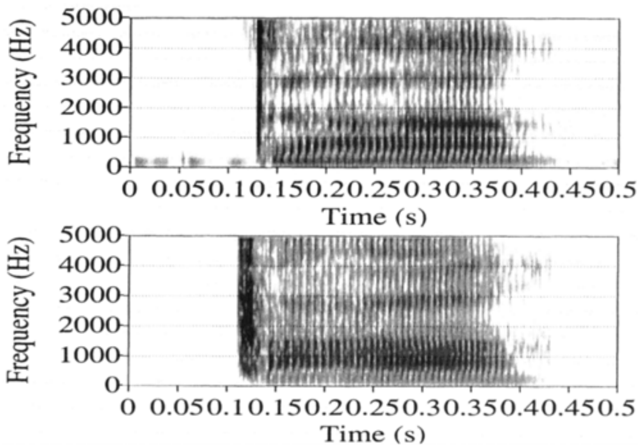


Figure 29. Spectrograms of roots containing diphthongs in voice quality: *ǂä̤ʱà* ‘plain’ (above) and *ǁɑ̀ˤá* ‘to hold down’ (below) (Subject KK)

the strong aspiration noise or uvularization noise that occurs before the onset of voicing in the vowel following guttural consonants clearly distinguishes them from guttural vowels. Thus, the Guttural OCP constraint active in the language assures that this cue to the different root types is maintained, by disallowing the co-occurrence of guttural consonants and vowels.

Silverman (1997) notes that languages containing different phasing patterns of laryngeal gestures usually adopt a more maximal phasing contrast in their inventory, before including a less optimal phasing contrast. The glottalized vowels seen in Figure 26 are phonetically diphthongs in voice quality, as seen by the diminution of energy in the third and fourth formant frequencies over the first half of the vowel, compared with the second half of the vowel. However, there is no phonological contrast, leaving the possibility of analyzing this as a fully glottalized vowel open. The contrast between glottalized consonants and vowels in Ju|’hoansi is, then, contra Silverman’s (1997) claim that languages will employ maximally contrastive phasing contrasts before less ideal

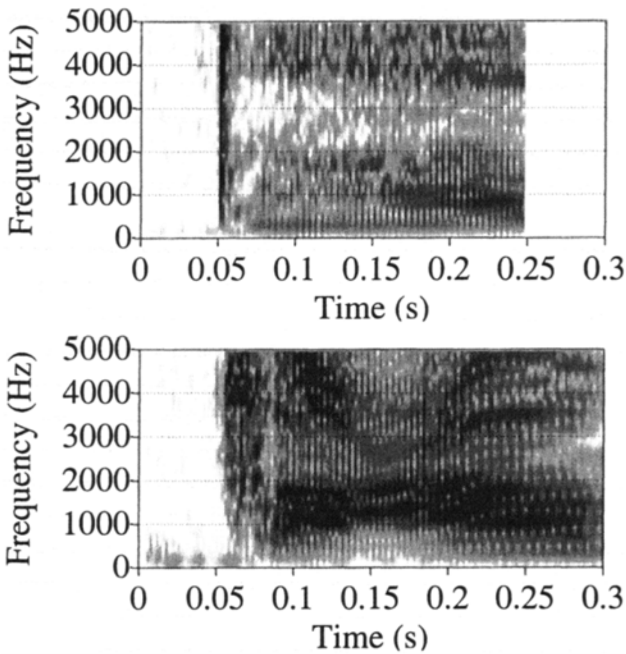


Figure 30. Spectrograms of roots containing diphthongs in voice quality: ǂȁʱà ‘plain’ (above) and ǁɑ̀ˤá ‘to hold down’ (below) (Subject NU)

phasing contrasts, since the language does not contain a glottalized vowel type where the glottalization is realized on the second half of the vowel, or the entire vowel. A contrast between a glottalized consonant and a glottalized vowel having glottalization realized over the second half of the vowel would be maximally contrastive, but this type of vowel does not exist. Comparing the spectrograms of a glottalized click-initial root in Figure 16 to the spectrogram of the root containing a glottalized vowel in the lowest panel of Figure 26, we see that the principal difference is in the extent of the glottalization, and the VOT of the consonants. However, Subject NU’s productions seen in Figure 31 do, however, seem to exhibit sufficient contrast in terms of the phasing of the epiglottalization. That is, since the epiglottalization is at its peak more towards the second quarter of the vowel than the first quarter of the vowel for this speaker's productions, the phasing does seem to be sufficiently discriminable. The Ju|'hoansi contrasts then note the import of understanding the phonological constraints active in a language in assessing the discriminibility of sounds within the inventory. That is, sounds do not need to be discriminable in contexts where they are blocked in the phonology from occurring. The Ju|'hoansi inventory also provides evidence that discriminibility of consonant and vowel contrasts, like discriminibility of consonant contrasts (Lindblom and Maddieson, 1988), are subject to sufficient discriminibility rather than maximal discriminibility.

Guttural vowels, just like guttural consonants, exhibit differences in fundamental frequency (F0) when compared with modal vowels, although the effect is largest with the epiglottalized vowels. Given the strong co-occurrence constraints found between guttural vowel types and tone patterns discussed in Section 3.5.2.2., it is not possible to get minimal pairs contrasting solely in tone. Therefore, tone patterns associated with guttural vowel types are presented pair wise, and in many cases use data from the recorded database described in Section 3.1.

Just as seen with the first formant frequency, fundamental frequency (F0) patterns on epiglottalized vowels are very different from those found on modal vowels, and the effect on F0 is the most extreme of all of the phonation types. Figure 31 provides mean F0 plots over the interval from the posterior release of the click consonant to the end of the vowel for each of the four vocalic phonation types. Both the raised F1 and lowered F0 values to modal vowels are consistent with this phonation type having the greatest constriction. These F0 values were computed directly from the marked period durations that had been hand-checked (described in Section 4.2.2), by taking the inverse of the average of the duration of three subsequent periods. The F0 values are plotted against the begin time of the middle of these three periods in Figure 31. The real time was used, subtracting the time of the posterior release of the click, in order to normalize for different begin times in the file. Data from female Subject NU is shown since the F0 differences are clearest in her speech, given her extremely large pitch range (200–450 Hz). Note that the root containing the modal vowel has a level Low lexical tone, while the other phonation types bear Super-Low tones.

The Low toned root containing the modal vowel displays a level F0 at about 300 Hz. All of the roots containing non-modal vowel types bear Super-Low tones on the first moraic position. As can be seen in Figure 31, the epiglottalized vowel is about 20–30 Hz lower in F0 over the first part of the vowel than the breathy vowel that bears the same lexical tone. The roots with fully epiglottalized vowels also exhibit a dipping F0 contour that is particularly strong over the first half of the vowel. The F0 then rises toward the end of the vowel. It is possible that the lowering and raising of F0 found in Ju|’hoansi roots with epiglottalized vowels is due to the complex interaction between lowering associated with the aryepiglottic constriction, and raising due to vocalic muscle contraction described by Fujimura and Lindqvist (1971). This might explain the minimal drop in H1-H2 found in glottalized vowels in the quantitative study presented in Part II. The rise in F0 found at the ends of roots of this type is interesting, given that most root types when recorded in isolation as these forms, exhibit marked F0 lowering and a lapse into breathiness, as was shown in Figure 2 above, which I attribute to utterance final position. Part of this disparity is due to the fact that that only the periodic portion of the vowel is shown in this figure. That is, the F0 values are only shown during the periodic portion of the vowel. While the entire vowel, including the last 50 msec. or so of breathiness found in utterance final position is shown in the individual pitch tracks provided in Section 2.3 above and in Figure 2. Still, there are additional effects of root type. This root type then exhibits a complex interaction between F0 requirements of the lexical tone, F0 requirements associated with the lexical phonation type, and F0 requirements of the higher-level prosodic context. As can be seen in Figure 31, there is a begin of F0 drop at 170 msec. into the roots.

The overall shape of the F0 pattern found on the breathy vowel is comparable to that associated with the modal vowel, although it is about 40 Hz lower than the modal vowel. This is expected given the difference in lexical tone in these tokens. Note that the Super-Low toned epiglottalized vowel is even lower in F0 than the Super-Low toned breathy vowel, despite the fact that they bear the same lexical tone. The glottalized vowel, transcribed with an L-H tone pattern, does show a consistently rising tone that far exceeds the gentler slope found in the epiglottalized vowel, consistent with roots of these types bearing rising tones. I will return to the F0 patterns found on epiglottalized vowels in the next section when I discuss the contrast between epiglottalized monophthongs and diphthongs having epiglottalization on the first vocalic position and modal voicing on the second vocalic position.

Figure 32 provides F0 traces of breathy vs. modal vowels that both contain Super-Low lexical tones produced by male Subject KK. These tokens are taken from the recorded database described in Section 3.1. The [Caah] transcription in the legend refers to [Ca^{ɦ}a^{ɦ}] roots, containing fully breathy vowels. The number of tokens is thus constrained by the type frequency of roots containing SL tones and modal and breathy vowels found in the known lexicon. There are thus only 9 tokens containing modal vowels, and 2 tokens containing breathy vowels, and initial voiceless unaspirated clicks, represented in Figure 32. As can be seen in the figure, there is just a slight difference in F0 (about 5–10 Hz) found between roots containing voiceless unaspirated consonants, and breathy vs. modal vowels bearing the same lexical tone, that occurs about 2/3 of the way through the vowel. There is also extreme overlap between the F0 values found over most of the duration of the vowel for the two root types. The F0 values are, however, quite distinct right after the click burst, showing that the breathy vowel phonation type masks the slight fall in F0 usually found at the beginning of the root in voiceless unaspirated click-initial roots (Figure 9). Notice also that the

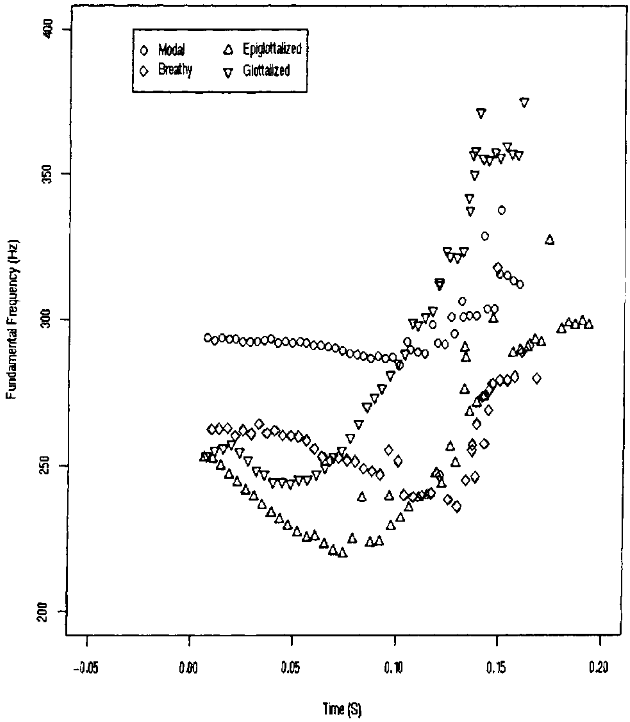


Figure 31. Mean F0 patterns found over 15 repetitions of the four-way vowel phonation types in the roots: ǂòà ‘sitting mat’, !ȁʱȁʱ ‘red crested korhaan’, !ȁˤȁˤ ‘iron’ and !ȁˀà ‘dry season’ (Subject NU)

breathy vowel roots are slightly longer in duration than the modal vowel roots. This same male subject’s (Subject KK) productions display an approximately 30 Hz difference between roots containing modal vowels, and roots containing epiglottalized vowels, as shown in Figure 33. There are 3 SL roots containing modal /a/ vowels and voiceless unaspirated clicks, and 34 Super-Low toned roots containing epiglottalized /a/ vowels, [ɑˤ], and voiceless unaspirated initial clicks. As can be seen in Figure 33, the epiglottalized roots in this figure display a dipping F0 pattern just as seen in Figure 31. Also, in

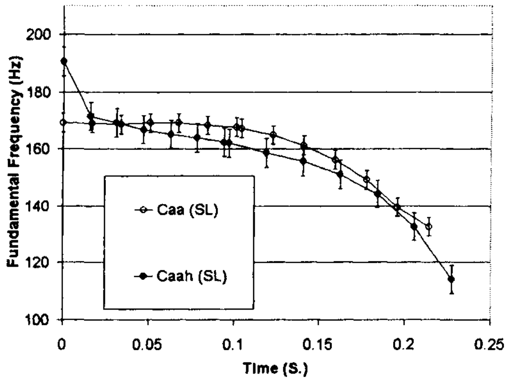


Figure 32. F0 patterns on Super-Low toned roots containing modal and breathy vowels (Subject KK)

Figure 33, the expected fall in F0 due to utterance final position is seen clearly. The roots containing epiglottalized vowels are longer, probably because it takes longer to realize the dipping F0 pattern associated with the vowel phonation type, and the fall in F0 associated with the prosodic position.

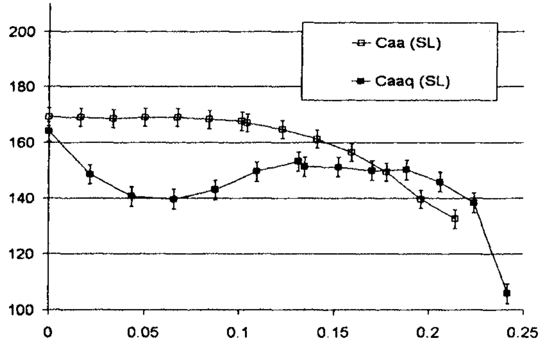


Figure 33. Super-Low root tone patterns on roots containing modal and epiglottalized vowels (Subject KK)

2.5. Parallels between Guttural Consonants and Vowels

In this section, I discuss the parallels between consonantal and vocalic phonation types in Ju|'hoansi. I interpret the contrast between voiced aspirated stops, breathy monophthongs and partially breathy diphthongs as a timing contrast. I discuss the differences previously noted between Ju|'hoansi voiced aspirated consonants and those found in Hindi, which have been better studied. I argue that the differences found are due to differences in the strength of the glottal opening gesture, which result in a period of voicelessness.

Kagaya and Hirose (1975) and Stevens (1998) describe the difference between voiceless and voiced aspirates as a difference in the timing of the glottal abduction and the oral constriction for a stop. In voiceless aspirates, the glottal abduction begins at the same time as the stop's oral constriction, and ends well after the oral closure is released. In voiced aspirates in Hindi, the glottal abduction starts either just before or right at the release of the oral constriction. Snyman (1970, 1975), Maddieson (1984), and Ladefoged and Maddieson (1996) note that Ju|'hoansi voiced aspirates differ from the Hindi ones in that the release is acoustically completely voiceless in Ju|'hoansi voiced aspirates. Ladefoged and Maddieson (1996) argue that the change in voicing within a single segment is universally prohibited, and thus Ju|'hoansi sounds must be clusters. While it is true that there is a period of voicelessness just following the release of the click in Ju|'hoansi voiced aspirates, part of the aspiration is definitely voiced, resulting in a shorter VOT lag for the voiced aspirates than the voiceless aspirates. Recall the 73 msec. vs. 45 msec. VOT lag values see in the spectrograms in Figure 10, and the VOT lag differences seen in the quantitative study in Part II. I attribute the period of voicelessness at the stop release (whether it be a pulmonic plosive or a velaric plosive) in Ju|'hoansi to a larger magnitude glottal opening than that found in the analogous Hindi voiced aspirated stops. Alternatively, the differences might be attributed to VOT differences associated with differences found in place of articulation, e.g. between click types and pulmonic stop place of articulation (Maddieson, 1997). The glottal abduction must start after the oral constriction is produced, just as it does in the Hindi stops, but the glottis must be abducted well before the oral constriction is released. This results in a period of voiced stop, followed by a period of voiceless stop, then a period of voiceless aspiration noise, and finally a period of voiced aspiration noise when the glottis is adducted again. The glottal area in Hindi is evidently small enough to allow voicing to be maintained throughout the consonant, while in Ju|'hoansi the glottis must achieve a greater peak glottal opening before the release of the oral constriction. This hypothesis is supported by the acoustic data that will be discussed in Part II. I analyzed these consonants as phonemic voiced aspirates, with a relatively large glottal opening gesture resulting in a period of voicelessness.

Breathy vowels are similar to voiced aspirated consonants in that the glottis is partially abducted during their production, allowing an increased level of airflow compared with modally voiced vowels. Breathy voiced vowels differ from aspirated consonants only in the timing of the glottal abduction with respect to the oral constriction of a preceding consonant, and perhaps in the magnitude of the glottal abduction gesture.

With these characteristics of the guttural specifications related to glottal abduction, the contrast between voiced unaspirated stops, voiced aspirated stops, fully breathy vowels, and partially breathy vowels can be seen as a four-way timing contrast involving the latency of the glottal abduction relative to the release of the oral constriction in the initial voiced stops and any subsequent adduction. That is, a voiced unaspirated click-initial root containing a modal vowel involves only a low magnitude glottal adduction gesture at the beginning of the consonant that is maintained throughout the vowel. Acoustically, the glottal abduction results in aspiration noise. Minimal sets ranging from earliest to latest timing of the glottal abduction/adduction sequence relative to the stop burst

[a] context		[o] context	
g!áá	‘rain’	g!óó	‘petrol’
g|ʱàà	‘place for drying meat’	g!ʱòó	‘to sit’
ǂȁʱà	‘plain’	|ȍʱò	‘to follow’
!ȁʱȁʱ	‘red crested korhaan’	‖ȍʱȍʱ	‘annual ground clover’

Table VIII. Data illustrating the four-way aspiration contrast

are provided in Table VIII. Spectrograms of the four-way aspiration contrast in the [a] context are provided in Figure 34.

Spectrograms showing the three different root types containing breathiness/ aspiration are provided in Figure 34, along with a root type containing unaspirated consonants and modal vowels for comparison. Unfortunately, voiced unaspirated click-initial roots containing partially breathy vowels were not collected in this study. Therefore, voiceless unaspirated click initials are used for roots containing the two breath vowel types. As can be seen, all three aspirated sounds exhibit noise in the 3000–5000 Hz frequency range over some interval within the vowel, which makes them very perceptually similar. This similarity is the reason why there are few known languages outside of Khoisan family that have both aspirated consonants and breathy vowels within their sound inventories. The Southeast Asian languages Suai and Wa are the only other languages that I am aware of that have both aspirated consonants and breathy vowels (Watkins, 1999; Abramson and Luangthongkum, 2001). All root type containing aspiration are, however, acoustically distinct. Notice the dynamic differences in the noise between 3000 and 5000 Hz in the three distinct root types containing aspiration. In the unaspirated consonant-initial root with the modal vowel, the third formant is clearly seen throughout the vowel. The root containing the initial voiced aspirated click displays a band of noise associated with the glottal abduction just after the release of the click consonant, but also displays noise in the third formant range throughout the remainder of the vowel. The noise is strongest just after the onset of voicing. Roots containing breathy vowels display noise just after the click burst, showing that the glottal abduction gesture begins at about the same time as in the voiced aspirated click-initial roots. The difference between the two root types is in the magnitude of the glottal abduction gesture. That is, the glottal abduction gesture is greater magnitude in the voiced aspirated click-initial roots than in the roots containing partially breathy vowels. The two root types also differ in the timing of the subsequent adduction gesture, with the adduction occurring earlier in the voiced aspirated click-initial roots.

The difference between fully and partially breathy vowels is in the timing of the subsequent glottal closure, with diphthongs reaching complete glottal closure by the end of the vowel, which may coincide with a smaller overall peak glottal opening. This timing difference can be seen via the timing of the noise patterns which are distinct in the partially and fully breathy vowels. In the fully breathy vowel, the third formant is somewhat clear at the beginning of the vowel, but becomes less clear toward the end of the vowel. In the partially breathy vowel, F3 is hardly visible before the midpoint, but becomes clearer over the second half of the vowel, similar to roots containing initial voiced aspirated clicks. However, the root containing the partially breathy vowel is differentiated from the root containing an initial aspirated click by the lack of aspiration noise before the onset of voicing, and thus the VOT differences found in the two root types. These dynamic differences will be elaborated on in the quantitative study reported in Part II.

Figure 35 provides spectrograms for a single female speaker, Subject NU. As was noted above with the partially epiglottalized vowel, we again see that this subject has a somewhat more clear voice quality just after the click release, and the breathiness reaches its peak in the second quarter of the vowel, showing that the glottal abduction gesture reaches its peak magnitude at the ¼ point of the vowel. This subject then allows more sufficient temporal contrast of the voice quality cues for breathiness, which would make Subject NU’s voiced aspirated click-initial roots and roots containing partially breathy vowels more perceptually distinct.

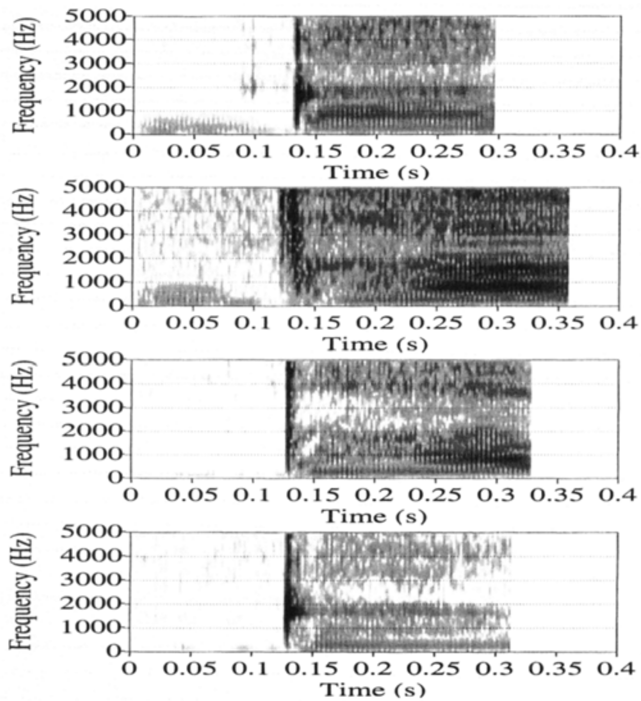


Figure 34. Spectrograms exhibiting the four-way aspiration contrast in the roots *g!áá* ‘rain’, *g|ʱàà* ‘place for drying meat’, *ǂȁʱà* ‘plain’, and *!ȁʱȁʱ* ‘red crested korhaan’ (top to bottom, Subject KK)

The noise differences are not as obvious in spectrograms as the relationship between the relative amplitudes of the first and second harmonics seen in spectra, which differ according to the voice quality of the vowel. The differences in the amplitude of the first and second harmonics are shown via spectra taken at the ¼ and ¾ points of the vowel for Subject KK in Figure 36. As can be seen, in the vowel following unaspirated consonants, there is a slight spectral roll-off, typical of modal vowels, which are presumed to have an approximately −6 dB per octave downward slope (Stevens, 1998). In the token shown here, there is an approximately 3 dB/Hz spectral roll-off. In the vowel following the voiced aspirated consonant,

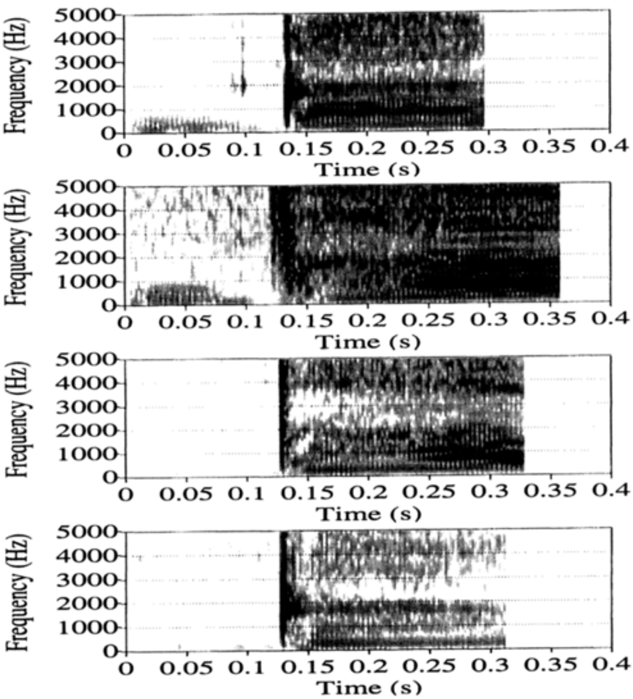


Figure 35. Spectrograms exhibiting the four-way aspiration contrast in the roots *g!áá* ‘rain’, *g|ʱàà* ‘place for drying meat’, *ǂä̀ʱà* ‘plain’, and *!ä̀ʱä̀ʱ* ‘red crested korhaan’ (top to bottom, Subject NU)

there is a slightly higher spectral slope at the ¼ point of the vowel. That is, the difference between H1 and H2 is greater in the vowel following voiced aspirated consonants than it is following a voiced unaspirated consonant. In this token, the spectral roll-off is about 8 dB/Hz. The larger H1-H2 value following voiced aspirated clicks relative to the value following voiceless unaspirated clicks, shows that there is still a larger open quotient after the onset of voicing in the vowel. In the partially breathy vowel, the difference between H1 and H2 is even greater at the ¼ point of the vowel (about 10 dB/ Hz) than that found in the vowel following the voiced aspirated click. The fully breathy vowel exhibits a value between the partially breathy vowel and the vowel following the voiced aspirated click at the quarter point of the vowel, with a value of about 6 dB/Hz.

At the ¾ points of the vowels, the differences are less marked among the different root types. In the voiced aspirated click-initial root, the spectral slope decreases by the ¾ point of the vowel, having a value of about 6 dB/Hz. In the root containing the partially breathy vowel, the ¾ point of the vowel also has about a 6 dB/Hz spectral slope value. In the root containing the fully breathy vowel, the slope increases towards the end of the vowel, and is larger at the ¾ point of the vowel than at the ¼ point of the vowel, having a value of approximately 8 dB/Hz. This suggests that roots containing fully breathy vowels are even more breathy at the end of the root than at the beginning of the root. It is also possible that the increase in breathiness at the end of the root is due to the additive effect of the prosodically-governed breathiness that occurs at the end of utterances in Ju|'hoansi.

Epiglottalization of consonants and vowels results from similar timing contrasts with respect to opening the glottis to generate frication at a narrow constriction created through a combination of larynx raising, pharyngeal constriction, and tongue root retraction. The difference between a root containing an epiglottalized consonant and a root containing an epiglottalized vowel is then mainly a difference in timing, although the degree of pharyngeal constriction achieved through the complex set of articulatory gestures is also greater for consonants than vowels.

Figure 37 provides spectrograms of the three root types containing epiglottalization, as well as a root containing no epiglottalization for comparison. In epiglottalized consonants, the larynx is raised and the tongue root and epiglottis are retracted in order to achieve maximal pharyngeal constriction, immediately upon release of the two constrictions associated with the click type, and prior to the onset of voicing in the vowel. The maximal point of constriction is achieved prior to the onset of voicing in the following vowel, as seen in the

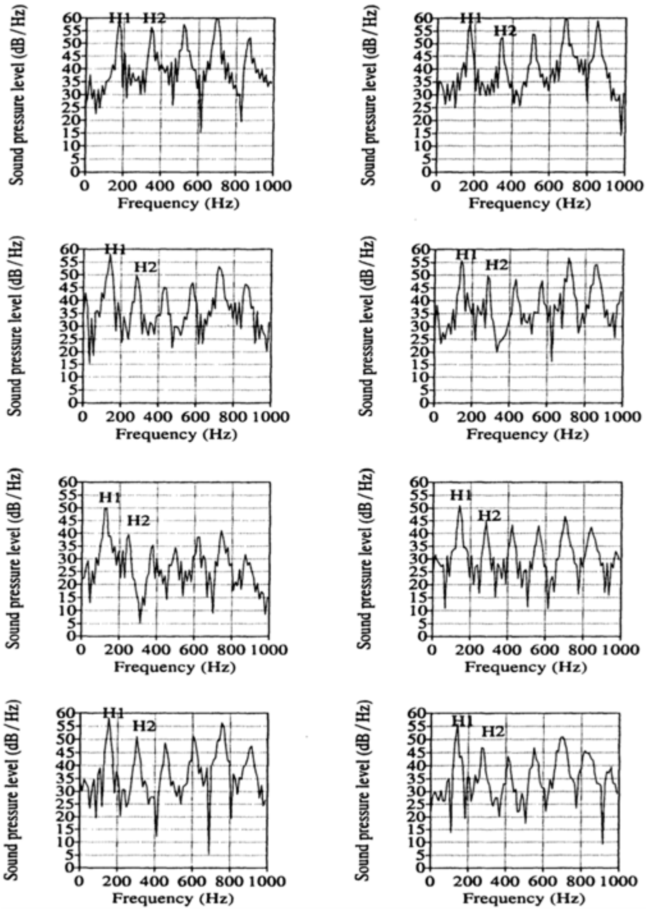


Figure 36. Spectra at ¼ point (left) and ¾ point (right) of vowels in roots displaying a four-way aspiration contrast: *g!áá* ‘rain’, *g|ʱàà* ‘place for drying meat’, *ǂäʱà* ‘plain’ and *!äʱäʱ* ‘red crested korhaan’ (top to bottom, Subject KK)

spectrogram in the second panel in Figure 37. In the production of partially epiglottalized vowels, these same articulatory gestures are implemented following the onset of voicing in the vowel, and maximal constriction is achieved at the ¼ point of

[a] context		[o] context	
ǁáá	‘to warm hands by fire’	!óó	‘older brother’
!ᴴàà	‘to dry out’	!ᴴòò	‘to hate’
ǁɑ̀ˤá	‘to hold down’	ǂòˤó	‘uterus’
!ɑ̀ˤɑ̀ˤ	‘iron’	!ȍˤȍˤ	‘pan’

Table IX. Data illustrating the four-way epiglottalization contrast

the vowel, as seen in the spectrogram in the third panel in Figure 37. The constriction is not as great in roots containing partially epiglottalized vowels, as seen by the fact that there is no cessation of voicing in any part of the vowel. In the production of fully epiglottalized vowels, the maximal point of constriction is at the end of the vowel, with the maximal point of constriction of the epiglottal sphincter coinciding with the cessation of voicing in the vowel. That is, the epiglottalization is timed so that the full stop-like constriction is achieved at the end of the vowel. There is little variability in the duration of this root type, showing that this timing is under phonetic control.

The timing of epiglottalization is consistent with an analysis where there is a four-way timing contrast in epiglottalization in Ju|’hoansi roots. Minimal quadruplets, with roots containing no epiglottalization, epiglottalization occurring between the click burst and the onset of voicing in the vowel, epiglottalization occurring over the first part of the vowel (which I analyze as being associated to the first mora), and epiglottalization occurring over the entire vowel, are plentiful. Table IX provides data illustrating two such minimal quadruplets, in the /a/ and /o/ contexts.

The Khoisan language !Xóõ contains two types of vowels containing constriction in the pharynx (Traill, 1985; Ladefoged and Maddieson, 1996). The Ju|’hoansi epiglottalized vowels are more similar to the strident vowels in !Xóõ than they are to the plain pharyngealized vowels found in that language suggesting that epiglottalized vowels are more parallel to epiglottalized consonants, than they are to uvularized consonants. Some support for the claim that vowels with pharyngeal constriction in Ju|’hoansi are epiglottalized (e.g. what Traill calls pharyngealized) rather than uvularized (e.g. strident vowel) is provided by the acoustic results provided in Part II. Esling (1999b), on the basis of fiberoptic evidence of his own articulations, claims that the contrast between pharyngeal and epiglottal consonants is more a matter of manner of articulation than of place of articulation, with the pharyngeals involving just narrowing of the pharynx and retraction of the tongue root, and epiglottals involving the formation of a stop-like constriction between the raised larynx, the retracted tongue root with concomitant epiglottal retraction, and pharyngeal constriction. Esling’s description of epiglottals is similar to the description of the articulation of strident vowels in the Khoisan language !Xóõ discussed in Traill (1985, 1986). The uvularized consonants in Ju|’hoansi correspond more to this frication caused by narrowing of the pharynx, and epiglottalized vowels correspond to the stops described by Esling, although the epiglottalization is clearly a vocalic phonation type contrast in this language. The acoustic effects found in the acoustic case study reported in Part II show differences that can be associated with such articulatory differences. That is, the spectral slope measure, H1-H2, shows relatively high values for uvularized consonants, and relatively low or negative values for epiglottalized vowels, like those found following epiglottalized consonants. The results are consistent with epiglottalized vowels being more comparable to the pharyngealized vowels in !Xóõ.

Spectrograms exhibiting the four-way contrast in epiglottalization are provided in Figure 36 for male Subject KK. As can be seen, the third formant is weak over the first part of the vowel in the epiglottalized click-initial root. The weakness of F3 can be attributed to guttural coarticulation from the initial consonant, but the low amplitude of F3 at the end of the vowel is due to breathiness that occurs at the end of utterances (shown in Part II). In the partially epiglottalized vowel, the third formant is weaker over the first half of the vowel, and F3 amplitude increases later in the vowel. The third

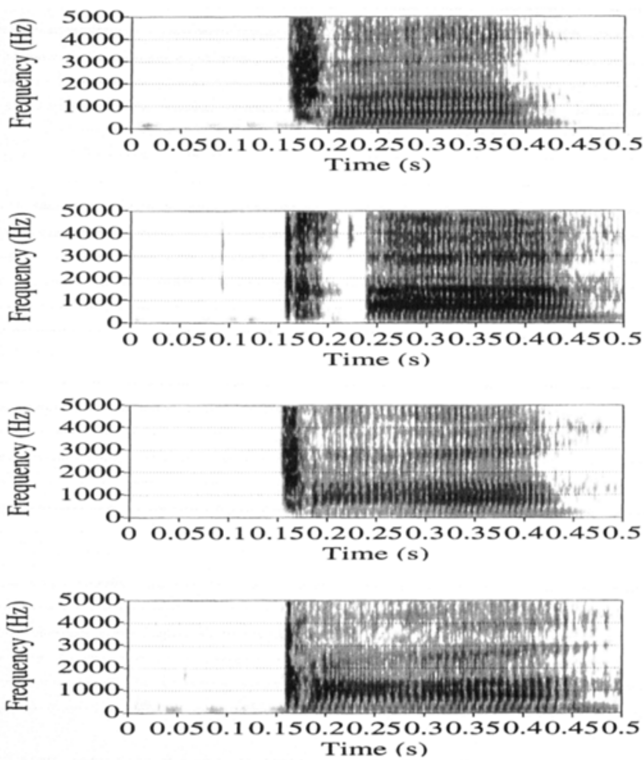


Figure 37. Spectrograms of roots exhibiting the four-way epiglottalization contrast in the roots ǁ*áá* ‘to warm hands by fire’, *!ᴴàà* ‘to dry out’, ǁàˤ*á* ‘to hold down’, and !ȁˤȁˤ ‘iron’ (top to bottom, Subject KK)

formant is faint throughout the entire vowel in the fully epiglottalized vowel, but becomes even weaker towards the end of the vowel. This provides evidence that the degree of constriction increases towards the end of fully epiglottalized vowels.

Figure 38 provides spectrograms of the four-way epiglottalization contrast for Subject NU. As can be seen, the whole vowel is lower amplitude in the vowel following the epiglottalized consonant in the second panel compared with the vowel following the voiceless unaspirated consonant in the upper panel. For Subject NU, there is a period of clear vowel before the sharp lowering of the frequency and amplitude of F3 over the second quarter of the vowel in the partially epiglottalized vowel token. Just as with the partially epiglottalized vowel, Subject NU's production of the fully epiglottalized vowel displays a period of clear vowel before the start of epiglottalization. Again, Subject NU displays better separation of timing patterns associated with vocalic and consonantal epiglottalization, and it is likely that listeners could perceive the distinction in her speech even if the consonant were not present in the signal (a task that is presumed to be more difficult based on Subject KK's productions). This prediction will be tested in a future perception study.

Figure 39 provides spectra over the first 1000 Hz at the ¼ and ¾ points of the vowels displaying the four-way contrast in epiglottalization. Notice that the differences between the first and second harmonics in the vowel following the voiceless unaspirated consonant (upper panel of Figure 38) exhibit a very slightly positive spectral slope (about 3.5 dB/Hz). That is, there is a small positive difference between the amplitude of the first harmonic and the second harmonic, which increases only slightly at the ¾ point of the vowel. In the vowel following the epiglottalized consonant, on the other hand, the difference between H1 and H2 is slightly negative (about −1 dB/Hz), which is low relative to the slightly positive slope found in the vowel following the unaspirated click in the upper panel. In the partially epiglottalized vowel, however, the difference between H1 and H2 is minimal, equaling about 0 at both the ¼ and ¾ points of the vowel, which is also low relative to the value associated with the vowel following the unaspirated consonant. In the fully epiglottalized vowel token, H1-H2 is negative throughout the vowel, as judged by the vlue of about 2 dB/Hz at the ¼ point of the vowel, though it becomes higher (about 3 dB/Hz) by the end of the vowel, which is very similar to the value found in the modal vowel

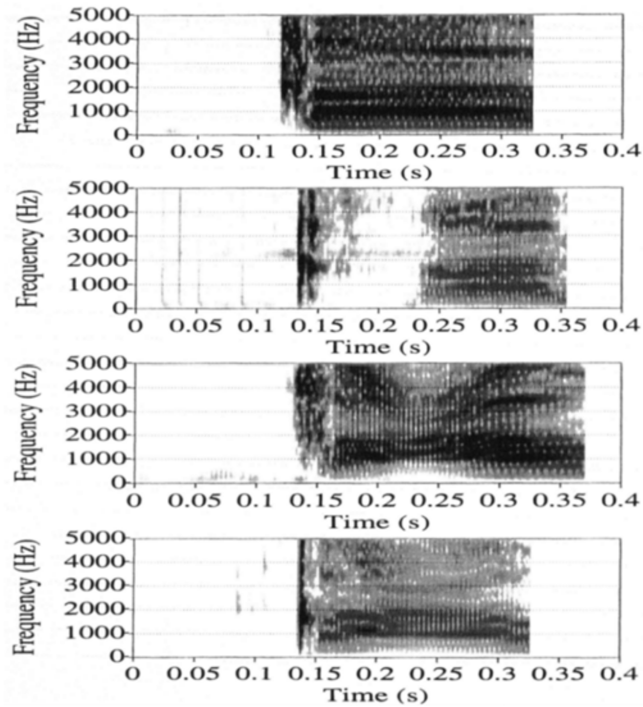


Figure 38. Spectrograms of roots displaying the four-way epiglottalization contrast: ǁ*áá* ‘to warm hands by fire’, *!*^{H}*àà* ‘to dry out’, ǁàˤ*á* ‘to hold down’, and *!*ȁˤȁˤ ‘iron’ (from top to bottom, Subject NU)

following the unaspirated click in the upper panel. Thus, we see a similarity between vowels following epiglottalized consonants on the one hand, and partially and fully epiglottalized vowels on the other. In all cases, the slope of H1-H2 is lower than the slope found in modal vowels following unaspirated consonants at the ¼ point of the vowel. This provides evidence that there is a constriction at or near the glottis in the production of epiglottalized consonants and vowels. The slight increase in spectral slope values at the ¾ points of the vowels is due to utterance final breathiness. The fact that the ¼ point of the vowel following the voiceless epiglottalized initial click is the lowest relative to the modal vowel provides further evidence that the greatest constriction is achieved in the epiglottalized consonant, than in either of the epiglottalized vowel types, at any part of the vowel.

Traill (1986) claims that there is epiglottal trilling in the production of strident vowels in !Xóõ. He provides fiberscopic data from his own productions, and claims that while native speaker data was not that clear, he can see trilling during the production. Edmondson, Esling, Harris, Shaoni and Ziwo (2000) show that there is aryepiglottic trilling associated with harsh voice in the Tibeto-Burman language Bai. Esling (1999b) claims, based on fiberscopic evidence from his own productions of pharyngeal consonants, that the lower pharyngeals described as epiglottal stops involve a constriction between the narrowed pharynx and raised larynx, and thus they should be interpreted as epiglottal stops. The uvularized consonants may also involve aryepiglottic trilling.

One difference between the Ju|’hoansi and !Xóõ vowel inventories is that there are no plain pharyngealized vowels in Ju|’hoansi as there are in !Xóõ. Thus, if !Xóõ velarized clicks are actually uvularized as they are in Ju|’hoansi (Miller-Ockhuizen, 2000) and |Gui (Nakagawa, 1996), then Xóõ can be viewed as having a parallel consonantal and vocalic inventory with uvularized consonants transcribed [!x] and strident vowels transcribed [a̰͇] both involving glottal opening, and uvular ejected consonants [!kχ’] (that I have reanalyzed as epiglottalized consonants in Ju|’hoansi) and pharyngealized vowels [ɑˤ] on the other hand, as both involving epiglottal constriction. Ju|’hoansi has no plain pharyngealized vowels, but rather, only epiglottalized vowels, which are more similar to the !Xóõ pharyngealized vowels.

F0 patterns are consistent with the analysis developed here, that partially epiglottalized vs. fully epiglottalized vowels contrast in timing. The differences in fundamental frequency found between epiglottalized monophthongs and diphthongs with epiglottal voice quality on the first mora, and modal voice quality on the second mora are shown in Figure 40. Notice that

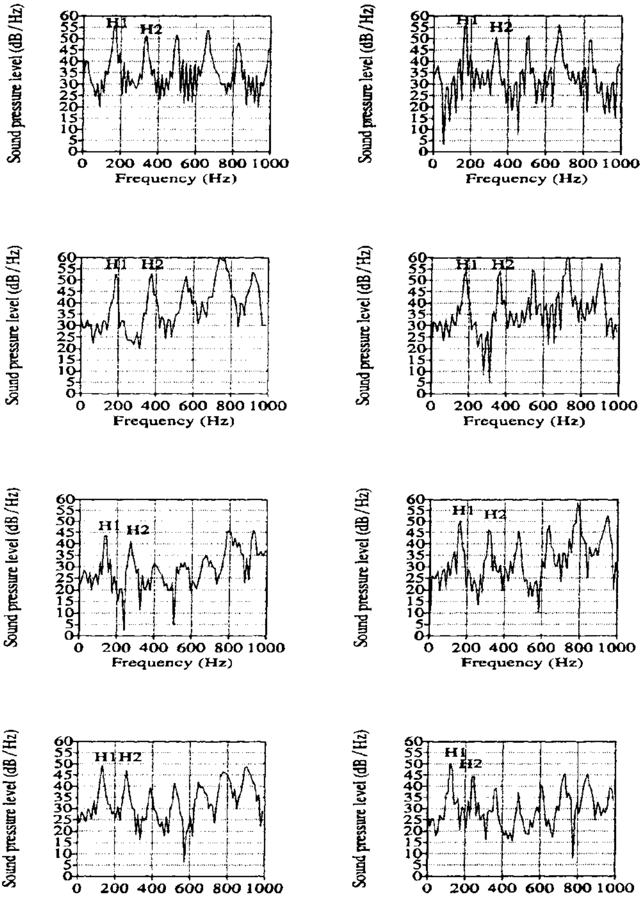


Figure 39. Spectra at ¼ point (left) and ¾ point (right) of the vowel in roots displaying a four-way epiglottalization contrast: ‖*áá* ‘to warm hands by fire’, *!*^{H}*àà* ‘to dry out’, ‖*à*ˤ*á* ‘to hold down’ and */ä*ˤ*ä*ˤ ‘iron’ (top to bottom, Subject KK)

the F0 associated with the fully epiglottalized vowel is lower throughout, although it also rises throughout the vowel, just as the partially epiglottalized vowel does. The F0 associated with the fully epiglottalized vowel levels out towards the end of the root, while the F0 associated with the partially epiglottalized vowel continues to rise throughout the entire duration of the root. The leveling of F0 at the end of the root in the fully epiglottalized vowels is probably what gives listeners transcribing these tones (including myself) the percept that these are level toned, with the dip in F0 in the first part of the root being caused by perturbation from the epiglottal voice quality rather than a separate lower tonal specification. Although the rise at the end of the root is small for fully epiglottalized vowels, it is rather large relative to the usual large drop in F0 that occurs at the end of the vowel. The continuous rise in F0 in the partially epiglottalized vowels, which remains after the offset of the largest spectral influences from epiglottal voice quality, signals that these roots have a rising tone pattern. The F0 patterns shown here for a single female speaker are consistent with the patterns found for the other three speakers. Lindqvist (1972) notes that vocalis muscle contraction is often used to counteract the F0 lowering effect that is concomitant with pharyngealization. This might explain the transcription of partially epiglottalized vowels, which are always rising in tone, with concomitant glottalization by Snyman (1975), and Dickens (1994).

If indeed vocalis muscle contraction is involved in the articulation of partially epiglottalized vowels, it might be tempting to assume that the low spectral slope values associated with such roots can be attributed solely to vocalis muscle contraction. However, the low spectral slopes found on epiglottalized vowels must be due in part, to the pharyngeal constriction, and not solely due to vocalis muscle contraction, since fully epiglottalized vowels also display similar low spectral slope values, and these do not involve an abrupt change in pitch characteristic of vocalis muscle contraction. I hypothesize that fully epiglottalized vowels do no involve any vocalis muscle contraction, and that vocalis muscle contraction is used to achieve rising tone during larynx raising.

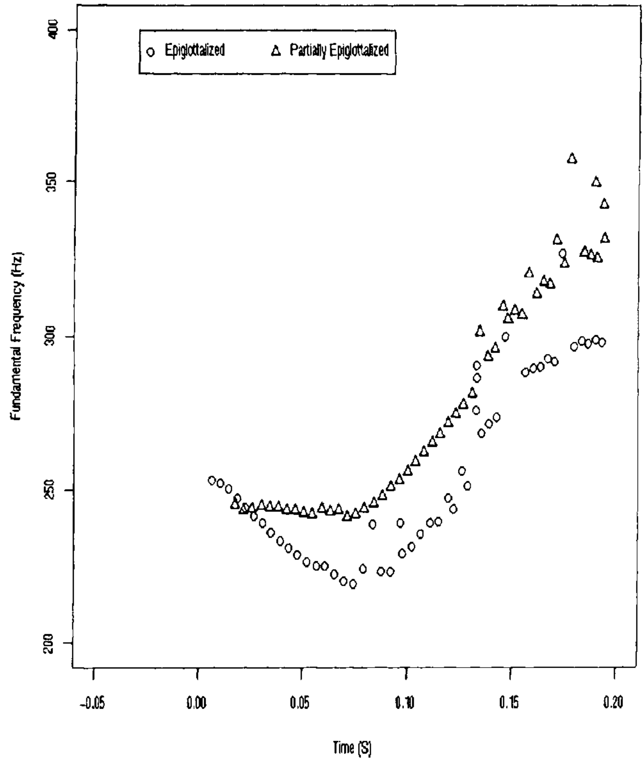


Figure 40. Mean F0 patterns found in roots containing partially and fully epiglottalized vowels: ‖ɑ̏ˤà ‘to hold down’ and !ɑ̏ˤɑ̏ˤ ‘iron’ (Subject NU)

The lowering of the amplitude of F3 found for all guttural vowel types indicates an increase in noise in the F3 range. For all guttural vowel types in Ju|’hoansi, this diminution of energy occurs over the first part of the vowel. Although the peak glottal gestures occur at the ¼ point of the vowel in roots containing guttural vowels, there are already increased or decreased spectral slope, formant and F0 values at the onset of the vowel for all roots containing guttural vowel types. This means that some of the noise overlaps with the C-V transition. Thus, guttural vowel types all involve an increase in noise within the third formant frequency range, which is similar to the noise found in the same interval of vowels following guttural consonants. Thus, the phonetic basis for the class of gutturals relevant to the Guttural OCP constraint is attributed to the presence of noise in the third formant frequency range, occurring in the C-V transition.

2.6.
Conclusion

In this chapter I have touched on the major phonetic characteristics that unite Ju|'hoansi guttural consonants and vowels in contrast with non-guttural consonants and vowels. I have shown that all guttural clicks in Ju|'hoansi have their acoustic cues concentrated in the tongue dorsum lag phase, and further, they all increase the noise in the following vowel through guttural coarticulation. This spectral noise is what groups guttural consonants into the same natural class with guttural vowels, which all contain increased aperiodic noise in the third formant range. Aspirated consonants and breathy vowels all involve the same spread glottal articulatory configuration, which results in the same acoustic characteristics of high spectral slope and the presence of noise in the third formant range. Epiglottalized vowels are unique in showing a dipping effect on the fundamental frequency, as well as all of the formant frequencies, with the most marked differences found in F1 and F3. Not all of these effects can be accounted for by a single articulatory mechanism, such as larynx raising or tongue root retraction. Rather, a more complex account on the lines of that found in Esling (1999b) is necessary. While there is a parallel contrast between pharyngeal and strident vowels in the related Khoisan language ! Xóõ (Traill, 1985; Ladefoged and Maddieson, 1996), Ju|'hoansi only has one type of epiglottalized vowel, which is articulated with a constriction in the pharynx. The Ju|'hoansi vowels are more parallel to the pharyngealized vowels in ! Xóõ. In the next chapter I describe the phonotactic constraints in Ju|'hoansi that target guttural consonants and vowels, and offer a new acoustically-based feature that unifies all of these sounds into a natural class. The feature I propose in Chapter 3, based on phonological evidence, is consistent with the acoustic phonetic properties of guttural sounds outlined in this chapter.

CHAPTER 3
Ju|’hoansi Phonotactics

3.0.
Introduction

In this chapter, I report the results of a database study, which investigates phonotactic patterns found over an 1878 root database, described in Section 3.1. I identify constraints on root size, as well as two major classes of phonotactic patterns found within Ju|’hoansi roots, positional constraints and co-occurrence constraints. All of the constraints involve either the natural class of gutturals, or the natural class of pharyngeals (consonants or vowels specified with the place feature [pharyngeal]). I propose a new acoustically based feature, [spectral slope], which marks all guttural consonants and vowels, and thus allows a unified representation of the natural class of guttural sounds. The phonetic basis of the class of [pharyngeal] sounds is articulatorily defined, as sounds involving pharyngeal constriction. In each section, after describing the distributional patterns and offering phonological analyses, I described how these constraints might aid in language processing.

There are two constraints on root size. That is, a root is minimally a bimoraic foot and maximally a trochee with a heavy initial syllable. These constraints are shown to be active in the phonology, as they motivate the parsing of longer loan-words into two or more prosodic words. The parsing of loan-words is also shown to be motivated by other positional and co-occurrence constraints. Strict constraints on word size aid the listener in parsing the incoming speech stream into word-size chunks.

There are two positional constraints, which I claim have a role in language processing. The first of these aligns every obstruent to the initial position of the root, leaving sonorants in medial position. The second constraint aligns a guttural feature to the initial vowel position within the root. Since guttural consonants are all obstruents, they are also limited to root-initial position. The analysis follows Beckman (1998) in many respects.

Four co-occurrence restrictions were identified via the database, all of which operate within the domain of the lexical root. First, I identify a new Guttural OCP constraint that prevents any two guttural consonants or vowels from co-occurring within the same root. Second, I identify a co-occurrence restriction between tone and gutturals, a third between the place features [pharyngeal] and [coronal], as well as a fourth, between [pharyngeal] and [high]. For many of these constraints, the co-occurrence restrictions for consonants and vowels are somewhat parallel. For other constraints, especially those between guttural features and tone, vowel patterns are categorical, and consonant patterns are more gradient. In each section, I first define the phonological constraints broadly, and offer a phonological analysis. I then offer some suggestions about the perceptual motivation of the Guttural OCP constraint and articulatory motivations for many of the guttural co-occurrence restrictions. The Guttural OCP, for instance, is argued to be motivated by a type of parsing ambiguity.

3.1.
Recorded Database

I recorded Dickens' (1994) entire Ju|'hoansi-English dictionary (which expands the number of roots found in Snyman's (1975) original dictionary) as produced by 6 speakers (3 female: DK, DX, NC; and 3 male: KK, KX, KG), and worked on eliciting new words not yet identified by either Snyman or Dickens. All words in the database were analyzed morphologically, and all apparently mono-morphemic roots were re-transcribed and entered into an electronic database. The database encodes the prosodic root shape, the root tonal pattern, the initial consonant type (based on manner features and airstream mechanisms), the initial consonant release type for stops and affricates (aspiration, uvularization, glottalization), the closure properties of the initial consonant (voicing, nasalization), the vowel quality in V_1 and V_2 (height and place features), and the vowel phonation type in V_1 and V_2.

After re-transcription was completed, the phonotactic patterns that were initially identified using Dickens' (1994) approximately 3600-word dictionary were rechecked, and exceptions to constraints under study were verified by listening again to the acoustic signal. A number of exceptions to the broad phonotactic patterns originally identified on the basis of Dickens' (1994) dictionary disappeared after some words that were originally thought to be mono-morphemic roots were identified as bimorphemic. The final database used in this study contains 1878 mono-morphemic roots.

Considerable work was undertaken to identify loan-words in the database, although it is extremely difficult to identify loan-words from other Khoisan languages, since words in these languages contain similar sounds, such as click consonants, and are prosodically and phonotactically similar. Also, the complete lack of dictionaries or small size of those dictionaries that are available for many of the other Khoisan languages makes it almost impossible to check potential source languages for cognates. Of the final list of 1878 native bimoraic roots, 291 were new roots identified by me during the course of my field work in 1997–1999. Given the heavy reliance on compounding and cliticization in creating new words in Ju|'hoansi, over 1000 new words were elicited in order to identify these 291 roots.

One way in which the database developed for this study expands the earlier dictionaries was a more systematic investigation of the loan-word status of words in Dickens

(1994). While some loan-words were identified in Dickens and Snyman, others were not identified. The current study shows that many of the roots that appear to violate phonotactic constraints in Ju|'hoansi are loan-words, and that these loan-words are often parsed into two or more prosodic words in order to satisfy the constraints.

The recorded nature of the database, as well as the re-transcription of the words allowed me to investigate fully the phonotactic patterns in Ju|'hoansi roots, and to identify the regularity of the patterns. While Snyman (1970) identified all of the contrastive sounds that occur in initial position in Ju|'hoansi, undertook a thorough phonemic analysis of the language, some of the allophones identified in his 1970 grammar were no longer transcribed in Snyman (1975), due to the reliance on phoneme economy in the development of this later work. Dickens (1994) notes some allophonic processes in his dictionary, but constraints are stated as tendencies, and there is no indication of the probability of the generalization, and the contexts that condition the various allophones are somewhat vague. For example, Dickens (1994:14) notes that orthographic 'ai' is pronounced as [əi] or [i], "especially after the clicks [|] and [ǂ]." Nowhere does he state that [əi] also occurs following uvularized clicks (his clicks with velar frication) and epiglottalized clicks (his clicks with velar ejection). Complete re-transcription of the words allowed me to make generalizations that I would not have been able to make based solely on Dickens' (1994) or Snyman's descriptions or lexicons (1970, 1975) with any confidence.

A further expansion from the previous dictionaries was the investigation of the morphology of all words in Dickens (1994). Although Snyman (1970, 1975) identifies the separate morphemes of complex words in his dictionary, not all bimorphemic words were labeled as such, leading Sagey (1986) to wrongly conclude that clicks occur root-medially in Ju|'hoansi (which she refers to as !Xũ based on Snyman's 1970 name). Dickens' (1994) dictionary does not mark the morphemic structure of roots at all, leaving new bi-morphemic roots identified by him unanalyzed. Since morphemic structure is crucial to understand the domain of phonological constraints, I analyzed the morphological structure of all words in Dickens (1994), and included only words determined to be roots in the database. A word was considered to be a root if it could not be broken down into separate morphemes.

Given the large number of contrastive consonant and vowel phonemes (89 consonants and 34 vowels), as well as seven root tone patterns, it is difficult to find minimal pairs that differ by only one feature. Thus, phonetic detail becomes extremely important. The generalizations made in this dissertation are considered to be robust, given that each root was re-transcribed by the author and the acoustic signal was available for rechecking any questionable exceptions or seemingly robust patterns, thus appreciably diminishing the chance that transcription error might result in false generalizations. Additionally, the reliability of the generalizations made is high because generalizations were made numerically over the entire database, and because recorded tokens were available from 3 different male and 3 different female speakers. Thus, reporting individual idiosyncrasies is avoided. Individual variation was transcribed, but only the majority pattern was entered into the database.

Throughout this chapter, I calculate the probability that a particular pattern will occur on any particular root in the database. I use the probability as a measure of how strong the pattern is. Probabilities also take word position into account where it is deemed to have a large effect. All phonological patterns in the languages of the world have some exceptions. While some of the lexical exceptions to the phonotactic patterns found in Ju|’hoansi can be identified as occurring only in loan-words, others can not. While loan-word sources were searched for in available Namibian and Botswana language dictionaries (e.g., RuKwangali, Otjiherero, Setswana, Naro, Khoekhoe, Afrikaans and English), some of the words do not derive from these languages. However, a full search is constrained by the availability of good dictionaries on under-studied Namibian and Botswana Bantu and Khoisan languages. It is likely that even after a full search for the source of words displaying irregular patterns, a small residue of words that do not follow robust phonotactic patterns would remain. Also, it is not clear that native speakers are always aware of the origins of words, so even loan-words are part of the lexicon. For these reasons, I use the probability of 0.9 that a root in the database will display a given phonotactic pattern as the cutoff point for deciding which patterns found in the lexicon should be formulated as “unviolable” phonological constraints, and which patterns reflect more gradient tendencies.

The 0.9 probability used in this dissertation would predict that patterns like English trochaic stress are not exceptionless generalizations, since Cutler and Carter (1987) found that 73% of the lexical word tokens in a computer-readable English dictionary containing over 33,000 entries contained strong initial syllables (meaning that they have either primary or secondary stress). Despite the low probability of words displaying trochaic stress, they found that English listeners were more likely to segment speech into separate words just before a strong syllable. Hammond (1999) posits the phonological constraint TROCHEE to account for the fact that feet are usually trochaic (having stress on the initial syllable of a foot), and offers other more highly ranked constraints that account for exceptions to the trochaic stress pattern. Thus, in the analyses that follow, I posit unviolable constraints for patterns that have a probability of at least .9 over the database used in this study and for patterns that display a lower probability, but for which higher ranked constraints can be identified that account for these lower probabilities. For patterns with probabilities lower than 0.9, and for which I cannot yet offer higher-ranked constraints that would explain the lower probabilities, I do not formulate constraints.

3.2.
Word Minimality and Maximality

In this section, I show that native roots are minimally a single heavy syllable and maximally bisyllabic with an initial heavy syllable. I argue that native roots constitute a single prosodic word and that minimality and maximality constraints on the size of the prosodic word result in all root shapes constituting a single foot. To support these claims, I draw on evidence from reduplication, phonetic duration, and loan-word adaptation strategies.

3.2.1. Word Minimality

As was noted in Chapter 2, native Ju|'hoansi roots are minimally bimoraic, and maximally trimoraic and bisyllabic. The lexical frequency found in the database for the different root shapes is provided in Table X. Trimoraic trochaic words are by far the lowest frequency native root-type, accounting for only 6% (116/1878) of native roots.

ROOT SHAPE	LEXICAL FREQUENCY IN DATABASE
CVV	911 (+6 *Loan-words*)
CVm/ŋ	177 (+1 *Loan-word*)
CVCV	674 (+38 *Loan-words*)
CVVCV	116
CVCVCV	*0 (49 Loan-words)*
TOTAL	1878 (+94 *Loan-words*)

Table X. Lexical frequency of prosodic root shapes

Bisyllabic roots contain at least two moras, given that each syllable must contain at least one mora. Monosyllabic roots can be monotonal (595/1088=55%) or bitonal (493/1088=45%). Reduplication patterns discussed by Miller-Ockhuizen (1999) provide clear evidence that the tone-bearing unit in Ju|'hoansi is the mora. While reduplication of monosyllabic roots is always complete, as shown in (1)(a), reduplication of bisyllabic roots is partial. The reduplicant is always a syllable, and in bisyllabic roots, the weight of the reduplicant is dependent on the tonal pattern of the base. If the base is monotonal, only the first syllable is copied, as shown in (1)(b). However, if the base is bitonal, both vowels in the base are copied, although the medial consonant is not, as in (1)(c). The disparity in the weight of the reduplicant is best described if the tone-bearing unit is the mora under the assumption that each tone is associated to a separate mora. Thus, bitonal syllables are necessarily bimoraic.

(a) Reduplication of Monotonal and Bitonal Bisyllabic Roots			
tĩ̀ĩ̀ʱ	'to be heavy'	tĩ̀ĩʱ.tĩ̀ĩ̀ʱ	'to cause to be heavy'
ʃə̀ò	'to be wide'	ʃə̀ò.ʃə̀ò	'to widen'
tòàⁿ	'to be finished'	tòàⁿ.tòàⁿ	'to cause to be finished'
!ə́úⁿ	'to be sick'	!ə́úⁿ.!ə́úⁿ	'to make ill'
g!ʱə̃̀ɪ̃̀ⁿ	'to move sideways'	g!ʱə̃̀ɪ̃̀ⁿ.g!ʱə̃̀ɪ̃̀ⁿ	'to move side to side'
n|ù ˀɪ́	'to be shiny'	n|ùˀɪ́.n|ùˀɪ́	'to cause to be shiny'
(b) Reduplication of Monotonal Bisyllabic Roots			
ǂʔʰù.βɪ	'to be crowded'	ǂʔʰù.ǂʔʰù.βɪ	'to cause to crowd'
gʱù.ɾì	'to become visible'	gʱù.gʱù.ɾì	'to cause to be visible'
ǂʰà.mà	'to take'	ǂʰàⁿ.ǂʰà.mà	'to cause to take'
xə́.βə́	'to sting (of skin)'	xə́.xə́.βə́	'to cause to sting'
mɑ̀ˤ. nì	'to speak a non-click language'	mɑ̀ˤ.mɑ̀ˤ.nì	'to cause to speak a non-click language'
gɔ̀ˤ.ɾò	'to rot'	gɔ̀ˤ.gɔ̀ˤ.ɾò	'to cause to rot'
(c) Reduplication of Bitonal Bisyllabic Roots			
tsχə̀.βí	'to grab'	tsχə̀î.tsχə̀.βí	'to grab forcefully'
|χə̃̀.nì	'to poke a hole'	|χə̃̀ɪ̃ⁿ.|χə̃̀nì	'to drill a hole'
‖χə̀.ɾí	'to fry'	‖χə̀í.‖χə̀.ɾí	'to cause to fry'
mɑ̀ˤ.ní	'to turn over'	mɑ̀ˤɪ́.mɑ̀ˤ.ní	'to cause to turn over'
nǂɑ̀ˤ.ɾó	'to find s.t.'	nǂɑ̀ˤó.nǂɑ̀ˤ.ɾó	'to find a lost object'

(1) Evidence from partial reduplication for the TBU=µ (Syllables are marked with a "." and the reduplicant is underlined)

Given that there is no duration contrast between monotonal and bitonal roots, we can assume that monotonal monosyllabic roots are also bimoraic. As shown in Miller-Ockhuizen (2001), roots that Dickens (1994) transcribed with short vowels have nearly the same duration as bitonal monosyllabic roots. All monosyllabic roots are nearly twice the duration of the second syllable of bisyllabic roots, as shown by the histogram taken from Miller-Ockhuizen (2001) provided in Figure 41. The duration of monotonal monophthongal roots is consistent with

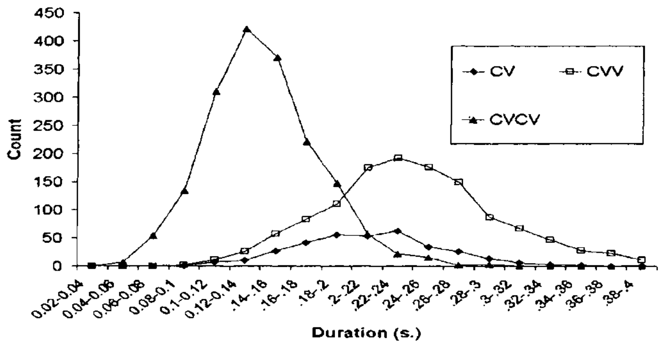


Figure 41. Histogram of the total rime duration of monotonal monophthongal vs. bitonal monophthongal monosyllabic roots and the final vowel of CVCV roots (taken from Miller-Ockhuizen, 2001)

such roots being bimoraic, prosodically parallel to bitonal roots. Thus, the smallest native roots are bimoraic syllables, which suggests that there is a minimality constraint in the language.

Further evidence for word minimality comes from augmentation of monomoraic loan-words. For example, Afrikaans and English loan-words with short vowels and final, non-moraic coda consonants are repaired with one of three strategies when they are adapted into Ju|’hoansi. When there is an illicit coda in the loan-word, vowel epenthesis creates a bisyllabic root as shown by the data in (1)(a); or else the illicit final consonant is lost, and the vowel is lengthened, as shown by the data in (1)(b). A third repair strategy involves insertion of a vowel into an initial cluster, breaking up the illicit cluster, and creating a bisyllabic word as shown in the data in (1)(c). While several of the words are loan-words from English into Afrikaans, they are provided since it is clear that Afrikaans is the source language. For example, in cases like *tube*, the voiceless [p] in the Ju|’hoansi word arises through final devoicing in the Afrikaans pronunciation. The native Afrikaans words are also given in such cases. The observation that coda consonants are never lost without lengthening and that short syllables are lengthened provides further evidence that there is a minimality constraint on roots.

	Afrikaans Word	Afrikaans Transcription	English Gloss	Ju|'hoansi Transcription
(a)	bal	[bɔl]	'ball'	bũ̀ʱɾù
	tube (<Eng.)	[tʃʰup]	'tube'	tʃʰúpə̀
	boer	[bu:r]	'farmer'	bùɾù
	hof	[ɦof]	'court'	hòfə̃̀
	saag	[sa:x]	'saw'	zàχà̃
	doek	[duk]	'scarf'	tṹkṹ
(b)	pap	[pʰap]	'porridge'	pʰáá
	soek	[suk]	'to look for'	tùù
	tii	[tʰiə]	'tea'	tʰíí
(c)	trou	[tʰrou]	'wedding'	tóɾó
	draad	[dra:t]	'wire'	tà̃ʱɾà
	stor (<Eng.)	[stɔr:]	'store'	tòɾà

(2) Loan-word adaptation strategies in monosyllabic roots

The phonological constraints enforcing minimality will be offered in Section 3.2.3, along with a unified analysis of optimal word size. First, I discuss evidence that motivates a maximality constraint in Section 3.2.2.

3.2.2. Word Maximality

Ju|'hoansi native roots are maximally bisyllabic, with the largest root type containing a heavy initial syllable and a light final syllable. The adaptation of bisyllabic loan-words with non-native clusters and trisyllabic loan-words provides evidence for the existence of the prosodic word (PrWd) and evidence that the PrWd is maximally a foot. This generalization has a probability of 1, since there are no roots in the language that are larger than two syllables, and realized as a single prosodic word.

The adaptation of trisyllabic loan-words involves parsing the material as two separate prosodic words. Prosodic word boundaries are marked with curly brackets. For example, the name of an educational consultant currently working in the Tsumkwe area of Namibia (the heart of the Ju|'hoansi area), is *Beverly*. The name is adapted into Ju|'hoansi as [{béé} {vìlì}].

Similarly, *Amanda*, is adapted as [{máán} {dàà}], with the initial vowel deleted, and both syllables lengthened. A number of loan-words that are parsed as two prosodic words are provided in (3). As can be seen, when these roots are parsed into two prosodic words, one of the syllables is lengthened. The prosodic structure is cued phonetically by insertion of a pause, final lowering, downstep of adjacent high tones, and manner of the initial consonant. I used these cues to transcribe the prosodic structure.

Whether the first prosodic word is monosyllabic and the second prosodic word is bisyllabic or vice versa, depends on which parse best satisfies the other phonotactic constraints discussed in the rest of this chapter. For example, the English word *petrol*, which has also been borrowed into Afrikaans, is parsed as [{p^{h}éé}{tòɾò}], but the Afrikaans word *knoop* ‘button’[7] is parsed as [{kónó} {mbèè}][8] or [{kónó} {bèè}]. These parses best satisfy the constraint that an obstruent should be aligned with the left edge of a prosodic word, as will be described in Section 3.3. In the adaptation of Afrikaans source words in which both the second and third syllables begin with obstruents, e.g., *ketting* ‘chain’, the source word is parsed into three separate prosodic words, [{káá} {$^!$t^{h}áá^n} {$^!$gáá}], with all three syllables lengthened, and the nasal consonant interpreted as nasalization on the vowel.

[5] The Afrikaans [r] is actually an apical trill (Donaldson, 1993).

[6] The word *store* is adapted by some speakers as [{sìì}{tòɾà}].

[7] Since pre-nasalized stops are not native to Afrikaans, its presence suggests that this word was acquired through a Bantu language, which has these sounds natively.

[8] This suggests that some speakers are starting to accept reanalyzed stops found in Bantu languages.

Afrikaans Word	Afrikaans Transcription	English Gloss	Ju|'hoansi Transcription
emmer	[ɛ́mr̩]	'bucket'	[{ʔə́m}{bɔ̀ɾè}]
appel	[ápl̩]	'apple'	[{ʔáá}{pə̀ɾə̀}]
farm	[fárm̩]	'farm'	[{fàɾà}{màà}]
tronk	[tʰróŋk]	'jail'	[{tóró}{kʰòè}]
kers	[kʰérs]	'candle'	[{kéré}{ˈsíí}]
knoop	[knúp]	'button'	[{kónó}{ˈbéé}]
petrol	[pʰɛ́trol]	'petrol'	[{pʰéé}{tʰòrò}]
papier	[pʰa(m)pʰʲɪr]	'paper'	[{kʰə̀m}{pʰə́ɾí}]
prys	[pʰrés]	'price'	[{pʰə̀ɾí}{tcìì}]
ketting	[kʰɛ́tʰŋ̩]	'chain'	[{káá}{ˈtʰáán}{ˈgáá}]
tafel	[tʰáfl̩]	'table'	[{táá}{fàɾà}]
patron	[pʰatʰrún]	'pattern'	[{bàá}{tʰóró}]
karton	[kʰartʰón]	'carton, box'	[{kàà}{tʰóón}{gàà}]
Satan	[sátʰən]	'Satan'	[{sáá}{tànà}]
saal kleedjie	[saːl kʰlèkʰi]	'Saddle cloth'	[{tə̀ɾè}{kìì}]
broek	[bruk]	'trousers'	[{búrú}{kʰòè}]
kruiva	[kréva]	'wheel barrow'	[{kə̀ɾí}{ˈbáá}]

(3) Loan-word adaptation strategies in longer roots

There is some variation among speakers as to which repair strategies they use to adapt loan-words. The younger speakers seem to prefer parses which maintain the root as a bisyllabic whole, and thus use deletion of whole syllables more often than older speakers. Older speakers are more likely to break up the loan-word into two or more prosodic words. For example, while older speakers parse *saraha* 'bait' as [{sàɾà}{háá}], the youngest speaker in this study produced it as [{tʃàà}{háá}]. Bilingual speakers in Tsumkwe are much more faithful to the original source words than the monolingual speakers in |Xoan n!huru, whose phonology I describe in this dissertation. That is, monolinguals adapt the words using Ju|'hoansi phonotactic constraints, while bilinguals tend to code-switch when producing the non-native words from a language they speak. Both age and mono vs. bilingualism affect the parsing of non-native roots. The fact that trisyllabic loan-words are parsed into separate prosodic words provides evidence that the prosodic word in Ju|'hoansi has a maximality condition as well as a minimality condition. That is, the prosodic word in Ju|'hoansi is maximally bisyllabic. Combining these conditions, it is clear that the prosodic word in Ju|’hoansi always equals a foot of some type, whether it be a heavy monosyllable or a bisyllabic trochee with either two light syllables or a heavy syllable followed by a light syllable. I refer to these trisyllabic loan-words in subsequent sections, as they provide important evidence that many phonotactic patterns in Ju|’hoansi operate over the prosodic word, rather than over the morphological category root.

3.2.3. Constraints on Word Size

Both minimality and maximality conditions on the prosodic word in Ju|’hoansi are captured by the two constraints in (4), which assure that the word is maximally a single foot (Miller-Ockhuizen, 1998). The constraint FOOT BINARITY (Prince, 1980; McCarthy and Prince, 1995a) requires that feet are binary either in terms of syllables or moras, and the constraint PRWD=FOOT (McCarthy and Prince, 1993) assures that the Prosodic Word will always equal a foot. McCarthy and Prince (1995a) note, citing work by Prince (1980), and that the constraint FOOT BINARITY, along with the prosodic hierarchy (Nespor and Vogel, 1986) enforces minimality without the need for a constraint such as PRWD=FOOT, because the PrWd dominates the foot. However, the constraint is independently necessary to ensure maximality conditions. Both constraints are inviolable in Ju|’hoansi, and thus assumed to be undominated. That is, they have a probability of 1.

PRWD=FOOT:	A prosodic word contains a single foot.
FOOT BINARITY:	Feet are binary under syllabic or moraic analysis.

(4) Constraints enforcing word size

3.2.4. Conclusion

In this section, I have shown that based on attested native root shapes found in the language, reduplication patterns and phonetic duration of monotonal and bitonal monsyllabic roots, as well as on augmentation found in loan-word adaptation, the prosodic word in Ju|'hoansi is equal to a foot. Both bimoraic monosyllables and bisyllabic roots that contain either two light syllables or a heavy syllable followed by a light syllable satisfy these constraints, as they are at least two moras and at most two syllables. I have offered the high ranking constraints PRWD=FOOT and FTBIN to account for both the minimality and maximality facts about roots.

3.3. Guttural Feature Specification

Many of the positional specification and co-occurrence restrictions found in Ju|'hoansi revolve around guttural consonants and vowels. It is therefore necessary to give a clear picture of which sounds belong to the natural class of guttural consonants and vowels based on their phonological patterning. As I show, this natural class motivates a new laryngeal feature that marks all guttural consonants and vowels.

Ju|'hoansi guttural consonants and vowels include all sounds that involve either a constriction at the pharynx or the larynx, as in glottalized and epiglottalized obstruents (that is, obstruents with a secondary constriction in the release portion between the raised larynx, the constricted pharynx, and the retracted tongue root), or a spread glottal posture, as is found with glottal and uvular fricatives, aspirated obstruents, and uvularized obstruents. Guttural vowels include sounds with either an extremely spread glottal posture, as in breathy vowels, or a constriction at the larynx or the pharynx, as in glottalized vowels and epiglottalized vowels. The complete list of guttural consonant and vowel types is provided in Table XI.

	Consonants	Vowels
Breathy	Aspirated Obstruents and Laryngeal fricatives	Breathy
Glottalized	Ejected Obstruents and Glottal Stop[9]	Glottalized
Uvular	Uvularized Obstruents and Uvular Fricatives	--
Epiglottal	Epiglottalized Obstruents	Epiglottalized

Table XI. Guttural consonants and vowels

In Chapter 2, I showed that all guttural consonants and vowels are unified acoustically in two ways. First, the noise that cues guttural consonants is located in the C-V transition, which makes the temporal domain of noise associated with guttural consonants overlap with the temporal domain of noise associated with guttural vowels. Second, spectral noise below 4000 Hz is found in all guttural vowels and in the C-V transition following guttural consonants (and this noise is often present throughout the entire following vowel). While the presence of this noise (measured in Chapter 5 via the gamnitude of the first rahmonic in the cepstrum) unifies guttural consonants and vowels, it does not make useful distinctions between gutturals that involve a constriction at the larynx or the pharynx and gutturals that involve a spread glottal posture. As was shown in Chapter 2, the relative height of the spectral slope distinguishes glottalized and epiglottalized sounds on the one hand, from aspirated and uvularized consonants, on the other hand. I thus propose that the binary feature [spectral slope], is specified as [high] for sounds articulated with a spread glottis, and as [low] for sounds articulated with a constricted glottis or a constriction in the pharyngeal cavity. Therefore, with regards to laryngeals, the new feature has the

[9] Recall from Chapter 2 that glottal stop is likely allophonic, being inserted before initial vowels. None of the words of this type in the database contain guttural vowels, but there are so few such words that one wonders if this is not just due the extremely low frequency of words of this type.

same marking as the traditional features [spread glottis] and [constricted glottis] (Halle and Stevens, 1971), or Lombardi's (1994) features [aspirated] and [glottalized]. That is, sounds that are traditionally specified for [spread glottis] are specified for the value [high], and sounds traditionally specified for [constricted glottis] are marked with the value [low][10].

Uvular fricatives and uvularized obstruents are similar to aspirated obstruents in that they are both articulated with a spread glottal configuration, and thus the acoustic correlate of high spectral slope results in the same feature values for [spectral slope] found for the traditional articulatory-based feature [spread glottis] for these sounds. However, a problem arises with the traditional feature [constricted glottis] with respect to the specification of epiglottalized consonants and vowels. Since epiglottalized consonants and vowels actually involve a constriction in the pharyngeal region, not at the glottis, marking these sounds as [constricted glottis] would not be phonetically accurate. However, as I have shown in Chapter 2, and will show in Chapter 5, these sounds have low spectral slope values similar to glottalized consonants and vowels. The feature must be restricted to laryngeal sounds in order to rule out the specification of fricatives, which also have increased airflow, and likely higher spectral slope values due to more spread glottal configurations.[11]

The feature values for the feature [spectral slope] must be interpreted in relative terms, since like all acoustic properties, for any given speaker and for any given language, the actual values that are marked with the value [high] of the feature [spectral slope], and the value [low] of the feature [spectral slope], may be somewhat different. However, data from a growing number of diverse languages (Huffman, 1987; Blankenship, 1997; Cho, Jun and Ladefoged, 2000; Löfqvist and McGowan, 1992) shows that sounds with a constricted

---

[10] This binary feature could also be interpreted as a separate node within Feature Geometry, with two dependent features, [high spectral slope] and [low spectral slope].

glottis consistently have low or negative acoustic spectral slope values relative to modal consonants and vowels, and sounds produced with a spread glottal configuration have consistently higher acoustic spectral slope values relative to sounds produced with an unmarked glottal configuration. The acoustic data provided in this dissertation show that sounds produced with a constriction near the glottis, as is found with epiglottalized consonants and vowels, display similar relatively low spectral slope values. The definition of the feature [spectral slope] is provided in (5).

> [spectral slope]:
>
> [high]: A sound that is specified for the value [high] of the [spectral slope] feature has a substantial interval during which there is audible and audibly periodic energy generated at the glottis, but a very steep spectral roll-off relative to the approximately −6 dB/octave observed in modal voice (see e.g. Figure 2.21 on p. 90 in Stevens, 1998) and a lower signal-to-noise ratio. In consonants specified for the [high] value of [spectral slope] feature, this interval of steep spectral roll-off is observed in the C-V transition.
>
> [low]: A sound that is specified for the [low] value of the feature [spectral slope] has a substantial interval during which there is audible and audibly periodic energy generated at the glottis, but a very shallow or negative spectral roll-off relative to the about −6 dB/octave observed in modal voice (see e.g. Figure 2.17. on p. 86 in Stevens, 1998). In consonants specified for the [low] value of the [spectral slope] feature, this interval of shallow or negative spectral roll-off is observed in the C-V transition.

[11] See Vaux (ms) for an argument that voiceless fricatives should be marked with the feature [+spread glottis], and voiced fricatives are marked with the feature [−spread glottis].

(5) Definition of new acoustically-based [spectral slope] feature

Table XII provides the featural specification of all of the guttural consonants and vowels in Ju|'hoansi. The feature values for [spectral slope] proposed here are applicable to sounds in other languages, as long as the fricated vs. constricted nature of pharyngeals and epiglottals (Esling, 1999b) is taken into account. All non-guttural consonants are unmarked for both the place feature [pharyngeal] (McCarthy, 1991, 1994) and the laryngeal [spectral slope] feature. The binary feature [spectral slope], combined with the privative place feature [pharyngeal], distinguishes just the contrastive sounds in the language, as well as providing the correct natural class of gutturals, which are all marked for one of the two values of [spectral slope]. Similarly, the fact that uvulars and epiglottals both cause vowel lowering and retraction, while laryngeals do not (described in Sections 3.4.4 and 3.4.5), is captured by the specification of uvulars and epiglottals as [pharyngeal] and laryngeals as placeless (following Steriade, 1987). As shown in Chapter 2, and as will be seen in Chapter 5, the acoustic spectral slope values in the C-V transition associated with guttural consonants and vowels marked with the same features are fairly parallel, despite the fact that glottalized consonants probably involve a much greater constriction than is found with glottalized vowels.

I define the place feature [pharyngeal] as in (6), which recognizes the different articulatory strategies that can be used to attain pharyngeal expansion or constriction (Lindau, 1979; Esling, 1999b), such as have been built into recent models of articulation in coordinative structures (Sproat and Fujimura, 1993; Browman and Goldstein, 1992). Miller-Ockhuizen (In progress) shows that epiglottalized vowels involve tongue root retraction and larynx lowering, while epiglottalized consonants involve some degree of tongue root retraction, but also a constriction made with the false vocal folds. That is, the articulation of uvular fricatives and epiglottalized consonants and vowels all employ different amounts of tongue root

	[pharyngeal]		Unmarked for [pharyngeal]	
[spectral slope] value	[high]	[low]	[high]	[low]
Aspirated consonants / Breathy vowels			X	
Glottalized consonants / Glottalized vowels				X
Uvularized consonants	X			
Epiglottalized consonants/ Epiglottalized vowels		X		

Table XII. Featural specification of guttural consonants and vowels

retraction, pharyngeal constriction, and larynx height (Lindau, 1979; Esling, 1999a, b; Miller-Ockhuizen, In Progress). Acoustically, this feature does not imply the presence of spectral noise in the source or any effect on the spectral slope of the vowel, but rather an increase in the first formant frequency, which is found to exist with Arabic pharyngeals and uvulars (Klatt and Stevens, 1969; Alwan, 1989; El-Halees, 1985; Zawaydeh, forthcoming) and has been shown to be present in epiglottalized vowels and formant transitions of pharyngeal consonants in Ju|'hoansi (see Chapter 2). My feature differs from McCarthy's (1994) feature with the same name in that it is not defined in terms of an oro-sensory region in order to include laryngeals. Esling (1999b) has shown that there are many different articulatory movements that are used to produce pharyngeals, and many of these are still poorly understood. However, all such articulations should be unified through the goal of pharyngeal narrowing and the associated rise in F1.

> [pharyngeal]: A sound that is specified with the feature [pharyngeal] involves a constriction in the pharyngeal region, through any combination of tongue root

retraction, pharyngeal narrowing, and larynx movement that results in a high first formant frequency value (either in the vowel or in the C-V transition for consonants) compared with non-pharyngeal sounds.

(6) Definition of the feature [pharyngeal]

The specification of laryngeals for the feature [RTR] in Rose's (1996) feature system or [pharyngeal] in McCarthy (1994) and as used here may depend on whether laryngeals are active in phonological processes such as vowel lowering (Cowell, 1964; McCarthy, 1994; Rose, 1996), or whether they are not active in such processes (Bessell, 1998; Bessell and Czaykowska-Higgins, 1992). Rose (1996) also proposes that the marking of laryngeals with the place feature [RTR] (my [pharyngeal]) is dependent on the presence of pharyngeals in the inventory.

Laryngeal consonants and vowels in Ju|'hoansi are treated as unmarked for the place feature [pharyngeal] based on the fact that they do not pattern with uvularized consonants, epiglottalized consonants, and epiglottalized vowels in causing vowel lowering and retraction, as will be shown in Section 3.4. 4. and 3.4.5. In fact, breathy and glottalized vowels do not block coarticulation of the vowel [a] to [ə] when it precedes high vowels as pharyngeals do, and aspirated and ejected consonants do not cause diphthongization of following front vowels as pharyngeal (uvular, uvularized, epiglottal, and epiglottalized) consonants do. Similarly, Ju|'hoansi laryngeals don't exhibit raised F1 values to the extent found with pharyngeals, and there is no evidence that there is a constriction in the pharynx. Of course, Ju|'hoansi epiglottalized vowels are specified for [pharyngeal] by virtue of the epiglottal constriction.

The model proposed here differs from McCarthy's (1994) model in that pharyngeals and laryngeals are unified through their laryngeal specification, rather then through their place specification. However, since pharyngeals are specified for the place feature [pharyngeal] here, in addition to the relevant [spectral slope] feature value, the model proposed here is not inconsistent with McCarthy’s model.

Halle’s (1995) model of guttural features is provided in Figure 42. My model differs from Halle’s in that it groups aspirates and ejectives together with the same feature, while voiced consonants are specified with the feature [voice]. Thus, my model predicts aspirated consonants and ejectives (as well as breathy vowels and glottalized vowels) should pattern together in some cases, separately from sounds specified for the feature [voice]. On the other hand, Halle’s (1995) model groups pharyngeals, aspirates and ejectives (and their vocalic counterparts) more closely than voiced consonants, which Halle specifies with the feature [slack vocal folds].

Halle (1995:1) claims that the features [spread glottis], [constricted glottis], [stiff vocal folds], and [slack vocal folds] are unified in that they are all executed by the larynx and that tongue root features and laryngeal features are unified under the Guttural node because they are next to each other in the vocal tract. The articulatory proximity of guttural articulators is undoubtedly true. It is the proximity of the constriction in the pharynx in epiglottalized consonants and vowels that causes acoustic results similar to those produced by constriction at the glottis. However, the vocal folds, which are responsible for voicing, are closer to the glottis than the lower part of the pharynx. Therefore, the grouping of pharyngeals (e.g., uvulars and epiglottals) more closely with aspirated consonants and ejected consonants than with voiced consonants cannot be based on articulatory proximity. Yet, there is clear phonological evidence (provided in Section 3.5) showing that these sounds pattern together, independent of voiced consonants and vowels.

The grouping of [high] and [low] values under a single binary feature predicts that aspirated stops and ejectives should pattern together phonologically, even in languages that lack pharyngeals. This prediction is borne out by the laryngeal co-occurrence restrictions between aspirated stops and ejectives in Cuzco Quechua and Peruvian Aymara discussed by MacEachern (1999). MacEachern claims that the restriction is

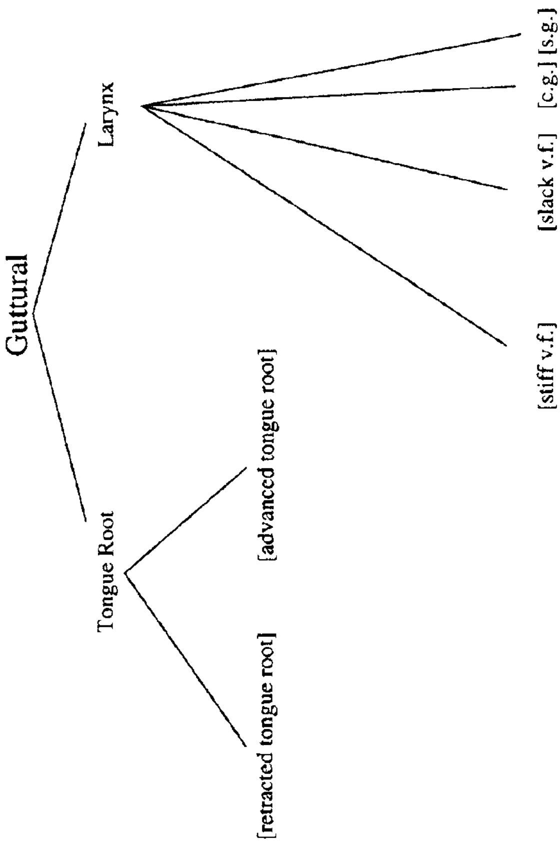


Figure 42. Halle's (1995) Guttural Node

based on acoustic similarity, with the similarity targeting voice onset times involved in these consonants. While her proposal works for the data she describes, VOT cannot be the basis of acoustic similarity in Ju|’hoansi, since VOT is not a relevant acoustic parameter for vowels. On the other hand, the co-occurrence restrictions MacEachern describes can be analyzed using the [spectral slope] feature proposed here, assuming the same type of Optimality Theoretic analysis she proposes to account for degree of similarity.

3.4. Positional Specification

The distribution of features on consonants and vowels in Ju|’hoansi is defined in terms of prosodic positions. Broadly stated, there are three phonological generalizations. First, prosodic word-initial position (which in native roots always corresponds to root-initial position) allows consonants with all manners of articulation and airstream mechanisms, while medial position licenses only sonorants (approximants, flaps and nasals), and final position licenses only nasal consonants. Second, place of articulation features on consonants are skewed in different prosodic positions. However, distributional facts regarding place of articulation are an artifact of the size of the obstruent vs. sonorant inventories at different places of articulation. The third generalization is that guttural features on both consonants and vowels always occur in the initial consonant and vowel positions. It turns out that the distribution of guttural features on consonants can be attributed to the distribution of manner of articulation features. Since obstruents only occur in initial position, and only obstruent consonants can bear guttural features, guttural features only occur in initial position. However, an alignment constraint is offered to account for the distribution of guttural vowels. In Section 3.4.1., I describe the phonotactic patterns with respect to manner of articulation and show that distributional patterns involving Place features are dependent on Manner features. In Section 3.4.2, I discuss the linkage of frequency of obstruent types to their inherent perceptual salience. In Section 3.4.3, I discuss the distributional patterns involving guttural features.

3.4.1. Positional Specification of Manner Features

Ju|'hoansi displays extremely strong distributional restrictions, whereby 89 consonant types occur in initial position, but only 4 consonant types occur in root medial position. Greenberg (1966) and Beckman (1998) noted that clicks occur only in initial position in Khoisan languages, but clicks are only a part of a much larger set of obstruents that occur only in the initial position of a prosodic word. I will first list the full set of initial consonants, and then list the medial consonants. Two of the medial consonants are nasals, and two of them are lenited sonorant versions of plain initial obstruents. The generalization that emerges is that all consonants are licensed in the initial position of a prosodic word, while obstruents do not occur in word-medial or word-final positions.

The set of initial consonants includes a rich set of guttural consonants that involve a constriction at the larynx or the pharynx in their release portion, as well as plain released pulmonic and velaric plosives, and sibilant and guttural fricatives. Nasals also occur in initial position, though as we shall see in the next section they are the least frequent in this position. The full set of pulmonic and velaric plosives with unmarked release properties that are observed in initial position are listed in Table XIII.

Sounds listed in parentheses here and throughout this dissertation, are very marginal based on their low frequency of occurrence. In the set of initial consonants, all of the words that begin with [w], [y], [f] and [v] can be identified as loan-words, and are listed in appendix B. However, in medial position, as we shall see shortly, some sounds occur in words that I am unable to identify as loan-words, but they are very rare.

	GLIDES	FRICATIVES	STOPS	AFFRICATES	CLICKS	NASALS
LABIAL	(w)	(f),(v)	p, b			m
DENTAL		s, z	t, d	ts, dts	|, g|	n,ŋ|
PALATAL	(y)	ʃ, ʒ		tʃ, dtʃ	ǂ, gǂ	ŋǂ
POST-ALVEOLAR					!, g! ‖, g‖	ŋ! ŋ‖
VELAR			k, g	k", g?		
UVULAR		χ				
GLOTTAL		ɦ	ʔ			

Table XIII. Full set of contrastive plosive types that are unmarked for guttural release properties and that occur in root-initial position

The language also contains a rich set of consonants with guttural release properties that only occur on initial consonants. These are listed in Table XIV. Notice that there are no dorsal nasal consonants in initial position. I attribute this to the combination of the low frequency of initial nasals of all places of articulation and the low frequency of dorsal consonants in all root positions.

Table XV lists the frequency of occurrence of sounds in initial position grouped by manner of articulation, merging velaric and pulmonic plosives, since these sounds are phonologically both [-continuant] obstruents. Note the extremely low frequency of sonorants in initial position (2%). In the next section, I will discuss differences in frequency of occurrence in initial position of velaric and pulmonic plosives, but that is not relevant to the phonological constraints on manner of articulation described in this section.

	CLOSURE PROPERTIES			
RELEASE PROPERTIES	VOICELESS		VOICED	
ASPIRATED (ORAL)	p^{h}, t^{h}, k^{h} ts^{h}, tʃ^{h}	ǀ^{h}, ǂ^{h}, ǃ^{h}, ǁ^{h}	b^{ɦ}, d^{ɦ}, g^{ɦ} ds^{ɦ}, dʃ^{ɦ}	gǀ^{ɦ}, gǂ^{ɦ}, gǃ^{ɦ}, gǁ^{ɦ}
ASPIRATED (NASAL)	—	ŋǀ^{h}, ŋǂ^{h}, ŋǃ^{h}, ŋǁ^{h}		ŋǀ^{ɦ}, ŋǂ^{ɦ}, ŋǃ^{ɦ}, ŋǁ^{ɦ}
GLOTTALIZED	ts’, tʃ’	ǀ^{ʔ}, ǂ^{ʔ}, ǃ^{ʔ}, ǁ^{ʔ}	dts’, dtʃ’	—
UVULARIZED	t^{χ},ts^{χ}, tʃ^{χ}	ǀ^{χ}, ǂ^{χ}, ǃ^{χ}, ǁ^{χ}	d^{χ},dz^{χ},dʒ^{χ}	gǀ^{ʁ}, gǂ^{ʁ}, gǃ^{ʁ}, gǁ^{ʁ}
EPIGLOTTALIZED	t^{ʜ}, k^{ʜ}	ǀ^{ʜ}, ǂ^{ʜ}, ǃ^{ʜ}, ǁ^{ʜ}	—	gǀ^{ʕ},gǂ^{ʕ}, gǃ^{ʕ}, gǁ^{ʕ}

Table XIV. Full set of guttural release properties on initial plosives

	OBSTRUENTS		SONORANTS	
CONSONANT TYPE	PLOSIVES (PULMONIC AND VELARIC)	FRICATIVES	SONORANTS (NASAL OR LIQUID)	TOTAL
LEXICAL FREQUENCY	1708	137	33	1878
PER CENT	91%	7%	2%	100%
	98%		2%	100%

Table XV. Lexical frequencies of root-initial consonant types in Ju|’hoansi

In contrast to the large number of sounds in initial position, there are only a few sounds that occur medially. These are coronal and labial sonorants. The full set of medial consonants is listed in Table XVI. Again, the sounds that occur in medial position with a frequency of less than 5% are listed in parentheses. The actual frequencies are provided in Appendix B.

The actual frequencies of sonorants and obstruents in medial position are provided in Table XVII. As we can see, the percentage of obstruents found in medial position is the same as the percentage of sonorants found in initial position. There is then a clear split for manner of articulation, with obstruents

	NASAL	SONORANT	FRICATIVE	PLOSIVES
LABIAL	m	β		
CORONAL	n	ɾ		
DORSAL	(ŋ)	(ɣ)		(nk, k)
UVULAR			(χ)	(kˤ)
GLOTTAL			(ɦ)	

Table XVI. Inventory of medial Ju|'hoansi consonants

Obstruents	Sonorants	Total
18	773	791
2%	98%	100%

Table XVII. Frequency of medial consonants in bisyllabic roots

occurring in initial position and sonorants occurring in medial position.

Focusing first on manner of articulation and ignoring the sounds in parentheses, we can state the generalization that consonants found in medial position are all sonorants. I analyze the sounds [β, ɾ] as sonorant allophones of the initial voiced obstruents /b/ and /d/, following Greenberg (1966) and Snyman (1970). There are several reasons underlying this assumption. First, [β, ɾ] are the only sounds in the language that only occur in medial position. In contrast, the sonorant nasals, [m, n] that occur in medial position, also occur in initial sonorant nasals, [m, n] that occur in medial position, also occur in initial position, and [m] occurs in final position as well. Second, the weakening of obstruents to sonorants is common cross-linguistically, and we find examples such as English flapping and Spanish spirantization. Third, there is evidence from loan-word adaptation, where Afrikaans or British English source words with a medial coronal or labial obstruent tend to be adapted with these variants. For example, the Afrikaans word *pampoen*[pʰəmpʰún] is assimilated as *pabu* [pʰə̀βú] in Ju|'hoansi, and the Setswana word *podi* is assimilated as *pari* [pə́ɾí]. The few exceptions include Afrikaans *Koppie* ‘cup’ which is Ju|’hoansi [kópì], and [bòpí] from Afrikaans *pop* ‘doll’. While the number of sonorants in initial position is very low, words with initial sonorants are not necessarily loan-words. The words themselves are basic vocabulary such as [màˤá] ‘to carry child on the back’, [mùˤnì] ‘to speak a non-click language’, [nə̀ɾ̃ì] ‘to be slow’, and [nòˤḿ] ‘to hook (a springhare)’, and the words are generally in accordance with all other phonotactic constraints described in this chapter.

Recall from the data in (3) in Section 3.2.2. that trisyllabic loan-words and bisyllabic loan-words that become trisyllabic through vowel epenthesis in adaptation, are parsed into two prosodic words, with the prosodic word boundary being placed before the medial obstruent. For example, the Afrikaans word *appel* is parsed as [{ʔáá}{pə̀ɾə̀}], and the Afrikaans word *knoop* ‘button’ is parsed as [{kónó} {!béé}]. In all of the recordings in the database used in this study, trisyllabic words are never parsed with the obstruent medially, yielding a form like * [{koo} {nobe}]. Based on this evidence, the phonotactic constraint given in (7) aligns every obstruent to the beginning of a prosodic word. As seen by Table XV, the probability of this constraint in .98. That is, 98% of all roots have initial obstruents. The probability of a medial consonant being a sonorant is also 0.98. It is likely that both generalizations constitute native speakers’ knowledge of the language. However, the constraint proposed in (7) is likely the most useful in parsing the incoming speech stream into words.

> ALIGN(OBSTRUENT, L; PRWD, L): The left edge of every obstruent corresponds to the left edge of some prosodic word.
>
> (7) Positional specification of manner features

The alignment constraint in (7) interacts with the constraints LINEARITY (McCarthy and Prince, 1995b) and IDENT [-SON] that are provided in (8).

/nə̀bò/	LINEARITY	ALIGN (OBSTRUENT, L;PRWD, L)	IDENT [-SON]
a. [bə̀nò]	*!		
b. [nə̀bò]		*!	
c. [nə̀βò]			*

Table XVIII. LINEARITY, ALIGN(OBSTRUENT, L; PRWD, L) >> IDENT [-SON]

IDENT[-SON]: Corresponding segments have identical values for the feature [sonorant].

LINEARITY: S_1 is consistent with the precedence structure of S_2 and vice versa.

(8) Additional constraint definitions (McCarthy and Prince, 1995b)

It is more important for an obstruent to occur at the left edge of the root than it is to preserve the [-sonorant] specification on the consonant. As a result, obstruents become sonorants in medial position. This is attributed to the high ranking of ALIGN (OBSTRUENT, L; PRWD, L) above IDENT [-SON] as shown below in Table XVIII for the input root /nə̀bò/ 'to gather wild food for a few days'. Crucially, the constraint LINEARITY 'No METATHESIS' (McCarthy and Prince, 1995b) is also ranked above IDENT [-SON] to rule out the possibility of metathesis causing a root like /nə̀bò/ 'to gather wild food for a few days' to be parsed as *[bə̀nò]. Instead, the correct surface form is [nə̀βò], which has two sonorants, and only violates low-ranked IDENT [-son]. There is no evidence for any ranking between LINEARITY and ALIGN(OBSTRUENT, L; PRWD, L).

The only consonants found in final position are nasals, which we know are moraic because they are tone bearing, as evidenced by the contrast between monotonal and bitonal CVm roots such as *ǃòm̀* 'dew' and *ǃòḿ* 'knot (in wood)'. I propose that it is the undominated constraints WEIGHT-BY-POSITION (Hayes, 1989; Morén, 1999), and *[-nasal]=μ provided in (9), that account for the observation that only nasal consonants occur as coda consonants. There are no non-moraic codas in the language. That is, these constraints have a probability of 1.

WEIGHT BY POSITION: Coda consonants are moraic.

*[-NASAL]=μ: Non-nasal consonants are not moraic.

(9) Manner constraints on PrWd-final position

In the next section, I will turn to the question of why certain consonants are disproportionately frequent in initial position, while others are disproportionately frequent in medial position. As I argue, the frequency of the consonant in any given position can be explained by taking into account the amount of acoustic modulation found between that sound and surrounding vowels. Sounds exhibiting strong modulation with a following vowel are better root-initial consonants, and consonants exhibiting little modulation with surrounding vowels are better medial and final consonants.

3.4.1.1. *Frequency of Initial Clicks Based on Perceptual Salience*

The frequency of different consonant types is, I would suggest, related to the perceptual salience of these sounds, measured subjectively based on the degree of modulation on any acoustic dimension that the consonant presents relative to adjacent vowels. Table XIX below provides the lexical frequencies of initial obstruents. The most important thing to note is the low frequency of pulmonic plosives compared with the higher frequency of velaric plosives, as has also been noted by Traill (1985).

Looking at the number of pulmonic and velaric obstruents in the Ju|’hoansi consonant inventory, we see that the number of

CONSONANT TYPE	VELARIC PLOSIVES (CLICKS)	PULMONIC PLOSIVES	FRICATIVES	TOTAL
LEXICAL FREQUENCY	1281	427	137	1845
PERCENT	69.43%	23.14%	7.43%	100%

Table XIX. Lexical frequencies of root-initial obstruent types by manner

CONSONANT TYPE	VELARIC PLOSIVES (CLICKS)	PULMONIC PLOSIVES	FRICATIVES	SONORANTS	TOTAL
LEXICAL FREQUENCY	47	32	6	(3)	88
PERCENT	53.4%	36.4%	6.8%	3.4%	100%

Table XX. Percentage of consonants by manner of articulation in the consonant inventory

obstruents of each type is provided in Table XX below. As we can see, clicks account for 53% of the total consonants in initial position, but they actually occur initially in 69.43% of roots, while pulmonic plosives, which account for 37% of the total consonants that occur in initial position, only occur in 23.14% of the roots. The percentage of root-initial clicks is thus disproportionately high, and the percentage of root-initial pulmonic plosives is disproportionately low. The low occurrence of fricatives in root-initial position (7%) is exactly proportional to the low number of fricatives in the inventory (7%). Sonorants are disproportionately low as previously noted.

I claim that the higher frequency of clicks is related to the fact that they offer more robust perceptual landmarks than pulmonic stops by virtue of having two releases and strong intensity modulation between themselves and neighboring vowels. Plosives and affricates have the second highest frequency, which, I claim, is due to their salient bursts and ability to bear guttural release features that enforce strong periodicity modulation. Sibilant fricatives and guttural fricatives are much lower in frequency, but the lower frequency is expected given the smaller number of fricatives in the inventory.

Traill (1997) has also shown the high perceptual salience of clicks through experimental investigation. In his study on the perception of !Xóõ consonants, clicks were almost never identified from formant transitions alone, while the non-click consonants showed high degree of correct identification from formant transitions when the bursts were cut out. Traill concludes from this that clicks are identified on the bases of the noise of the stop burst alone. Based on these perceptual data, Traill concludes that the burst is more important in the perception of click type than in pulmonic obstruent place of articulation. This might also be due to the large number of guttural contrasts that occur in the transitions and obscure place information in the clicks that have these release properties. Thus, the attention to the bursts for place of articulation information might show a language specific effect similar to the way Korean listeners are more tuned in to the transition that they need to focus their attention on for the identification of the three way contrast between tense, lax, and aspirated stops (Hume et. al., 1999). That is, the presence of a large set of guttural release properties on clicks may cause listeners to attend more to the stop burst noise than to transitions, which are obscured by the strong guttural release properties such as pharyngeal constriction found in uvularized and epiglottalized consonants. The high-frequency noise present in click bursts might also make them stand out better against noise in the C-V transition associated with the various guttural release properties (e.g., aspiration, uvularization, glottalization and epiglottalization), and would therefore provide an advantage for velaric plosives over pulmonic plosives, which have weaker lower-frequency bursts. The weaker pulmonic noise-bursts would be masked more by the noise associated with guttural phonation types than stronger click noise-bursts.

It is the combination of intensity and frequency modulation which makes clicks perceptually superior. We see that even though sibilant fricatives in Ju|’hoansi contain frequency modulation somewhat parallel to frequency modulation to clicks, clicks are still by far the most frequent initial consonants in the language. Sibilant fricatives are not much more frequent than guttural fricatives, which do not show much modulation between themselves and the following vowel at all. Thus, the intensity modulation itself is not what makes clicks better prosodic word edges than other Ju|’hoansi consonants. Rather, the intensity modulation makes the noise bursts stand out even more from the following vowel than bursts found on pulmonic plosives, and the loud noise-bursts focus the listener’s attention on the high-frequency energy found in them. That is, Traill’s (1997) claim that clicks are enhanced stop consonants could explain why clicks are the most frequent initial consonant, followed by pulmonic plosives. The phonotactic constraint that aligns an obstruent to root-initial position and the salience of different obstruent types (particularly in combination with guttural release types that dominate the language inventory) determines their relative frequency.

Nasal non-click consonants [m, n, ŋ] are the only consonants that can be found in all three positions of the word. However, they are also the least frequent sounds in both initial and medial positions. This is in stark contrast to the fact that nasals are the only licit consonants found in final position of roots. This could be related to the increased perceptibility of nasals in coda position as found in perceptual experiments reported in Winters (2000, 2001).

However, there could also be an historical explanation. Roots with final nasals are only preceded by the vowels [ə] and [ɔ], while monosyllabic roots with open syllables and the final syllable of bisyllabic roots can have all modal vowel contrasts. The only other context where the vowels [ə] and [ɔ] occur is in the first syllable of bisyllabic roots that have the high vowels [i] and [u] in the second syllable. Also, while there are many roots that contain ə-u (or əu – u sequences) and ɔ-u sequences with medial [β, ɾ, n], there are no such roots with medial [m]. As was suggested to me by David Odden, these two facts could be related, if high vowels were lost after [m]. This would explain the observation that [m] is the only coda consonant and that Cə mu and Cɔmu roots are lacking in the language. Haacke (1999) suggests that both final [m] and [n] in Khoekhoe arise through loss of final vowels.

Place of articulation features also display skewed distributional patterns, but the distributional regularities based on place are all proportional to the number of obstruent types in the inventory that bear a particular place and the observation that obstruents account for over 95% of initial consonants. As shown in Figure 43, which plots frequencies of place in initial position against frequency of place in medial and final positions, corono-dorsals occur only in initial position. Coronals are fairly frequent in both initial and medial positions, but never occur in final position. On the other hand, labials are very infrequent in initial position, much more frequent in medial position, and occur almost exclusively in final position. Uvulars and dorsals are infrequent in all positions.

The frequency information regarding corono-dorsals can be explained by the alignment constraint in (6), since all corono-dorsals in the language are obstruents, being either clicks or coronal obstruents with uvular release features. The constraint thus captures the fact that these obstruents are always aligned to the left edge of a prosodic word. The fact that labial sounds display extremely low frequency in initial position can be accounted for by the fact that there are only 4 labial obstruents in the inventory, [p, b, p^h, $ɓ^ɦ$], compared with 76 coronal obstruents (see Tables II and III). Even if click consonants are not counted as coronals, there are still 27 coronal obstruents. The smaller number of labial obstruents in the inventory leads to the expectation that there should be fewer labials in initial position. The increased frequency of labials in medial position is due to the fact that there are no corono-dorsal sonorants, and

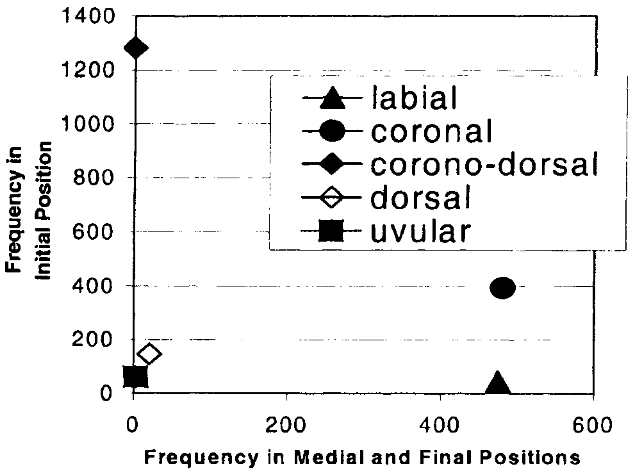


Figure 43. Plot of frequency of place features by root position

uvular and dorsal consonants are very low frequency overall, leaving coronals and labial sonorants to fill in the gaps.

The constraints in Section 3.3.1. above, which account for the observation that all coda consonants are moraic nasals, do not account for the generalization that the majority of final nasals are labials (there are only a few dorsal nasals and no coronal nasals). The low frequency of dorsal nasals in coda position is parallel to the low frequency of dorsal nasals in medial position. The lack of coronal nasals in final position is odd, given the allowance of moraic nasals in the adaptation of loan-words seen above in Section 3.2. and the fact that coronals occur with relatively high frequency in both initial and medial positions. The absence of final coronal nasals may be explained if nasalized vowels were shown to be related to coronal nasal stops historically. Most other analyses of Khoisan languages are parallel to my analysis of Ju|’hoansi, positing nasalized vowels, but only labial and dorsal coda nasals (e.g., Traill’s 1985 analysis of !Xóõ, and Visser’s 1998 analysis of Naro). Nakagawa (1996) posits final labial and coronal nasals but no dorsal nasals, in addition to nasalized vowels, for |Gui. Clearly this is an area ripe for a diachronic study.

3.4.2. Positional Specification of Gutturals

Guttural consonants and vowels are only found in the initial syllable. In the case of consonants, the lack of gutturals in the medial consonant position can be accounted for by the fact that all guttural consonants are obstruents, and obstruents are aligned to the left edge of a prosodic word, as described in Section 3.3.1. Guttural vowels, which include breathy, glottalized, and epiglottalized vowels are only licensed in the initial mora of a root, whether that mora is in the initial syllable of a bisyllabic root or the first mora of a bimoraic monosyllabic root. I offer a constraint that refers only to vowels and assume that the skewed distribution of guttural consonants is due to the alignment of obstruents to prosodic word-initial position offered in the previous section. Further consideration may lead to a unified analysis of the distributional patterns of guttural consonants and vowels.

First, in (10), I provide data showing the three different types of guttural vowels found in Ju|'hoansi monosyllabic roots, and the contrast between partially and fully breathy vowels, as well as between partially and fully epiglottalized vowels. Glottalized vowels are analyzed as glottalized in the first mora and modal on the second mora.

Table XXI shows that vowels in the second moraic position of monosyllabic roots (CV_1V_2) and vowels in the second moraic position of the initial syllable in bisyllabic roots ($CV_1V_2CV_3$) never bear [spectral slope] features that are different from the feature specification found in the initial moraic position of both of these root types. Only roots with non-guttural initial consonants are included in the table, since guttural vowels and guttural consonants never co-occur within the same root, as will be shown in Section 3.5.

BREATHY VOWELS

Partially Breathy		**Fully Breathy**	
g‖aʱò	‘to sprout’	tȁʱȁʱ	‘indigenous citrus fruit’
g|ȍʱò	‘to be dwarf-like’	zȁʱȍʱ	‘grass species’
		gȍʱȍʱ	‘flower’

EPIGLOTTALIZED VOWELS

Partially Epiglottalized		**Fully Epiglottalized**	
gǂàˤá	‘spleen’	!ȁˤāˤ	‘iron’
ǂāˤò	‘dent’	!ȍˤāˤ	‘honey guide’
ǂòˤó	‘pipe resin’	|ȍˤōˤ	‘gall’
nǂàˤáⁿ	‘to throw (liquid)’	n‖òˤòˤ	‘territory’

GLOTTALIZED VOWELS

tȁˀàⁿ	‘to win, beat’
ǂàˀú	‘to be cold’
g|ȍˀò	‘ostrich egg-shell beads’

(10) Three guttural vowel types in monosyllabic roots

μ_1 ↓ μ_2 →	MODAL	GLOTTALIZED	BREATHY	EPIGLOTTALIZED	NASAL	TOTAL
MODAL	255	0	0	0	81	336
GLOTTALIZED[12]	83	0	0	0	0	83
BREATHY	44	0	86	0	10	140
EPIGLOTTALIZED	104	0	0	48	26	178
TOTAL	486	0	86	48	117	737

Table XXI. Co-occurrence of guttural vowels within the same syllable of CV_1V_2 and CV_1V_2CV roots with non-guttural initial consonants

There is clear phonological evidence from the process of reduplication in Ju|’hoansi that indicates that both partially and fully epiglottalized vowels in Ju|’hoansi are vowels, and thus both Cvˤv roots and Cvˤvˤ roots are monosyllabic. As I have shown in Figures 29 and 30 in Chapter 2, the phonetics of epiglottalized vowels clearly indicate that epiglottalization is a vocalic voice quality feature, in accordance with Ladefoged and Maddieson's (1996) claim that these segments are vowels in !Xóõ. However, Ladefoged and Maddieson (1996) also note the similarity of epiglottalized vowels to pharyngeal approximants found in some Arabic dialects. Recall from Section 3.2. that in partial reduplication the weight of the reduplicant is dependent on the tonal pattern of the base. If the base is bitonal, the vowels from each of the two syllables are copied, but the medial consonant is not copied in roots like [‖χə̀.ɾí] 'to fry', resulting in the reduplicated form [‖χə̀í.‖χə̀.ɾí] 'to cause to fry'. If the base is monotonal, as in the root [ŋǂʱə́ɾɪ] 'to clean'; only the initial vowel is copied, yielding the causative [ŋǂʰə́.ŋǂʰə́.ɾɪ] 'to clean thoroughly'. This results in minimal pairs in reduplicated forms of roots like [*màˤ. nì*] 'to speak a non-click language' and [*màˤ.ní*] 'to turn over' reduplicating as [*màˤ.màˤ.nì*] 'to cause to speak a non-click language', and [*màˤi.màˤ.nɪ*] 'to cause to turn over.' If the epiglottal vowel quality were instead a pharyngeal approximant, it would not have been copied. Thus, the causative reduplication patterns in Ju|'hoansi provide clear evidence that roots with epiglottalized vowels of the type listed in (1) above are monosyllabic.

Guttural vowels are also restricted to the first syllable of a root, and the vowel in the final syllable is always modal. The data in (11) show the full set of contrasts involving guttural vowels in bisyllabic roots.

---

[12] In Miller-Ockhuizen (2001) I show that roots transcribed by Dickens (1994) as Cəˀm roots are always bitonal, while Cəm roots are always monotonal. Also, the duration of Cəˀm roots are proportionally longer to the duration of monotonal Cəm roots, as bitonal Caa roots are to monotonal Caa roots. I suggest that CVˀm roots may not be glottalized. Future research will investigate whether the acoustic correlates found with glottalized vowels in this dissertation are also present in CVˀm roots.

Breathy Vowels

kə̏ʱ.ɾȅ	'to praise'	n‖ȍʱ.βȍ	'orphan'
g!ȍʱ.βȅ	'split (in seam)'	dȕʱ.βì	'rash'

Epiglottalized Vowels

nǂɑ̀ˤ.ɾó	'to find something'	nǂɔ̀ˤ.mì	'to twist'
mɑ̀ˤ. nì	'to speak a non-click language'	mɑ̀ˤ.ní	'to turn over'
!ɔ̏ˤɑ̏ˤ.ɾà	'hawk, falcon'	dɔ̀ˤɑ̀ˤ.ɾá	'to be thin'
n‖ɔ̀ˤɑ̏ˤ.ɾȁ	'to squeeze'	nǂɔ̀ˤɑ̀ˤ.ɾá	'black crow'
n‖ɔ̏ˤò.βù	'to wade'	!ɔ̀ˤó.βá	'to fasten on'

Glottalized Vowels

dȁˀà.βà	'to get a fright'	g!ùˀú.βú	'to swell, rise'

(11) Guttural vowels in initial syllable of bisyllabic roots

σ_1 ↓ σ_2 →	MODAL	GLOTTALIZED	BREATHY	EPIGLOTTALIZED	TOTAL
MODAL	313	**0**	**0**	**0**	313
GLOTTALIZED	29	**0**	**0**	**0**	29
BREATHY	27	**0**	**0**	**0**	27
EPIGLOTTALIZED	79	**0**	**0**	**0**	79
TOTAL	448	0	0	0	448

Table XXII. Co-occurrence of guttural vowels in different syllables of CVCV and CVVCV roots with non-guttural initial consonants

The frequency of guttural vowels in the second syllable of bisyllabic roots in the database is provided in Table XXII. This table clearly shows that guttural vowels never occur in V_2 position of bisyllabic roots.

Fully epiglottalized and fully breathy vowels in monosyllabic roots, or in the first syllable of bisyllabic roots, are assumed to have a single [spectral slope] feature specification. That is, the [spectral slope] feature is aligned to the first mora and spreads to the second mora within the same syllable, but is blocked from spreading to a mora in the second syllable. Roots with a specification linked to the second mora, such as roots with an initial clear vowel in the first moraic position, and a guttural vowel in the second moraic position, do not exist.

In order to capture the fact that the [spectral slope] feature is always aligned to the initial moraic position of the root, I posit the constraint in (12). This constraint captures the fact that guttural vowels always occur in the first moraic position of the root, and that the vowel in the second moraic position is always modal, unless it bears the same guttural feature as the initial vowel. The observation that guttural consonants do not occur in medial position is accounted for by the fact that there are no guttural sonorants. However, the generalization that guttural features always align to the first syllable of the root, whether they attach to a consonant or a vowel can not easily be captured. If the first mora of a root can be shown to be prosodically strong on other grounds, than one might adopt an analysis in terms of strength. This constraint has a probability of 1.

ALIGN([SPECTRAL SLOPE]$_V$, L; μ_1, L): Align a [spectral slope] feature on a vowel to the initial moraic position within the prosodic word.

(12) Positional specification of a [spectral slope] feature

3.4.3. Conclusion

In this section, I have shown that all obstruents are aligned to the left edge of a prosodic word in Ju|'hoansi. Since native roots always correspond to a prosodic word, this means that obstruents will be aligned with the left edge of the root. However, we saw that loan-words that contain more than one obstruent are often divided into several prosodic words, satisfying the constraint by creating a new prosodic word boundary to align the obstruent to within the root. In native roots and loan-words with two obstruents that are shorter than a single foot, medial obstruents are weakened, becoming sonorants in order to satisfy the constraint by shedding the [-sonorant] feature. It seems that both weakening and splitting the word into two different prosodic words are both plausible repair strategies for roots containing illicit medial obstruents. The distributional regularity of place features (Figure 43) has been shown to be proportional to the number of consonants at that given place of articulation bearing the manner features licensed in that position (Tables XIX and XX). Guttural features also only occur in the initial syllable of a root. The fact that guttural consonants occur at the beginning of the root is due to the fact that all guttural consonants are obstruents, and obstruents are aligned to the left edge of the root. Guttural features on vowels also only occur in the initial syllable position.

3.5.
Co-occurrence Restrictions

Given the extremely large inventory of sounds found in Ju|'hoansi, there is also a large number of co-occurrence restrictions on the distribution of these sounds in roots. I will focus on co-occurrence restrictions in Ju|'hoansi that involve either all gutturals, or a subset of gutturals, pharyngeals, (uvulars and epiglottals). That is, I focus on constraints that target the laryngeal feature [spectral slope], and constraints that target the place feature [pharyngeal]. In section 3.5.1., I discuss co-occurrence restrictions found between two guttural features within the prosodic word. In Section 3.5.2, I discuss co-occurrence restrictions between gutturals and tone. I argue that both of these constraints target laryngeal [spectral slope] features specified on all gutturals, and not place features. In Section 3.5.3., I discuss co-occurrence restrictions between [pharyngeal] consonants and the place feature [coronal], and in

Section 3.5.4., I discuss co-occurrence restrictions between [pharyngeal] consonants and vowels and vowel height features. Both of these constraints provide evidence that laryngeals are not specified for the [pharyngeal] place feature, and thus support the claim that the Ju|'hoansi guttural co-occurrence constraint must be targeting laryngeal features.

3.5.1. Laryngeal Co-occurrence Restrictions

Ju|'hoansi displays co-occurrence restrictions between guttural consonants and vowels within the same root that provide important evidence regarding the featural specification of the natural class of gutturals. McCarthy (1994) and Rose (1996) have proposed that gutturals are unified by place of articulation features [pharyngeal] and [RTR], respectively. McCarthy (1994) and Rose (1996) suggest that the place features in the oral cavity are grouped under a separate Oral node to account for the independent patterning of pharyngeals. Halle (1995) proposes that there is a separate Guttural node with daughter Tongue Root and Laryngeal nodes, as was shown in Figure 41. I have proposed that Gutturals are all marked with an acoustically-based laryngeal feature [spectral slope], and that pharyngeals (uvulars and epiglottals) are also specified for the place feature [pharyngeal], while Ju|'hoansi laryngeals are placeless[13]. My model therefore differs from Halle's (1995) model in two distinct ways. First, and most importantly, it predicts that there should be processes that affect pharyngeals (uvulars and epiglottals), aspirates, and ejectives, but not voiced consonants. Conversely, Halle's model predicts that any process targeting gutturals should also affect voiced consonants, since voiced consonants are grouped with aspirates and ejectives under a Laryngeal node. In Halle's model, the laryngeal node is a sister to the Tongue root node with the dependent [RTR] feature that is specified on pharyngeals. Secondly, my model predicts that [pharyngeal] consonants should also pattern with other place features in processes targeting a general place node, but laryngeals should not. I argue that the Guttural co-occurrence constraints found in Ju|’hoansi roots are targeting laryngeal [spectral slope] features, and not place of articulation features, since there are no co-occurrence restrictions found between other place of articulation features.

I first provide evidence that guttural consonants and vowels pattern as a natural class with respect to a guttural co-occurrence constraint. As shown in Table XXIII, guttural consonants do not co-occur with guttural vowels. There are only 16 exceptions to this constraint in the nearly 1900 distinct mono-morphemic roots that have been identified by field linguists (including myself) working on the language. As can be seen in the chart, modally voiced consonants and nasal consonants do not participate in the class of guttural consonants targeted by the phonotactic constraints in Ju|’hoansi, and like plain unaspirated voiceless consonants, they can co-occur with any guttural vowel. This is consistent with Rose’s proposal that gutturals are unified by the presence of the place feature [RTR], as well as McCarthy’s (1994) proposal that gutturals are all united based on the presence of the place feature [pharyngeal]. It is also consistent with my proposal that gutturals are unified by the presence of laryngeal [spectral slope] features that are independent of the laryngeal feature [voice]. However, it is not consistent with Halle’s (1995) model that groups pharyngeals together with all laryngeals.

Recall from Chapter 2 that Ju|’hoansi contrasts modally voiced and voiceless consonants, in roots like [g!óó] ‘petrol’ and [!óó] ‘older brother’, [g!ű́ű́] ‘water’ and [!ű́ű́] ‘name’, as well as voiceless and voiced aspirates as in [g!ʰə́ní] ‘small non-poisonous snake species’ and [!ʰə́ní] ‘fan palm’. Voiceless and voiced nasal aspirates also contrast in roots like [ŋǂʰə̀m̀] ‘to eat’, and [ŋǂʱə̀m̀] ‘spider’. Therefore, whether binary or privative features are used to mark these contrasts, there must always be

[13] I do not deny that laryngeals in some languages may be specified for [pharyngeal] (See Rose, 1996 for related discussion).

some laryngeal specification involved. I assume that voiced consonants are marked with the feature [voice], following Lombardi (1994). Notice that of the 16 counter examples, 12 of these involve sounds with opposing spectral slope feature values, e.g., there are 11 aspirated consonants co-occurring with epiglottalized vowels, and one example of an aspirated consonant occurring with a glottalized vowel. These examples involve a consonant marked for the [high] value of [spectral slope] occurring with a [low] value of spectral slope on the vowel. Only 2 out of the 16 counter examples involve two [high] specifications of [spectral slope] (e.g., roots containing aspirated consonants and breathy vowels), and an additional 2 counter-examples involve two [low] specifications (e.g., roots containing ejected consonants and glottalized vowels). The lower frequency of more similar consonants and vowels supports my claim that the Guttural co-occurrence restrictions in Ju|'hoansi are perceptually-based. If the constraint were articulatorily-based, there would be no explanation for this distribution.

I now turn to an investigation of place co-occurrence restrictions in Ju|'hoansi in order to see if the guttural co-occurrence restrictions proposed here could also be interpreted as part of a larger set of place co-occurrence restrictions. While there are apparent place co-occurrence restrictions, they can be attributed to the positional constraints discussed above in Section 3.4. and the frequency of different manner types at particular places of articulation. Recall from Section 3.4. that the distribution of place of articulation features in consonants is also skewed in different positions. Corono-dorsals are the most frequent in initial position, followed by coronals, with labials being the least frequently attested initial consonants. This corresponds to the opposite pattern in medial position, with labials and coronals being the most frequent. In final position, primarily labials are found. While dorsal and guttural initial consonants rarely co-occur with dorsal and guttural medial consonants, this is expected given the overall low frequency of these consonants in both positions.

		Non-Guttural Vowels	Guttural Vowels			
	C1 ↓ V1 →	Modal	Breathy	Glottalized	Epiglottalized	Total
Non-Guttural Consonants	Voiceless	265	70	32	84	451
	Voiced	203	71	31	65	370
	Nasal	155	25	23	70	273
Guttural Consonants	Aspirated	254	2	1	11	268
	Ejected	169	0	2	0	171
	Uvularized	214	0	0	0	214
	Epiglottalized	132	0	0	0	132
	Total	1392	168	89	230	1879

Table XXIII. Co-occurrence of guttural consonants and vowels

The low frequency of labials co-occurring in initial position is proportional to the low frequency of labial obstruents compared with other places of articulation. That is, since only 21 out of 1072 disyllabic roots begin with labials, we would expect only 2% of the 182 roots, or 3.6 roots with medial labials to have initial labials; we find that there are 3 roots that have both labial initial and medial consonants in the database. It is also striking that there are no labial initial roots that co-occur with a final labial consonant, even though labial is the only place of articulation found finally, with dorsal nasals only occurring in a few roots. This is not that unusual given that labial initials are so rare. Given that there are only 119 roots that have nasal coda consonants, we would expect only 2% of these roots, or 2 roots, to contain initial labials and final labials. With these small number statistics, the difference between 0 and 2 occurrences is not likely to reach significance. Given that all skewed distributions of place of articulation features can be accounted for by the frequency of different places of articulation in the obstruent and sonorant inventories, I conclude that there are no independent place of articulation co-occurrence restrictions in Ju|'hoansi phonology.

Given the fact that pharyngeals (uvulars and epiglottals) pattern together with laryngeals and the fact that there are no across-the-board place of articulation co-occurrence restrictions in the language, I propose the Guttural co-occurrence constraint in (13). The constraint is stated in terms of the [spectral slope] feature, which was claimed above in Section 3.3. to mark all guttural consonants and vowels. Based on Table XXIII, this constraint has a probability of. 99.

$$*\{\text{ [spectral slope] [spectral slope]}\}_{\text{PrWd}}$$

(13) Guttural Co-occurrence Constraint

Since pharyngealized (uvularized and epiglottalized) consonants and epiglottalized vowels are all specified for a laryngeal [spectral slope] feature, this constraint straightforwardly rules out the co-occurrence of all guttural consonants and vowels without the need for an additional Guttural node, as proposed in both Halle's (1995) and McCarthy's (1994) feature-geometric models. The Guttural OCP constraint in Ju|'hoansi provides clear evidence for the laryngeal nature of pharyngeal consonants and vowels broadly construed. Notice that voiced consonants are also not predicted to participate in the constraint since voicing is not marked by the [spectral slope] feature, based on the fact that voicing has only a very minor effect on the spectral slope values of the following vowel (as shown in Chapter 2 and Part II). Roots containing initial voiced unaspirated consonants and guttural vowels are plentiful, such as [g‖àʱàʱ] ‘to press down’, [gǂàˀá] ‘wide’, and [g‖àˤàˤ] ‘aunt’.

There is independent evidence (discussed in Sections 3.5.3 and 3.5.4) that laryngeals are placeless, based on the fact that they neither cause vowel lowering nor retraction as is found with uvularized and epiglottalized consonants and epiglottalized vowels in Ju|’hoansi. While Rose’s (1996) proposal that the Place node is split into Oral and Pharyngeal nodes could account for this fact if laryngeals were specified with the Pharyngeal node, but not a [pharyngeal] or [RTR] feature. The specification of glottalized vowels and breathy vowels with secondary place features is unwarranted based on their phonological behavior, as I do not know of any cases where purely breathy or glottalized vowels display co-occurrence restrictions with vowel height or backness in any language.

In this section, I have shown that co-occurrence restrictions between guttural consonants and vowels are extremely strong. I have offered a constraint in terms of the feature [spectral slope], a laryngeal feature that unifies pharyngeals (uvulars and epiglottals) and laryngeal consonants and vowels. Since voiced consonants are not specified for the [spectral slope] feature, statement of the constraint in terms of the [spectral slope] feature correctly predicts that voiced consonants are not subject to the constraint. This is superior to Halle’s (1995) feature geometric tree where the split between laryngeal features and the feature [RTR] (specified on pharyngeals—uvulars and epiglottals), would predict that aspirated and glottalized consonants should pattern with voiced consonants, and the entire class of gutturals must also include voiced consonants. I have also shown that it can not be place features that are being targeted in this constraint, since there are no other place of articulation co-occurrence restrictions found in the language, and there is independent evidence in the language that laryngeals are placeless.

The fact that voiced consonants do not pattern with guttural consonants in the Guttural OCP constraint in Ju|’hoansi suggests that this constraint might have perceptual grounding.

As I showed in Chapter 2, cues to voicing occur during the closure, and this differentiates them from guttural consonants, which all have cues in the release portion. Additionally, as is built into the acoustic feature I have proposed, guttural consonants and vowels all have high or low spectral slopes compared with modal vowels and voiceless and voiced consonants. From an articulatory point of view, it is much more difficult to motivate a closer connection between pharyngeal constriction and glottal constriction or spreading, than between voicing and spread or constricted glottal configurations. Articulatory grounding of the OCP constraint might explain why opposing [spectral slope] features could not co-occur, but would predict that there should be no co-occurrence restrictions found between sounds showing similar articulation, such as aspirated consonants and breathy vowels.

3.5.2. Tone and Guttural Co-occurrence Restrictions

Both guttural consonants and vowels in Ju|'hoansi display co-occurrence restrictions with tone, although the co-occurrence restrictions with guttural vowels are exceptionless, and the patterns found with guttural consonants are more gradient. Guttural vowels only link to moras that contain the two lowest tones. This is not surprising given the lowered fundamental frequency found with these vowel types relative to modal vowels in roots that bear the same lexical tone level, as was shown in Chapter 2 for epiglottalized vowels, which have the strongest lowering. Guttural consonants show more skewed distributions with different root tonal patterns, with voiceless guttural consonants co-occurring more often with the two higher tones and voiced guttural consonants co-occurring more often with the two lowest tone levels.

3.5.2.1.
Description of Root Tonal Patterns

There are seven tone patterns in Ju|’hoansi. Four of the tone patterns are level, and two are bitonal. All six of these patterns occur on monosyllabic roots, as shown in the data in (14), and the seventh pattern is a falling tone (H-L), which only occurs on bisyllabic roots. I will assume throughout this dissertation that single level tones combine to form more complex root tone patterns in the tradition of Goldsmith (1976) (Miller-Ockhuizen (1998) offers a full account of why the missing tonal patterns do not occur, but such discussion is beyond the scope of this dissertation).

Super-Low (SL)		Low (L)	
kə̏ʱȕʱ	‘to light a fire’	|ˣə̀ì	‘to spit’
nǁɔ̏m	‘to bewitch’	ǂɔ̀m	‘to distribute’
High (H)		Super-High (SH)	
n|óá	‘to cook’	!ə̋ű	‘to move house’
ǂə́m ‘to wrap up’		n!ə̋ŋ	‘to put down’
Super-Low – Low		Low - High	
nǁʱȕì	‘to take’	g|ˤùí	‘to twist’
nǁə̏m̀	‘to dance (of women)’	n!ə̀ḿ	‘to be poor’

(14) Six tone patterns on Ju|’hoansi monosyllabic roots

Words exemplifying the seven tone patterns found on bisyllabic roots are shown in (15). The additional H-L pattern is often found on roots that can be identified as loan-words from the Khoisan languages Naro and Khoekhoegowab, where they are native patterns. However, not all such words have been identified as loan-words.

<u>Super-Low</u>
n‖ȍʱ.βo 'orphan'
<u>High</u>
g!ú.βú 'to bubble up'
<u>Super-Low - Low</u>
!ȁ.βɪ 'to color'
<u>High - Low</u>
dtʃ'áɦà 'edible, plum-shaped fruit'[15]

<u>Low</u>
ǂʰà.nà 'to be scraped'
<u>Super-High</u>
g!ő.βɛ̋ 'hare path'
<u>Low - High</u>
n|ò.βá 'to walk fast'

(15) Seven tone patterns on Ju|'hoansi bisyllabic roots

T1 ↓ T2→	SL	L	H	SH	TOTAL
SL	175 (+ 3)	314 (+ 3)	0	0	489 (+ 6)
L	0	636 (+ 13)	305 (+ 6)	0	941 (+ 19)
H	0	21 (+ 11)	371 (+ 6)	0	392 (+17)
SH	0	0	0	57 (+3)	57 (+ 3)
TOTAL	175 (+3)	971 (+ 27)	676 (+ 12)	57 (+ 3)	1879 (+ 45)

Table XXIV. Co-occurrence of tones within roots

The number of roots exemplifying each logically possible combination is summarized in Table XXIV. The number of loan-words exhibiting each pattern is listed in parentheses. As shown above in Section 3.2, roots are mainly bimoraic, with all roots having minimally two moras. However, recall from Section 3.2 that there are 118 trimoraic native roots. As noted in Miller-Ockhuizen (1999), these roots are also bitonal, just as with bimoraic roots. Tritonal patterns are non-existent except on a few of the trisyllabic loan-words (Miller-Ockhuizen, forthcoming).

3.5.2.2 Tone and Guttural Vowel Co-occurrence Restrictions

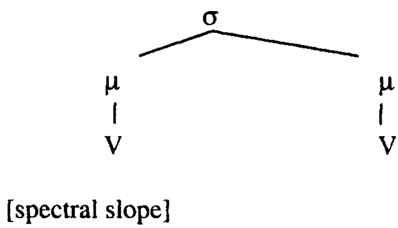
There are co-occurrence restrictions between tones and guttural vowels. First, consider the dependencies between tone and guttural vowel types shown by the frequencies of root tone patterns that bear each of the distinct guttural vowel types. As can be seen in Table XXV, the distributions of tonal patterns that occur with guttural vowels are all skewed. Roots with partially epiglottalized, partially breathy, and glottalized vowels are always bitonal. Roots with fully breathy and fully epiglottalized vowels are always level toned, bearing one of the two lowest tone levels.

	V_1V_2	$V_1{}^{\text{ʕ}}V_2$	$V_1{}^{\text{ʕ}}V_2{}^{\text{ʕ}}$	$V_1{}^{\text{ɦ}}V_2$	$V_1{}^{\text{ɦ}}V_2{}^{\text{ɦ}}$	$V_1{}^{\text{ʔ}}V_2$	TOTAL
SL	15	0	**21**	0	**47**	0	83
L	81	0	**25**	0	**38**	0	144
H	67	0	0	0	0	0	67
SH	20	0	0	0	0	0	20
SL-L	44	**58**	0	**36**	0	**14**	152
L-H	29	**47**	0	**8**	0	**69**	153
TOTAL	256	105	46	44	85	83	619

Table XXV. Tone and guttural vowel co-occurrence patterns on CVV and CVVCV roots with non-guttural initial consonants

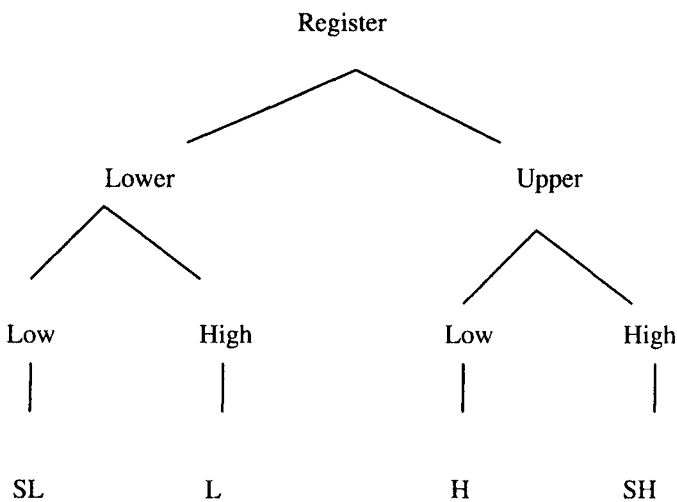
The generalization that emerges is that a mora specified for [spectral slope] may not also be specified for one of the two upper tones. Note that this is under the assumption that the first mora of glottalized, partially breathy and partially epiglottalized vowels is specified for the feature [spectral slope], as shown in (16). Of course, in fully breathy and fully epiglottalized vowels, both moras are specified for this feature.

[15] This is one of the few roots that contains a medial guttural consonant included in Table XXXI in Appendix B.



(16) Representation of glottalized, partially breathy and partially epiglottalized vowels

The generalization that a mora specified for [spectral slope] may not be specified for H or SH tones can be captured using Register features (Yip, 1993). Register features separate four tones into Upper and Lower registers, with each register having High and Low features, as shown in (17), with the four Ju|'hoansi tones written below the features.



(17) Register Tone Features (Yip, 1993)

Ju|'hoansi H and SH tones are both in the Upper register, and thus the co-occurrence restriction can be formulated as a constraint on the [spectral slope] feature and [upper] register feature, as in (18). Register features cannot, however, account for the fact that two tones that occur within the same root are always adjacent on the tonal scale (see Miller-Ockhuizen, 1998 for an analysis of optimal root tone patterns).

> *Guttural/[upper]: A vowel bearing a [spectral slope] feature can not bear an [upper] register tonal feature.

> (18) Tone and Guttural Vowel Co-occurrence Constraint

3.5.2.3. Tone and Guttural Consonant Co-occurrence Restrictions

The co-occurrence of guttural consonants and root tone patterns is provided in Table XXVI, which provides the number of observed roots (O) that bear a certain root tone pattern, the expected number of roots (E) that should bear that pattern, and the ratio of the observed over expected frequencies (R) of each tone by each initial consonant type. All roots that contain guttural vowels are left out of the table, since they cannot bear a guttural feature on the consonant due to the Guttural OCP constraint described in Section 3.5.1. Roots containing the H-L tone pattern are left out due to the low lexical frequency of roots bearing this tone pattern, and SH toned roots are merged with H toned roots, in order to avoid cells with observed frequencies (O values in the table) less than five, given the very low frequency of SH toned roots overall. In this type of analysis, judgments about the frequency of cells with observed frequencies (O values) less than five are not very reliable. Pulmonic and velaric consonants with the same release properties are also grouped together, as are voiceless and voiced epiglottalized consonants, and voiceless and voiced uvularized consonants in order to increase the O and E values. This was possible only because there were no noticeable differences between the frequencies of voiceless and voiced uvularized consonants, and voiceless and voiced ejectives. Voiced and voiced aspirated consonants were not merged, given the noticeable differences in the frequencies of these consonant types with different root tone patterns. Notice that the difference in co-occurrence patterns with voiced vs. voiceless aspirated consonants with H initial vs. L initial tone patterns is predictable from the difference in voicing in the aspirated portion of the consonants that immediately precedes the vowel. Likewise, the similar patterning of voiceless and voiced ejected and an epiglottalized consonant, and voiceless and voiced uvularized consonants, is predictable from the fact that all of these consonant types are voiceless during the release or tongue dorsum lag phase (Section 2.3.3.2).

Shaded cells have ratios of observed to expected frequencies (R values) of less than or equal to 0.6, which shows that the patterns are highly under-represented. Cells with text in bold have expected to observed frequencies (R values) over 1.5, which shows that the pattern is highly over-represented. The main co-occurrence patterns found with non-guttural consonants are the under-representation of voiced consonant initials with the L-H tone pattern and the over-representation of nasal consonant initials with the SL-L tone pattern. Ejectives are under-represented with SL toned initial roots, and voiced aspirates are under-represented with the two highest tones, but over-represented with the SL initial root tone patterns (both SL and SL-L patterns). There is no under- or over-representation with respect to uvularized consonants. A Pearson chi-square test is significant at $p<.001$, which means that these patterns would not be very likely to have arisen from chance alone.

The consonant-tone co-occurrence patterns are not nearly as skewed as those reflecting dependencies between tone and vocalic phonation types. In order to see if they play a role in synchronic phonology, however, would require adapting behavioral measures that have been used only with literate speakers up to now (e.g., Vitevitch et al., 1996; Hay et al., in press; Frisch & Zawaydeh, 1997; Yoneyama, 2001) to

	INITIAL C	SL	SL – L	L	L – H	H / SH	TOTAL
GUTTURALS	VOICELESS ASPIRATES (tʰ, |ʰ, ŋ|ʰ)	O=11 E=11 R=1.0	O=27 E=22 R=1.2	O=65 E=72 R=0.9	O=12 E=22 R=0.5	O=68 E=57 R=1.2	183
	VOICED ASPIRATES (dʱ, g|ʱ, n|ʱ)	**O=10 E=4 R=2.5**	**O=17 E=8 R=2.0**	O=29 E=28 R=1.0	O=9 E=8 R=1.1	O=6 E=22 R=0.3	71
	GLOTTALIZED (ts’, dts’, |ʜ, g|ʕ)[15]	O=8 E=17 R=0.5	O=13 E=35 R=0.4	O=140 E=116 R=1.2	O=26 E=35 R=0.7	O=107 E=91 R=1.2	294
	UVULARIZED (tᵡ, |ᵡ, dᵡ, g|ᵡ)	O=7 E=12 R=0.6	O=13 E=26 R=0.5	O=108 E=85 R=1.3	O=26 E=25 R=1.0	O=60 E=66 R=0.9	214
NON-GUTTURALS	VOICELESS (t, |)	O=21 E=15 R=1.4	O=29 E=31 R=0.9	O=96 E=109 R=0.9	O=27 E=31 R=0.9	O=86 E=80 R=1.1	259
	VOICED (d, g|, ɦ)[16]	O=9 E=12 R=0.8	O=27 E=24 R=1.1	O=55 E=80 R=0.7	**O=46 E=24 R=1.9**	O=64 E=62 R=1.0	201
	NASAL (n, ŋ|)	O=13 E=9 R=1.4	**O=38 E=18 R=2.1**	O=52 E=61 R=0.9	O=17 E=18 R=0.9	O=34 E=48 R=0.7	154
	TOTAL	79	164	545	163	425	1376

Table XXVI. Tone and consonant phonation type co-occurrence patterns (O=Observed frequency, E=Expected Frequency, R=Ratio of observed to expected frequency)

speakers of a language that is only now developing a writing system, and for which there are very few literate speakers (not enough to get the requisite number of subjects needed for psycholinguistic experiments). experiments). The development of analogous tasks for Ju|’hoansi is beyond the scope of this dissertation.

[15] Epiglottalized sounds were categorized with glottalized sounds for the purpose of this chart, although they do not involve glottal closure. Thus, this is actually a perceptual categorization, grouping everything that is marked with the [low] value of the feature [spectral slope] together.

3.5.3. [pharyngeal] and Other Place Co-occurrence Restrictions

The backing and lowering caused by epiglottalization and uvularization results in articulatory incompatibilities with front vowels. This is apparent in the set of diphthongs found in the language. Table XXVII provides the lexical frequency of all of the diphthongs found in the language, including partially epiglottalized and partially breathy vowels. This table includes all of the vowels occurring in CV_1V_2 roots and the first syllable of CV_1V_2CV roots. Roots with initial consonants bearing guttural release properties are omitted from this table, given the co-occurrence restrictions found between guttural consonants and vowels described in Section 3.5. However, back clicks [!,ǁ], which are specified for [pharyngeal], are included in the table, and all of the 17 modal diphthongs, the 6 partially breathy diphthongs, the 4 fully breathy diphthongs and the 6 glottalized vowel tokens containing the diphthongs in vowel quality, [əe] and [əi], occur in roots containing initial back clicks ([!] and [ǁ]. All of the other 29 roots that contain the diphthongs [əe] and [əi] occur in roots containing diphthongs in epiglottalization, and thus these vowels can be analyzed as underlying /!/, with lowering and retraction associated with epiglottalization.

Roots with initial back clicks and diphthongs with epiglottalized vowels are the only segments that co-occur with the cross-height, cross-place diphthongs [ɑˤi], [ɑˤe], [ɑˤo], [ɑˤu], [ɔˤe], and [ɔˤi]. This distributional evidence motivates the description of the low and mid cross-height, cross-place epiglottalized diphthongs as allophones of the diphthongs [əi], [əe], [əo], [əu], [oe], and [ui]. That low back allophones occur with epiglottalized vowels is not surprising, given the raising of F1 and lowering of F2 and F3 shown to exist in epiglottalized low back vowels when compared with modal low

[16] The phone [ɦ] is categorized here as a plain voiced unaspirated obstruent, given the fact that it occurs freely with guttural vowels.

σ V_1 V_2	V_1V_2	$V_1{}^{ʕ}V_2$	$V_1{}^{ʕ}V_2{}^{ʕ}$	$V_1{}^{ɦ}V_2$	$V_1{}^{ɦ}V_2{}^{ɦ}$	$V_1{}^{ʔ}V_2$	TOTAL
[əe] / [ɑe]	2	11	0	0	0	2	15
[əɪ] / [ɑɪ]	15	18	0	6	4	4	47
[əo] / [ɑo]	13	17	0	6	16	1	53
[əu] / [ɑu]	40	15	0	3	5	4	67
[oa]	22	0	6	5	14	5	52
[oe] / [ɔe]	7	7	0	3	3	0	20
[ui] / [ɔi]	20	10	0	2	2	4	38
TOTAL	119	78	6	25	44	20	292

Table XXVII. Lexical frequency of Ju|’hoansi diphthongs

back vowels. Notice that glottalized vowels and breathy vowels do not occur with these extra back allophones, showing that there is no vowel raising accompanying [spectral slope] features. Rather, it is the place feature [pharyngeal] specified on epiglottalized vowels, that causes their incompatibility with front vowels. Thus, this data also provides evidence that breathy and glottalized vowels are not specified for [pharyngeal] place.

As mentioned above, consonants that bear a [pharyngeal] place feature also cause diphthongization on a following front vowel. The consonants that cause this backing are uvular fricatives, uvularized plosives, and back clicks. The two classes of clicks, are provided in (19). The back clicks might now be renamed pharyngeal clicks.

Front Clicks	|	Dental Click
	ǂ	Palatal Click
Back Clicks	!	Post-Alveolar Click
	||	Post-Alveolar Lateral Click

(19) Two Classes of Clicks (Miller-Ockhuizen, 2000; Johnson, 1993)

As shown by the data in (20) (taken from Miller-Ockhuizen, 2000), there are front vowels and back vowels following front clicks and plain coronal consonants in Juǀ'hoansi.

(a) Both Front and Back Vowels Following Front Clicks

Phonetic Transcription	English Gloss	Phonetic Transcription	English Gloss
Front Vowels			
ǀ^{ʔ}èè	'axe'	nǀ^{ɦ}èé	'laughter'
nǀè'é	'one'	gǂȉì	'wrist'
ǂȅ^{ɦ}è^{ɦ}	'malaria'	ǂȅè	'spur'
Back Vowels			
ǀám	'to be thirsty'	ǀȍ^{ɦ}è^{ɦ}	'horse'
ǀȕnì	'leg rattle'	ǂám	'to wrap around'
ǂò^{ɦ}à^{ɦ}	'to copy'	ǂúǂú	'black ant species'

(b) Both Front and Back Vowels Following Plain Front Consonants

Front Vowels			
tȉì	'to shelter s.o.'	tè'é	'coqui francolin'
ts^{h}ȉì	'to laugh'	ts'èè	'small'
dʃ^{ɦ}ìí	'carry on shoulders'	ʃȅè	'to return'
sȉì	'to laugh at'	séé	'to look after'
Back Vowels			
t^{h}úɾú	'to change into s.t.'	to^{ɦ}m	'to hunt'
ts^{h}ùù	'to vomit'	tsȍ^{ɦ}à^{ɦn}	'civet cat'
sàù^{n}	'fine maize meal'	ʃúú	'to lie down'

(20) Back and front monophthongs following coronal pulmonic consonants and front clicks

In contrast to this, the data in (21) show that only back vowels and cross-place diphthongs follow back clicks, uvularized clicks of both classes, and uvularized coronals.

(a) Back Vowels and Diphthongs following Back Clicks

g!àà	‘rain’	!óβó	‘red-billed francolin’
g!ùù^{n}	‘pestle’	ǁəβè	‘hunger’
gǁóré	‘early afternoon’	gǁùú	‘meteor’
g!ə̏ì	‘blue wildebeest’	!ə̏ì^{n}	‘to kick’
ǁə̏ì$^{\text{ɦ}}$	‘to lead’	gǁəí	‘to coagulate’

(b) Back Vowels and Diphthongs following Multiply Articulated Affricates

t$^{\chi}$áá	‘to hit’	dʒ$^{\text{ʁ}}$áá	‘to swim’
t$^{\text{H}}$áá	‘to feel weak’	t$^{\chi}$óɾó	‘red crested korhaan’
tʃ$^{\chi}$òò	‘to swell’	t$^{\text{H}}$úɾí	‘to peer’
tʃ$^{\chi}$ə́í	‘to sing, dance’	ts$^{\chi}$ə̀í.ts$^{\chi}$əβí	‘to grab’
d$^{\text{ʁ}}$ə́í	‘frog’	ts$^{\chi}$ə̀ì^{n}	‘monitor lizard’

(21) Cross-place diphthongs following back [pharyngeal] clicks and uvularized consonants

Based on this evidence, the diphthongs [əi] and [əe] can be seen as allophones of /i/ and /e/ that occur following back clicks, uvularized consonants and epiglottalized consonants. Note that roots with the epiglottalized dorsals [k^{H}] were included as plain back dorsal consonants in Miller-Ockhuizen (2000), although here they are seen as uvularized consonants that bear the feature [pharyngeal]. Thus, while there were eight words listed in Miller-Ockhuizen (2000) that had plain back initials with following [əi] diphthongs, five of these words contain initial epiglottalized or uvularized initial consonants. This leaves only three words with plain back consonants and cross-place diphthongs in the database, [kə̀^{ʔ}ín] ‘to melt, dissolve’, [kə́íʃé] ‘very, much’, and [kə̀ìnkə̀ìə̀n] ‘to hush (a child).’ Recent field-work has shown that there is variation with plain back consonants as to whether the diphthong [əi], or the front vowel [i] is present. Thus, the word [kə́íʃé] ‘really’ is variably pronounced as [kíʃé]. This variation may be due to inter-dialect variability or inter-language variability, since the medial consonant [ʃ] is also not a usual medial phone in Ju|’hoansi, but does occur as a regular medial consonant in the neighboring

Naro language.[18] Plain dorsal consonants seem to be ambiguous as to whether they are [pharyngeal] or not, given the inter-speaker and intra-speaker variability of productions of words like [kə́íʃé] /[kíʃé] ‘really’. Based on the possibility of producing the [i] allophone of /i/ following plain dorsals, I treat them as unmarked for the feature [pharyngeal] in this dissertation.

It is important to note that other guttural consonants, such as aspirated consonants and ejectives, co-occur with both back vowels and front vowels, as shown by the data in (22). All of the data are non-click coronal consonants and front clicks, since these are the only sounds that allow a contrast between front and back vowels, due to their primary place of articulation. Aspirated and ejected consonants also never co-occur with the cross-place diphthongs [əi] and [əe]. This observation provides evidence that laryngeals (e.g., aspirated and ejected consonants) are not specified for [pharyngeal] in Ju|’hoansi.

tz’íí^{n}	‘louse’	tz’úú^{n}	‘nose’
ts^{h}ìì	‘terrapin’	ts^{h}úú^{n}	‘fart’
|$^{?}$íí	‘axe’	ǂ$^{?}$úú	‘joining piece between arrow shift and tip’

(22) Front and back vowels co-occur with laryngeal (aspirated and glottalized) consonants

Given the further back articulation of post-alveolar clicks (Thomas-Vilakati, 2000) and the phonological patterning of back clicks with other [pharyngeal] consonants, I propose that back clicks in Ju|’hoansi are marked with the feature [pharyngeal]. This also extends to other Khoisan languages

[18] During the course of my fieldwork at |Xoan N!huru, Namibia, it was revealed to me that there are numerous loan-words from Naro in the language, which the people attributed to the frequent intermarriage with Naro people before the creation of the South African homeland, *Bushmanland*, via the Odendaal commission in 1969 (Hitchcock,

such as !Xóõ that have the same co-occurrence restrictions described here, known as the Back Vowel Constraint (BVC) (Traill, 1985; Miller-Ockhuizen, 2000). Given my proposal that back clicks are specified for [pharyngeal], the co-occurrence constraint against uvularized consonants and following coronal vowels and the constraint against the simultaneous specification of epiglottalized vowels with the [coronal] place feature can then be unified by the constraint in (23). This rules out the co-occurrence of the place feature [pharyngeal] and the place feature [coronal] within the same or different V-place positions within a syllable, assuming Clements and Hume's (1995) theory. Plain uvular fricatives and plosives with uvularized frication are subject to the constraint, if we assume that these consonants are all specified for [pharyngeal] within the V-place position. This secondary [pharyngeal] place specification is independently necessary, since primary pharyngeal clicks [qʘ], [q|], [qǂ], [q!] and [q‖] in !Xóõ are not subject to the BVC and co-occur with both front and back vowels (Miller-Ockhuizen, 2000). These positions might be seen as representing the timing of uvularization on consonants which always occurs within the C-V transition, even in back clicks, which have been shown by Miller-Ockhuizen (2000) to have a raising effect on the first formant (F1) into the C-V transition and the following vowel, just as is found with other uvular consonants. Notice that this constraint does not block the co-occurrence of [pharyngeal] and [coronal] on a single consonants, since there is a rich set of coronal consonants with

---

1996). Before the creation of the *Bushmanland* and *Hereroland* homelands, both Naro and Ju|'hoansi people moved freely over the entire area which includes both of these territories. After the creation of homelands, Tsumkwe Ju|'hoansi were more confined to *Bushmanland,* while Gobabis Ju|'hoansi (also known as Gobabis !Kung) and Naro were confined to *Hereroland.* This slowed the intermarriage between Tsumkwe area Ju|'hoansi and Naro. The Naro people living within *Bushmanland,* adopted the Ju|'hoansi language and culture, as this opened the way to gaining traditional hunting rights in the area, which were only afforded to Tsumke Ju|'hoansi and their descendants by the South African government.

uvularized releases such as [tχ] as well as uvularized clicks, such as [|$^{\chi}$] that also bear a coronal feature.

*{ [pharyngeal] $_{\text{Vplace}}$ [coronal]$_{\text{Vplace}}$}$_{\sigma}$:
[pharyngeal] and [coronal] can not be specified on the same or different v-place within a syllable.

(23) The Back Vowel Constraint

While this constraint probably reflects the universal articulatory difficulty involved with the production of front vowels with pharyngeal narrowing, it can be violated since Even, a Tungus language spoken in North-Central Siberia, maintains a contrast between front and back pharyngealized vowels (Novikova, 1960 cited in Ladefoged and Maddieson, 1996:306). While it may turn out that the Tungus pharyngealized vowels have less extreme pharyngeal narrowing and tongue root retraction than the Khoisan ones, which would explain the compatibility of front vowels and tongue root retraction in that language, we know that sounds that bear the same phonological features can be phonetically different in different languages. It is therefore necessary to know relevant phonetic properties in order to judge whether a phonetically based co-occurrence constraint is likely to be active (see e.g., Steriade 1999).

There is a co-occurrence restriction found between labial consonants and guttural release properties. Table XXVIII shows the co-occurrence of guttural release properties with different places of articulation of the initial consonant. Note that there is a complete absence of labial consonants with guttural release properties.

The absence of initial labials with guttural release features is perhaps due to the low salience of noise bursts associated with labial consonants. Since guttural release features often mask the formant transitions associated with the initial constriction (particularly in the case of the uvularized consonants), the cue for stop place is the remaining frequency of the noise burst. Since labials have weak formant transitions into a vowel, the

INITIAL C PLACE	LABIAL	CORONAL	CORONO-DORSAL	DORSAL	GUTTURAL 19	TOTAL
TOTAL	0	123	572	41	43	779

Table XXVIII. Co-occurrence of place features with guttural release properties

presence of strong guttural release properties that often mask the formant transitions of the stops may have made Khoisan listeners more attentive to the frequency of the noise bursts.

3.5.4. [pharyngeal] and Vowel Height Co-occurrence Restrictions

Ju|'hoansi consonants and vowels specified for [pharyngeal] place both have a lowering effect on the vowel height of high back vowels, and there are also co-occurrence restrictions with vowel height, that change phonemic front vowels into cross-height diphthongs. Epiglottalized vowels, marked with the place feature [pharyngeal] are never high vowels. This can be seen most clearly in the set of diphthongs listed above in Table XXV. As can be seen, cross-height diphthongs only occur when the initial vowel is epiglottalized. Therefore, the initial vowels in these diphthongs can be viewed as underlying high. That is, the diphthong [ɔˀi] is underlying /uˀi/, and differs from its surface modal counterpart not only in voice quality, but also in the lowering of the initial vowel, making it more distinct from the diphthong [ui].

[19] Guttural is used to refer to all laryngeal and pharyngeal consonants in Ju|'hoansi. The data in (21) shows that laryngeals are not marked for the feature [pharyngeal].

Bessel (1998) and Rose (1996) discuss incompatibilities with consonants that they mark as [RTR] (and I would specify for [pharyngeal]) and preceding and following high vowels in Salishan and Semitic languages. However, there has thus far been no discussion of vowel height co-occurrence restrictions with pharyngealized vowels in the phonological literature, although there have been discussions of height restrictions involved in ATR harmony, which involves a similar articulatory posture. Archangeli and Pulleyblank (1994) assume that there is a single binary feature [+/− ATR], and propose the grounding [−ATR] / HI Condition, which states that “If [−ATR] then [−high].” I state the co-occurrence constraint for Ju|’hoansi in (24) below, which is basically a translation of their condition using the monovalent feature [pharyngeal], rather than the feature [− ATR].

* {[pharyngeal]$_{\text{Vplace}}$ [high]$_{\text{Vplace}}$}$_{\sigma}$

A sound specified for the place feature [pharyngeal] is not [high].

(24) Incompatibility of [pharyngeal] and [high] Features

Notice that this constraint, like all others, is violable, since the Tungus language Even has a contrast in vowel height that is also maintained in pharyngealized vowels (Novikova, 1960 cited in Ladefoged and Maddieson, 1996:306). Again, this is probably linked to the phonetic details of the articulation of these vowels in Even.

3.6. Functional Unity of Constraints

The positional constraints discussed in Section 3.4 and the co-occurrence constraints discussed in Section 3.5, are functionally unified in the sense discussed in Kisseberth (1970), in that they all promote acoustic modulation over the root on some dimension. That is, the positional constraint that aligns an obstruent to the left edge of a prosodic word, discussed in Section 3.4.1, ensures that the majority of roots contain an initial obstruent and that the majority of roots contain a medial sonorant. The alignment constraint thus assures modulation within the root on at least one dimension in bisyllabic roots. In roots containing initial pulmonic stops, this is amplitude modulation, in roots containing initial sibilant fricatives it is frequency modulation, and in roots containing initial velaric plosives (e.g., click consonants) it is both frequency and amplitude modulation. The high proportion of obstruents to sonorants within the consonant inventory assures that the most contrasts are found in the initial consonant position. The positional constraint discussed in Section 3.4.2., which aligns a guttural feature to the initial vocalic position, assures that the most vocalic contrasts are found in vowel-initial position. The proportion of obstruents to sonorants at different places of articulation within the consonant inventory assures that modulation between initial and medial consonants in bisyllabic roots often occurs in terms of both place and manner of articulation, resulting in both frequency and periodicity modulation. Modulation of all types helps eliminate parsing ambiguity.

Since obstruents are the only consonants in roots that can bear guttural features, the alignment of an obstruent to the initial position of the word also means that initial position bears the most consonant contrasts. The alignment of guttural features to the initial vocalic position also assures that the initial syllable contains the most contrasts. These two constraints together thus assure that the size of the cohort is quickly decreased and allow the listener to access the word more quickly. Additionally, the bursts and guttural features licensed on initial consonants help focus listeners’ attention to the word-initial position, and so makes the word boundary more salient.

Modulation between the initial consonant and vowel is always implicitly present, given the amplitude modulation found between pulmonic obstruents and vowels in most languages, and the amplitude and frequency modulation occurring in click-initial roots. Even roots transcribed by

Dickens and Snyman as vowel initial begin with a glottal stop, and thus exhibit amplitude modulation. The modulation found between clicks and following vowels seems to be of a different type than that between initial pulmonic stops and vowels. That is, click-vowel sequences do not involve low amplitude bursts followed by high energy vowels. The high-frequency energy found in click bursts makes these sounds more similar to sibilant fricatives, but the transient stop-bursts are clearly characteristic of stops more generally. Thus, it seems that clicks are so salient because the very effective frequency modulation found in sibilant vowel sequences, is amplified in the extremely loud click noise-bursts.

Kawasaki (1982) and Ohala (1992) have claimed that C-V modulation constitutes an auditory motivation for C-V syllables being the most common syllable type found cross-linguistically, and they claim that the syllable is the universal unit of modulation. However, the alignment of obstruents to the left edge of the root, and sonorants to the middle, results in modulation over the entire prosodic word rather than over the syllable, opening the possibility that modulation could be found over any prosodic constituent cross-linguistically. This suggests that while modulation may be a universal property of languages, the acoustic dimensions employed in modulation and the domain of modulation, are language-specific. The dimensions employed and their domain are mediated by language-specific phonotactic constraints.

A closer look at monosyllabic roots shows that modulation is also achieved over adjacent vowels within a single syllable. The breakdown of roots by different types of contour between the two vowel positions in monosyllabic roots is provided in Table XXIX. Note that just by considering roots with contours in the rime on a single dimension, such as roots exhibiting contours in tone (e.g., [ǂèè] ‘spur’), roots with traditional diphthongs (e.g., [ǂóá] ‘heart’), roots with diphthongs in voice quality (e.g., [gǂɑ̀ˤá] ‘spleen’), and roots exhibiting contours in nasality (e.g [!ɔ́m̃] ‘leg’), almost 70% of monosyllabic roots in the language contain a contour on some dimension within the rime. This contour is in addition to the C-V intensity

Contour Type	Lexical Frequency in Database (Per cent of Roots)
Cv_1v_2	349 (32%)
Cv̀v́	62 (5.7%)
$Cv^{ʕ}v$, $Cv^{ɦ}v$, $Cv^{ʔ}v$	170 (15.6%)
Cvm	177 (16.3%)
Total	758 (69.6%)

Table XXIX. Lexical frequency of roots with different contour types

modulation discussed by Kawasaki (1982) and Ohala (1992) found over the initial syllable in most roots, and the frequency modulation found with clicks and affricates and following vowels found in the majority of roots in the language, which are expressed via syllable structure constraints. Traditional diphthongs involve frequency modulation, tonal contours involve fundamental frequency modulation, diphthongs in voice quality involve periodicity modulation, and contours in nasality involve modulation in amplitude, as well as F1 bandwidth and frequency.

Table XXIX only includes roots with a contour on a single dimension within the rime. The percentages of roots that contain contours on two or more dimensions are listed in Table XXX. If we look at roots with contours in the rime on two or more dimensions, we find that an additional 22.6% of known monosyllabic roots contain contours on two or more dimensions. 6.3% of all known roots contain both a tonal contour (e.g., either SL-L, L-H or H-L root tone patterns) as well as a contour in voice quality (e.g., they are either partially breathy or partially epiglottalized). Another 6% of the known roots in the language contain traditional diphthongs, such as [əi] , [əu], [ao], [oa], as well as a contour in tone. Another 10.3% of roots contain contours on three dimensions, vowel quality (e.g., traditional diphthongs), tone, and voice quality (diphthongs in voice quality). We can now see that 92.2% of the roots in the language exhibit modulation on at least one

Contours in tone & voice quality (monophthongal)	68 (6.3%)
Diphthongs (vowel quality) with tonal contours (modal voice)	65 (6%)
Diphthongs with tonal contour and voice quality contours	113 (10.3%)
Total	246 (22.6%)

Table XXX. Lexical frequency of C-V contour types

additional acoustic dimension, above and beyond C-V amplitude modulation, and additional frequency modulation found in click-initial roots that account for most of the roots in the language.

This leaves only 84 of the total 1088 monosyllabic roots (7. 8% of the known lexicon) that have no contour whatsoever. However, if we were to consider contours between initial consonants and vowels as well, such as contours on roots between guttural specification on consonants and vowels such as [!^{H}áá] ‘heart’ and [gǂòˤòˤ] ‘food plant’, roots with contours in nasality, whereby the initial consonant is non-nasal, and the vowel is nasalized in roots such as [!áá^{n}] ‘side’, it is likely that there would be no root that did not contain a contour in the specification of some feature. That is, there would be no root lacking acoustic modulation on at least one dimension. Thus, although only the Guttural Obligatory Contour Principle constraint was proposed here, it is clear that a much larger role of contour (and its acoustic manifestation of modulation) is at play in determining the structure of the Ju|’hoansi lexicon. This is in accordance with the view set out by Frisch et. al. (1997).

All of the constraints discussed in this chapter are functionally unified in that all of the constraints mediate a single functional goal, namely that roots must display acoustic modulation, This larger functional goal is cross-linguistic, but it has little predictive power. That is, there is no way to predict what type of modulation will be employed in a specific language or in a specific context, and different languages use different strategies for achieving modulation. This grandiose constraint on language can thus be considered external to the phonology, in the respect that it does not determine what features will be targeted by phonotactic constraints in different languages or what types of phonotactic constraints may or may not exist. The database study reported in this chapter follows up on Frisch et. al’s (1997) study of Arabic, by investigating how much of the lexicon of a particular language displays modulation, and which dimensions are involved. With the dimensions of acoustic modulation only identified for Arabic and Ju|’hoansi, we do not know the full range of modulation types used cross-linguistically. Results reported here suggest that modulation is achieved through a variety of acoustic cues. Phonotactic constraints are language specific, but their functional motivation is universal.

3.7.
Conclusion

In this chapter, I have outlined some basic aspects of Ju|’hoansi word prosody. I have shown that native Ju|’hoansi roots consist of a single prosodic word that corresponds to a single foot, but in loan-word adaptation, longer words are often divided into separate prosodic words in order to satisfy several constraints that operate over the domain of the prosodic word. There are two positional constraints, namely a constraint that aligns an obstruent to the left edge of a prosodic word and a constraint that aligns guttural features to the left edge of a prosodic word. There are also several co-occurrence restrictions. The most notable of these is a constraint that blocks the co-occurrence of any two guttural features within the same root. I have claimed that this restriction has its basis in perception. The co-occurrence constraints involving uvularized and epiglottalized consonants described in Sections 3.6, 3.7, and 3.8, are due to articulatory incompatibilities found between pharyngeal narrowing and high tones (Elgendy, 1982), pharyngeal narrowing and front vowels, and pharyngeal narrowing and high vowels (Ladefoged and Maddieson, 1996). These co-occurrence restrictions also result in the increased perceptual distinctiveness found between uvularized and epiglottalized consonants and epiglottalized vowels and their modal counterparts, causing roots to contrast in at least two phonological features (and thus two types of acoustic modulation).

Part II of this dissertation reports on an acoustic case study of several voice quality cues that are relevant in describing the dimension of acoustic similarity targeted by the perceptually-based Guttural OCP constraint in Ju|’hoansi. I assess the level of acoustic similarity found between Ju|’hoansi guttural consonants and vowels. In Chapter 4, the methodology is laid out. The results of the acoustic study, and the implications of the results for the possible perceptual bases of the Guttural OCP constraint in Ju|’hoansi are provided in Chapter 5. Voiced consonants only show a slightly lower level of periodicity following them than do voiceless consonants. The general contour of periodicity in vowels following voiced plosives parallels that found following voiceless consonants, but all guttural consonants and vowels show decreased HNR, as well as either very high or very low spectral slopes compared with modal vowels. Thus, the acoustic properties of guttural consonants and vowels correctly define the natural class of gutturals, which is targeted by several of the phonotactic constraints described in this chapter.

Part II

A Quantitative Acoustic Case Study: The Perceptual Grounding of the Guttural OCP

CHAPTER 4
Methods for Acoustic Case Study

4.0. Introduction

In this chapter, I describe the methodology used in the acoustic case study presented in Chapter 5, which was designed to test the hypothesis that the Guttural OCP in Ju|'hoansi (described in Section 3.5.1) has its basis in the phonetic similarity of guttural consonants and vowels and in the temporal overlap of the acoustic cues associated with guttural consonants and guttural vowels. The dimension of aperiodicity is measured through both the Pitch Perturbation Quotient (PPQ) measure of time-domain aperiodicity (jitter) (Davis, 1976) and through the frequency-based measure of aperiodicity of the gamnitude of the first rahmonic peak (deKrom, 1993, 1995; Hillenbrand, Cleveland and Erickson, 1994; Qi and Hillman, 1997). The dimension of spectral slope is measured via the difference between the amplitude of the first and second harmonics (H1-H2).

4.1. Materials for Acoustic Study

Numerous studies have investigated the voice quality cues associated with breathy vowels (Bickley, 1982; Huffman, 1987; Ladefoged, Maddieson and Jackson, 1988; Kirk, Ladefoged and Ladefoged, 1993; Stevens and Hanson, 1995; Hanson, 1997). A few of these also looked at glottalized vowels (Javkin, et. al, 1987; Blankenship, 1997), but no previous studies have investigated voice quality cues associated with pharyngealized or epiglottalized vowels. Similarly, laryngeal coarticulation, whereby vowels following laryngeal consonants bear acoustic voice quality cues usually associated with non-modal vowels, has also been shown to exist in Swedish (Löfqvist and McGowan, 1992; Gobl and Ní Chasaide, 1999), French, German (Gobl and Ní Chasaide, 1999), English (Löfqvist and McGowan, 1992), Italian (Gobl and Ní Chasaide, 1999), Korean (Cho, Jun and Ladefoged, 2000), as well as Navajo and Tagalog (Blankenship, 1997). However, investigation of acoustic voice quality cues associated with pharyngealized and epiglottalized consonants has never been undertaken, although pharyngeal coarticulation has been shown to exist in Arabic using articulatory measures (Elgendy, 1982), as well as through acoustic formant transitions (Alwan, 1989; Obrecht, 1968). Additionally, Gobl and Ní Chasaide (1999) have shown that the extent of laryngeal coarticulation associated with laryngeal consonants is language-specific, since the amount of coarticulation is much stronger in Swedish than it is in Italian, while it is rarely present in German and French. The current study investigates the degree of guttural C-V coarticulation, as well as voice quality cues associated with guttural vowels.

One reason for the lack of studies investigating voice quality cues associated with pharyngealized vowels and pharyngealized consonants may be due to the methodological difficulty of measuring the spectral slope values of non-low vowels. Since pharyngealized vowels usually are characterized by changes in the first formant frequency compared with modal vowels (Ladefoged and Maddieson, 1996; Shahin, 1997; Bessell, 1998; Lindau, 1979) and since marked first, second and third formant transitions following pharyngealized consonants also exist (Klatt and Stevens, 1969; Delattre, 1971; Butcher and Ahmad, 1987; Zawaydeh, forthcoming), it is difficult to control for vowel quality differences between pharyngealized and modal vowels, even when the same vowel phoneme is used. As noted by Ní Chasaide and Gobl (1997), amplitudes of harmonics are dependent on both formant frequency locations and the fundamental frequency, in addition to glottal aperture. Thus, the amplitude of the second harmonic can be boosted in the speech of people with either a high fundamental frequency, and/or a low first formant frequency, due to its location within the bandwidth of F1. This is one of the reasons why data with the vowel /a/ were chosen for the wordlist used in this study.

In order to avoid problems associated with boosting the second harmonic, words were chosen for this study that contain both low and mid vowels. Both vowel types contain fairly high first formant frequencies, allowing a good separation between the fundamental frequency and the first formant frequency, although the acoustic space is found to be smaller overall for one male speaker who has a lower F1 and a higher F0. For the low vowel [a], boosting of the second harmonic is completely avoided. As we shall see in Chapter 5, the second harmonic of the mid vowel [o] does get minimally boosted for all speakers, but the boosting effect is small enough and the spectral slope differences between the vowel types are large enough, that the results are still robust. The use of [a] and [o] vowels for all contexts also allows minimal pairs to be obtained for all classes of guttural consonants and vowels with minimal differences in formant values across the different types. Such control could not be found for high vowels, since epiglottalized vowels are only non-high back vowels, and uvularized consonants cause lowering of F1 and F2 in high vowel contexts (Miller-Ockhuizen, 2000, Section 3.5.4 of this dissertation).

Fundamental frequency has also been controlled to the extent possible by choosing like lexical tones. As noted in Chapter 1, Ju|’hoansi is a four tone language. Words were chosen that bear SL or L tone levels to the extent possible. However, it is impossible to control completely for F0 differences given the lowering effect on F0 associated with epiglottalized vowels (see Figure 26 in Section 2.3.5). The complete wordlist used in this study is provided in Appendix A.

4.2.
Data Collection and Preparation

Recordings were made with a SONY Digital Audio Tape (DAT) recorder, and an AKG C451E condenser microphone in order to assure the good low-frequency response necessary for spectral investigation and inverse filtering used in the computation of epochs. The wordlist in Appendix A consists of all click-initial words, since clicks offer a robust acoustic landmark associated with the posterior release and since the use of click-initial words allows me to compare laryngeal coarticulation associated with click-initial consonants to previous studies that all involved pulmonic consonants. Non-click consonants could not be included in the study for comparison within the same language, because their low lexical frequency means that it is difficult to find minimal pairs across all consonant and vowel phonation types. The posterior release of the click is used as the beginning of the vowel in clicks with no marked release types, since it always occurs somewhat after the anterior release. Four speakers' productions of the word list in Appendix A are used, with two of the subjects being female (DK and NU), and two being male (KK and KB). Fifteen productions of each word type in Part I, containing three different vowels (/a/, /o/ and /ao/) were recorded, yielding 45 tokens of each vowel phonation type for each speaker. Fifteen productions of each word in part II, subdivided by two vowel contexts (/a/ and /o/), were recorded, resulting in 30 tokens of each type. One difficulty associated with the use of words containing clicks is that recording at a low enough volume to avoid clipping the high intensity part of the click consonants, results in low intensity in the vocalic region. This low intensity, coupled with the lack of periodicity in the signals associated with the phonemic contrasts in the language, may have decreased the lack of success the ESPS *epochs* program had in marking pitch epochs. The full wordlist was recorded in full before moving on to the next repetition. Since I am fluent in the language, I produced the words myself in order to tell the subjects what word I was looking for, and the meaning was discussed each time to ensure that subjects were producing the correct word.

4.2.1. Labeling

After data were collected, they were labeled using the ESPS XLABEL utility associated with XWAVES. The labels used are NASAL (beginning of pre-nasalization), VOICE (beginning of closure voicing), AR (anterior release of the click), PR (posterior release of the click), TR (transient release of the click), EP (end of the periodic portion of the vowel), and V (end of the vowel). A sample labeled token is provided in Figure 44.

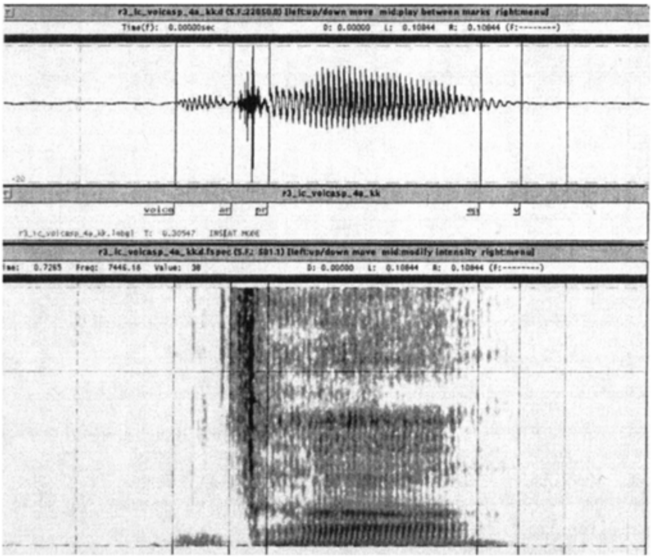


Figure 44. Labeled token of a single production of the root *g|ʱàà* ‘place for drying meat’ produced by male Subject KK

Jitter refers to the variation in the length of pitch periods within a given window, and thus uses pitch periods as the primary data. This makes the accuracy of the measure heavily dependent on the accurate determination of pitch period beginnings and endings. Since measures of fundamental frequency are also very difficult to obtain in non-periodic speech, fundamental frequency was also computed directly from the epoch marks as the inverse of the distance between two pitch epochs averaged over a particular window, since epoch marks were checked by hand for all of the data.

4.2.2.
Pitch Period Determination

Pitch periods were determined as the distance between two adjacent epoch marks. Epochs are the times within each period when the vocal folds obtain complete closure. Epochs were obtained from the residual signal obtained through inverse filtering (Talkin, 1989). Epochs were determined through analysis of only the voiced portions of the signal (i.e., voiceless portions of the signal are gated out) following Davis (1976). This was done by first computing the reflection coefficients of the vowel over a Hanning window of 20 msec., with a step size of 5 msec and 24 LPC coefficients. Probability of voicing, obtained from the results of the ESPS utility *get_f0* was used to gate out unvoiced portions of the signal. The probability of voicing was computed with a step size of 5 msec between the posterior constriction of the click (or the transient release) (PR or TR labels) and the end of the vowel (V label). The residual signal, which approximates the second glottal flow derivative, was computed using the reflection coefficients, and the unvoiced portions of the signal (as defined by having a probability of voicing value of 0) were masked before computing the locations of the pitch epochs. Standard *epochs* program parameters were used, except that the allowed jitter level was decreased from 0.1 to 0.01. Systematically varying the *jitter* parameter, and re-running the program on the epiglottalized vowels that exhibited the highest degree of jitter optimized the identification of epochs and determined the necessary increase. When the higher degree of jitter was allowed as is standard in epoch analysis, pitch doubling was often found, because of the diplophonic nature of the epiglottalized vowels. That is, given the high fundamental frequencies of Ju|’hoansi speakers and the diplophonic signals found in epiglottalized vowels, the level of allowed jitter needed to be constrained more than it typically is with signal processing in European languages that do not show either of these properties.

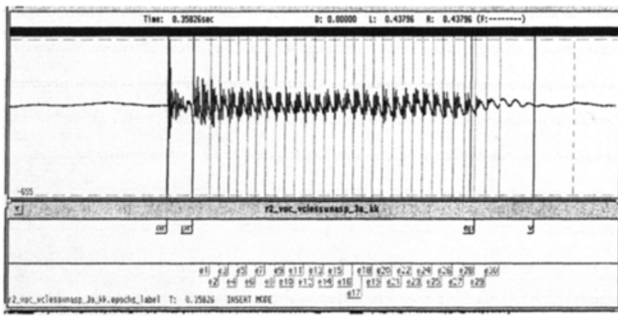


Figure 45. Sample epoch markers on epiglottalized vowel token in the root *!àˁàˁ* ‘iron’

After epochs were computed, I inspected each token visually, and adjusted epoch markers where necessary. Epoch markers were never added before the first epoch marker, except in cases where the first epoch mark was so late in the signal that the program was obviously missing periodic portions of the signal. In cases where the first automatically detected epoch mark occurred before my subjectively marked end of the transient release (TR label), I moved the TR mark back to match the location of the first epoch. However, in the majority of tokens, the first epoch was identified computationally after my more subjective TR mark (the human eye does not place such stringent requirements on what counts as the same pattern as the *epochs* program parameters does). The first epoch mark was used rather than just the TR label, as this provides a more objective measure of the voice onset time, especially in voiced aspirated clicks, where the end of aperiodicity after the aspiration is difficult to judge. The first epoch within the entire signal could not be used to rule out epochs found during the pre-voicing of the click or spurious epochs found during the epiglottalized releases of clicks. A sample file showing epoch markers is provided in Figure 45. Each vertical line represents an epoch location. The time between each two epochs represents a single pitch period. Note that ill-fitted epoch markers outside of the PR or TR and EP labels were ignored in checking the data, as the calculation of jitter only took place between the PR or TR and EP labels.

4.3. Acoustic Measures of Periodicity

Much of the literature on acoustic periodicity has treated aperiodicity in the time domain and in the frequency domain as capturing the same acoustic dimension. However, as noted by Pinto and Titze (1990), time domain aperidocity and frequency domain aperiodicity are capturing slightly different acoustic properties. It is easy to gloss over the differences because most sounds that exhibit aperiodicity in the time domain also exhibit aperiodicity in the frequency domain. Pinto and Titze suggest that frequency domain measures of perturbation are more reliable, since they are more closely related to the fundamental frequency. Qi and Hillman (1997) note that frequency domain measures are also easier to calculate. Time domain measures of aperiodicity are heavily dependent on the methods used to determine pitch periods. Since non-modal voices exhibit variation in the shape of the period from period-to-period, it is difficult to find accurate pitch-tracking algorithms. That is, it is difficult to simultaneously overlook controlled period-to-period shape variation associated with phonemic phonation type contrasts and correctly skip random variation in pitch shape. By using a high quality program to identify pitch epochs, and through thorough hand checking of the data, I can be sure that the pitch periods were correctly identified, and that any jitter found in the data is real variation in the length of pitch periods, rather than an artifact of mis-identification of periods. As noted by Qi and Hillman (1997), time domain measures are less dependent on windowing effects than frequency measures are. Thus, comparison of both time domain and frequency domain measures should yield a more comprehensive picture of overall aperiodicity within the signal.

4.3.1.
Spectral Slope (H1-H2)

A general measure of spectral slope, the difference between the amplitude of the first harmonic and the amplitude of the second harmonic (H1-H2), is used in this study. H1-H2 has been claimed by Cho, Jun and Ladefoged (2000) and Blankenship (1997) to correspond to the open quotient of the vocal folds. That is, when the vocal folds are open during most of the glottal cycle, the spectrum is dominated by the energy at the fundamental frequency (first harmonic), with the second harmonic being much lower in amplitude. When the vocal folds are closed most of the time, as they are in glottalized vowels, the spectrum has more energy in the higher harmonics than it does at the fundamental. Modally voiced vowels exhibit a pronounced spectral roll-off (generally estimated at 6 dB per octave), but the roll-off is not as sharp as that found with breathy vowels. Breathy vowels in Gujerati and !Xóõ were shown by Bickely (1982) to have steeper downward spectral tilt in the region between 0 and 1000 Hz for breathy vowels than for modal vowels. Ladefoged, Maddieson, and Jackson (1988) and Kirk, Ladefoged and Ladefoged (1993) showed similar results by looking at the difference between the amplitude of the fundamental and that of the harmonic with the highest amplitude in the first formant. Jackson et al (1985a, b) argue that H1-H2 is a good, general measure of spectral slope, because H2 is generally a good reflection of the entire spectral slope above H1. Thus, I expect that in Ju|’hoansi, breathy vowels and vowels following aspirated consonants would also have a higher H1-H2 value, while vowels following glottalized consonants and glottalized vowels should have a lower H1-H2 value. Since epiglottalized vowels and uvularized consonants are poorly understood, it is not clear what values they are predicted to have.

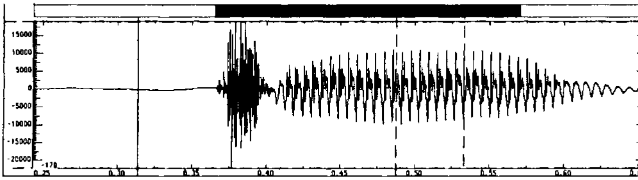


Figure 46. Waveform with the window for the calculation of the spectrum at the midpoint of the vowel in Figure 40 marked

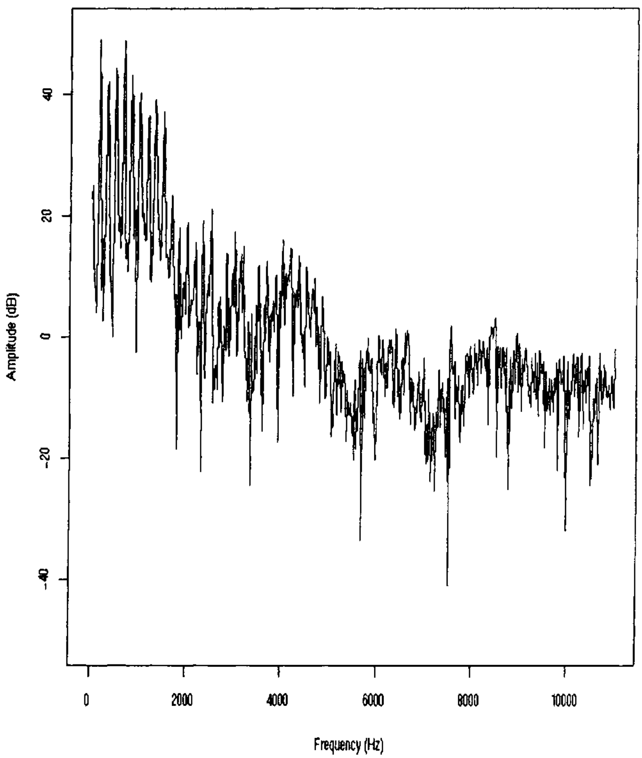


Figure 47. Spectrum at the midpoint of the vowel of a single production of the word ǁáá 'to warm hands by fire' (Subject KK)

Spectra were computed using Fast Fourier Transform (FFT) with 1024 points, resulting in a 46.6 ms window, and 10 coefficients appropriate for a sampling rate of 22050 Hz. The first window started flush with the posterior release of the click in non-released clicks (PR label), or at the end of the transient release in clicks with guttural release properties (TR label). A step size of 10 ms was used, until there was less than 46.5 ms left in the vowel to serve as the relevant window for computation of the spectrum.

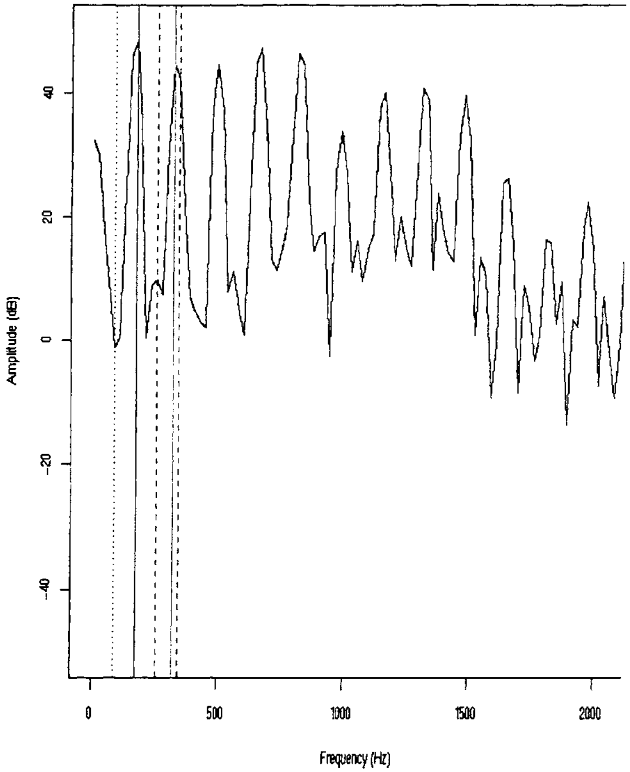


Figure 48. Method of determining frequency bin for H1 and H2

The waveform of an unaspirated consonant is provided in Figure 46. The window over which the spectrum was calculated at the midpoint of the vowel is marked.

A sample spectrum calculated at the window shown in Figure 46 is provided in Figure 47.

The amplitudes of the first and second harmonics were estimated by computing the average duration of pitch periods over the same window that the spectrum was calculated over. The first harmonic is estimated as the inverse of the average duration of the periods within the window, and the estimated second harmonic is twice the estimated first harmonic. Figure 48 shows the first 100 frequency bins in a sample spectrum calculated at the midpoint of the vowel of the word *ǁáá* ‘to warm hands by fire’ produced by male Subject KK. The figure shows the windows used around the estimated frequency bins for the first and second harmonics, and the peak frequencies chosen by the peak-picking algorithm used within the R statistics package.

The GNU R statistics package was used so that each token could be visually inspected while the peaks were being identified, to assure correct identification of peaks. The algorithm also computed the difference between the amplitudes of the first and second harmonic peaks.

4.3.2.
Harmonics-to-Noise Ratio (HNR)

Several measures of harmonics-to-noise ratio (HNR) have been used in the clinical literature. Qi (1992) and Qi et. al (1995) compare the amplitude of the entire shape of the waveform from one period to the next. This method is very reliable, but is also time consuming. When substantial jitter is present in the waveform, it is difficult to compare the shape of the cycle from one cycle to the next. Normalization of the length of the periods must first be accomplished.

DeKrom (1993, 1995), Hillenbrand et. al. (1994), and Qi and Hillman (1997) use the quefrency of the first rahmonic peak in the cepstrum as a robust measure of fundamental frequency.

They then feed the inverse of the estimate of the fundamental into a routine to calculate the locations of the rahmonics, which are then liftered from the signal to obtain an estimate of the noise component of the signal. The HNR is then expressed as a ratio of the amount of energy in the harmonic component to the amount of energy in the noise component.

Blankenship (1997) uses the gamnitude of the first rahmonic peak in the cepstrum as a measure of aperiodicity. This is in fact a measure of HNR. When Fast Fourier Transform is applied to the spectrum, an inverse cepstrum is the result. The first rahmonic corresponds to the fundamental frequency because this component is the lowest frequency that is present in the complex spectral wave. The gamnitude (analog of amplitude) of the rahmonic peak tells us how well correlated the amplitude of the fundamental is with the amplitude of all of the other harmonic components in the closed system. Thus, a low gamnitude peak tells us that the harmonics are not very similar in amplitude across the spectrum, while a high gamnitude peak tells us that the harmonics are all fairly consistent in amplitude within the window under study. Lower HNR can result from high or low spectral slopes, or from general low amplitude harmonics within any range of the spectrum.

In order to create the cepstral files, I used the ESPS utility *fftcep* that is associated with X-Waves, which first takes the Fast Fourier Transform of the waveform, producing a log-magnitude spectrum. Then the program takes the fast Fourier transform of the log-magnitude spectrum to produce a cepstrum. The method used here differs from Blankenship's (1997) and Hillenbrand et. al.'s (1994) methods that compute the cepstrum from the un-normalized spectrum, and follows DeKrom's (1993) method of computing the cepstrum from a log-magnitude spectrum. Figure 49 shows a sample cepstrum computed from a log-magnitude spectrum following the method employed in this dissertation. The cepstra shown here are centered at 0, similar to DeKrom's (1993) cepstra, due to the use of the log magnitude spectrum. The use of the log magnitude spectrum removes the need for normalization to the

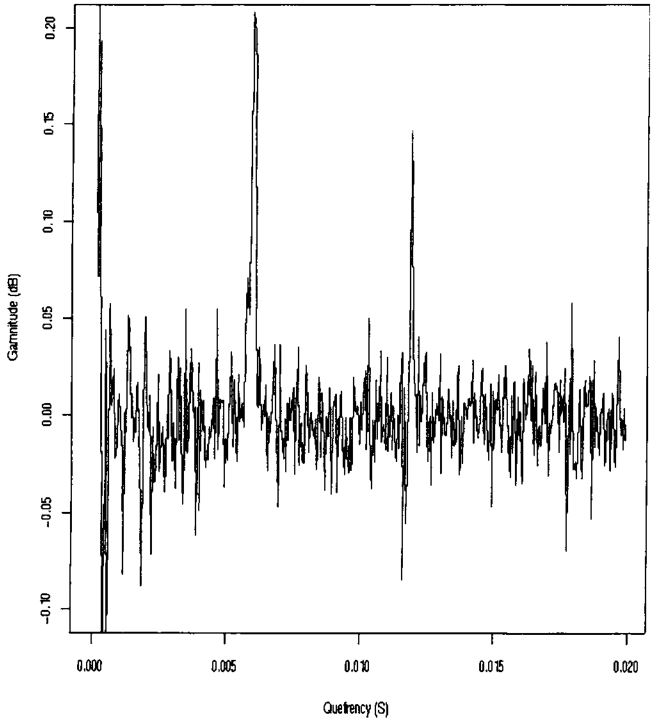


Figure 49. Cepstrum computed over log magnitude spectrum at the midpoint of the vowel in a production of the root word ǁáá 'to warm hands by fire' (Subject KK)

mean of the cepstral gamnitude values used by Blankenship (1997) or the difference between the gamnitude of the cepstral peak and a linear regression of the cepstral gamnitude values used by Hillenbrand et.al. (1994).

The spectra and cepstra were computed with 10 coefficients, and 1024 points, appropriate for a sampling rate of 22050 Hz, using a Hamming window. This resulted in a 46.5 ms window size. Only the real part of the cepstrum was saved. A series of cepstra were computed at a step size of 10 ms from the beginning of the vowel (either the posterior release of the click if it is an unreleased consonant (PR label), or the end of the transient release in the case of a consonant with a non-modal release property (TR label), until the end of the periodic portion of the vowel (EP label). Whatever extraneous part of the signal remained after taking the first 46.5 msec of the vowel, and then stepping through the vowel at 10 ms intervals, was unanalyzed at the end of the vowel. Moving in this direction assures that the first window over which the cepstrum is computed is exactly aligned with the beginning of the vowel. Since non-modal phonation types all occur at the C-V transition in Ju|'hoansi, it is critical to not lose any information right at the beginning of the vowel. Note that the maximum gamnitude of the cepstral peaks in the data used in this study is only 0.2 dB, which is a result of the rather low recording level used to avoid clipping of the click bursts. As was noted in Chapter 2 (Figure 7), clicks have much higher-frequency bursts than the vowels that follow them. This means that the recording level has to be kept rather low, in order to avoid clipping the click bursts, resulting in rather low amplitude vowels.

The pitch period durations computed over the interval between pitch epochs were used to estimate the quefrency of the first rahmonic in the cepstrum at each window. Since the quefrency is the inverse of the frequency, the average pitch period should theoretically correspond exactly to the quefrency of the first rahmonic peak. The average pitch period was computed over the same windows used to calculate the vowel cepstra, including all periods with initial epochs that start after the beginning of the window and epochs that end before the end of the window. These estimates were then fed into a peak-picking program within the R statistics plotting package, and the peak gamnitude found within a 30 frame window around the estimated peak was determined to be the exact 1st rahmonic peak for each window. Figure 50 shows the window around the estimated quefrency as well as the actual peak quefrency that was identified for the cepstrum displayed in Figure 49 above. Peak picking was checked visually, and epoch locations that the estimates were based on in the

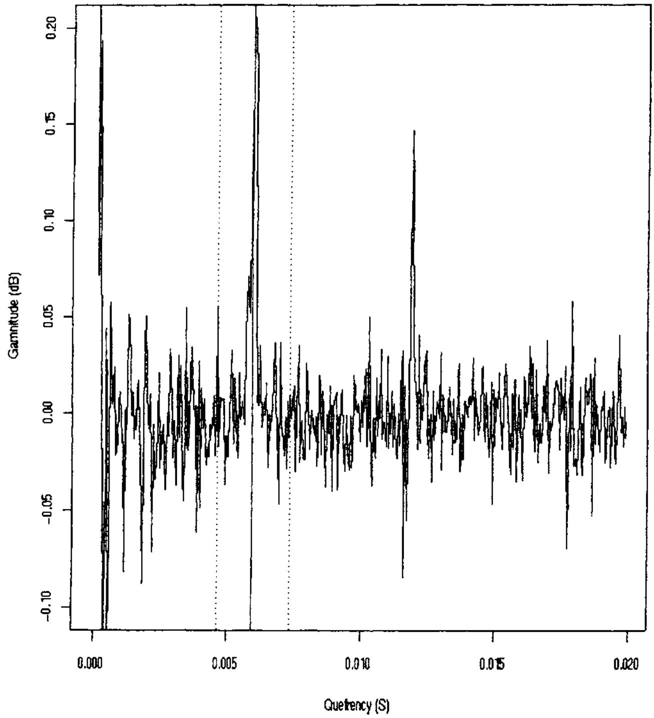


Figure 50. Window surrounding estimated quefrency of first rahmonic peak (marked by dotted lines) and the chosen peak within the window (marked by the solid line)

waveform were visually inspected again for cases where there was any discrepancy. The peak-picking program was then re-run to fix errors. Given the initial hand checking of the epochs, there were very few cases where such errors occurred. In this way, really robust 1st rahmonic cepstral peaks were determined every 10 msec. throughout the duration of the vowel.

The real time of the 1st rahmonic peak is the beginning of the window used to calculate the cepstrum and estimate the averaged pitch period. The times were normalized by subtracting the time of the posterior release of the click from the window time, so as to be able to compare across different files which have different start times, thus allowing any differences in begin times to be associated solely with voice onset time differences. VOT differences were not normalized for, as VOT is an important cue for different consonant release types, and it allows us to see what portions of the vocalic signal are comparable within different consonant and vowel phonation types.

4.3.3. Jitter

The term jitter refers to variation in the length of pitch periods from cycle to cycle. The variation that is sought is random, not changes in pitch period found due to phonologically controlled pitch excursion. Acoustic jitter is caused by changes in vocal fold vibration patterns from cycle to cycle. There may be several different articulatory properties that are associated with jitter. Jitter may be the result of differences in the lengths of the vocal folds from period to period, which causes differences in the duration of the period, or it may be due to glottal closure at irregular time intervals, or due to laryngeal postures that make the vocal folds thick and tense, so that there is non-linear fluctation in the phase of the opening of different parts of the glottis. As there are many distinct ways of lowering and raising fundamental frequency (Honda, 1995), there are probably also many articulatory factors that contribute to acoustic jitter.

The measure of Pitch Perturbation Quotient (PPQ) (Davis, 1976) is used as a measure of mean rectified jitter (Pinto and Titze, 1990) in this study. Jitter was calculated over the first half of the vowel, taking either the posterior release of the click in modally released consonants (PR label), or the end of the transient release associated with the click in guttural click types (TR label) as the beginning of the vowel. The midpoint of the vowel is the end of the window. It was discovered that controlling for lexical tone, and using only level and not rising lexical tones controlled for a lot of the fundamental frequency changes throughout the vowel. However, as was shown in Chapter 2 (Figure 3) there was still a large fall in fundamental frequency at the end of the vowel, and this fall accounted for the majority of change in fundamental frequency within the level toned roots used in this study. Using only the first half of the vowel allowed me to minimize the changes in fundamental frequency over the window being evaluated. The duration of each of the periods, defined as the duration between each two subsequent epoch marks within the window, was calculated. These period durations were used as the raw data for computing jitter.

The actual jitter value was computed using the measure of Pitch Perturbation Quotient (Davis, 1976). The formula for PPQ is provided in (24). In the formula, **To** stands for the duration of the extracted pitch periods, and N stands for the number of extracted pitch periods found in the window. PPQ is considered far more robust against pitch changes (Pinto & Titze, 1990) than earlier measures, such as Relative Average Perturbation (Koike, 1973). The relative success comes from the degree of smoothing present in the formula, which is accomplished over 5 windows. PPQ was calculated over the first half of the vowel, as well as over the entire vowel.

$$PPQ = \frac{\frac{1}{N-4}\sum_{l=1}^{N-4}\left|\frac{1}{5}\sum_{r=0}^{4}\mathrm{T}o^{(l+r)} - \mathrm{T}o^{(l+2)}\right|}{\frac{1}{N}\sum_{l=1}^{N}\mathrm{T}o^{(l)}}$$

(25) Formula for PPQ (Pitch Period Perturbation Quotient)

Initial and peak fundamental frequency values were calculated from the pitch epochs (the inverse of the duration of the period) in order to determine the degree of F0 change over the vowel. The initial F0 was calculated as the duration between the first two pitch epochs after the posterior release of a modally released click (PR label), or the first two pitch epochs after the transient release of a guttural click (TR label). The peak F0 was calculated as the shortest period duration found between any two epoch marks within the first half of the vowel for the first jitter value and over the entire vowel for the full vowel jitter measure. The final F0 was calculated as the inverse of the duration between the last two epoch marks before the EP label (end of the periodic portion of the vowel). Results show that all of the guttural consonant phonation types investigated in this study cause perturbations in the F0 of the vowel. While most effects are local (10–20 msec), voiced aspirated consonants exhibit a very large F0 excurses over the first half of the vowel. Since changes in F0 over the first portion of the vowel associated with guttural release types cannot be completely controlled for, jitter results are plotted against the total change in F0 found over the first half of the vowel (the same window used to calculate PPQ).

Chapter 5 reports on the results of the measures described in this Chapter, and discusses the relevance of the results to the acoustic similarity of guttural consonants and vowels in Ju|'hoansi, relating these to the perceptual basis of the Section constraint found in the language.

CHAPTER 5
Guttural Vowels and Guttural Coarticulation

5.0.
Introduction

In this chapter, I report the results of an acoustic case study designed to test the hypothesis put forth in earlier chapters, that the Guttural OCP constraint in Ju|'hoansi has its basis in perception. That is, I explicitly test the hypothesis that all guttural vowels display a high degree of noise in their C-V transition and that guttural consonants display similar spectral noise in the C-V transition through guttural coarticulation. I investigate the dimensions of periodicity represented by the gamnitude of the first rahmonic peak (Hillenbrand et. al., 1994) and the PPQ measure of jitter (Davis, 1976), as well as the dimension of spectral slope represented by the difference between the amplitudes of the first and second harmonics (H1-H2) (Bickley, 1982).

The periodicity dimension and the spectral slope dimension exemplify the two different types of parsing ambiguity: fatal similarity of cues and mutual cancellation of cues. Periodicity within the frequency domain, represented by the gamnitude of the first rahmonic peak, nicely captures the acoustic similarities across all gutturals. Since the degree of aperiodicity found on vowels following guttural consonants is in some cases as large as that associated with guttural vowels, the only way to differentiate roots containing a guttural vowel from a root containing a guttural consonant would be through lengthened VOT differences associated with the consonants. If both a guttural consonant and a guttural vowel were to co-occur within the same root, it would be impossible to differentiate such a root from a root containing only a guttural consonant. Thus, the co-occurrence of guttural consonants and guttural vowels within a single root is ruled out via the phonotactic constraint described in Section 3.5.1.

On the spectral slope dimension, some pairings of guttural consonants and vowels are acoustically similar, motivating the same dissimilatory restriction as the aperiodicity values. Other pairings of consonants and vowels have conflicting slope values. For example, breathy vowels and vowels following uvularized consonants both display increased spectral slope values, while glottalized vowels and vowels following glottalized consonants exhibit very small or even negative spectral slope values. The co-occurrence of consonants and vowels that should be cued by opposing spectral slope values during the same temporal interval at the beginning of the root would lead to masking if one of the cues were much stronger than the other, or canceling out of cues if they were about the same magnitude. In either case, this would lead to decreased perceptibility of the two contrastive units.

The results of both the cepstral measure of periodicity and the H1-H2 measure of spectral slope also support the pairings between guttural consonant and vowel types proposed in Chapter 2. That is, aspiration on consonants and breathy voicing on vowels are both characterized by high spectral slope values and increased aperiodicity in the frequency dimension. Similarly, epiglottalization on consonants and epiglottalization on vowels, as well as glottalization on both consonants and vowels, are all characterized by low spectral slope values and increased aperiodicity in the frequency dimension during the C-V transition. These results are also consistent with the transcription of epiglottalized consonants and vowels as described in Chapter 2 of Part I.

The dynamic movement from high spectral slope values to unmarked spectral slope values is consistent with the analysis of diphthongs exhibiting breathy voice quality followed by modal voice. Diphthongs that I have represented as having epiglottalization on the first mora and modal voice on the second mora exhibit low spectral slope values over the first half of the vowel and unmarked spectral slope values over the second half of the vowel. The noise levels represented by the gamnitude of the first rahmonic in the cepstrum are also consistent with these vowels being diphthongs.

This chapter is organized as follows: in Section 5.1, I provide the results of the HNR measure, in Section 5.2, I provide results of spectral slope investigation and in Section 5.3, I provide the jitter results. In each section, I first discuss the results associated with guttural coarticulation of guttural consonants, followed by discussion of results associated with phonemic guttural vowel contrasts and the contrasts between monophthongal and diphthongal voice quality contrasts. Finally, I discuss the parallelism shown on each acoustic dimension between guttural consonants and vowels that support the view outlined in Chapter 2, that the differences between roots containing consonants and roots containing vowels, lie primarily in VOT.

5.1. Harmonics-to-Noise Ratio

The gamnitude of the first rahmonic peak within the cepstrum is used as a measure of the harmonics-to-noise ratio (HNR) within this dissertation. If the first rahmonic is low in gamnitude, this means that higher harmonics do not stand out clearly. They either have low amplitude relative to noise (Low HNR) or they are simply low in gamnitude due to period-to-period fluctuation in frequency. In contrast, high gamnitude signals that the higher harmonics are high in amplitude relative to noise present in the signal (High HNR). Gamnitude of R1 is expected to be low for breathy vowels because F0 dominates and higher harmonics don’t stand out well above the noise. Gamnitude of R1 should be low for glottalized vowels because period-to-period fluctuation in length leads to lower amplitudes across the spectrum (as well as greater bandwidth for each harmonic). Guttural C-V coarticulation results in low gamnitude values in vowels following guttural consonants. The low gamnitude values resulting from this coarticulation are similar to the low gamnitude values associated with guttural vowels. Thus, there is little contrast between vowels following guttural consonants and guttural vowels on this dimension. Of course, the consonants and vowels remain distinct given the VOT associated with the guttural consonants.

5.1.1.
Guttural Coarticulation

Vowels following voiceless guttural consonants all display upward sloping gamnitudes of the first rahmonic peaks over the first part of the vowel, as shown in Figure 51, which plots the median gamnitude of the first rahmonic peak within the cepstrum on the vertical axis and real time on the horizontal axis for female Subject NU.

Vowels following voiced guttural consonants also display an upward slope in the gamnitude of the first rahmonic over the initial part of the vowel, as can be seen in Figure 52. Additionally, voiced aspirated consonants and voiced uvularized consonants that behave as gutturals are much lower in gamnitude initially than the vowels following unaspirated consonants. While vowels following voiced unaspirated consonants also display somewhat lower gamnitudes than are found following voiceless unaspirated consonants, the difference is not as marked as that found with the guttural consonants.

Both voiceless and voiced aspirated-initial roots start much lower in gamnitude than the unaspirated consonant-initial roots. Both root types rise in gamnitude throughout the vowel, with the voiceless aspirated-initial roots surpassing the voiceless unaspirated-initial roots in gamnitude about 1/3 of the way through the root. The voiced aspirated-initial roots take longer to rise, but probably attain peak gamnitude at around the same time as the voiceless aspirated-initial roots. The voiceless uvularized consonant-initial roots start out at

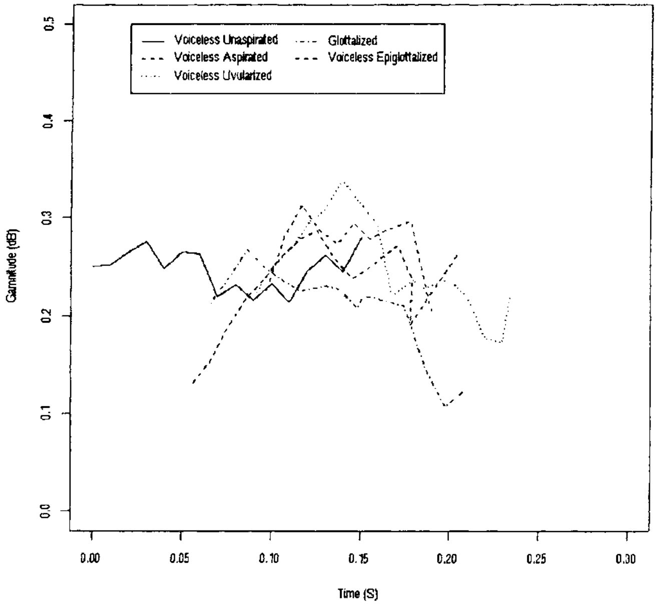


Figure 51. Median gamnitude of the first rahmonic peak (vertical axis) plotted against time (horizontal axis) in voiceless consonant-initial roots for female Subject NU in the [a] context

about the same gamnitude as found with the voiceless unaspirated roots, but they rise in gamnitude over the first half of the root and fall again at the end. This contrasts with voiceless unaspirated roots, which maintain a stable gamnitude level throughout the vowel. The voiced unaspirated-initial roots also maintain a steady gamnitude level throughout the root, although they are slightly lower in overall gamnitude than their voiceless counterparts. The voiced uvularized consonants are also fairly steady in gamnitude throughout the root, falling slightly throughout the root, but they are much lower in gamnitude throughout their duration when compared with voiceless and voiced unaspirated consonant-initial roots. The glottalized click-initial roots also rise slightly in gamnitude over the initial part of the root, and then fall towards the end, following the general pattern found with all guttural consonant-

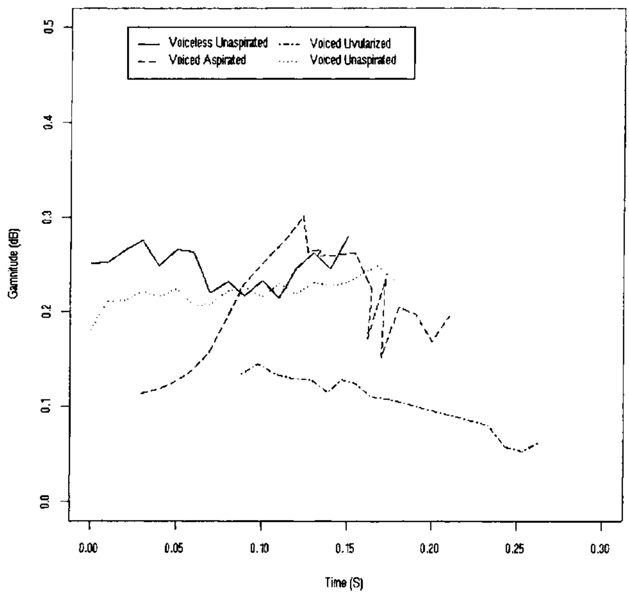


Figure 52. Median gamnitude of the first rahmonic peak (vertical axis) plotted against time (horizontal axis) in voiced consonant-initial roots for female Subject NU in the [a] context

initial roots. In general, the guttural consonant-initial roots all rise from a lower gamnitude at the beginning of the root, or stay at a very low gamnitude throughout the root as found with the voiced uvularized consonant-initial roots. In contrast, the non-guttural voiceless and voiced unaspirated initial roots maintain a steady relatively high gamnitude throughout the root, although the voiceless unaspirated stops display the highest gamnitude initially.

Figure 53 displays the median gamnitude of the first rahmonic peak throughout the duration of the vowel for the other three subjects' productions of both voiceless and voiced consonant-initial roots. For subjects DK and KK the results are even clearer than they are for Subject NU. All of the voiceless guttural consonant-initial roots start much lower in gamnitude

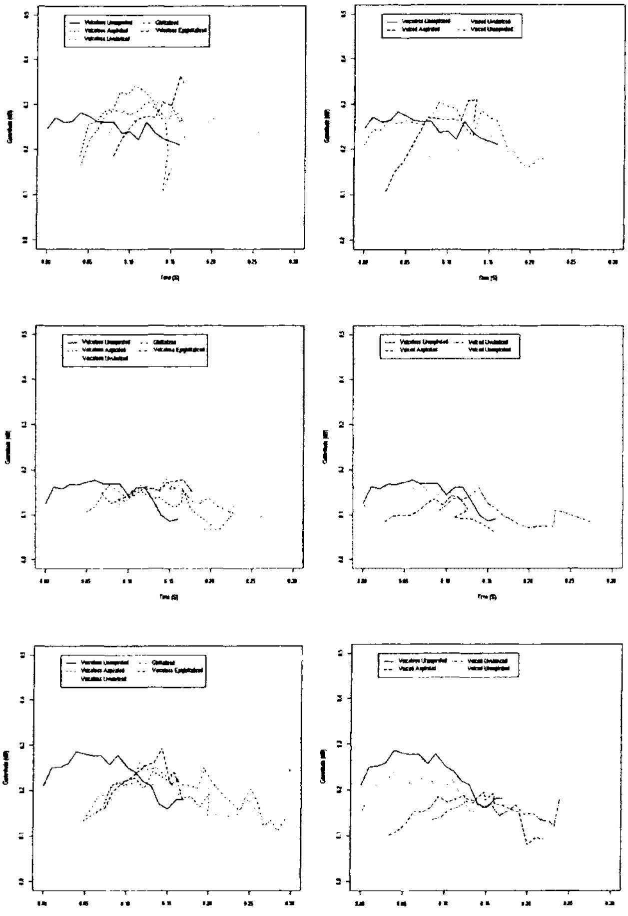


Figure 53. Median gamnitude of the first rahmonic peak plotted against time in voiceless (left) and voiced consonant (right) initial roots for subjects DK (top), KB (center) and KK (bottom) in the [a] context

much later in the root. The low gamnitude values found in the initial part of the vowel following all guttural consonants are the most relevant to my hypothesis that auditory similarity is the basis of the Guttural OCP constraint; as this is in the same temporal domain where the noise associated with guttural vowels is located. The voiced aspirated and voiced uvularized-initial roots both start out with low gamnitude values, and rise throughout the root, although the rise is more subtle in the voiced uvularized initial-roots than it is in the voiced aspirated-initial roots.

The results for Subject KB are also in agreement with those found with other subjects, although the different consonant types are packed closer together in the acoustic space because of his generally rough voice quality.

5.1.2. Guttural Vowels

Guttural vowels all exhibit lower gamnitude values for the first rahmonic peak in the cepstrum, while modal vowels exhibit higher gamnitude values. Figure 54 provides a plot of the first rahmonic peak gamnitude values on the vertical axis against time on the horizontal axis for the guttural vowel contrasts found in Ju|'hoansi produced by female Subject NU. Notice that NU's fully breathy and fully epiglottalized vowels have lower gamnitude peaks throughout their duration than the modal vowels. The partially breathy vowels are very similar in aperiodicity to the fully breathy vowels throughout the first half, but rise in periodicity to the level found for modal vowels by the end of the root. In fact, they overshoot the level of periodicity found in modal vowels at the very end of the root. The pattern for partially epiglottalized vowels is similar to that in fully epiglottalized vowels over the first half of their duration, and they overshoot the level of periodicity found in modal vowels by the end of the root. The acoustic properties of these vowels are consistent with them being diphthongs in voice quality. The glottalized vowels start out as very periodic at the beginning of the vowel, and then exhibit a dip in periodicity in the center of the vowel, before rising in periodicity by the end of the vowel. This dip in periodicity corresponds to the dip in F0 and RMS amplitude associated with NU's production of these vowels. As with the diphthongs,

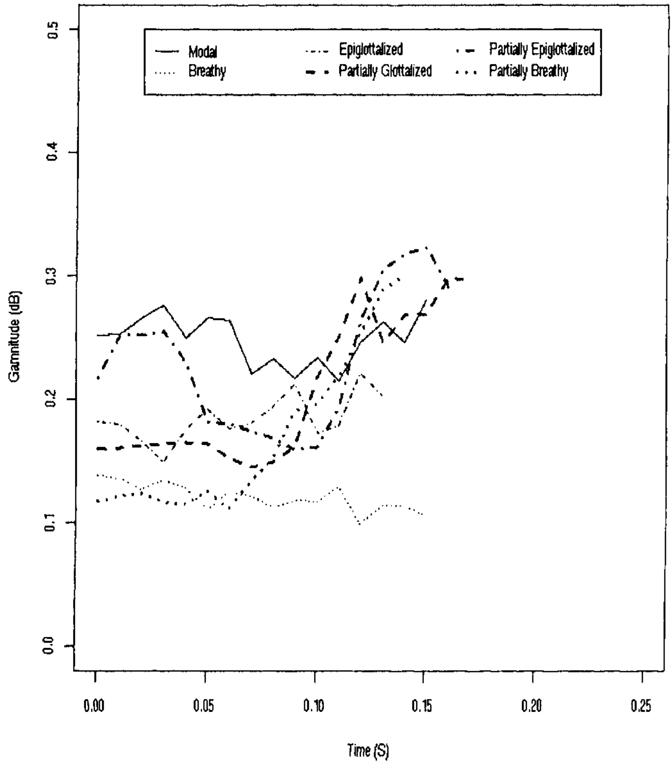


Figure 54. Median gamnitude of the 1st rahmonic peak in the cepstrum (vertical axis) plotted against time for the Ju|'hoansi guttural vowel contrasts in the [a] context (Subject NU)

the glottalized vowels also overshoot the level of periodicity found with modal vowels.

Figure 55 displays results for all of the guttural vowel types in the [o] context. As can be seen, the gamnitude values of the first rahmonic peak are very similar in the [o] context to the results seen in Figure 54 in the [a] context. Therefore only the [a] context will be shown from here on out.

Similar results are found for guttural vowel type contrasts with the other three subjects, as shown by the plots in

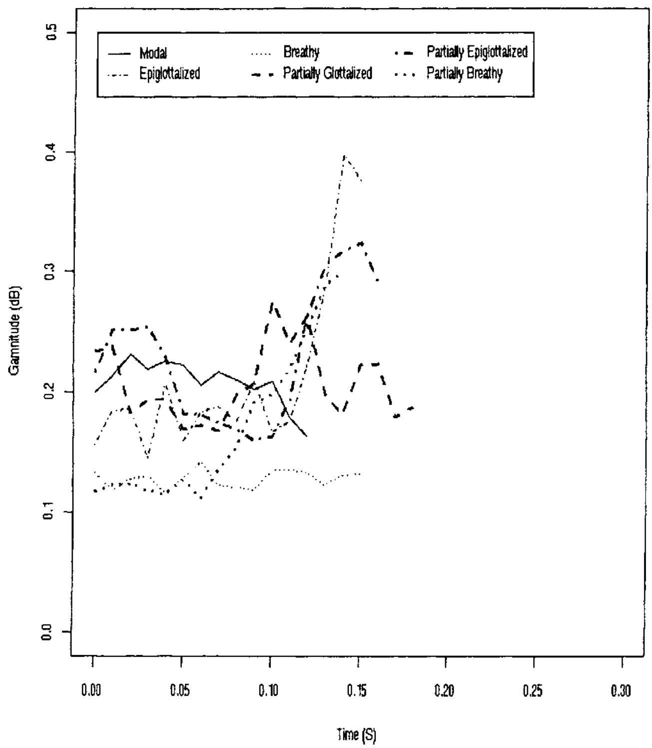


Figure 55. Median gamnitude of the 1st rahmonic peak in the cepstrum (vertical axis) plotted against time for the Ju|'hoansi guttural vowel type contrasts in the [o] context (Subject NU)

Figure 56 in the [a] context. The range of HNR values employed by male Subject KB is much smaller than that utilized by Subject NU, making it more difficult to see the clear separation between the different vowel types. Again, this is because of the generally rough quality of KB's voice. Notice that all of his vowel types are very low in gamnitude compared with the other subjects, even though the recording level was the same. Subject KB still maintains the same relationship between levels of periodicity and vowel type found with the other subjects. His entire range is just smaller.

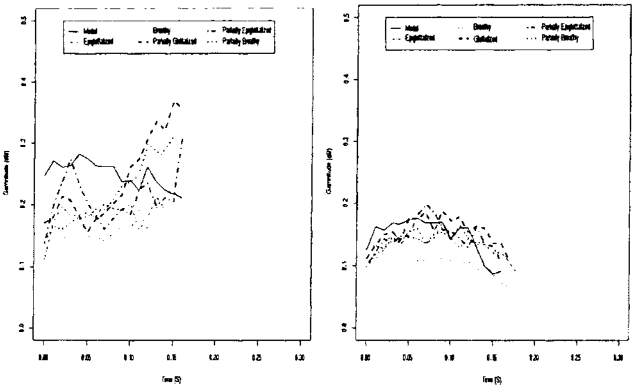


Figure 56. Median gamnitude of the 1st rahmonic peak in the cepstrum (vertical axis) plotted against time for the Ju|'hoansi guttural vowel contrasts for Subjects DK (top left) & KB (top right) in the [a] context

Male Subject KK's productions exhibit a similar pattern, as seen in Figure 57. For this subject, all of the roots containing guttural vowels cluster together very closely, showing that a low HNR is a nice unifying characteristic of the feature guttural. These separate nicely from modal vowels. Notice that all of the vowel types except for the diphthongs in voice quality fall in gamnitude at the end of the root, showing that they are less periodic. In contrast, the diphthongs rise in periodicity over the last half of the root.

5.1.3. Parallels between Guttural Consonants and Vowels

The measure of HNR shows striking parallels between guttural consonants and vowels. In particular, the temporal patterns of noise associated with aspirated consonants and breathy vowels are rather similar. Recall that in Chapter 2, I suggested that aspirated consonants and breathy vowels can be seen as involving the same articulatory patterns and the same acoustic manifestations of these patterns with the difference between them being mainly one of timing of the abduction of the

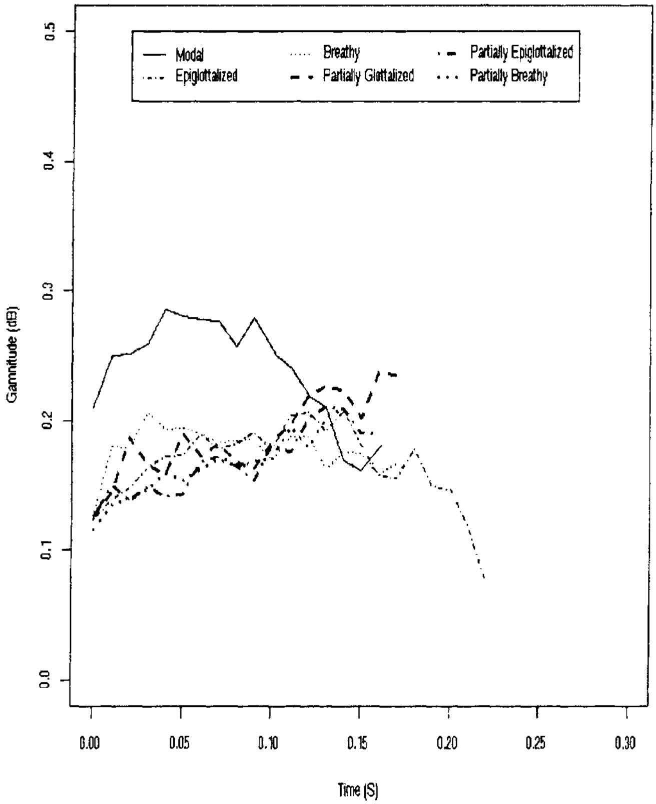


Figure 57. Median gamnitude of the 1st rahmonic peak in the cepstrum (vertical axis) plotted against time for the Ju|'hoansi guttural vowel type contrasts for Subject KK in the [a] context

glottis. Voicing on consonants is treated as marked, while voicing on vowels is treated as unmarked. I suggested that Ju|'hoansi has a four-way contrast in aspiration, with voiceless aspirated consonants having the largest magnitude of glottal abduction with the peak abduction timed at the release of the velaric cavity in the click, the voiced aspirates having a smaller magnitude abduction timed to start at about the release of the posterior constriction forming the velaric cavity of the click, and with breathy vowels having the abduction concurrent with the vowel. Partially breathy vowels differ from fully breathy vowels in that the adduction of the glottis starts part way through the vowel, and full adduction is reached well before the end of the vowel. In the production of fully breathy vowels, the glottis remains sufficiently abducted throughout most of the duration of the syllable to allow aspiration noise to occur.

A similar temporal contrast exists between epiglottalized consonants and vowels. Uvularized consonants and epiglottalized vowels differ in the location of the constriction, with epiglottalized vowels having a lower constriction than uvularized consonants, and the glottal postures also being different.

The HNR associated with the four-way aspiration contrast is shown in Figure 58 for all four subjects. Female Subject NU shows the greatest similarities, with the HNR associated with both voiceless and voiced aspirated consonants and partially breathy vowels being almost exactly parallel over the entire duration of the vowel. One remaining difference between the aspirated consonants and partially breathy vowels is the Voice Onset Time (VOT), which also differs for voiced aspirated and voiceless aspirated consonants. However, if both an aspirated consonant and a partially breathy vowel were to co-occur in the same root, such a root would be virtually undistinguishable in terms of HNR from a root that contained only an aspirated consonant and a “modal” vowel. Voiced and voiceless aspirated consonants are distinguishable by the subtle VOT difference seen here, from the fundamental frequency differences shown in Chapter 2, and from the presence of stronger aspiration noise associated with aspiration in consonants during the C-V transition. However, partially breathy voiced vowels also rise in pitch and therefore would not likely be very distinguishable from voiced aspirated consonant-initial roots based on fundamental frequency either. Subjects DK and KK don’t show quite as wide spacing between modal vowels and breathy vowels or vowels following aspirated consonants, but the overall patterns are

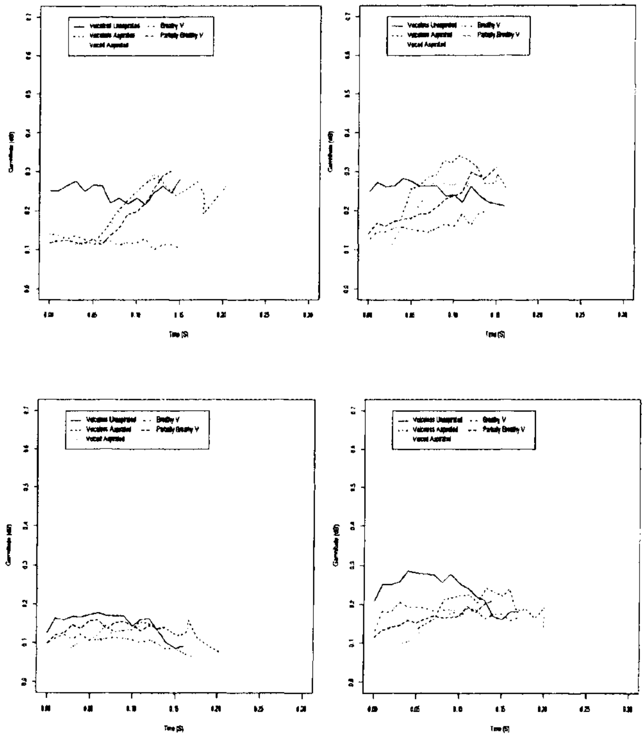


Figure 58. Median gamnitude of the 1st rahmonic peak in the cepstrum (vertical axis) plotted against time for the Ju|'hoansi four-way aspiration contrast in the [a] context: subjects NU (upper left), DK (upper right), KB (lower left) and KK (lower right)

very similar. Both subjects' productions display a smaller difference in HNR between partially and fully breathy vowels than is found in NU's speech, which results in a slightly better separation between the noise levels associated with aspirated consonants and breathy vowels. Subject KB also shows strong similarities between all of the types, although it is difficult to see with the small acoustic space he employs here.

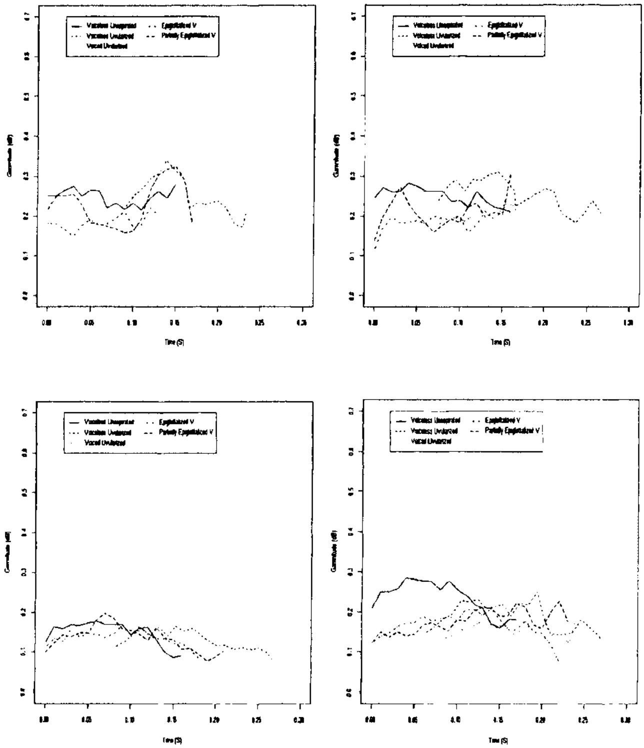


Figure 59. Median gamnitude of the 1st rahmonic peak in the cepstrum (vertical axis) plotted against time for the Ju|'hoansi four-way epiglottalization contrast in the [a] context: subjects NU (upper left), DK (upper right), KB (lower left) and KK (lower right)

The HNR results for the four-way contrast in epiglottalization are provided in Figure 59. The results are consistent with the idea that this is a timing contrast. For all four subjects, HNR associated with voiceless uvularized consonants and partially epiglottalized vowels are quite parallel. The parallelism in periodicity in the frequency domain between voiceless uvularized consonants, and partially epiglottalized vowels is striking for Subject NU's productions. Subjects DK and KK's patterns also show striking similarities. Again the really long VOT's associated with uvularized consonants are evident here. If voiceless uvularized consonants and epiglottalized vowels were to co-occur in the same root, the only cues that would allow a listener to differentiate them would be the long voiceless noise interval associated with the uvularization, and the pitch. However, as I showed in Chapter 2, the fundamental frequencies found in roots with voiced uvularized consonants and partially epiglottalized vowels are rather parallel as well.

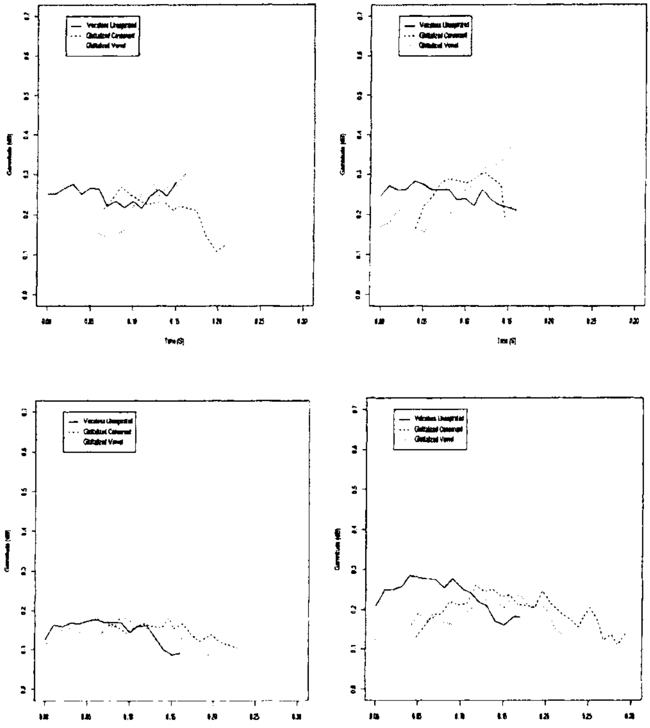


Figure 60. Median gamnitude of the 1st rahmonic peak in the cepstrum (vertical axis) plotted against time for the Ju|'hoansi two-way glottalization contrast in the [a] context: subjects NU (upper left), DK (upper right), KB (lower left) and KK (lower right)

One last parallel between consonants and vowels is the resemblance between glottalized consonants and glottalized vowels. The HNR associated with these contrasts for all four subjects are shown in Figure 60. For this two-way contrast, subjects DK, KB and KK show the highest degree of acoustic similarity on the HNR dimension. Subject NU shows the greatest separation in terms of HNR.

5.1.4. Conclusion

The HNR measure of the gamnitude of the first rahmonic peak captures nicely the aperiodicity in the frequency domain found in all of the Ju|’hoansi guttural consonants and vowels. Guttural vowels are much less periodic than non-guttural (modal) vowels. Similarly, guttural consonants cause following vowels to be aperiodic through coarticulation over a large portion of the vowel. All four subjects’ productions in this study display similar patterns with modal vowels being the most periodic, followed by glottalized vowels, epiglottalized vowels and breathy vowels. As for coarticulatory effects, vowels following voiceless unaspirated consonants are the most periodic throughout their duration, with voiced unaspirated-initial roots showing a slightly lower level of periodicity that also remains stable throughout the root. Roots with guttural consonant-initials all display lower levels of periodicity during the initial portion of the root.

This dimension offers the most robust measure of acoustic similarity found between guttural consonants and vowels in Ju|’hoansi. Hillenbrand et. al. (1994) found that the gamnitude of the first rahmonic peak in the cepstrum of breathy vowels accounted for a large proportion of the variance in breathiness ratings performed by English listeners. This gives indication that in breathy vowels at least, the measure of the first rahmonic gamnitude is perceptually salient in English. Other measures of periodicity combined with this measure accounted for 80% of the variance. Thus, it is likely that Ju|’hoansi speakers use the reliable cue of periodicity in their perception of the guttural vs. non-guttural contrast. It is this dimension that likely accounts for the perceptual similarity of gutturals that is targeted by the Guttural OCP constraint in the language. Future research will investigate the perception of voice quality in Ju|’hoansi, to determine the cue weighting of this and other cues associating with guttural consonants and vowels. For instance, the spectral slope measure discussed in the next section shows deviation from flat spectral slope for all guttural vowels and all vowels following guttural consonants, and thus suggests that spectral slope is probably an additional dimension targeted by the OCP constraint.

5.2. Spectral Slope

The second measure of acoustic similarity of gutturals in Ju|’hoansi is a measure of spectral slope (H1-H2). This measure also captures a dimension of acoustic similarity that separates guttural consonants and vowels from non-guttural consonants and vowels. The measure is, however, much more localized for guttural consonant types than the gamnitude of the first rahmonic peak in the cepstrum, which showed effects throughout the entire duration of the vowel. It also provides evidence for a different type of perceptual difficulty. Since aspiration and breathiness display more steeply falling spectral slopes than are found with modal vowels and unaspirated consonants, and glottalized and epiglottalized vowels and glottalized and uvularized consonants display lower spectral slopes than modal vowels and vowels following unaspirated consonants, the co-occurrence of parallel guttural consonants and vowels within the same root would also result in masking of the more transient cues associated with the consonant. However, the co-occurrence of opposing cues might result in either masking of consonants and vowels that display the weaker cue, or complete loss of cues associated with each type, resulting in an unmarked voice quality.

5.2.1.
Guttural Coarticulation

Pharyngeal and laryngeal coarticulation is also found in the spectral slope dimension. This coarticulation is parallel to laryngeal coarticulation in many respects, and shows that the acoustic voice quality cues present in laryngeals are also present in pharyngeals (uvulars and epiglottals), offering further phonetic motivation for the class of gutturals in Ju|’hoansi.

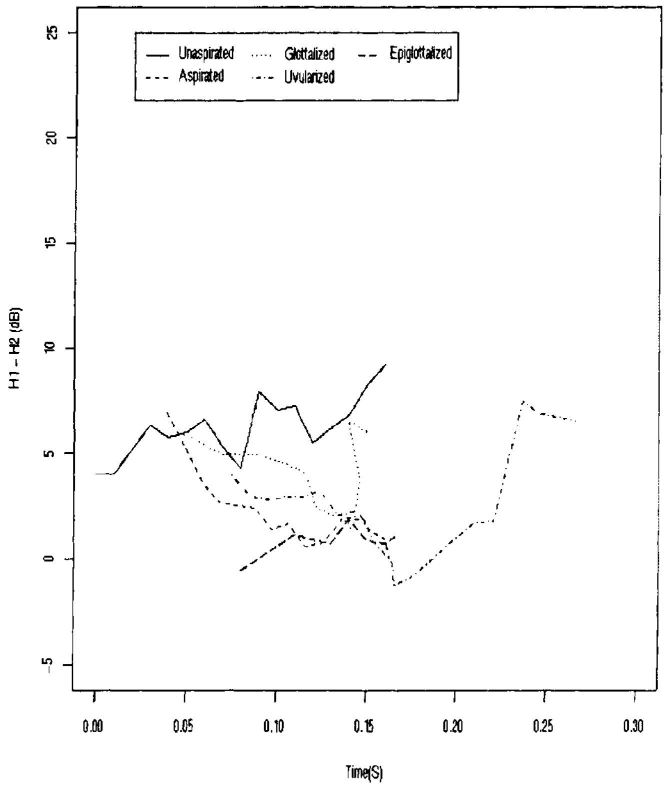


Figure 61. Median spectral slope values associated with voiceless consonant-initial roots in the [a] context (Subject DK)

Figure 61 displays the spectral slope values on vowels following root-initial voiceless guttural consonants. First, notice that the vowel following the voiceless aspirated consonant has a slightly higher spectral slope at the beginning of the vowel, and that the slope falls quite a bit lower than the modal vowel toward the end of the vowel. All other guttural consonant-initials display this lower level of H1-H2 than is found in modal vowels. For vowels following glottalized and epiglottalized consonants, this is as predicted. However, for uvularized consonants this is a bit surprising, since we would expect such consonants to have a larger open quotient during the fricated portion than modal vowels, given the high amount of airflow used in their production (Miller-Ockhuizen, In progress). We see that the uvularized consonants do fall throughout their duration, and the smaller H1-H2 values found after the transient release, may be due to articulatory overshoot as speakers attempt to quickly close the glottis after the fricated portion of the consonant is completed. However, this articulatory overshoot explanation doesn't explain why glottalized consonants display a similar pattern.

A similar pattern can be found for the other three subjects as well, as shown in Figure 62. Interestingly, on the spectral slope dimension, Subjects KB and DK use the largest acoustic space, while Subject KK uses the smallest space. This is opposite of the spacing found in the gamnitude of the first rahmonic peak measure. For Subject KB, the larger space on this dimension might help cue the contrasts that are not marked very well in HNR given his overall rough voice quality. Subject KK, on the other hand, has a clear ringing voice, and his productions exhibit a good separation on the cepstral measure. Thus, there is evidence that cue trading may be occurring on the dimension of periodicity measured via the gamnitude of the first rahmonic peak in the cepstrum and the dimension of spectral slope, measured via the measure of H1 and H2.

The spectral slope values associated with voiced guttural consonants are similar for Subject DK, who has a higher spectral slope value for voiced Figure 63. The effect is slightly greater for the voiced aspirates than for the voiceless aspirates.

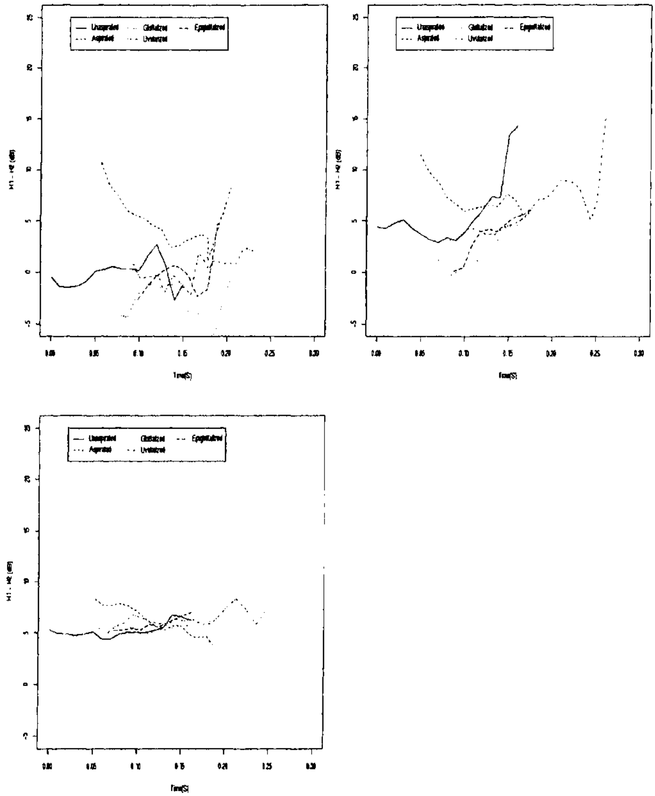


Figure 62. Median spectral slope values associated with voiceless consonant-initial roots in the [a] context: Subjects NU (top left), KB (top right) and KK (below)

The vowels following uvularized consonants and glottalized consonants do not seem to be much different on this dimension from the non-guttural unaspirated voiceless and voiced consonant-initial roots. The other three subjects' productions display higher spectral slope values for voiced aspirated initial roots, and voiced uvularized initial roots than unaspirated roots throughout the duration of the vowel. Thus it appears that spectral slope is a stronger acoustic correlate of guttural coarticulation involving voiced gutturals than it is for guttural coarticulation involving voiceless guttural consonants. In order to achieve pre-voicing, the glottis must be fairly adducted until right up to the oral constrictions involved in the clicks. This may result in later opening of the glottis to achieve aspiration and uvularization noise, which could also coincide with later closing of the glottis and increased acoustic correlates of the glottal aperture associated with the release properties on the following vowel.

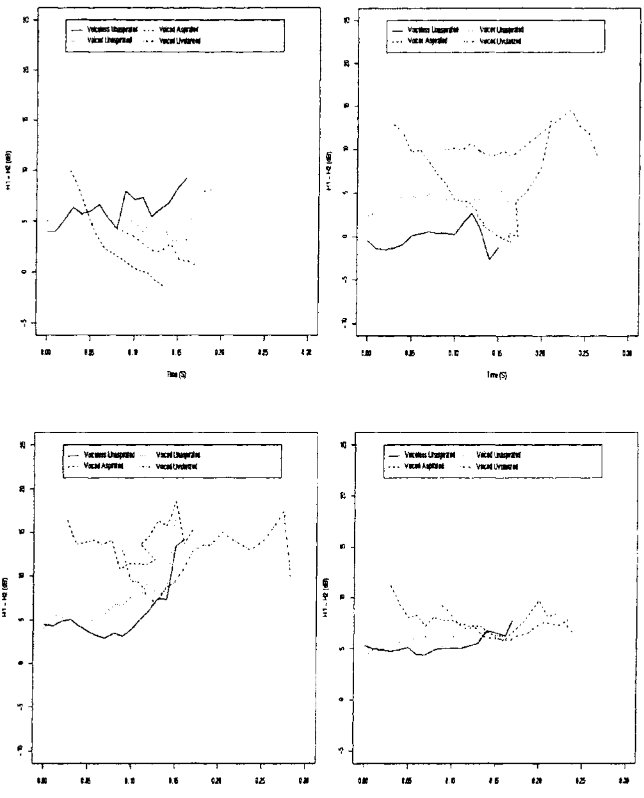


Figure 63. Median spectral slope values associated with voiced consonant-initial roots in the [a] context for subjects DK (top left), NU (top right), KB (bottom left) and KK (bottom right)

5.2.2.
Guttural Vowels

The guttural vowels exhibit a much greater degree of separation on this measure than was found in vowels following guttural consonants, again showing that acoustic effects associated with guttural coarticulation are much more transient than the more stable similar cues associated with phonemic guttural vowels. Notice in Figure 64, which shows the spectral slopes associated with guttural vowels produced by female Subject NU, the breathy vowels exhibit a greater spectral slope value than modal vowels over their entire duration. Epiglottalized vowels also display higher spectral slope values than modal vowels, although it is not as markedly different. Partially breathy vowels have higher spectral slope values over the first half of the vowel than modal vowels do, but they fall throughout the vowel and have a spectral slope value that reaches the spectral slope associated with modal vowels by the end of the vowel. The partially epiglottalized vowels, as well as the glottalized vowels, have negative spectral slope values throughout their duration. Thus, even though glottalized vowels only show a lower degree of periodicity in the center of the vowel with a periodic beginning and ending as measured by the gamnitude of the first rahmonic peak in the cepstrum, the vocal folds do appear to be more closely constricted throughout the vowel than they are in the production of modal vowels, assuming that the H1-H2 measure is a good indicator of open quotient (Bickley, 1982; Blankenship, 1997; Stevens, 1998; Cho, Jun and Ladefoged, 2000).

Similar patterns are found with the other three subjects. As shown in Figure 65, the breathy vowels have much higher spectral slopes than the modal vowels for all three subjects. Subject DK, who has the next widest space available, also shows a clear separation between fully breathy vowels and partially breathy vowels, with partially breathy vowels having spectral slope values that are even lower than modal vowels over the second half of the vowel. Subject KB also exhibits a nice degree of separation between modal and breathy vowels, but

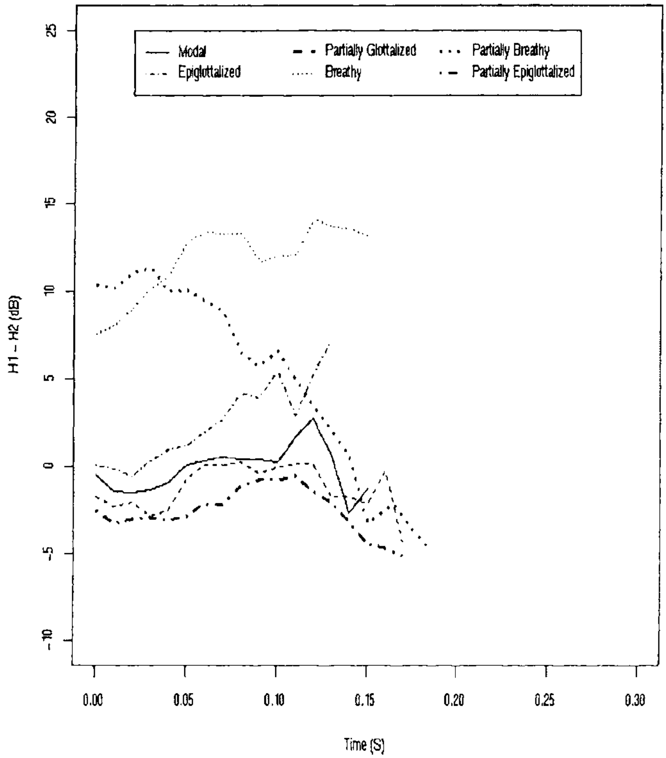


Figure 64. Median spectral slope values associated with guttural vowel contrasts on roots with voiceless consonant-initials in the [a] context (Subject NU)

only a small separation between fully and partially breathy vowels at the end of the roots. His speech shows an increase in spectral slope values at the end of the root for all vowel types under study, and thus suggests that he may tend to lapse into breathiness at the end of utterances. Since the roots here are produced in isolation they constitute an utterance. Subject KK again exhibits the smallest degree of separation on this measure, which suggests that he doesn't use this cue dimension much to signal contrasts. Still, the breathy vowels and modal

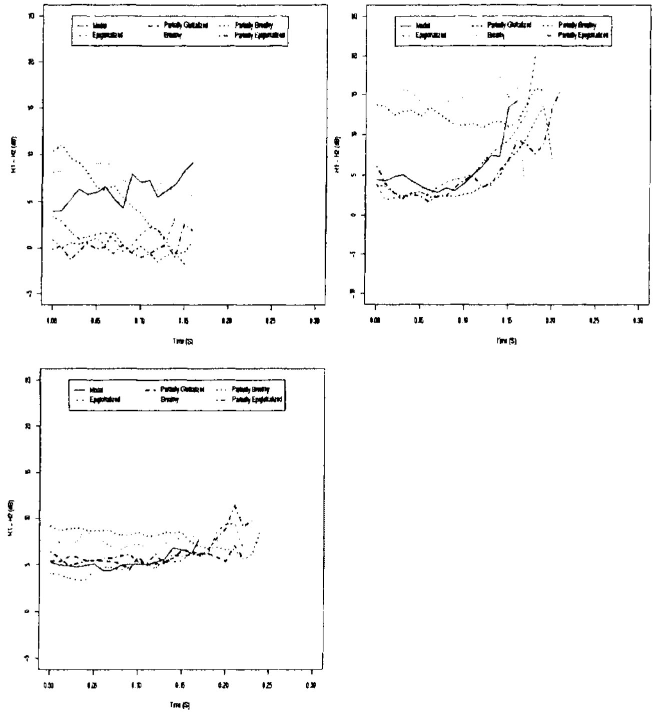


Figure 65. Median spectral slope values associated with guttural vowel types for subjects DK (top left), KB (top right) and KK (bottom) in the [a] context

vowels are differentiated in the same direction found with the other subjects. That is, breathy vowels of both types exhibit higher spectral slope values than modal vowels, and glottalized vowels exhibit lower spectral slope values. Epiglottalized vowels for Subject KK do however, display slightly higher spectral slope values than modal vowels, as they are for Subject NU. Thus, there seems to be the most individual variation in spectral slope values associated with the articulation of epiglottalized vowels.

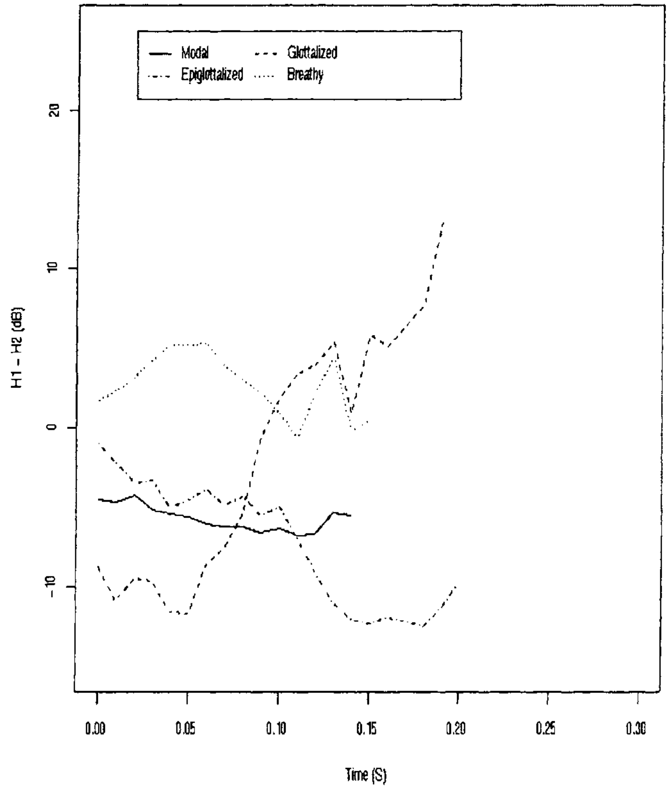


Figure 66. Median spectral slope values associated with guttural vowel type contrasts in the [o] context (Subject NU)

Unlike the other measures, vowel quality does have an effect on spectral slope. As seen in Figure 66, the general difference between H1 and H2 is lower over all guttural vowel types in the [o] context, being very low or even negative for all guttural vowel types. This is probably due to boosting of the second harmonic from the first formant, which is lower in the [o] context than in the [a] context. The same relationship between the different vowel types still holds even though there is probably boosting of the second harmonic from the formant. Glottalized vowels here show a large increase in spectral slope throughout the vowel that was not found in the [a] context for this subject, or for any of the other subjects in the study. It is not clear what the cause of this might be.

5.2.3. Parallels between Guttural Consonants and Vowels

The spectral slope results are consistent with the approach laid out in part I, viewing breathy vowels and aspirated consonants as differing mainly in timing. The spectral slope (H1-H2) values throughout the vowel are displayed in Figure 67 for all four aspirated roots produced by female Subject NU. For this female subject, the high spectral slope values at the beginning of the fully breathy vowels shows that the glottis is already widely abducted at the release of the voiceless unaspirated consonant, but the positive slope of the line from the point of the posterior release of the click throughout the midpoint of the vowel shows that the peak glottal abduction is not reached until the midpoint of the vowel. The glottis then stays at peak abduction throughout the remainder of the vowel duration. The partially breathy vowel, the vowel following a voiced aspirated consonant, and the vowel following a voiceless aspirated consonant, all have nearly the same spectral slope patterns throughout the duration of the vowel. They all start out with highly positive slope values and fall throughout the vowel, showing gradual adduction of the glottis throughout the vowel. One main difference is the VOT, with the vowel being voiced immediately after the posterior release of the consonant in the breathy vowels, but having a voice onset time delay for both types of aspirates. In addition to VOT differences, both aspirated consonant-initial roots have falling spectral slope values over the last 30 msec. of the vowel, showing that the glottis starts opening before the very end of the vowel. Additionally, as was shown in Chapter 2, voiced aspirated consonants have a large lowering effect on fundamental frequency, while the other types of aspirates do not.

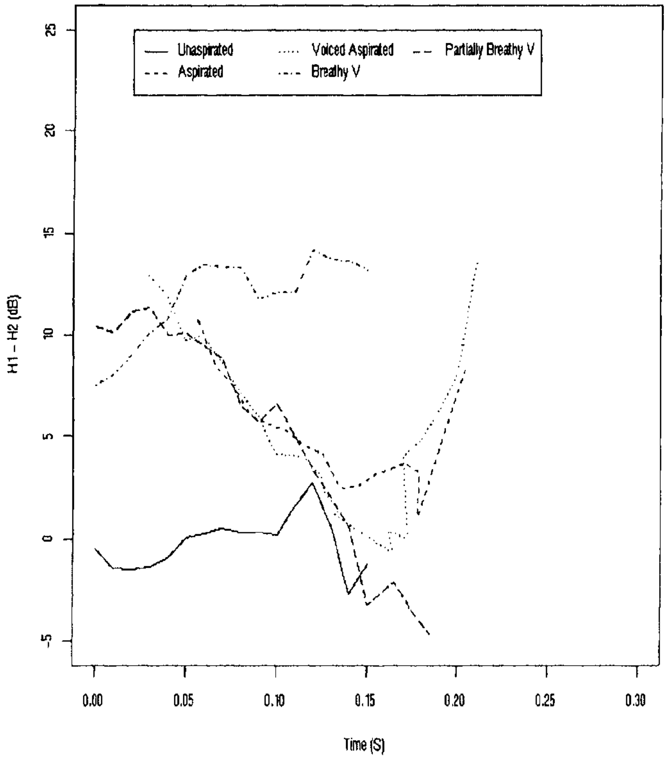


Figure 67. Median spectral slopes values associated with the four-way aspiration contrast (Subject NU)

The other three subjects exhibit similar patterns, which are displayed in Figure 68. The other female subject, DK, shows a good contrast between fully breathy vowels and partially breathy vowels. Vowels following the two aspirated consonants and the partially breathy vowels show a similar shape throughout the root, just as they do in Subject NU’s productions. Interestingly, in order to attain this nice separation in glottal abduction, the glottis achieves a higher degree of adduction at the end of the root in vowels following both types of aspirates and in breathy vowels, than it ever achieves

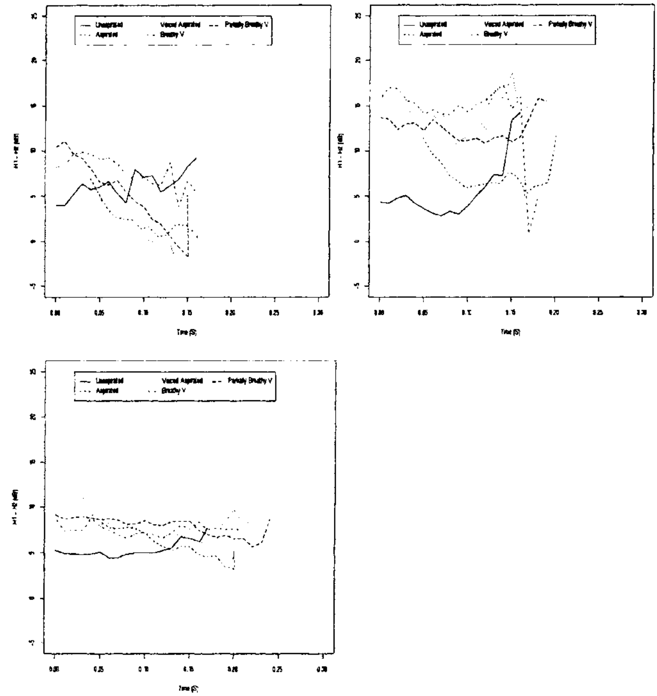


Figure 68. Median spectral slope values associated with the four-way aspiration contrast: subjects DK (top left), KB (top right), and KK (bottom)

throughout the root in roots having voiceless unaspirated-initial consonants.

Subject KB’s productions also display interesting individual patterns. In his productions, roots beginning with voiced aspirated-initial consonants attain a higher magnitude of glottal abduction throughout the root than is found in vowels following voiceless aspirated-initial consonants. Roots with voiced aspirated-initial consonants maintain a widely abducted glottal aperture throughout the root. Partially breathy and fully breathy vowels are still nicely distinguished in the direction predicted, with the partially breathy vowels falling in spectral slope values throughout the root. Subject KB's productions of voiceless unaspirated consonant-initial roots with modal vowels also shows a high degree of breathiness at the end of roots. Subject KK shows little separation in all of the different types, although all of the aspirated sounds display a higher spectral slope, which implies a more abducted glottis throughout most of the duration of the root.

As expected, the results for the four-way epiglottalization contrast aren't quite as clear as those for the four-way aspiration contrast. The results for all four subjects are shown in Figure 69. Subject NU's productions display lower spectral slope values for partially epiglottalized vowels than for fully epiglottalized vowels.

The contrast between partially epiglottalized vowels and fully epiglottalized vowels is even harder to interpret. The contrast between partially and fully epiglottalized vowels for Subject NU is seen by partially epiglottalized vowels having more negative spectral slope values at the beginning, which rise in the middle of the vowel, and fall again toward the end, while fully epiglottalized vowels start out with low spectral slope values and rise slightly toward the end of the root. Assuming that spectral slope values are correlated with pharyngeal or laryngeal opening (it can not be only open quotient because epiglottals do not involve a constriction at the glottis), the partially epiglottalized vowels display greater constriction in the middle of the vowel.

The two vowel types aren't differentiated as well for subjects DK and KB. In their productions, the fully epiglottalized vowels seem to have both a wider glottal and pharyngeal opening at the beginning of the vowel, with the constriction either at the pharynx or at the glottis (or both) increasing toward the end of the vowel. The two vowel types are also differentiated at the end of the vowel, with the fully epiglottalized vowels displaying higher spectral slope values at the end of the vowel. This could be due to the general trend for all roots to become somewhat breathy at the end of the root. Since this breathiness is likely more of a property of the ends of phrases than of the ends of roots, this may just mean that there

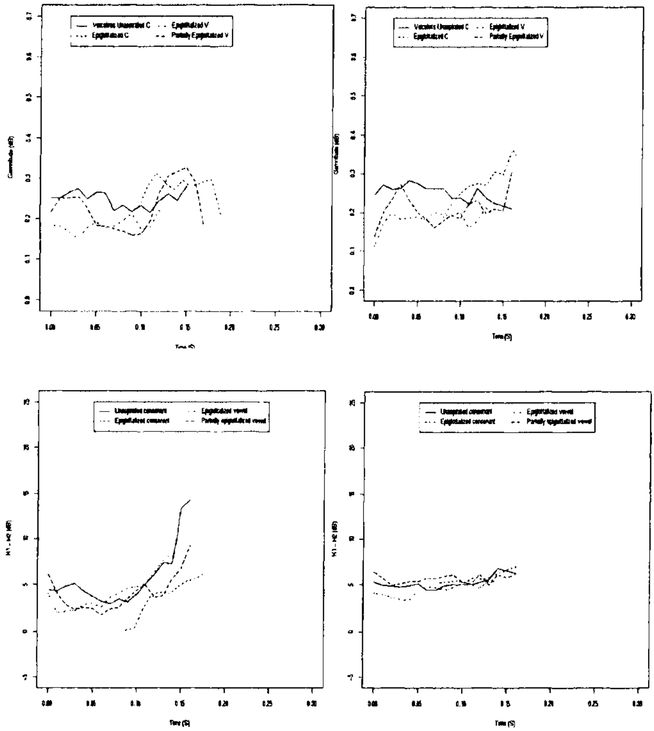


Figure 69. Median spectral slope values associated with the four-way epiglottalization contrast: subjects NU (top left), DK (top right), KB (bottom left), and KK (bottom right)

is time in fully epiglottalized vowels to show this phrasal property as well, while in partially epiglottalized vowels the dynamic demands of the articulation of the diphthong are too many to allow for phrasal breathiness to be realized. For Subject KK’s productions, we find the opposite relationship between the two vowel types, with fully epiglottalized vowels displaying lower spectral slope values at the beginning of the root than are found with partially epiglottalized vowels. The spectral slope values at the end of the two root types are quite similar for Subject KK.

The results show that epiglottalized consonants are more similar to the epiglottalized vowels on the spectral slope measure than plain uvularized consonants are, particularly for subjects DK and KB. For subjects NU and KK, there isn't much difference between the spectral slope values of vowels following plain voiceless uvularized consonants and vowels following epiglottalized consonants. The data is supportive of epiglottalized consonants and vowels both being characterized as [low spectral slope], and uvularized consonants being specified for [high spectral slope].

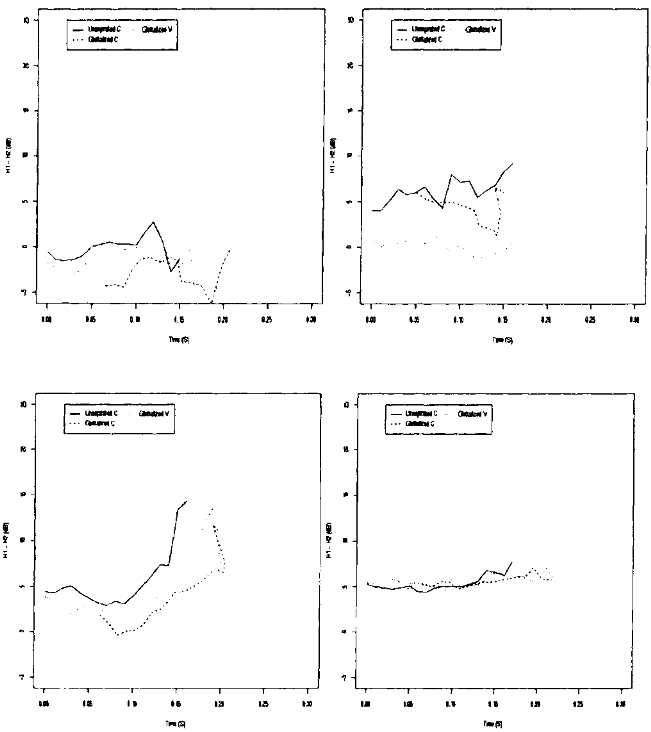


Figure 70. Median spectral slope values associated with the two-way glottalization contrast: subjects NU (top left), DK (top right), KB (bottom left), & KK (bottom right)

There is only a two-way contrast with glottalization, relating to the paradigmatic contrast between roots with glottalized consonants and modal vowels on the one hand, and consonants with unmarked release properties and glottalized vowels on the other hand. The spectral slope results shown in Figure 70 offer a good contrast between the two root types, while the direction of differences is not consistent across subjects. For Subject NU, vowels following glottalized consonants display lower spectral slope values than are found in glottalized vowels. In Subject DK's productions, the glottalized vowels exhibit lower spectral slope values throughout their duration than vowels following glottalized consonants. Subject KB's productions are more similar to Subject NU's productions, with the spectral slope values in vowels following glottalized consonants being more negative than those associated with glottalized vowels. It is difficult to ascertain any differences in KK's speech on this dimension. The individual differences don't seem to be attributable to differences in gender or voice quality.

5.2.4. Conclusion

Given the consistent categorization of guttural and non-guttural vowels by the measure of H1-H2 and the relative stability of the measure over different vowel contexts, it is likely that this dimension is also a part of the basis of the Guttural OCP constraint in Ju|'hoansi. However, given that the guttural vowels have opposing values for H1-H2, it cannot be said that that the guttural vowels are more acoustically similar to each other than they are to the modal non-guttural vowels. In fact, the data show that modal vowels are more similar to glottalized and epiglottalized vowels than they are to breathy vowels on this dimension. However, opposing values on the same cue dimension are also problematic from a perceptual point of view, as well as from an articulatory point of view. That is, if spectral slope is indeed correlated with open quotient (the percentage of a glottal period that the vocal folds remain open), then it would be rather difficult from an articulatory point of view to have an aspirated consonant and a glottalized vowel co-occur within the same root, as this would require simultaneous narrowing and widening of glottal aperture. From a perceptual point of view, if a sound that had a high spectral slope value (e.g., aspirated consonant or breathy vowel) co-occurred within the same temporal position as a sound with a low spectral slope (e.g. a glottalized vowel or a glottalized consonant), the two values would likely cancel each other acoustically, resulting in a relatively flat spectral slope similar to a modal vowel or a vowel following an unaspirated consonant.

5.3. Jitter (PPQ)

In this section, I report the results of the Pitch Perturbation Quotient (PPQ) analysis of Ju|'hoansi guttural and non-guttural vowels, as well as vowels following guttural consonants vs. vowels following non-guttural consonants. Although PPQ is the least susceptible to changes in fundamental frequency of available jitter measures in the literature, given the smoothing used (Pinto and Titze, 1990), the data here show that it is still quite susceptible to large changes in fundamental frequency. The measure was unable to discriminate between different phonation types even though the changes in fundamental frequency were minimized within the window used. The results show that all guttural vowels exhibit a higher degree of jitter than modal vowels, and vowels following guttural consonants also show an increased amount of jitter from vowels following non-guttural consonants. However, such increases in jitter are mostly attributable to increased changes in F0 throughout the root. Therefore, the PPQ measure of jitter is plotted against the total change in F0 over the same window (the first half of the vowel).

Although the wordlist in this study was controlled for lexical tone by choosing level SL and L toned roots wherever possible, there are still large changes in fundamental frequency throughout the root. Most of these are attributable to a large fall in F0 at the end of all roots. However, there are also some local effects on fundamental frequency by the guttural release properties of consonants, and large changes in F0 that accompany epiglottalization on vowels. Also, recall that glottalized vowels always occur on roots with either of the two rising tone patterns (SL-L or L-H), so it was impossible to find level-toned roots containing glottalized vowels. In order to minimize the total change in F0 in the window used to computed the PPQ measure, the jitter was computed over only the first half of the vowel. This factors out the drop in F0 found at the end of a root when it is uttered in isolation due to a prosodically conditioned boundary tone. The total change in F0 throughout the first half of the vowel was calculated as the change in F0 between the first point in the window and the peak F0 within the window, summed with the change in F0 between the peak F0 within the window and the final F0 within the window. The F0 for this purpose was calculated as the inverse of the length of the pitch period. The peak F0 is therefore the inverse of the shortest period duration within the window. The window used begins at the PR or TR label marking the posterior release of the click, or the transient release of the click in guttural consonants, and ends at the midpoint of the vowel calculated from the beginning of the window through the end of the periodic portion of the vowel (EP label).

Figure 71 shows the amount of PPQ found in all four guttural vowel types for female Subject NU. Note that the PPQ values never exceed 0.05. This is because the parameter setting for the ESPS program *epochs* cut off the amount of allowed jitter at 0.05 seconds. As noted in Chapter 4, this low allowed jitter amount was necessary in order to avoid pitch doubling due to diplophonia in epiglottalized vowels. The vertical axis shows the PPQ value, and the horizontal axis shows the total change in F0, over the first half of the vowel. The diagonal line is a regression line, which shows that there is a strong correlation between change in F0 and PPQ. That is, the amount of jitter present in the vowel is dependent in part on the amount of change in F0 over the same interval. Thus, it

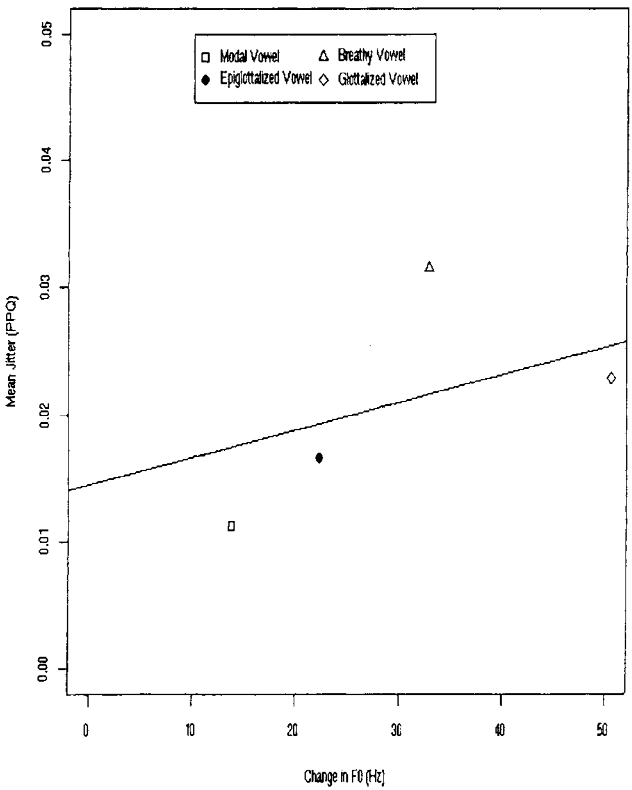


Figure 71. Mean change in F0 against mean PPQ associated with guttural vowel type contrasts over the 1st half of the vowel (Subject NU)

appears that jitter, the random variation in period length, is not a cue signaling guttural vowel type contrasts in Ju|'hoansi, or a cue that groups guttural vowels as opposed to non-guttural vowels.

Looking at the jitter associated with the different consonant types under study in Figure 72 for the same female subject, we find a similar situation. Again, the mean values all fall close to the regression line, showing that the amount of PPQ present in

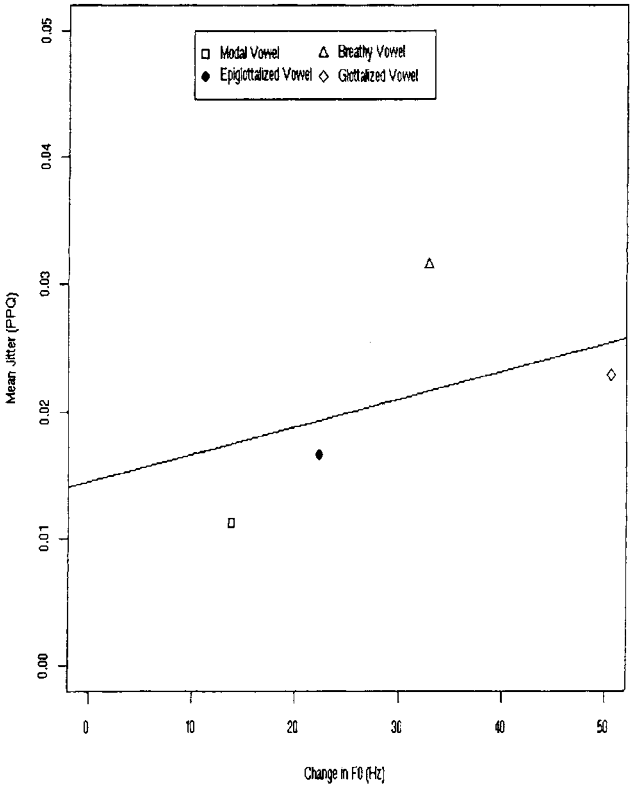


Figure 72. Mean change in F0 against mean PPQ over the first half of the vowel for guttural consonant contrasts (Subject NU)

the signal is highly dependent on the amount of F0 change in the same interval. Note that the voiced aspirated consonant-initial roots have the largest change in F0 over the first half of the vowel. For the unaspirated and aspirated consonants, the voiced series both have a larger change in F0 over the first half of the vowel than in the voiceless series. However, the opposite is the case with the uvularized consonants.

The results for the other three subjects display some variation. Figure 73 plots the change in F0 against the PPQ for

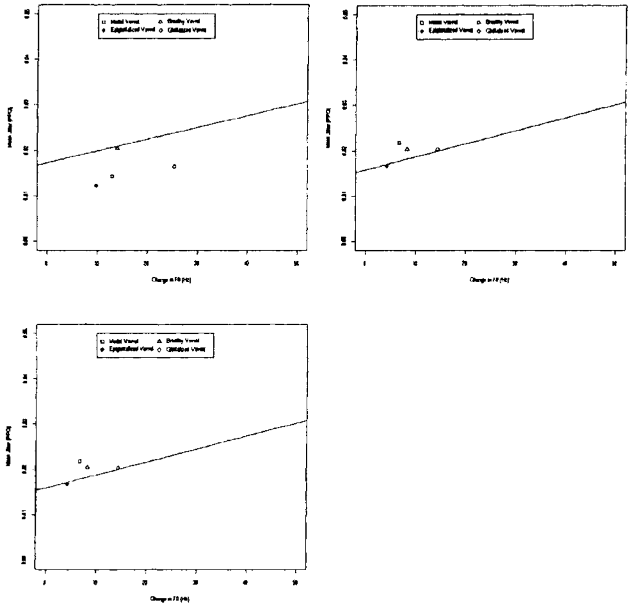


Figure 73. Mean change in F0 plotted against mean PPQ over the first half of the vowel associated with guttural vowel contrasts: subjects DK (left), KB (center), and KK (right)

the guttural vowel types under study for the other three subjects. For the other female Subject DK, the mean points cluster around the diagonal line, showing that most of the period-to-period duration variation is due to change in F0 and not random jitter.

Graphs showing the jitter associated with the guttural consonant types for the other three subjects are provided in Figure 74. As can be seen, the means for all consonant types fall close to the regression line, indicating that PPQ is highly dependent on Change in F0 over the window.

The PPQ measure of jitter is the least susceptible to changes in F0 throughout the window being evaluated (Pinto and Titze, 1990). However, the results here show that even this measure, which involves smoothing over 5 cycles, is still too dependent on changes in fundamental frequency, and thus probably is not

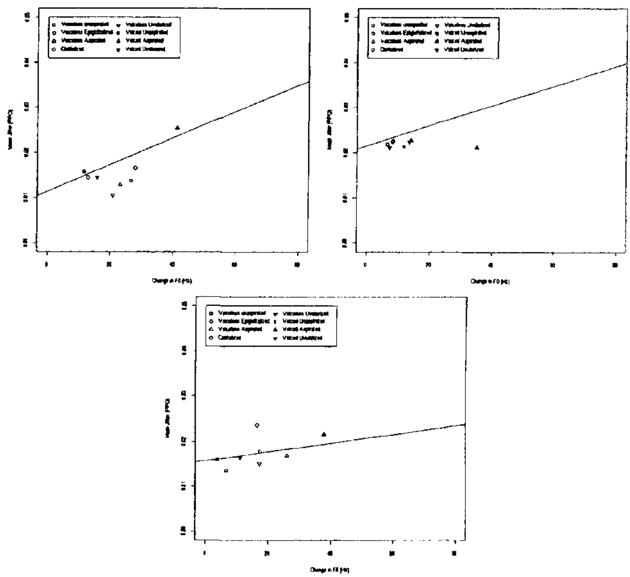


Figure 74. Mean change in F0 against PPQ over the 1st half of vowel following guttural consonants: subjects DK (top left), KB (top right) and KK (bottom)

a good measure of linguistic phonation type contrasts. We have seen that even in level toned roots used in this study, the small F0 changes associated with guttural consonant type contrasts are too great to allow the measure to be useful. A smaller window that started after the F0 rise associated with guttural consonant types would be too small to give useful results and would also possibly be too late to catch any jitter caused by C-V coarticulation. This study is another case where jitter is not present to cue differences in phonation type, such as the contrast between creaky and modal tones in Burmese (Javkin et. al, 1987) and in some of the phonation type contrasts in Hmong (Huffman, 1987).

In all root types, we have seen that Ju|’hoansi men display increased jitter compared with Ju|’hoansi women. This would be an extremely important finding since Klatt and Klatt (1990) and Hanson (1997) have shown that female English subjects tend to have more breathy voice qualities than male English speakers. For the two male subjects in this study, there are other available explanations for this fact. Namely, KB has a rougher overall voice quality, which could be an individual speaker characteristic, rather than one associated with gender. While male Subject KK has a clear ringing voice quality, his speech does exhibit smaller changes in F0 throughout the window compared with the two female subjects in this study. This is interesting, since his F0 for L and H tones are as high as those for female Subject DK, but he still seems to employ a smaller pitch range.

5.4.
Conclusion

The acoustic results in this chapter have thus provided two acoustic cues that would contribute to a perceptual account for the basis of the Guttural OCP constraint found in Ju|'hoansi. Periodicity is a stable cue that is present in all guttural vowels and in vowels following all guttural consonants. Furthermore, the cues extend quite a ways into the vowel. Periodicity is then a likely candidate for the acoustic and perceptual similarity of all guttural consonants and vowels in the language, and the most likely perceptual basis for the Guttural OCP constraint found in the language. Transient co-articulatory cues found on vowels following guttural consonants would likely be masked by the stronger, more stable cues associated with the guttural vowels, decreasing the perceptibility of the consonants. Additionally, guttural vowels all display rising or falling spectral slope values, compared with non-guttural vowels that display fairly flat spectral slope values, and vowels following guttural consonants also exhibit similar co-articulatory effects on the spectral slope values of following vowels. Cues to breathy vowels and aspirated consonants would thus be very similar, as would cues to glottalized consonants and vowels. The co-occurrence of these acoustically similar sounds would also lead to masking of the transient cues for the consonant.

However, aspirated consonants and glottalized vowels display conflicting cues on the spectral slope dimension, and these are also blocked from co-occurring. Conflicting cues are also perceptually problematic, in that their co-occurrence within a single root would lead to canceling out of the cues for both sounds, and thus decrease the perceptibility of each type of sound. Perceptual experiments will be undertaken to assess the use of these cues by native speakers to identify guttural consonants and vowels, to show how strongly weighted the individual cues are. Of course, the role of articulatory impossibility also rules out the co-occurrence of consonants and vowels that are produced with different glottal apertures.

CHAPTER 6
Conclusion

This dissertation has broadly outlined several phonotactic constraints in the Khoisan language Ju|’hoansi and shown how the phonetic bases of these constraints can be captured in terms of acoustic modulation. I have also offered a detailed acoustic case study, investigating the spectral properties associated with guttural consonants and vowels in the language and shown how the Guttural OCP constraint can be explained in terms of these properties. That is, if two sounds that contain similar harmonics-to-noise ratios and similar or opposing spectral slope values were to co-occur in the same temporal position of the same root, spectral cues to one of the sounds would be lost. While perceptual studies are necessary to confirm the hypotheses, the acoustic properties themselves show that the probability of masking is very strong. While the spectral properties associated with consonants are often more transient and smaller, guttural consonant contrasts are also cued by the presence of stronger noise during the C-V transition, as well as larger fundamental frequency effects that are not found as sharply in guttural vowels. These additional cues in guttural consonants make it likely that it is the consonants that would be the sounds perceived if both sounds were to co-occur within the same root. Future research will provide more detailed acoustic studies of the interactions between fundamental frequency and guttural consonants and vowels, and the first formant frequency values associated with guttural consonants and vowels. It is expected that these acoustic studies will offer explanations for the co-occurrence restrictions found between tone and guttural consonants and vowels, vowel height and vowel frontness with guttural consonants and vowels.

The plethora of co-occurrence constraints between consonants and vowels, on many different acoustic dimensions in Ju|'hoansi, helps enforce acoustic modulation within the root, which should help listeners identify the beginnings and ends of roots, perform segmentation, and parse both moras of a bimoraic root. Recall that 66% of all roots in Ju|'hoansi begin with a guttural feature in either the first consonantal or vocalic position of the root. These guttural features not only offer modulation in terms of intensity and spectral noise over the root, but they also offer robust fundamental frequency modulation within the root through the interaction of laryngeal posturing and fundamental frequency.

There are many directions in which this research will be extended in the future. Additional areas of modulation, such as first formant frequency and fundamental frequency patterns over the root, will be submitted to quantitative well-controlled investigations in order to determine more precisely the manner of modulation found within these domains. Investigations in the articulation of guttural contrasts in Ju|'hoansi will also be undertaken to garner a better understanding of the articulatory mechanisms involved in fundamental frequency control in F0 modulation. Investigations into the articulation of Ju|'hoansi pharyngeals will be undertaken to understand the bases of the BVC patterns discussed in Chapter 3. Finally, other OCP constraints in languages of the world will be investigated in order to gain a better understanding of the different types of modulation that are enforced through phonotactic constraints found in different languages. There is a wide array of acoustic parameters that could be involved in the auditory bases of OCP constraints cross-linguistically, and while the perceptual bases of OCP constraints have been suggested since it's inception (Leben, 1973; Goldsmith, 1976), the phonetic bases of most OCP constraints have yet to be fully investigated.

Appendix A.
Wordlist for Acoustic Case Study

Guttural Vowels

	Voiceless unaspirated click initial	Voiced unaspirated click initial
1.	**Fully breathy vowel**	
	a. !ȁʱȁʱ 'red crested korhaan'	a. g‖ȁʱȁʱ 'to press down'
	b. ‖ȍʱȍʱ 'small annual clover'	b. g!ȍʱȍʱ 'dust, smoke'
2.	**Fully epiglottalized vowel**	
	a. !ȁˤȁˤ 'iron, steel'	a. g‖àˤàˤ 'aunt'
	b. !ȍˤȍˤ 'pan'	b. g!ȍˤȍˤ 'male'
3.	**Glottalized vowel**	
	a. !ȁˀà 'dry season'	a. gǂàˀá 'wide '
	b. tʃòˀó 'to be unconscious'	b. g!òˀó 'cough'
4.	**Partially breathy vowel**	
	a. ǂȁʱà 'plain'	
	b. |ȍʱò 'to follow'	
5.	**Partially epiglottalized vowel**	
	a. ‖ȁˤà 'to hold down'	
	b. ǂòˤó 'uterus'	

Guttural consonants

1. **Voiceless unaspirated click**
   a. ǁáá 'to warm hands by fire'
   b. ǂòà 'sitting mat'
   c. !óó 'older brother'
2. **Voiceless aspirated click**
   a. ǂ^{h}àà 'path'
   b. ǀ^{h}óạ́ 'to rub away'
   c. !^{h}òò 'forcibly'
3. **Voiced unaspirated click**
   a. g!àà 'rain'
   b. g!òà 'wet leaf'
   c. g!óó 'petrol'
4. **Voiced aspirated click**
   a. gǀ^{ɦ}àà 'place for drying meat'
   b. gǂ^{ɦ}òà 'dog'
   c. g!^{ɦ}òó 'to sit'
5. **Voiceless epiglottalized click**
   a. !^{ʕ}àà 'to dry out'
   b. ǁ^{ʕ}òà 'to work'
   c. !^{ʕ}òò 'to hate'
6. **Voiceless uvularized click**
   a. ǂ$^{\chi}$àà 'moist sand'
   b. ǂ$^{\chi}$òà 'to pacify'
   c. !$^{\chi}$òò 'to be unsuccessful'
7. **Voiced uvularized click**
   a. g!^{ʁ}àà 'to take out'
   b. g!^{ʁ}òà 'knee'
   c. gǀ^{ʁ}òò 'to give something to someone'
8. **Glottalized click**
   a. !^{ʔ}àà^{n} 'to catch up'
   b. !^{ʔ}òà^{n} 'yawn'
   c. !^{ʔ}òò 'leadwood tree'

Appendix B.
Place co-occurrence tables

Initial C → Medial C ↓	Labial	Coronal	Corono-Dorsal	Dorsal	Guttural	Total
No C2	9	148	424	50	5	636
Labial [β, m]	3	48	120	9	2	182
Coronal [r, n, s, ʃ]	9	59	143	29	6	246
Dorsal [ŋ, ɣ]	0	1	3	1	0	5
Guttural [ɦ, χ]	0	0	2	1	0	3
Total	21	256	692	90	13	1072

Table XXXI. Co-occurrence of initial and medial place features on roots with unmarked release properties

INITIAL C → FINAL C ↓	LABIAL	CORONAL	CORONO-DORSAL	DORSAL	GUTTURAL	TOTAL
Labial [m]	0	22	72	17	3	114
Dorsal [ŋ]	0	0	5	0	0	5
Total	0	22	77	17	3	119
INITIAL C (GUTTURAL RELEASE) → FINAL C ↓	**LABIAL**	**CORONAL**	**CORONO-DORSAL**	**DORSAL**	**GUTTURAL**	**TOTAL**
Labial [m]	0	7	46	3	0	56
Dorsal [ŋ]	0	0	2	0	0	2
Total	0	7	48	3	0	58

Table XXXII. Co-occurrence of initial place features on guttural-released initial consonants and medial place features

Index

References

Abercrombie, D. (1967) *Elements of General Phonetics,* Edinburgh, Edinburgh University Press.

Abramson, A.S. and Luangthongkum, T. (2001) Phonation types in Suai, *JAS 110(5).*

Alwan, A. (1989) Perceptual cues for place of articulation for the voiced pharyngeal and uvular consonants, *JASA 86(2),* 549–556.

Archangeli, D., and Pulleyblank, D. (1994) *Grounded Phonology,* Boston, MIT Press.

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

Beckman, J.N. (1998) *Positional Faithfulness,* Unpublished Ph.D. dissertation in linguistics, The University of Massachusetts, Amherst.

Beddor, P.S. (1993) The perception of nasal vowels, in *Phonetics and Phonology, Volume 5: Nasals, Nasalization, and the Velum* (Huffman, M. K. and Krakow, R.A., editors), 171–192.

Bessell, N. (1992) *Towards a Phonetic and Phonological Typology of Post-Velar Articulation,* unpublished Ph.D. thesis, University of British Columbia.

——. (1998) Phonetic aspects of retraction in Interior Salish, in *Salish Languages and Linguistics: Theoretical and Descriptive Perspectives* (Czaykowska-Higgins, E., and Kinkade, M.D., editors), 125–152.

——and Czakowska-Higgins, E. (1992) Interior Salish evidence for placeless laryngeals, *Proceedings of NELS 22*, GLSA, University of Massachusetts, 35–49.

Bickley, C. (1982) Acoustic analysis and perception of breathy vowels, *Speech Communication Group Working Paers 1, Research Laboratory of Electronics,* MIT, 71–82

Bladon, A. (1986) Phonetics for hearers, in *Language for Hearers* (McGregor, G., editor), Oxford, Pergamon, 1–24.

Blankenship, B. (1997) *The Time Course of Breathiness and Laryngealization in Vowels,* unpublished UCLA dissertation in linguistics.

Bradshaw, M. (1999) *A Crosslinguistic Study of Consonant-Tone Interaction*, unpublished Ohio State University dissertation in linguistics.

Brent, M.R., and Cartwright, T.A. (1996) Distributional regularity and phonotactic constrains are useful for segmentation, *Cognition 61,* 93–125.

Browman, C.P., and Goldstein, L.M. (1992) Articulatory phonology: an overview, *Phonetica 49,* 155–180.

Butcher, A., and Ahmad, K. (1987) Some acoustic and aerodynamic characteristics of pharyngeal consonants in Iraqi Arabic, *Phonetica 44,* 156– 172.

Cho, T., Jun, S., and Ladefoged, P. (2000) An acoustic and aerodynamic study of consonants in Cheju, *Speech Sciences,* March 2000, 109–137.

Christman, S.S., and DePaolis, R.A. (1996) Sonority as a constraint on word identification processes, *Applied Psycholinguistics 17,* 273–291.

Clements, G.N., and Hume, E.V. (1995) The internal organization of speech sounds, in *The Handbook of Phonological Theory* (Goldsmith, J., editor), Oxford, Blackwell Publishers, 245–306.

Cowell, M. (1964) *A Reference Grammar of Syrian Arabic,* Washington, DC, Georgetown University Press.

Cutler, A., and Butterfield, S. (1992) Rythmic cues to speech segmentation: evidence from juncture misperception, *Journal of Memory and Language 31,* 218–236.

Cutler, A., and Carter, D.M. (1987) The predominance of strong initial syllables in the English vocabulary, *Computer Speech and Language 2,* 133–142.

Davis, S. (1993) Pharyngealization in Palestinian Arabic: the implication for feature theory, in *Proceedings of WCCFL* (Head, J., editor), Stanford, Center for the Study of Language and Information, 108–119.

——. (1995) Emphasis spread in Arabic and grounded phonology, *Linguistic Inquiry 26(3),* 465–498.

Davis, S.B. (1976) *Computer Evaluation of Laryngeal Pathology Based on Inverse Filtering of Speech (Speech Communications Research Laboratory Monograph 13),* New York, National Institute of Health and The Voice Foundation.

DeKrom, G. (1993) A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals, *Journal of Speech and Hearing Research 36,* 254–266.

——. (1995) Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments, *Journal of Speech and Hearing Research 38,* 749–811.

Delattre, P. (1971) Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English, *Phonetica 23,* 129–155.

Dickens, P. (1994) *English-Ju|'hoan/Ju|'hoan-English Dictionary (Quellen zur Khoisan-Forschung 8),* Köln, Rüdiger Köppe Verlag.

Doke, C.M. (1926) *The Phonetics of the Zulu Language (Bantu Studies Special Number),* University of the Witwatersrand Press, Johannesburg, reprinted 1969, Krauss Reprint, Nandeln, Liechtenstein.

Donaldson, B.C. (1993) *A Grammar of Afrikaans,* Mouton, New York.

Edmondson, J.A., Esling, J., Harris, J.G., Shaoni, L., and Ziwo, L. (2000) The aryepiglottic folds and voice quality in the Yi and Bai languages: laryngoscopic case studies, *Minzu Yuwen 5.*

Elderkin, E.D. (1988) Patterns of sound in Northern Khoisan, *African Languages and Cultures 1(2),* 123–148.

Elgendy, A.AM. (1982) Acoustic data on the consonants and vowels in Egyptian Arabic, Paper from the Institute of Linguistics, Stockholm University, Stockholm, Sweden.

——. (2001) *Aspects of Pharyngeal Coarticulation,* Amsterdam, *LOT 44,* Netherlands Graduate School of Linguistics.

El-Halees, Y. (1985) The role of F1 in the place-of-articulation distinction in Arabic, *Journal of Phonetics 13,* 287–298.

——. (1986) Does the epiglottis function as an articulator in the production of pharyngeal sounds, in *Logopedics and Phoniatrics: Issues for Future Research: Proceedings of the XXth Congress of the International Association of Logopedics and Phoniatrics,* 3–7 August, 1986, Sasakawa Hall and Miyako Inn Tokyo, Tokyo, Japan. Bulletin of the Research Insitute of Logopedics and Phoniatrics.

Esling, J.H. (1999a) Voice quality settings of the pharynx, *Proceedings of ICPhS 1999,* San Francisco, 2449–2452.

——. (1999b) The IPA categories 'pharyngeal' and 'epiglottal.' *Language and Speech 42(4),* 349–372.

Friederici, A.D., and Wessels, J. (1993) Phonotactic knowledge of word boundaries and its use in infant speech perception, *Perception and Psychophysics 54(3),* 287–295.

Frisch, S., Broe, M., and Pierrehumbert, J. (1997) Similarity and phonotactics in Arabic, unpublished manuscript, Northwestern University.

Fujimura, O. (1992) Phonology and phonetics—a syllable-based model of articulatory organization, *JASJ (E) 13,* 38–48.

——and Lindqvist, J. (1971) Sweep-tone measurements of vocal-tract characteristics, *JASA 49,* 541–558.

Gerber, S.E. (1974) *Introductory Hearing Science,* London, W.B.Saunders Company.

Goad, H. (1991) [Atr] and [Rtr] are different features, in *Proceedings of WCCFL 10* (Bates, D., editor), Stanford, Center for the Study of Language and Information, 517–529.

Gobl, C., and Ní Chasaide, A. (1999) Voice source variation in the vowel as a function of consonantal context, in *Coarticulation: Theory, Data, and Techniques* (Hardcastle, W.J., and Hwelett, N., editors), Cambridge, Cambridge University Press, 122–143.

Goldsmith, J. (1976) *Autosegmental Phonology,* New York, Garland Press, 1979.

Greenberg, J. (1966) *The Languages of Africa (Research Center for the Language Sciences 25),* Bloomington, Indiana University.

Haacke, W.H. (1999) *The Tonology of Khoekhoe* (Nama/Damara), Cologne, Rudiger Koppe Verlag.

Halle, M. (1995) Feature geometry and feature spreading, *Linguistic Inquiry 26(1),* 1–46.

——and Stevens, K. (1971) A note on laryngeal features, *RLE QPR 101,* MIT, 198–312.

Hammond, M. (1999) *The Phonology of English: A Prosodic Optimality-Theoretic Approach,* New York, Oxford University Press.

Hanson, H. (1997) Glottal characteristics of female speakers: acoustic correlates, *JASA 101(1),* 466–481.

Hay, J., Pierrehumbert, J., and Beckman, M. (In press) Speech Perception, Well-Formedness, and the Statistics of the Lexicon, *Papers in Laboratory Phonology VI,* Cambridge, Cambridge University Press.

Hayes, B. (1989) Compensatory lengthening in Moraic phonology, *Linguistic Inquiry 20,* 253–306.

Hayward, K.M., and Hayward, J.R. (1989) Guttural: arguments for a new distinctive feature, *Transactions of the Philological Society Volume 87(2),* 179–193.

Herzallah, R.S. (1990) Aspects of Palestinian Arabic Phonology: A Non-linear Approach, unpublished Ph.D. dissertation, Cornell University.

Hillenbrand, J., Cleveland, R.A., and Erickson, R.L. (1994) Acoustic characteristics of breathy voice quality, *Journal of Speech and Hearing Research, Volume 37,* 769–778.

Honda, K. (1995) Laryngeal and extra-laryngeal mechanisms of F0 control, in *Producing Speech: Contemporary Issues* (Bell-Berti, F., and Raphael, L.J., editors), New York, American Institute of Physics, 215–232.

Huffman, M.K. (1987) Measures of phonation in Hmong, *JASA 81,* 495–504.

Hume, E., Johnson, K., Seo, M., Tserdanelis, G., and Winters, S. (1999) A cross-linguistic study of stop place perception, *ICPhS XIV Proceedings,* 2069–2072.

Jackson, M., Ladefoged, P., Huffman, M., and Antonanzas-Barroso, M. (1985a) Measures of spectral tilt, *UCLA WPP 61,* 72–125.

——. (1985b) Automated measures of spectral tilt, *UCLA WPP 62,* 77–88.

Javkin, H., Antonanzas-Barroso, N., and Maddieson, I. (1987) Digital inverse filtering for linguistic research, *Journal of Speech and Hearing Research 30,* 122–129.

Johnson, K. (1997) *Acoustic and Auditory Phonetics,* Cambridge, Blackwell.

——. (1993) Acoustic and auditory analyses of Xhosa clicks and pulmonics, *UCLA WPP 83,* 33–45.

Kagaya, R. (1978) Soundspectrographic analysis of Naron clicks—a preliminary report, *Annual Bulletin RILP 12,* Faculty of Medicine, University of Tokyo, 113–125.

——and Hirose, H. (1975) Fiberoptic, electromyographic, and acoustic analyses of Kindi stop consonants, *Annual Bulletin of RILP 12,* Faculty of Medicine, University of Tokyo, 113–125.

Kawasaki, H. (1982) An acoustical basis for universal constraints on sound sequences, unpublished Ph.D. dissertation, University of California, Berkeley.

Kirk, P.L., Ladefoged, J., and Ladefoged, P. (1993) Quantifying acoustic properties of modal, breathy, and creaky vowels in Jalapa Mazatec, in *American Indian Linguistics and*

Ethnography in Honor of Laurence C. Thompson (Mattina, A., and Montler, T., editors), *University of Montana Occasional Papers in Linguistics 10,* 435–450.

Kisseberth, C.W. (1970) On the functional unity of phonological rules, *Linguistic Inquiry 1,* 291–306.

Klatt, D.H., and Klatt, L.C. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers, *JASA 87(2),* February 1990.

Klatt, D.H., and Stevens, K.N. (1969) Pharyngeal consonants, *MIT RLE QPR 93,* 207–216.

Koike, Y. (1973) Application of some acoustic measures for the evaluation of laryngeal disfunction, *Studia Phonologica 7,* 17–23.

Ladefoged, P. (1983) The linguistic use of different phonation types, in *Vocal Fold Physiology: Contemporary Research and Clinical Issues* (Bless, D. M, and Abbs, J.H., editors), San Diego, College Hill Press.

——. (1997) Linguistic phonetic descriptions, in *The Handbook of Phonetic Sciences* (Hardcastle, W.J., and Laver, J., editors), Cambridge, Blackwell, 589–618.

——and Maddieson, I. (1996) *The Sounds of the World's Languages,* Cambridge, Blackwell Publishers.

——, Maddieson, I., and Jackson, M. (1988) Investigating phonation types in different languages, in *Vocal Physiology: Voice Production, Mechanisms, and Functions* (Fujimura, O., editor), New York, Raven Press.

——and Traill, A. (1994) Clicks and their accompaniments, *Journal of Phonetics 22,* 33–64.

Laufer, A., and Condax, I.D. (1981) The function of the epiglottis in speech, *Language and Speech 24,* 39–61.

Leben, W.R. (1973) *Suprasegmental Phonology,* unpublished Ph.D. dissertation in linguistics, MIT.

Leehiste, I. (1970) *Suprasegmentals,* Cambridge, MA, MIT Press.

Lindau, M. (1979) The feature expanded, *Journal of Phonetics 7(2),* 163–176.

Lindblom, B., and Maddieson, I. (1988) Phonetic universals in consonant systems, in *Language, Speech, and Mind* (Hyman, L.M., and Li, C.N., editors), London and New York, Routledge.

Lindblom, B. and Sundberg, J.F. (1971) Acoustical consequences of lip, tongue, jaw, and larynx, *JASA 50(4),* part 2, 1166–1179.

Lindqvist, J. (1972) A descriptive model of laryngeal articulation in speech, *KTH QSPR 2/3,* 1–9.

Lisker, L., and Abramson, A.S. (1964) A cross-language study of voicing in initial stops: acoustical measurements, *Word 20,* 384–422.

Löfqvist, A., and McGowan, R.S. (1992) Influence of consonantal envelope on voice source aerodynamics, *Journal of Phonetics 20,* 93–110.

Lombardi, L. (1994) *Laryngeal Features and Laryngeal Neutralization,* New York, Garland Publishers.

MacEachern, M.R. (1999) *Laryngeal Cooccurrence Restrictions,* New York, Garland Publishers.

——. (1997) Laryngeal similarity effects in Quechua and Aymara, in *The Proceedings of the Fifteenth West Coast Conference on Formal Linguistics* (Agbayani, B., and Tang, S., editors), Cambridge, Blackwell, 337–351.

Maddieson, I. (1984) *Patterns of Sounds,* Cambridge, Cambridge University Press.

——. (1997) Phonetic universals, in *The Handbook of Phonetic Sciences* (Hardcastle, W.J., and Laver, J., editors), Cambridge, Blackwell, 619–639.

McCarthy, J.J. (1986) OCP effects: geminiation and antigemination, *Linguistic Inquiry 17,* 207–263.

——. (1991) Semitic Gutturals and distinctive feature theory, in *Perspectives on Arabic Linguistics III* (Comrie, B., and Eid, M., editors), Amsterdam, Benjamins, 63, 191–233.

——. (1994) The phonetics and phonology of Semitic pharyngeals, *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III,* (Keating, P., editor), Cambridge, Cambridge University Press, 191–233.

McCarthy, J.J., and Prince, A.S. (1993) Prosodic Morphology I: Constraint Interaction and Satisfaction, manuscript. University of Massachusetts and Rutgers University.

——(1993) Generalized alignment, in *Yearbook of Morphology 1993* (Booij, G., and Van Marle, J., editors), Dordrecht, Kluwer, 79–153.

———(1995a) Prosodic morphology, in *The Handbook of Phonological Theory* (Goldsmith, J.A., editor), Cambridge, Blackwell, 318–366.

——(1995b) Faithfulness and reduplicative identity, in *The Prosody Morphology Interface* (Kager, R., van der Hulst, H., and Zonneveld, W., editors), ROA-60, Rutgers Optimality Archive.

Miller-Ockhuizen, A. (In progress) The articulation of uvularized consonants and epiglottalized vowels in Ju|'hoansi, manuscript, Cornell University.

——. (Forthcoming) Northern Khoesan tonology, to appear in *The Khoesan Languages* (Vossen, R., editor), Curzon, Surrey, England.

——. (2001) Contrastive vowel-length and variable weight reduplicative templates in Ju|'hoansi, in *Cornell Working Papers in Linguistics 18, Khoisan: Syntax, Phonology, and Contact* (Bell, A., and Washburn, P., editors), Ithaca, NY, CLC Publications, 104–125.

——. (2000) C-V coarticulation and complex consonants: evidence for ordering in click place gestures, in *Proceedings of LP '98: Item Order in Language and Speech* (Fujimura, O., Joseph, B., and Palek, B., editors), Prague, Charles University Press.

——. (1999) Reduplication in Ju|'hoansi: tone determines weight, in *Proceedings of NELS 29* (Pius, T., Hirotani, M., and Hall, N., editors), 261– 276.

——. (1998) Towards a unified decompositional analysis of Khoisan lexical tone, in *Language, Identity, and Conceptualization among the Khoisan (Quellen zur Khoisan-Forschung 15)* (Schladt, M., editor), Köln, Rüdiger Köppe Verlag, 217–243.

Morén, B. (1999) *Distinctiveness, Coercion, and Sonority: A Unified Theory of Weight,* New York, Routledge.

Nakagawa, H. (1996) An outline of |Gui phonology, *African Studies Monography, Suppl. 22,* December 1996, 101–124.

Nespor, M., and Vogel, I. (1986) *Prosodic Phonology (Studies in Generative Grammar 28),* Dordrecht, Foris Publications.

Ní Chasaide, A., and Gobl, C. (1997) Voice source cariation, in *The Handbook of Phonetic Sciences* (Hardcastle, W.J., and Laver, J., editors), Cambridge, Blackwell, 427–461.

Novikova, K.A. (1960) *Ocherki Dialektov Evenskogo Jazyka,* Izdatjelstvo Akademii Nauk SSSR, Moscow.

Obrecht, D. (1968) *Effects of the Second Formant on the Perception of Velarized Consonants in Arabic,* Mouton, The Hague.

Ohala, J.J. (1992) Alternatives to the sonority hierarchy for explaining segmental sequential constraints, *Papers from the Parasession on the Syllable,* Chicago, Chicago Linguistics Society, 319–338.

——and Ohala, M. (1993) Phonetics of nasal phonology, in *Phonetics and Phonology: Nasals, Nasalization, and the Velum* (Huffman, M.K., and Krakow, R.A., editors), New York, Academic Press.

Pierrehumbert, J. (In press) Why phonological constraints are so coarse-grained, in SWAP special issue, *Language and Cognitive Processes 16 5/6* (McQueen, J., and Cutler, A., editors), 691–698.

Pinto, N., and Titze, I. (1990) Unification of perturbation measures in speech signals, *JASA 87(3),* 1278–1289.

Prince, A. (1980) A metrical theory for Estonian quantity, *Linguistic Inquiry 11,* 511–562.

Qi, Y. (1992) Time normalization in voice analysis, *JASA 92(5),* 2569–2572.

——and Hillman, R.E. (1997) Temporal and spectral estimations of harmonics-to-noise ratio in human voice signals, *JASA 102(1),* 537–543.

——, Weinberg, B., Bi, N., and Hess, W.J. (1995) Minimizing the effect of period determination on the computation of amplitude perturbation in voice, *JASA 97(4),* 2525–2532.

Rose, S. (1996) Variable laryngeals and vowel lowering, *Phonology 13,* 73–117.

Sagey, E. (1986) Representation of Features and Relations in Non-linear Phonology, MIT dissertation in linguistivs.

Sands, B. (1991) Evidence for click features: acoustic characteristics of Xhosa clicks, *UCLA WPP 80,* 6–37.

Shahin, K.N. (1997) Postvelar Harmony: An Examination of Its Bases and Crosslinguistic Variation, unpublished Ph.D. dissertation, University of British Columbia.

Shipman, D.W., and Zue, V.W. (1982) Properties of large lexicons: implications for advanced isolated word recognition systems, *ICASSP 82,* 546–549.

Silverman, D. (1997) *Phasing and Recoverability,* New York, Garland Publishing.

Snyman, J.W. (1970) *An Introduction to the !Xũ Language,* Cape Town, A.A. Balkema.

——. (1973) Die Vroeë Stemaanset as ‘n Fonetiese Segment by sommige Fone van Zu|’hoasi, *Taalfasette 16,* 43–61.

——. (1975) *Zu|’hoasi Fonologie en Woordeboek (Communication of the University of Cape Town School of African Studies 37),* Cape Town, A.A. Balkema.

——. (1977) The interrupted juxtaposed vowels of Zu|'hoasi, in *Khoisan Linguistic Studies 3. A.S. I. Communication #6* (Traill, A., editor), 93–106.

——. (1978) The clicks of Zu|'hoasi, in *Proceedings of the 2nd African Languages Congress of UNISA* (Baumback, E.J.M., editor), Mucklenenk for the University of South Africa (UNISA), Pretoria, 144–168.

——. (1997) A preliminary classification of the !Xuu and Zu|'hoasi dialects, in *Namibian Languages: Reports and Papers* (Haacke, W.H.G., and Elderkin, E.D., editors), Köln, Rüdiger Köppe Verlag, 21–106.

——. (1999) The phonetic description of the Zu|'hoasi clicks: a confusion of sounds?, in *African Mosaic: Festschrift for J.A.Louw* (Finlayson, R., editor), Pretoria, UNISA, 92–117.

Sproat, R.W., and Fujimura, O. (1993) Allophonic cariation in English /1/ and its implications for phonetic implementation, *Journal of Phonetics 21,* 291– 311.

Steriade, D. (1987) Locality conditions and feature feometry, *Proceedings of NELS 17,* 595–618.

——. (1999) Paradign uniformity and the phonetics-phonology boundary, in *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon* (Broe, M., and Pierrehumbert, J., editors), Cambridge, Cambridge University Press.

Stevens, K.N. (1998) *Acoustic Phonetics,* Cambridge, MA, MIT Press.

—ùand Hanson, H.M. (1995) Classification of glottal vibration from acoustic measurements, in *Vocal Fold Physiology: Voice Quality Control* (Fujimura, O., and Hirano, M., editors), San Diego, Singular Publishing Group, 147–170.

Talkin, D. (1989) Voicing epoch determination with dynamic programming, *JASA 85,* Supplement 1, S149.

Thomas-Vilakati, K.D. (1999) Coproduction and Coarticulation in IsiZulu Clicks, unpublished UCLA dissertation in linguistics, Los Angeles.

Traill, A. (1985) *Photeic and Phonological Studies of !Xóo Bushman (Quellen zur Khoisan-Forschung J),* Hamburg, Helmut Buske Verlag.

——. (1986) The laryngeal sphincter asa phonatory mechanism in ! Xóo Bushman, in *Variation, Culture, and Evolution in African Populations (Papers in Honor of Hertha de Villiers)* (Singer, R., and Lundy, J.K., editors), Johannesburg, Witwatersrand University Press.

——. (1992) A confusion of sounds: the phonetic description of !Xũ clicks, in *African Linguistic Contributions (Festschrift für Ernst Westphal)* (Gowlett, D. R, editor), Pretoria, Via Afrika Limited, 345–362.

——. (1993) Doke's description of clicks, paper presented at the 7th ALASA conference, Johannesburg.

——. (1994a) The perception of clicks in !Xóo, *JALL 15*, 161–174.

——. (1994b) Place of articulation features for clicks: anomalies for universals, in *Studies in General and English Phonetics in Honor of Professor J.D.O'Connor* (Lewis, J.W., editor), London, Pergamon Press.

——. (1997) Linguistic phonetic features for clicks: articulatory, acoustic, and perceptual evidence, in *African Linguistics at the Crossroads: Papers from Kwaluseni* (Herbert, R.K., editor), Köln, RüdigerKöppe Verlag, 99– 117.

Tserdanelis, G., and Hume, E. (2000) Nasal assimilation in Sri Lankan Portuguese Creole: implications for markedness, paper presented at the Montreal Ottawa Toronto Workshop on Phonology, Toronto, February 4–6.

van der Hulst, H., and Mous, M. (1992) Transparent consonants, in *Linguistics in the Netherlands 1992* (Box-Bennema, R., and van Hout, R., editors), Philadelphia, Benjamins, 101–112.

Vaux, B. (ms.) The Laryngeal Specification of Fricative, Harvard University.

Visser, H. (1998) The phonological system of Naro, in *Language, Identity, and Conceptialization among the Khoisan (Quellen zur Khoisan-Forschung 15)* (Schladt, M., editor), Köln, Rüdiger Köppe Verlag.

Vitevitch, M.S., Luce, P.A., and Charles-Luce, J. (1996) Phonotactic and metrical influences on adult ratings of spoken nonsense words, *Proceedings of the International Conference on Spoken Language Processing*, Philadelphia, PA, 85–85.

Vroomen, J., van Zon, M., and deGelder, B. (1996) Cues to speech segmentation: evidence from juncture misperceptions and word spotting, *Memory and Cognition 24(6)*, 744–755.

Watkins, J. (1999) CQ of laryngeal gestures and settings in Wa, in *Proceedings of the 14th International Congress of Phonetics Sciences*, San Francisco, August 1999, 1017–1020.

Winters, S. (2000) Turning phonology inside out: testing the relative salience of audio and visual cues for place of articulation, *OSU WPL 53*, 168–199.

——. (2001) VCCV perception: putting place in its place, *OSU WPL 55*, 70– 87.

Yip, M. (1993) Tonal register in East Asian languages, in *The Phonology of Tone: The Representation of Tonal Register* (van der Hulst, H., and Snider, K., editors), New York, Mouton de Gruyter, 476–494.

Yoneyama, K. (2001) Phonological Neighborhoods and Phonetic Similarity in Japanese Word Recognition, unpublished Ohio State University dissertation in linguistics.

Zawaydeh, B. (Forthcoming) The interaction of the phonetics and phonology of gutturals, to appear in *Laboratory Phonology VI* (Broe, M., and Pierrehumbert, J., editors), Cambridge, Cambridge University Press.