A phonetic and phonological description of Ao: a Tibeto-Burman language of Nagaland, north-east India

Coupe, A. *A phonetic and phonological description of Ao: A Tibeto-Burman language of Nagaland, north-east India.*
PL-543, xix + 137 pages. Pacific Linguistics, The Australian National University, 2003. DOI:*10.15144/PL-543.cover*

Also in Pacific Linguistics

Justin Watkins, 2002, *The phonetics of Wa: experimental phonetics, phonology, orthography and sociolinguistics.*

Robert S. Bauer, ed., 2002, *Collected papers on Southeast Asian and Pacific languages.*

John Bowden, 2001, *Taba: description of a South Halamahera Austronesian language.*

Elisabeth Patz, 2002, *A grammar of the Kuku Yalanji language of north Queensland.*

Joel Bradshaw and Kenneth Rehg, eds, 2001, *Issues in Austronesian morphology: a festschrift for Byron W. Bender.*

Fay Wouk and Malcolm Ross, eds, 2001, *The history and typology of western Austronesian voice systems.*

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at The Australian National University. The Editorial Board of Pacific Linguistics is made up of the academic staff of the school's Department of Linguistics. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

Pacific Linguistics 543

A phonetic and phonological description of Ao: a Tibeto-Burman language of Nagaland, north-east India

A.R. Coupe

Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia

First published 2003

National Library of Australia Cataloguing-in-Publication entry:

Coupe, Alexander Robertson
A phonetic and phonological description of Ao: a Tibeto-Burman language of Nagaland, north-east India

Bibliography.
ISBN 0 85883 519 3

1. Ao language – Phonetics. 2. Ao language – Phonology.
I. The Australian National University. Research School of Pacific and Asian Studies. II. Title. (Pacific Linguistics 543).

495.415

Typeset by Jeanette Coombes
Cover design by Emily Brissenden
Printed by Ciril's Printers, Fyshwick, Canberra

Dedicated to the memory of my grandfather
Captain David Ross Dean
1894 to 1984

Table of contents

List of maps, figures and tables

Maps

Figures

Tables

Acknowledgements

This book is a revised version of a thesis originally submitted at The Australian National University in total fulfilment of the requirements for the degree of Master of Arts. I am grateful to a number of people who have helped in its preparation and subsequent revisions. Phil Rose in particular deserves special mention for everything he has taught me about phonetics and phonology, and for undertaking the supervision of my MA thesis. David Bradley and Carol Genetti both read the entire manuscript and gave detailed comments and suggestions on earlier drafts, and various insights have been gained from discussions with Sasha Aikhenvald, Graham McKay and other members of the Research Centre for Linguistic Typology, La Trobe University. I remain solely responsible from any remaining errors or misinterpretations.

I would also like to express my thanks to my principal Ao consultants Mr Moa Aier and Mr Chuba Longchar and their families for sharing their language with me. Moa gave up many of his weekends so that I could work with him in Sydney while he was studying there, and in Mokokchung Chuba was always ready to give me his time and thoughtful explanations, or to record texts on my behalf.

I thank my extraordinary extended family in Nagaland, Meghalaya and Australia for continually supporting me in my endeavours. Lastly and above all, I would especially like to thank my wife Pavitra Gurung for benignly tolerating my obsession with linguistics, and for sharing my aspirations. Without Pavitra, none of this would have eventuated.

Abbreviations and conventions

The following abbreviations and conventions are used in notation and to gloss grammatical categories of examples. The variety of the Mongsen dialect of Ao spoken in Waromung village will be referred to in the text as WM.

Symbols:

/ /	phonemic transcription
[ ]	phonetic transcription
_#	word-finally
-	morpheme boundary
~	free variation
C	consonant
G	glide
Co	coda
V	vowel
σ	syllable
$]_\sigma$	syllable boundary
.	syllable boundary
ω	word
$]_\omega$	word boundary

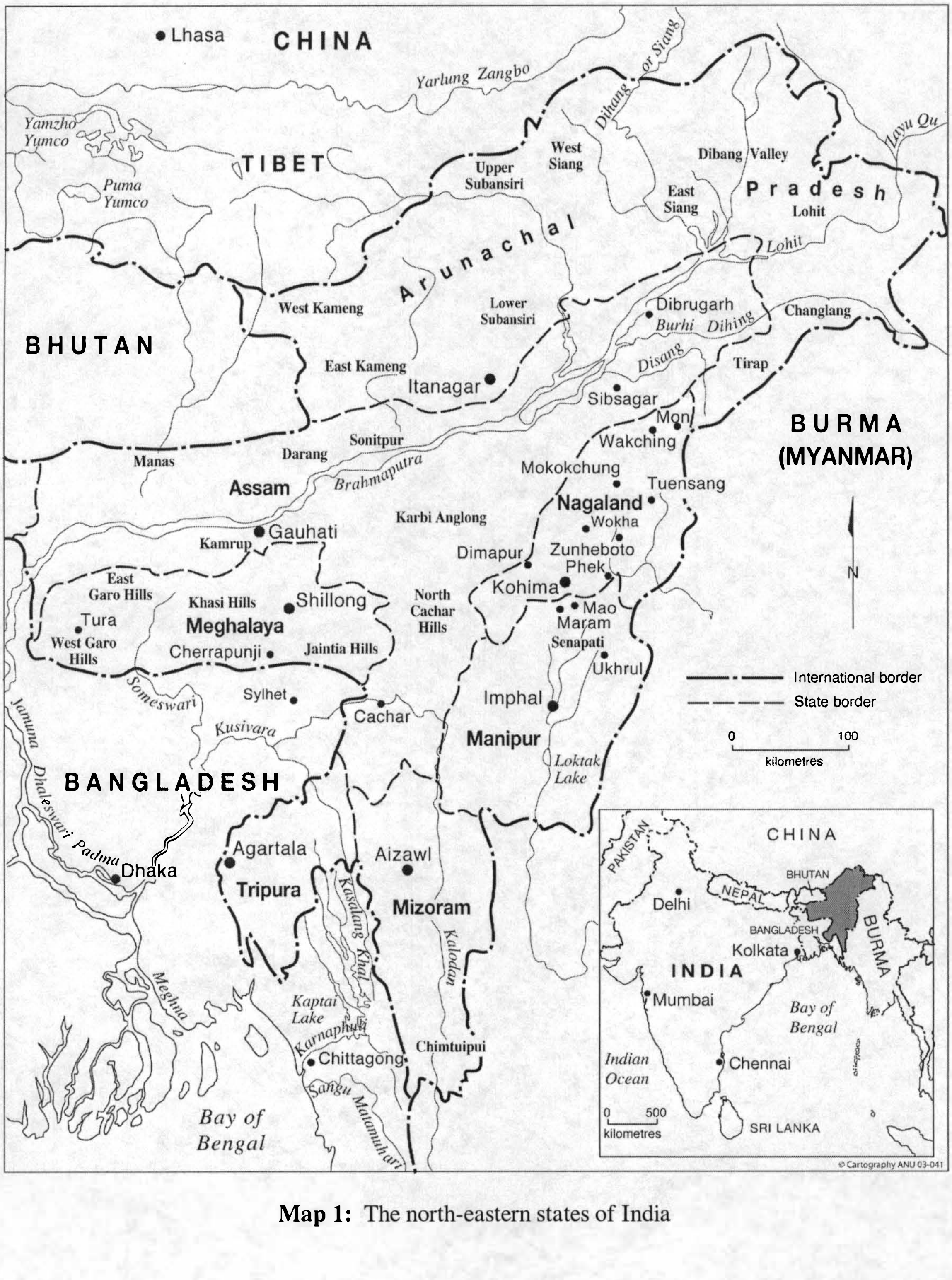
Feature abbreviations:

ant	anterior
asp	aspirated
cons	consonantal
constr	constricted
cont	continuant
cor	coronal
nas	nasal
obstr	obstruent
oral	oral
son	sonorant
syll	syllabic

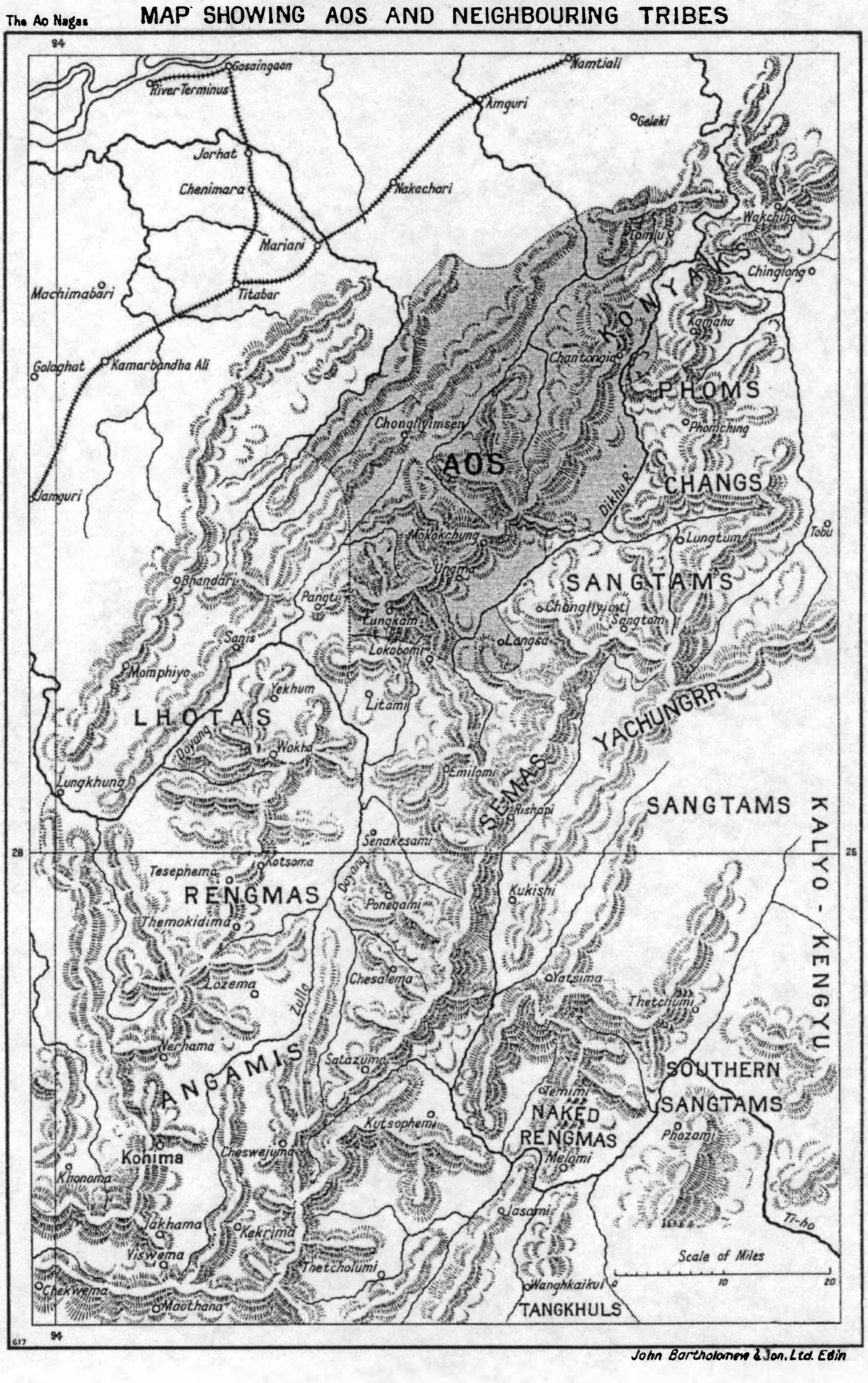
Morphological abbreviations:

1	first person
2	second person
3	third person
ADM	admonitive mood
AGT	agentive case
ANT	anterior
CAUS	causative
COLL	collective
DAT	dative case
DEC	declarative
DET	determiner
DIM	diminutive
DIST	distal demonstrative
DU	dual
EXC	exclusive
FEM	feminine
FREQ	frequentative aspect
IMPER	imperative mood
INC	inclusive
INDEF	indefinite
INST	instrumental case
IRR	irrealis mood
MASC	masculine
MOD	modality
NCV	narrative converb
NML	nominalising suffix
NPF	nominal prefix
NS	nominal suffix

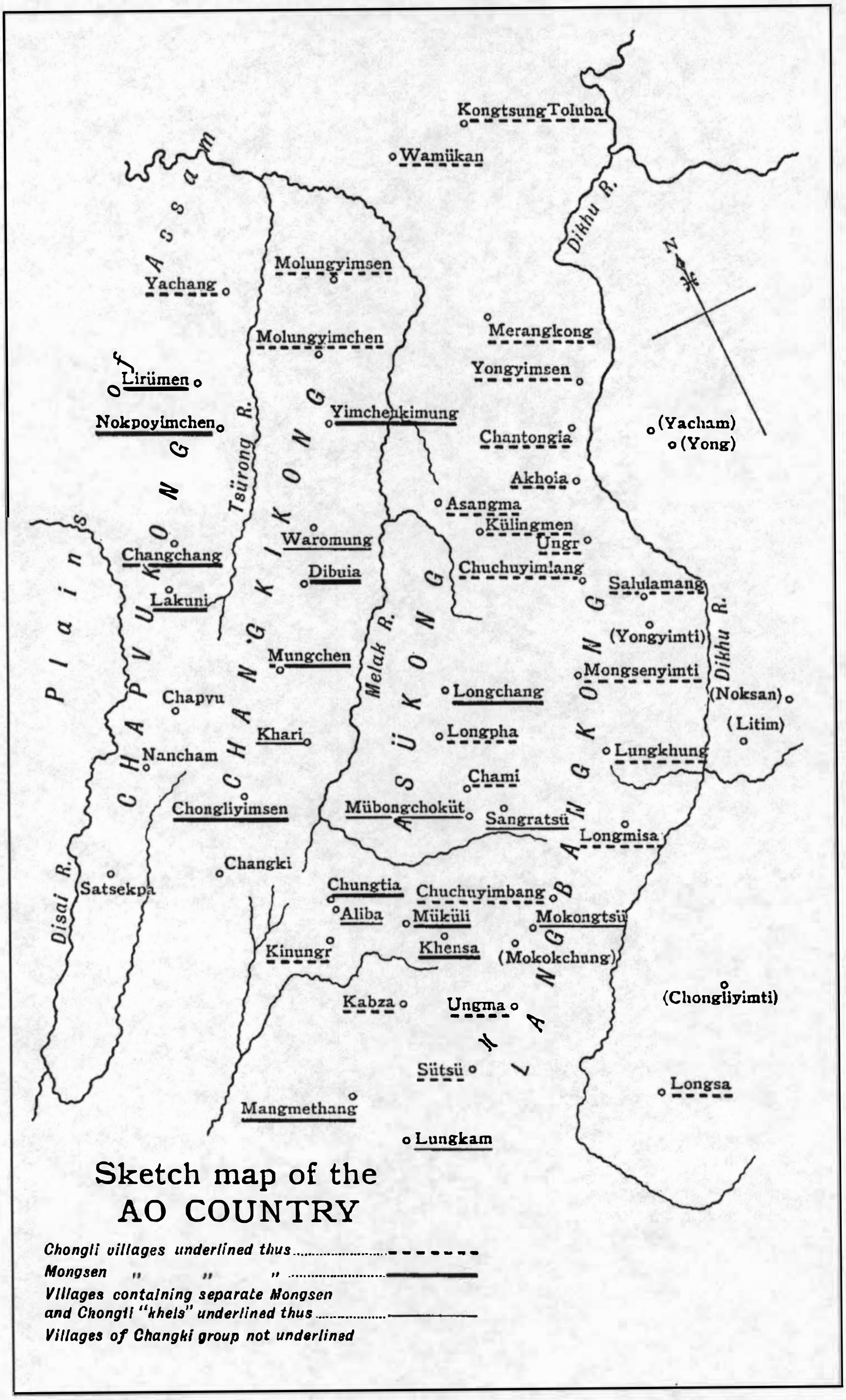
ONOM	onomatopoeia
PAST	past tense
PCV	progressive converb
PERF	perfective aspect
PL	plural
POSS	possessive
PRES	present tense
PROG	progressive aspect
PROH	prohibitive mood
RECIP	reciprocal
SG	singular
SPEC	specifier demonstrative
TERM	terminative suffix



Map 1: The north-eastern states of India



Map 2: Distribution of Naga tribes neighbouring the Ao territory of northern Nagaland (from Mills 1926, map facing p.1)



Map 3: Location of Ao villages (from Mills 1926, map facing p.4)

1 *Introduction*

1.1 Aim

The aim of this work is to provide a comprehensive description of the phonetic and phonological features of Ao, a Tibeto-Burman language spoken in the hill state of Nagaland, located in north-east India. The description is primarily based upon the data of three native speakers, and the language of investigation is a variety of the Mongsen dialect spoken in Waromung village, situated in the Mokokchung district.

1.2 Motivation

The Kuki-Chin-Naga (hereafter KCN) languages of Tibeto-Burman remain scantily described for a number of reasons, the most apparent being that continued political strife, insurgency and war in the north-eastern states of India, where the majority of these languages are spoken, has resulted in a central government policy of restricting access by foreigners to areas deemed to be sensitive. Nor has much work of a substantial nature been carried out by Indian linguists, for whom entry permits are much easier to obtain. Consequently little work using modern methods of linguistic description has been done on the Tibeto-Burman languages of the north-east. What limited data are available in the form of old word lists are often inconsistently transcribed and omit important information such as tonal contrasts; early amateur grammars rarely delve deeper into syntax than the grammatical categories of Latin. The relationship of the KCN languages (as well as many others) to other branches of Tibeto-Burman is still unclear because we lack good descriptions of these languages and relatively recent classifications continue to be based on dubious or incomplete data, much of which were collected over half a century ago.

This is the first extensive acoustic description of a language belonging to the KCN branch of Tibeto-Burman. Originally it was my intention to write a sketch grammar of the Mongsen dialect of Ao. However at the outset I needed to establish whether this was a tone language, and if it was, how many tones were contrastive, and how I should represent them. What started as a minor concern soon became a topic of major focus, eventually culminating in this monograph. A comprehensive grammatical description of Ao awaits, but in the meantime it is hoped that this work will help to shed a little more light on the segmental and suprasegmental features of one member of a largely undescribed branch of Tibeto-Burman.

Coupe, A. *A phonetic and phonological description of Ao: A Tibeto-Burman language of Nagaland, north-east India.*
PL-543, xix + 137 pages. Pacific Linguistics, The Australian National University, 2003. DOI:*10.15144/PL-543.1*

1.3 Contents overview

This book is organised into seven chapters and two appendices. Some background information on the language and its speakers, the geographical setting, previous research, and the genetic affiliation of Ao is presented in the remainder of this chapter. The linguistic description commences in Chapter 2 with an analysis of the phonotactic structure of the syllable. It may seem strange not to commence with a description of the phonemes, but beginning with phonotactics has become a Tibeto-Burman descriptive tradition with good reason. Matisoff (1973:1) embarks upon *The grammar of Lahu* with the statement that 'In Lahu, as in all the languages of the Sino-Tibetan family, the most fruitful point of departure for phonological analysis is the syllable'. Mazaudon (1974:28) echoes Matisoff's sentiments in positing that '

described', and in the same passage she cites Burling (1969:19) who, speaking of Karen, states: 'As in most Tibeto-Burman languages and in many others in Southeast Asia, the Karen syllable can be given a central place in a phonological description'. The consensus of opinion acknowledging the pivotal role of the syllable in Tibeto-Burman provides strong encouragement to follow in the footsteps of my predecessors. Following a linear and metrical description of syllable structure, the chapter then presents a treatment of vowel sequences and syllable merger in word formation processes and argues for treating the glottal stop as a prosodic element before concluding with an analysis of syllable-final stops.

Chapter 3 presents an auditory description of the segmental phonology, outlines approaches to the analysis of phonation types, discusses linguistic variation, and suggests a practical orthography. It also provides a foundation for the instrumental investigation of voice onset time in the stop and affricate series and for the acoustic analysis of the vowel formants and phonation types in Chapter 4. Chapters 5 and 6 respectively present an auditory and acoustic analysis of the tone system. Two appendices follow the synopsis of Chapter 7: Appendix A lists raw fundamental frequency and rhyme duration values calculated for words used in the instrumental analysis of the tone system; Appendix B consists of word lists gathered from two native speakers that might be of use for comparative studies.

1.4 Framework for description

All religions seem to have something of value to offer, but perhaps none exclusively has all the answers, hence the profusion of different beliefs. The same might be said of other belief systems, such as models of linguistic description. This book therefore adopts an eclectic approach to the phonological description of Ao and draws upon a number of different models according to what is judged to offer the greatest expediency and explanatory power for a particular topic, bearing in mind that an overriding purpose of the work is to make the contents maximally accessible.

The syllable is described in terms of distinctive feature theory and metrical structure, and syllabification processes are accounted for from within the Autosegmental paradigm. The latter is also used in conjunction with a prosodic analysis for describing the distribution of the glottal stop. At the end of Chapter 2 I briefly consider the theory of neutralisation and the archiphoneme in attempting to account for coda constituents. The description of the segmental phonology in Chapter 3 is based on phonemic principles with statements of allophonic realisations. Phonemes are also described in terms of distinctive features, which struggle to deal with contrasts in the vowel phoneme inventory and require the introduction of an additional place feature.

The source-filter theory of speech production (Fant 1960) implicitly underlies the acoustic analysis of vowel formants, phonation types and the tonal system. This theory is of a different order to the multifarious theories of phonology however, because it is the received theory. Thus in contrast to the many models available for phonological description, the issue of choice of a particular model for acoustic description does not arise.

1.5 The language and its speakers

1.5.1 Location

Ao is spoken in the Mokokchung district of Nagaland, a north-eastern state of India (see Map 1). The state of Nagaland consists of hilly to mountainous country formed by spurs that extend southward from the main Himalayan range and rise to a maximum altitude of 3826 metres south of the Ao region; this forms a great watershed whose rivers flow westward into neighbouring Assam and ultimately into the Brahmaputra River. The majority of Ao villages are found in commanding hilltop positions at an altitude of approximately 600 to 1800 metres above sea level on four roughly parallel ranges running down the western side of Nagaland (see Maps 2 and 3). The traditional Ao country lies between the Dikhu River at the eastern boundary and the Disai River at the western boundary. Villages are located geographically between 26°12'N and 26°45'N, and from 94°18'E to 94°46'E.

1.5.2 Dialects

Speakers of the Ao language traditionally recognise two main dialects, viz. Chungli and Mongsen. Another variety known as Changki is spoken in a few villages on the western Changkikong and Chapvukong ranges and is reported by Mills (1926:11) to be most closely related to Mongsen. Some additional varieties of Ao spoken in Yacham, Tengsa and Longla villages located east of the Dikhu River are given dialect status in the literature (Mills 1926; Marrison 1967) and are reported to exhibit contact effects from the neighbouring languages of Phom and Chang. Mills (1926:333) describes these trans-Dikhu varieties of Ao as sub-varieties of Chungli.

The Chungli variety spoken in the village of Molungyimchen and its colony Molungyimsen became the prestige dialect of Ao under the influence and patronage of the American Baptist missionaries in the late 19th and early 20th centuries. It has a Roman orthography and a Bible translation, it is taught in local schools up to the tenth grade, and it is the language used in Ao Baptist Church services, even those conducted in Mongsen villages. Consequently most Ao are able to speak this standard variety of Chungli; however, it is less common for a native speaker of Chungli to also be fluent in Mongsen unless he or she is from a village that has both Chungli- and Mongsen-speaking communities living within the same village, and even then there is no guarantee that such a speaker will be competent in both dialects.

The 1981 Census of India reports that approximately 105,000 people claim Ao as their first language but does not give a breakdown of this figure according to dialect. Map 3 (from Mills 1926) records 25 Chungli villages, 13 Mongsen villages, and 7 villages with both Mongsen and Chungli *khels* [administrative wards] within the same village. Of the latter group, Mokongtsü (Mokokchung), Khari (*sic.*) and Waromung villages are now wholly Mongsen-speaking, but I can personally vouch for the continued existence of bidialectalism in Longkhum, Sangratsü and Mübongchokiit villages. One might hazard a conservative guess from these numbers that roughly 40% of the Ao group, ca. 40,000 people, speak the Mongsen dialect as their first language.

Each of the two main dialects subsumes a number of varieties that demonstrate varying degrees of phonological, morphological and lexical divergence. A preliminary survey suggests that every village speaks its own variety; native speakers report that the unique village-specific characteristics of each variety serve as shibboleths to identify their speakers' villages of origin. Tonal contrasts, however, appear to be quite constant across varieties, at least within the Mongsen dialect. A comparison of two Mongsen varieties (Waromung and Khensa) spoken in villages located on opposite sides of the Ao territory demonstrates an extremely high degree of tonal correspondence, both in auditory impressions of pitch shape and number of tones. Presumably varieties of the Chungli dialect also demonstrate a similar degree of tonal uniformity, although this has not yet been personally attested.

1.5.3 Previous work on the Mongsen dialect of Ao

Aside from Coupe (1998), most of the contents of which appears in this book, the segmental and tonal phonology of the Mongsen dialect has not been previously described. A short description of the grammar of the Mongsen dialect was done by Mills (1926:332–369), who sketched an outline in a chapter of his excellent anthropological monograph. This gives a brief description of the phones with reference to the sounds and orthography of English; a description of lexical tone is omitted in the text and is absent in the transcriptions. Grammatical categories and clause types are briefly discussed and illustrated with examples, and the chapter also includes some comparative word lists of five varieties of Ao. Grierson (1967:281–283; 292–327, first published 1903) also provides comparative vocabulary lists of varieties of the Mongsen and Chungli dialects. Shafer (1974) gives extensive comparative word lists of varieties of Ao, presumably gathered from the same sources as Grierson (1967) that were compiled during the colonial era. A word list by Mills (Typescript n.d.) is reported by Marrison (1967) to exist in the library of SOAS, University of London. I have not had access to this. The most extensive work to date is the comparative study of Naga languages by Marrison (1967); his information on the Mongsen dialect is based upon the above-mentioned material gathered by Mills.

1.5.4 Previous work on other varieties of Ao

The Chungli variety of Ao has attracted the greatest attention of past government administrators, missionaries and linguists, the first grammar being written by an American Baptist missionary, Mrs E.W. Clark (1981, first published 1893). This constitutes forty-nine pages of grammatical description, thirty-eight pages of illustrative phrases and an English-Ao dictionary of ninety-four pages. The section on grammar consists of a discussion of word classes, verbal inflections and aspectual suffixes. Lexical tone is not mentioned or transcribed, and the description is Latinate in approach. Her husband, the Reverend E.W. Clark, also published an Ao-Naga dictionary of Chungli (1990, first published 1911). The dictionary (977pp.) is useful for its detailed descriptions but does not indicate tone, is inconsistent in its transcriptions and does not identify the parts of speech of its entries. A brief grammatical description of Chungli is also included in Grierson (1967:265–283) and is based on Clark (1981).

Wolfenden (1929) makes reference to the prefixes of Chungli in a comparative study of Tibeto-Burman morphology; the source for this work is also Clark (1981) and Clark (1990). The previously mentioned study of Marrison (1967) makes lexical, morphological and syntactic comparisons of all the Naga languages for which there is available data; this work is mostly based upon secondary sources, but is also augmented by original field research on Chungli as well as a number of other Naga languages. The most recent publications dealing with the Chungli dialect are Gowda (1972), a phonetic reader (60pp.) written for the purpose of teaching pronunciation to non-native speakers, and Gowda (1975), which gives a brief sketch of the Chungli dialect (71pp.) with chapters on phonology, morphology and syntax.

1.5.5 Genetic affiliation

There have been a number of different classifications dealing with the place of the so-called Naga languages within the Tibeto-Burman family. Four proposed classifications will be briefly outlined here. Figures 1.1–1.4 of this section are limited to demonstrating the position of Ao only in Tibeto-Burman and are not comprehensive.

The classification of Shafer (1955; 1974) is based on colonial era sources and posits six main divisions of Sino-Tibetan — Sinitic, Daic, Bodic, Baric, Burmic and Karenic — with the latter four divisions constituting what is generally accepted as the Tibeto-Burman family. Each division subsumes sections that split into hierarchically arranged branches and lower level units. The Naga languages fall under Shafer's Burmic and Baric divisions. Specifically, the Ao language is located in the Northern Naga branch of Shafer's Kukish section of the Burmic Division, together with Lotha, Yimchunger, Sangtam and Rong [Lepcha]. Hale (1982:5–6) reports that Shafer's methods have received criticism from a number of quarters and consequently have gained little following. Recently Bradley (1997) has reassessed Shafer's proposed Burmic grouping and has reclassified specific languages of this division in the light of subsequent research.

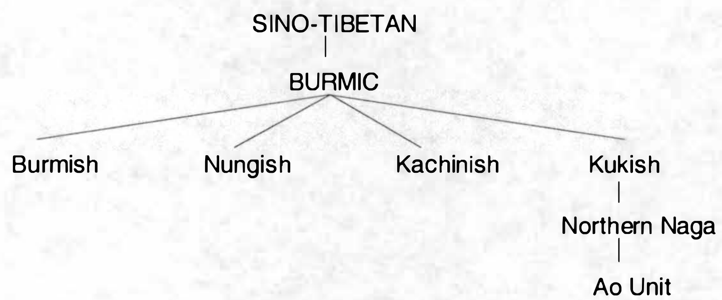


Figure 1.1: Classification of Ao in relation to Sino-Tibetan by Shafer (1955; 1974)

The influential classification of Benedict (1972) excludes Tai languages (Shafer's Daic) completely from Sino-Tibetan and classifies Karen as being coordinate with Tibeto-Burman. This forms a higher level grouping called Tibeto-Karen. Instead of the traditional family tree model, genetic relationships holding within Tibeto-Burman are demonstrated schematically via a helical model of seven primary nuclei: Tibetan-Kanauri (Bodish-Himalayish); Bahing-Vahu (Kiranti); Abor-Miri-Dafla (Mirish); Kachin; Burmese-Lolo (Burmish); Bodo-Garo (Barish); and Kuki-Naga (Kukish). Benedict considers Kachin to be at the crossroads both linguistically and geographically, with transitions to the satellite sub-groupings.

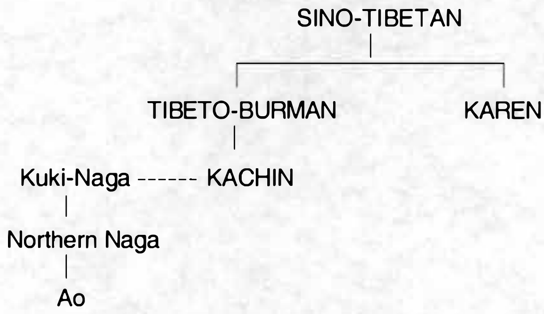


Figure 1.2: Classification of Ao in relation to Sino-Tibetan by Benedict (1972)

Benedict identifies two main sub-types of Naga languages in the Kuki-Naga nucleus: Northern Naga, of which Ao is posited to be a member, and Southern Naga. The Naga languages spoken in the north of Nagaland, in the Tirap and Changlang districts of Arunachal Pradesh, and in adjacent regions of Burma (often referred to generically as 'Naked Naga', or as Konyak languages in the literature) are placed in the Barish supergroup with Bodo-Garo. The easternmost languages of this group, Moshang and Shangge (both also known as Tangsa), are reported to demonstrate contact effects with Kachin.

The classification of Marrison (1967) is limited to the Naga languages. This recognises three types — A, B and C — subsuming groupings sharing phonological, morphological, syntactic and lexical similarities. Ao is classified as belonging to Type B1, which includes Yacham-Tengsa, Ao Chungli, Ao Mongsen and Sangtam.

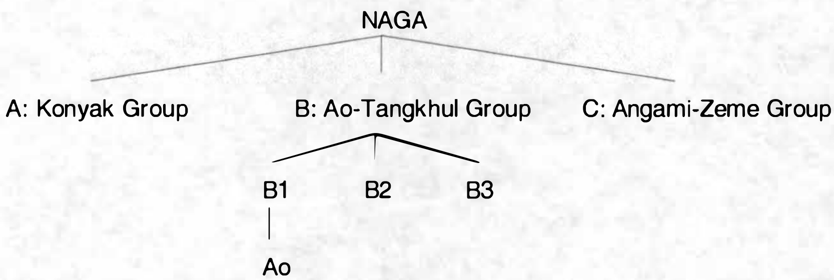


Figure 1.3: Classification of Naga languages by Marrison (1967)

At the level of the three main groupings, Type A consists of the northern Konyak languages and has affinities with the Kachin and Boro groups. These languages also demonstrate features suggestive of a closer relationship with Tibetan than with Burmese. Type B, the Ao-Tangkhul group, is spoken in the centre and south-east of the Naga Hills. These languages are reported to have mixed features, making their precise affiliations and origins difficult to determine. The last group, Type C, demonstrates affinities with Manipuri (Meithei), the Kuki languages and, to a lesser degree, Burmese.

The most recent classification is Bradley (1997), which proposes four main groups of Tibeto-Burman languages: North-eastern India (or Sal languages, after Burling 1983), Western, South-eastern and North-eastern. The revision of relevance here is the treatment of the Naga languages.

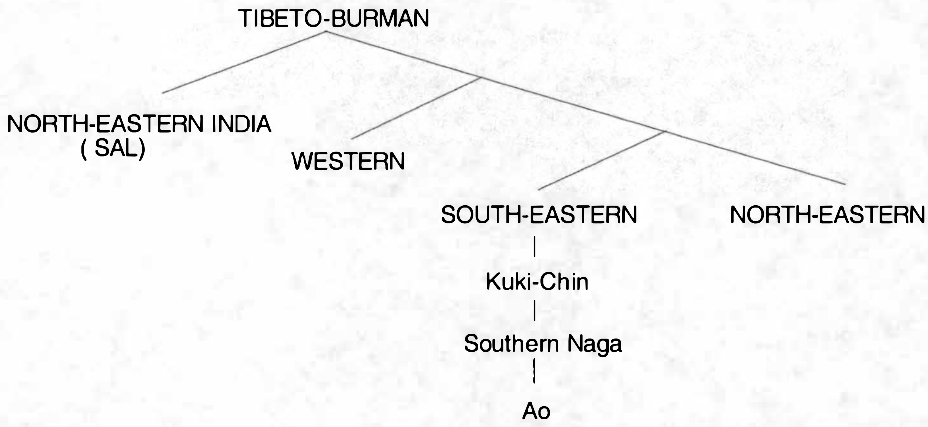


Figure 1.4: Classification of Tibeto-Burman languages by Bradley (1997)

Some of these Naga languages are classified as belonging to the North-eastern India division and others are classified as members of the South-eastern division, with shared lexical and morphosyntactic features suggestive of a link between the two divisions. The North-eastern India division consists of the Northern Naga languages (corresponding to Marrison's Type A and Shafer's Nagish section of the Baric division), and the languages of the South-eastern division form a sub-grouping of Kuki-Chin (including languages of Shafer's Kukish and Benedict's Kuki-Naga). Under Kuki-Chin are subsumed a number of lower level groupings: Southern Naga (of which Ao is a member), the Kuki languages of Manipur and neighbouring regions, and sub-groupings of various Chin languages, as well as a few languages with mixed characteristics such as Karbi (Mikir).

1.5.6 Data sources

Data were gathered from speakers of Waromung Mongsen (henceforth WM), a variety of the Mongsen dialect spoken in Waromung village. The bulk of the corpus in the form of word lists and other elicited data was provided by two tertiary-educated males in their early thirties (MA and CL). Data provided by a tertiary-educated female speaker (CA) were additionally used for the analysis of voice onset time of stops and affricates, and for the phonological analysis. Half of the corpus was collected during a period of fieldwork in Nagaland between December 1996 and February 1997, and the other half was collected in Australia between 1996 and 1998. Data also comes from texts narrated by two female native speakers in their sixties in Nagaland and from texts narrated by MA in Australia.

It is pertinent to mention that all of the native speakers who provided me with data originate from or reside in the Chungli ward of Waromung village. The fact that the two wards of Waromung village are named after the two main dialects of Ao suggests that bidialectalism once existed synchronically in the village; the logical corollary of this is that the former inhabitants of the Chungli ward once spoke the Chungli dialect as their first language. It appears that this may have influenced the present-day Mongsen now spoken in the Chungli ward of Waromung village to some extent. Evidence in support of this claim is presented in Chapter 3.

2 *Phonotactics*

2.1 Introduction

This chapter describes the phonotactics of WM. It commences with a description of the linear and metrical structure of the WM syllable in §2.2 and §2.3 respectively and gives examples of representative monosyllabic words. Section 2.4 discusses co-occurrence restrictions operating within the rhyme, and processes of syllable merger and syllabification are described in §2.5. Syllabic patterns of WM words and the phonological status of nominal prefixes are discussed in §2.6. The simultaneous segmental and autosegmental functions of the glottal stop are described in §2.7. Here arguments are presented for analysing the glottal stop as an autosegment whose domain of application is the word or syllable, rather than treating it as a segmental phoneme. Lastly, in §2.8 two theoretical approaches to the analysis of stop codas are presented, with a discussion of the relative merits of each approach.

2.2 Linear syllable structure

The canonical WM syllable minimally consists of an obligatory vowel and a tone with up to three optional elements having the following linear structure:

(2.1) (C_1) V_1 (V_2) (C_2) T

where

C_1 = [–syll]

V_1 = [+syll]

V_2 = [–syll, –cons]

C_2 = [+cons, –strid, –lat, +oral]

T = any tone

The phonotactic statement of (2.1) entails a sonority sequencing constraint. Where two vowels of differing qualities occur tautosyllabically, the first vowel forms a sonority peak and functions as the [+syll] component of the sequence, and the second vowel in the sequence functions as the [–syll] component. This can be viewed as a type of language-specific Obligatory Contour Principle (OCP) applying to sonority conditions within the syllable, a rule originally proposed by Leben (1973) to account for restrictions on sequences of identical tonemes, and later extended to account for restrictions on sequences of adjacent segments by McCarthy (1986, 1988). The OCP applies in WM to prevent two [+syll] elements occurring within the same syllable; it therefore accounts for the non-occurrence of long vowels in the data. If two vowels of identical quality occur in juxtaposition as a result of morpheme concatenation, then resyllabification processes will ensure that each vowel is associated with the nucleus of an independent syllable, or alternatively, one of the vowels will be deleted. Examples of syllable merger and resyllabification are presented in §2.5.

The syllable formula of (2.1) licenses eight possible syllable types, represented by the following monosyllabic words.

V_1	/à/	'one'	C_1V_1	/nì/	'I'
V_1V_2	/āʉ/	'axe'	$C_1V_1V_2$	/tʃʰái/	'play.PAST'
V_1C_2	/ə́ɹ/	'sew.PAST'	$C_1V_1C_2$	/ɹūŋ/	'burn.PAST'
$V_1V_2C_2$	/áuk/	'pig'	$C_1V_1V_2C_2$	/tsʰə́uk/	'wash.PAST'

The phonemes of WM will not be discussed until Chapter 3; a preview of the inventory is therefore provided below in Table 2.1. Non-syllabic offglides of phonetic diphthongs occurring tautosyllabically are signalled in phonemic transcriptions by the absence of a superscript diacritic; for example, the phonetic diphthong of [ai̯^{55}] 'dog' has only one tone associated with its nucleus and is therefore represented phonemically as /ái/. Phonetic pitch is signalled by superscript numbers from 1 to 5, with 1 representing the lowest pitch; tonemes are transcribed phonemically with the diacritics listed in Table 2.1.

Table 2.1: Waromung Mongsen phonemes

Consonant	Vowel	Toneme	Prosodic Phoneme
p pʰ t tʰ k kʰ			
ts tsʰ tʃ tʃʰ	i ʉ u	/H(igh)/ σ́	ʔ
s z h	ə	/M(id)/ σ̄	
m n ŋ	a a̰	/L(ow)/ σ̀	
ɹ l j w			

2.3 Metrical syllable structure

Represented metrically, the WM syllable has the following hierarchical structure. The labels C_1, V_1, V_2 and C_2 of the terminal nodes in Figure 2.1 below stand for the distinctive feature matrices of each constituent stated in the formula for linear syllable structure of (2.1) above. Optional constituents are enclosed in parentheses.

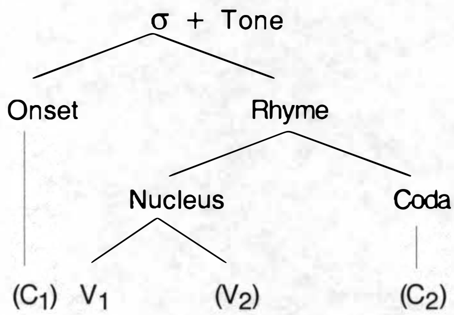


Figure 2.1: Metrical structure of the Waromung Mongsen syllable

Utterances consist of one or more syllables, which in turn minimally consist of an optional onset plus a rhyme. The rhyme contains an obligatory nucleus and an optional coda, and the nucleus of every syllable is associated with an obligatory tone whose domain of relevance extends over the syllable. The onset slot (C_1) allows a maximum of one consonant and can be filled by any of the twenty consonant phonemes, whereas the coda slot (C_2) is limited to just the rhotic /ɹ/, the nasals /m, n, ŋ/ and the stops /p, t, k/. An alternative treatment of allowable finals could posit archiphonemes /P, T, K/ instead of unaspirated stops. Consideration is given to both these analyses of coda constituents in §2.8. If, through the resyllabification of concatenated morphemes, a choice must be made between a consonant being associated with either the onset of a following syllable or the coda of a preceding syllable, then the Maximal Onset Principle (stated explicitly in Selkirk (1980:359), and implicitly in Hooper (1972) and Kahn (1976)) applies to ensure that the onset slot is given precedence.

The nucleus consists of one or two vowels, viz. V_1 (V_2), with the V_1 slot constituting the head and the optional V_2 slot a dependent element. Any vowel may occur in the V_1 slot, but only four segments may occur in the V_2 slot: these are /i, ʉ, u, a/, which are predictably realised as [i̯, ʉ̯, u̯, a̯] respectively, due to the OCP. Tautosyllabic vowel sequences consequently form phonetic falling diphthongs because the dependent vowel always constitutes the non-syllabic offglide of a complex nucleus. Vowel sequences may also be formed across morpheme boundaries as a result of morpheme concatenation; for example, /ātā/ 'wai
'go-ANT'. These fusions result in the same nucleic pattern of head plus non-syllabic offglide due to the OCP and consequently also form phonetic falling diphthongs. Attested vowel sequences are set out in Table 2.2 below. V_1 elements are in the leftmost vertical column and V_2 elements are in the uppermost horizontal row. Ticks indicate sub-syllabic phonetic diphthongs encountered in lexical roots and grammatical morphemes. Shaded cells represent phonetic diphthongs formed by the fusion of syllables across morpheme boundaries. Chequered cells represent collocation restrictions imposed by the OCP. Thus out of twenty possible V_1V_2 permutations, only six occur tautosyllabically and only four can form *via* syllable fusion.

The least restricted of the vowels is the low central vowel /a/. This may co-occur in the V_1 slot of a complex nucleus with any of the high vowels in V_2; it also allows /w, j/ onsets to precede a complex nucleus forming across a morpheme boundary e.g. /jā-ʉ/ 'hea
/wā-ʉ/ 'go-IRR', or occurring tautosyllabically e.g. /jàuŋ/ 'cen
restrictions are on the high rounded vowels; these are totally prohibited from occurring in the V_1 slot in complex nuclei. The permitted sequences, then, are /iu, ia, əʉ, əu, ai, aʉ, au/.

Table 2.2: Attested tautosyllabic vowel sequences in Waromung Mongsen

	i	ʉ	u	ə	a
i			✓		✓
ʉ					
u					
ə			✓		
a	✓	✓	✓		

In addition to the treatment outlined above, there are a number of other approaches to the analysis of the constituents of the rhyme. The main issues are the status of the element labelled V_2 in Figure 2.1 above, and whether the nucleus or the coda should be considered to be branching. Some alternative approaches are exemplified by (i–iv) of Figure 2.2 below; this is followed by an appraisal of the pros and cons of each approach. The alternative analyses are (i) simple nuclei with hiatus; (ii) complex nucleus with unit diphthong; (iii) simple nucleus and complex coda; and (iv) the approach adopted in this book: complex nucleus with vowel sequence (C = consonant; Co = coda; G = glide; N = nucleus; R = rhyme; V = vowel).

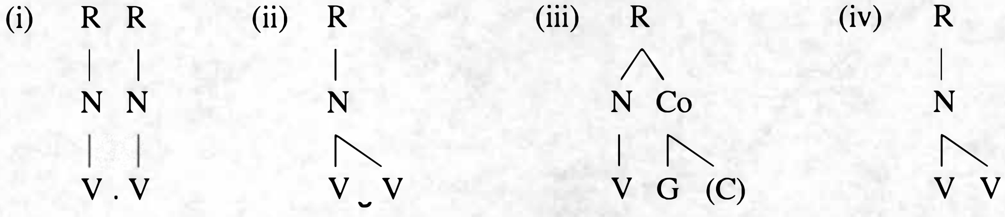


Figure 2.2: Options for the analysis of rhyme constituents

Option (i): simple nuclei with hiatus

The first alternative assumes that when two vowels occur juxtaposed, each forms the nucleus of an independent syllable and an intervening hiatus separates the two syllables. This assumption is immediately challenged by the existence of words such as /āuŋ/ 'jungle' or /áiʔ/ 'blood', whose rhymes have a duration approximately equivalent to words of just one nucleic element, such as /ùŋ/ 'see.PAST' or /à/ 'one'. Secondly, the above-mentioned words with vowel sequences are heard to have just one tone that is pronounced without a hiatus, furthermore suggesting that the sequence is tautosyllabic. The only situation in which (i) holds is when the OCP applies to prevent juxtaposed [+syll] vowels of identical quality occurring tautosyllabically as a result of word formation processes. Under these circumstances it is always the case that the vowels are associated with different tones, the most common sequences in the corpus being /Low-Mid/. Consequently when the syllabification represented by (i) does occur, it serves to prevent two tones being associated with just one nucleus, thereby creating a contour tone out of two level tones, and it also prevents the formation of a long vowel. This is interpreted as a constraint imposed by the OCP under a specific set of conditions, examples of which are presented in detail in §2.5.3.

Option (ii): complex nucleus with unit diphthong

The second alternative assumes that two vowels occurring in the nucleus form unit diphthongs. Substitution tests ultimately demonstrate this to be empirically falsifiable. To draw an analogy, whereas neither the initial stop nor the strident release of the affricate /ts/ can be substituted for any other consonant in the phoneme inventory of WM, thus proving that /ts/ is a unit phoneme consisting of two components associated with one terminal node (C_1) in the syllable, the same cannot be said of tautosyllabic vowel sequences. The lack of syntagmatic cohesion of complex nuclei in WM is demonstrated by the attested vowel sequences of Table 2.2 above, and by the following substitution possibilities. Starting with the word /áiʔ/ 'blood', the second element can be exchanged for a high central rounded vowel to produce /áʉʔ/ 'rat', and with the addition of an onset, the second element can be substituted for yet another vowel to form /hàuʔ/ 'yes'. A similar result obtains for substitution tests involving the first vowel of a sequence. Ignoring the irrelevant differences in the onset, the first vowel of /tsʰə́uk/ 'wash.PAST' may be substituted for a low central vowel to form /tʃʰàuk/ 'keep.PAST'. Thus option (ii) lacks appeal essentially on empirical grounds, but also because this approach necessarily entails an unjustifiable increase in the size of the phoneme inventory at the expense of economy of description.

Option (iii): simple nucleus and complex coda

The third option assumes a branching coda rather than a branching rhyme and treats the non-syllabic element as a glide. The fundamental problem with this approach is that one is forced to accept the fact that Glide-Consonant sequences can occur in the coda, as in /tsʰə́wk/, but not in the onset. With the exception of the typologically rare Australian languages, which generally allow the maximum coda to be more complex than the maximum onset (Evans 1995:742), languages of the world show a tendency to place greater phonotactic restrictions on complex coda formation than on complex onset formation, particularly in those languages that have contrastive tone or register. Although this is not recognised as a cross-linguistic universal of syllable formation in a survey by Greenberg (1978), the implication is nevertheless convincingly demonstrated in a number of related and unrelated languages of the South-East Asian region for which descriptions of the syllable are available; for example, Thai (Henderson 1951), Cambodian (Henderson 1952), Khasi (Rabel 1961) and Burmese (Okell 1969). All of these languages allow sequences of either Consonant-Glide or Consonant-Consonant in the onset, yet none allows any sequences whatsoever to occupy the coda. The point of this observation is that if the WM syllable template prohibits Consonant-Glide sequences in the onset, then it is even more unlikely that it would allow mirror-image Glide-Consonant sequences to occur in the coda.

Option (iv): complex nucleus with vowel sequence

The last option treats two tautosyllabic vowels as a sequence consisting of syllabic and non-syllabic vocalic segments immediately dominated by the nucleus node. The preference for option (iv) is justified by the following considerations. Firstly, a branching nucleus is consistent with phonotactic constraints on sequences of vowel segments and coda formation; it is also consistent with substitution tests involving constituents immediately dominated by the nucleus. Secondly, it does not result in an increase in the size of the vowel phoneme inventory, and so offers the greatest economy of description. Thirdly, the durations of rhymes with tautosyllabic vowel sequences are comparable to that of rhymes with single vowels. And lastly, option (iv) is consistent with resyllabification processes, a discussion of which is presented in the next section.

Having reviewed the recent literature, I cannot find any convincing arguments to support the assumption that an analysis positing a branching coda is preferable to one positing a branching nucleus. Goldsmith (1990:109) appears to be the only one who explicitly insists on treating the nucleus as a single obligatory position, although perhaps the main motivation for his claim is that it is consistent with CV Phonology and the Autosegmental paradigm, rather than being unassailably based on well-founded language universals. In contrast to Goldsmith, it appears that a good number of phonologists automatically assume a branching nucleus (for example, Cairns & Feinstein (1982); Lapointe & Feinstein (1982); Selkirk (1982); Anderson (1984), and Durand (1990)), although none actually says why.

In any event, this book is not the appropriate forum in which to attempt to resolve issues pertaining to the universal characteristics of metrical syllable structure. My investigation of different avenues of analysis of the WM syllable suggests that complex codas are less well motivated for this particular language, and that greater ease of description is therefore afforded by the theoretical assumption of branching nuclei.

2.4 Co-occurrence restrictions in the rhyme

Particularly tight syntagmatic constraints hold between the immediate constituents of the rhyme. Whereas there are no restrictions on allowable C_2 segments after simple nuclei other than those specified in the phonotactic statement of (2.1), Table 2.3 below demonstrates the extremely limited number of consonants that may fill the C_2 slot if the V_2 slot is filled by an offglide. Attested phonological vowel sequences are tabled in the left column, permitted C_2 consonants are tabled in the top row and ticks at coordinates indicate attested rhymes. Only some complex nuclei with /u, a/ occurring as the offglide element permit a syllable-final stop, and even greater restrictions apply to the offglides of complex nuclei that permit a syllable-final nasal. Phonotactic restrictions on the formation of the maximally specified rhyme can be captured by the generalisation that the only allowable coda is a velar consonant.

Table 2.3: $V_1V_2C_2$ syntagmatic constraints in Waromung Mongsen

	m	n	ŋ	p	t	k	ɹ
ai							
aʉ							
au			✓			✓	
iu							
ia			✓			✓	
əu						✓	

2.5 Syllable merger and resyllabification processes

In §2.2 it was mentioned that the OCP imposes a collocation constraint on two [+syll] segments occurring tautosyllabically. The juxtaposition of syllabic vowel segments typically results from morpheme concatenations involving verb stems and tense/aspect/modality affixes; these trigger two types of processes, namely diphthongisation and vowel deletion, depending upon the vowel segments involved and the type of affix. I should add a caveat that the analysis of merger and resyllabification is based on a very small amount of elicited data in the form of verb paradigms collected from a single native speaker (CL). The analysis would no doubt benefit from additional data collected from a range of native speakers so that the rigour of the generalisations can be assessed.

Diphthongisation and vowel deletion can be best accounted for using the Autosegmental and Metrical paradigm (Goldsmith 1976, 1990). This abandons the concept of a unidimensional, linear arrangement of segments (as originally proposed in Generative Phonology, which did not acknowledge the syllable) in favour of a multidimensional structure of autonomous tiers. Segments and autosegments such as tones are linked by association lines to positions in the skeletal tier, with the latter providing the framework for the internal structure of the syllable. For a segment to be phonetically realised, it must be linked by an association line to a position in the skeletal tier. Figure 2.3 below presents a schematic representation of the syllable demonstrating the skeletal (or CV) tier, the tonal tier and the segmental tier linked by association lines intersecting at the skeletal core. This is often likened to the spine of a spiral notebook, a metaphor acknowledged by Goldsmith (1990:281) as being first proposed by Halle. Symbols on the segmental tier stand for the feature matrices of their respective segments. If desired, the metrical structure of the syllable could be incorporated on an additional plane.

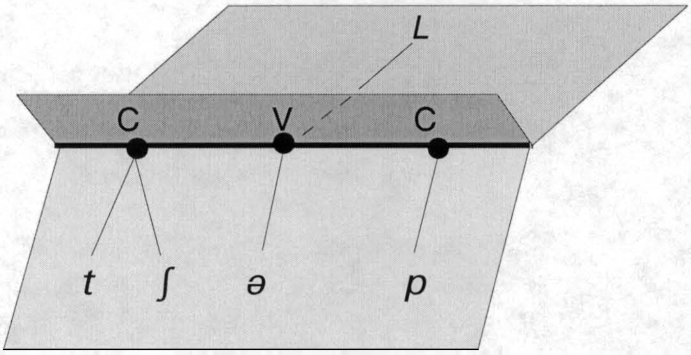


Figure 2.3: Autosegmental 'spiral notebook' representation of the features of the word /tʃə̀p/ 'weep.PAST'

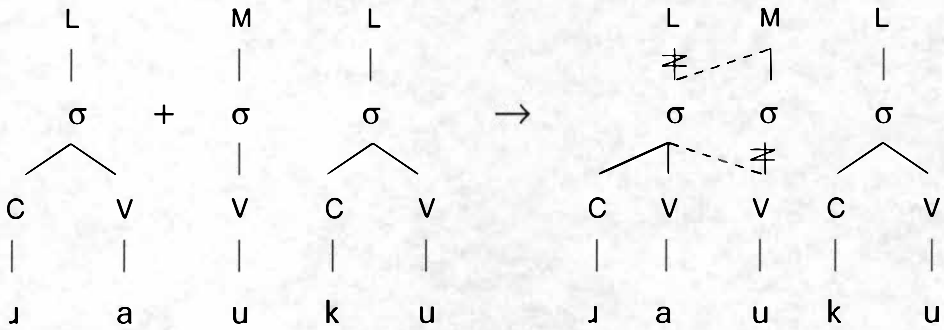
Autosegmental phonology can also be used to account for the affricates of WM. Affricates can be decomposed into a sequence of a stop onset followed by a friction-generating release associated with one position on the skeletal tier. Because the features of both components are associated with just one skeletal position, they are analysed as constituting unit phonemes, viz. /ts/, /tsʰ/, /tʃ/ and /tʃʰ/. To take this one step further, an aspirated affricate might additionally be decomposed into supralaryngeal gestures located on the segmental tier and the laryngeal feature of aspiration located on an autonomous tier, with all feature values linked by association lines to one skeletal position that functions as a single segment.

Returning to the previous discussion, processes resulting from syllable mergers can be considered to be dissimilation strategies employed by WM to avoid violations of the OCP. This is achieved either through the de-linking of a syllabic segment and its reassociation to a neighbouring syllable, or through the outright deletion of a segment violating the OCP. Some of these processes interact with the particular tone a syllable bears. Examples of both processes are presented below.

2.5.1 Diphthongisation

Diphthongisation is only found to occur with morpheme strings involving suffixes. If the juxtaposed vowels constitute one of the sequences identified in Table 2.2, then the second vowel of the sequence loses its syllabic status and forms the non-syllabic offglide component of a phonetic diphthong. Some examples are /ɹà/ 'come' + /ūkù/ ANT → [ɹau̯^{33}ku^{11}] 'come-ANT', and /sə̄/ 'die' + /ʉ̄/ IRR → [səʉ̯^{33}] 'die-IRR'. The process of diphthongisation is represented schematically as follows in (2.2). To simplify the representation, the internal constituency of the syllable as exemplified by Figure 2.1 is omitted.

(2.2)



Resyllabification →

In (2.2) above, the following resyllabification rules apply:

- The OCP deprives the second syllable of its syllabic status and the now [–syll] vowel segment of that syllable is associated with the initial syllable to form the offglide of a phonetic diphthong.
- The /Low/ tone is de-linked from the initial syllable.
- The /Mid/ tone of the second syllable spreads leftward and docks onto the initial syllable.

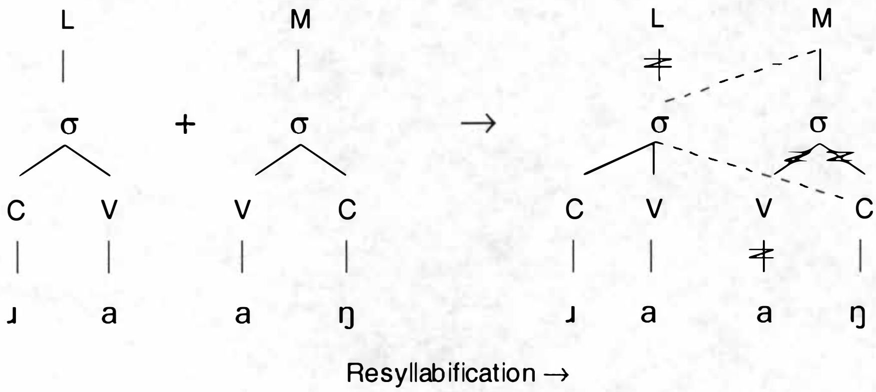
Despite the superficial similarity of the level tones of WM to the level tones of many African languages, some major differences exist. Firstly, Hyman and Schuh (1974:88–90) report that tones are much more likely to extend to the right than to the left in African languages, the motivation being phonetic assimilation. In WM, however, tones tend to spread to the left. Judging from the examples in the above-mentioned article, another characteristic of African languages is the widespread occurrence of contour tones developing out of the extension of a level tone to a syllable bearing a tone of a different pitch level, for example /H/ + /L/ → /H-L/. This type of tonal coalescence is unknown in WM word formation processes. If two underlying tones are juxtaposed as a result of morpheme concatenation and diphthongisation, then one of the tones will always undergo de-linking from its syllable.

2.5.2 Vowel deletion involving suffixes

Vowel deletion prevents V_iV_i sequences (that is, identical vowels) occurring sub-syllabically and affects all vowels except for /i/ (no suffixes with an initial high front vowel were encountered in the corpus) and /ʉ/. Verbal suffixes in which /ʉ/ occurs as the first linear segment are subject to an allomorphic realisation when occurring before a [+round] segment. A discussion of this is presented in §3.8.

If a V_iV_i sequence forms across a morpheme boundary as a result of morpheme concatenation, then the second vowel of the sequence is deleted and any remnant constituents on the segmental tier reassociate with the initial syllable to form a new syllable. Some examples are /ɹà/ 'come' + /-āŋ/ IMPER → [ɹaŋ^{33}] 'come-IMPER, and /sā/ 'say' + /-āitāu/ PROG → [sai^{33}tau̯^{33}] 'say-PROG'. Vowel deletion is represented as follows:

(2.3)



In (2.3) above, the juxtaposition of identical [+syll] segments across a morpheme boundary triggers the following resyllabification processes:

- The [+syll] vowel of the second syllable is deleted due to the OCP applying at the segmental level, thus depriving the second syllable of its syllabic status: $V_i \rightarrow \emptyset \,/\, V_i\_$
- The /Mid/ tone of the second syllable spreads leftward and docks onto the initial syllable.
- The /Low/ tone is de-linked from the initial syllable.
- The remnant coda of the second syllable is reassociated to the initial syllable.

Since the vowels of a sequence are identical in quality, it is difficult to claim with any certainty which vowel of the sequence is deleted. However we might determine this by analogy to the previously described process of diphthongisation. Recall that the second vowel of the sequence loses its syllabic status as a result of syllable merger in (2.2). Because of this, it is assumed that the second vowel of the underlying sequence is similarly affected in the process of vowel deletion demonstrated by (2.3), but in the latter case the juxtaposition of identical vowels across a morpheme boundary results not just in the loss of syllabicity of one of the segments, but also the complete dissociation of the affected vowel from its skeletal position. Segments not linked to the skeletal tier by an association line are consequently not realised.

Alternatively, it might also be argued that the second syllable retains its syllabicity and it is instead the first syllable that loses its vowel. Under such circumstances we would have to assume that the remnant onset of the first syllable associates with the second syllable, and thus there is no requirement for tone spreading.

A consequence of using an autosegmental representation is that one is forced to make a choice between these alternatives. Although there is no compelling evidence to claim convincingly that either of these are necessarily correct interpretations of vowel deletion and merger processes, in (2.2) we can directly observe that both the /Mid/ tone and the vowel of the second syllable associate in a leftward direction to the neighbouring syllable, which also happens to be the root. Because languages tend to do things symmetrically, the same direction of association is assumed to occur in vowel deletion involving suffixes. This is only possible if the second vowel of the underlying sequence is assumed to be deleted.

From the limited data at hand, it appears that tone sandhi changes operating across morpheme boundaries in word formation processes are motivated by the underlying tone carried by the suffix, and not by the tone of the stem-final syllable with which it is concatenated. Most suffixes in WM carry a mid tone; this spreads leftward and associates with the final syllable of a juxtaposed verb stem regardless of that syllable's underlying tone; for example, a high, mid or a low tone on the final syllable of a verb stem is always realised as a mid tone before the mid tones of the anterior marker /-ūkù/, the reciprocal/collective marker /-tə̄p/, the imperative marker /-āŋ/ and the perfective marker /-tʃūk/, suggesting that this tonal environment results in a neutralisation. An analysis of tone sandhi involving other suffixes awaits additional data.

2.5.3 Vowel deletion involving prefixes

Vowel deletion involving prefixes is discussed separately because prefixation can produce different results to suffixation where vowel sequences formed across syllable boundaries are concerned.

Two of the three verbal prefixes found in WM are the negative /mə̀-/ and the prohibitive /tə̀-/, both of which have a schwa nucleus and bear a /Low/ tone. Some examples of these prefixes marking verbs are /mə̀-lī-là/ 'NEG-stay-NEG.PAST' and /tə̀-tùʔ/ 'PROH-dig'. If a sequence of vowels involving the schwa of one of these prefixes forms across a morpheme boundary, then vowel deletion will always occur if (a) the tone of the prefix and the tone of the syllable to which it is affixed are identical, and (b) the initial vowel of the root to which the prefix adjoins is also a schwa. As it is impossible to determine which vowel of the sequence is deleted, a rule accounting for vowel deletion must be written in the transformational format of Structural Description (S.D.), Structural Index (S.I.), and Structural Change (S.C), as proposed by Kenstowicz and Kisseberth (1979:370–377). To illustrate, word formation processes involving the verb root /ə̀mə̄t/ 'hold' and the prohibitive and negative prefixes can be represented in this format as follows:

(2.4) /tə̀- + ə̀mə̄t/ → [tə^{11}.mət^{33}] 'PROH-hold'

	t	ə̀	+	ə̀	m	ə̄	t
S.D.	C	V		V	C	V	C
S.I.	1	2		2	3	4	5
S.C.	1		2		3	4	5

/mə̀- + ə̀mə̄t + -là/ → [mə^{11}.mət^{33}.la^{11}] 'NEG-hold-NEG.PAST'

	m	ə̀	+	ə̀	m	ə̄	t	l	à
S.D.	C	V		V	C	V	C	C	V
S.I.	1	2		2	3	4	5	6	7
S.C.	1		2		3	4	5	6	7

If (a) does not hold but (b) does, then vowel deletion does not take place and the vowels of the prefix and the initial syllable of the verb root form syllabic peaks of their own syllables. In other words, the two syllables are uttered independently with a very brief hiatus and consequently each retains its original tone. From the limited data at hand, it appears that the tonal mismatch blocks the process of vowel deletion:

(2.5) /tə̀- + ə̄ɹ/ → [tə^{11}.əɹ^{33}] 'PROH-sew'
/mə̀- + ə̄ntʃūk + -là/ → [mə^{11}.ən^{33}.tʃuk^{33}.la^{11}] 'NEG-discard-NEG.PAST

If neither (a) nor (b) holds, then vowel deletion does not take place. A verb root-initial [+high] vowel will cause an assimilation of the schwa of the prefix, but the two syllables involved maintain their syllabic independence and retain their separate tones:

(2.6) /tə̀- + īpī/ → [ti^{11}.i^{33}.pi^{33}] 'PROH-sleep'
/mə̀- + īpī + -là/ → [mi^{11}.i^{33}.pi^{33}.la^{11}] 'NEG-sleep-NEG.PAST'

/tə̀- + ūkjā/ → [tu^{11}.uk^{33}.ja^{3}] 'PROH-swell'
/mə̀- + ūkjā + -là/ → [mu^{11}.uk^{33}.ja^{33}.la^{11}] 'NEG-swell-NEG.PAST'

The assimilation process observed in the examples of (2.6) can be captured by the following representation:

(2.7)

αT   βT
| |
V   V

ə   [+high]

Assimilation does not apply to verb roots with an initial low vowel /a/, depending upon one's approach to the analysis. In (2.8) below, neither (a) nor (b) holds; however this results in vowel deletion instead of assimilation of the schwa to the low vowel /a/ and the maintenance of its syllabic independence. An autosegmental analysis might therefore posit the following series of steps in word formation:

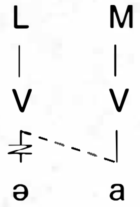
- The schwa of the prefix is deleted.
- The /Low/ tone of the prefix spreads rightward to the root-initial syllable.
- The /Mid/ tone is de-linked from the root-initial syllable.
- The remnant onset of the prefix now becomes the onset of the adjacent root-initial syllable.

(2.8) /tə̀- + ātsə̄/ → [ta^{11}.tsə^{33}] 'PROH-look'

/mə̀- + ātsə̄ + -là/ → [ma^{11}.tsə^{33}.la^{11}] 'NEG-look-NEG.PAST'

Alternatively, we could presume that an assimilation of the prefix's schwa *does* occur, but it is then subject to the OCP constraint and is subsequently deleted, in which case the steps involved in word formation listed above would be preceded by the following assimilation rule:

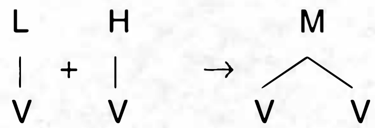
(2.9)



This would then be followed by the deletion rule applying to *a* instead of to *ə* to derive the output. Regardless of the means taken to achieve this end, the output demonstrates that a constraint applies to prohibit the realisation of VV sequences in word formation when the second vowel of the sequence is [+low].

Note that resyllabification involving prefixes operates phonologically in the opposite direction to that presumed for diphthongisation involving suffixes. This suggests that what both processes have in common is for the operations at both the segmental and suprasegmental levels to take place in a direction towards the root.

Tone sandhi operating across morpheme boundaries appears to occur less frequently with these two verbal prefixes than with the verbal suffixes. The only change observed in the tones of the prohibitive mood marker /tə̀-/ and the negative marker /mə̀-/ occurs with the juxtaposition of an underlying high tone on a verb root. When the prohibitive or negative prefix is attached to a verb root with an underlying high tone, the output is a mid tone on both prefix and root, and if the root is disyllabic, then a perseverative tone sandhi effect is observed to extend across the prosodic domain of both syllables:



(2.10) /tə̀- + tsə́uk/ → [tə^{33}.tsəuk^{33}] 'PROH-wash'

(2.11) /tə̀- + átʰúʔ/ → [ta^{33}.tʰuʔ^{33}] 'PROH-vomit'

The third verbal prefix of WM is the admonitive mood marker /āsáʔ-/, which is used to express a partially prohibitive meaning equivalent to 'don't VERB too much'. It has underlying mid and high tones respectively on its two syllables, with the high tone of the final syllable triggering some interesting tone sandhi changes in a juxtaposed root-initial syllable. If the initial syllable of a verb root is underlyingly low, then it is displaced by the rightward-spreading high tone of the admonitive prefix:

$$\begin{array}{ccccccc} \text{H} & & \text{L} & & \text{H} & & \text{L} \\ | & + & | & \rightarrow & | & & \cancel{|} \\ \text{V} & & \text{V} & & \text{V} & & \text{V} \end{array}$$

(2.12) /āsáʔ- + nàk/ → [a³³.sa⁵⁵.nak⁵⁵] 'ADM-scratch'

But an underlying mid tone on the initial syllable of a verb root is unaffected by the high tone of the admonitive prefix:

$$\begin{array}{ccccccc} \text{H} & & \text{M} & & \text{H} & & \text{M} \\ | & + & | & \rightarrow & | & & | \\ \text{V} & & \text{V} & & \text{V} & & \text{V} \end{array}$$

(2.13) /āsáʔ- + tə̄n/ → [a³³.sa⁵⁵.tən³³] 'ADM-sing'

However, if the initial syllable of a verb root has an underlying high tone, then it is realised as a mid tone in the environment of the high tone of the admonitive marker. This tone sandhi change also appears to have a perseverative effect on the second syllable of a disyllabic root:

$$\begin{array}{cccccccc} \text{H} & & \text{H} & & \text{H} & & \text{M} & \\ | & + & | & \rightarrow & | & & | & \\ \text{V} & & \text{V} & & \text{V} & & \text{V} & \text{(V)} \end{array}$$

(2.14) /āsáʔ- + tsə́uk/ → [a³³.sa⁵⁵.tseuk³³] 'ADM-wash'

(2.15) /āsáʔ- + mə́tʃʰáp/ → [a³³.sa⁵⁵.mə³³.tʃʰap³³] 'ADM-suck'

(2.16) /āsáʔ- + átʰúʔ/ → [a³³.sa⁵⁵.a³³.tʰuʔ³³] 'ADM-vomit'

The nature of tone sandhi in WM and perhaps in other KCN languages is of sufficient complexity and interest to warrant a comprehensive investigation as a separate subject of research, but it will not be pursued to any deeper level in this book.

2.6 Nominal prefixes in polysyllabic words

In a selection of 200 WM lexical items based on the Swadesh word list, 21% of the corpus comprised monosyllabic words, 62% comprised disyllabic words and 17% comprised trisyllabic words. In addition to those disyllables that are formed from monosyllabic roots — for example /pāŋ-kə̄p/ 'lip', a compound noun derived from the bound nominal roots of /tə̄-pāŋ/ 'NPF-mouth' and /tə̄-kə̄p/ 'NPF-skin' — a large number of disyllabic words conform to a (C)V.CV(C) pattern, the first syllable of which is a prefix. The latter two examples are representative of that phonotactic pattern.

The most commonly occurring nominal prefix is composed of a dental stop and a schwa, viz. /tə-/, which also derives verbal nouns and adjectives in Ao; Wolfenden (1929:133) traces this via Sikkim Tibetan *te* 'that' to an Old Tibetan demonstrative *de*. Another nominal prefix consisting solely of the vowel /a-/ is found to occur on bound noun stems, as in /á-sáʔ/ 'NPF-meat' and /á-ŋáʔ/ 'NPF-fish'. In contrast to /tə-/, this prefix is not involved in derivational processes of word formation and does not express any semantic content. It is restricted to occurring on bound monosyllabic nominal roots where it forms disyllabic words.

The vowel of the /tə-/ prefix is often 'coloured' by the vowel(s) of the root, such as in the word /tə̄-kūlūk/ 'NPF-brain', sometimes realised phonetically as [tu^{33}ku^{33}luk^{33}], although the bulk of examples demonstrate that this is not a consistent enough phenomenon to justify it being recognised as vowel harmony at the phonological level of description. Being a schwa, it is inherently unstable and regularly assimilates in height and place to a following semi-vowel, for example [ti^{33}jiŋ^{33}], derived from /tə-/ plus the lexical root /jīŋ/ 'root'. Examples in the corpus suggest that it undergoes deletion when occurring in juxtaposition with a low vowel as a result of affixation, such as in [ta^{33}ɹu^{33}], derived from /tə-/ plus the intransitive verb root /āɹū/ 'good'.

The prefix /tə-/ is most frequently found as the initial syllable of body part and kinship terms, for example /tə̄-pūk/ 'NPF-belly', /tə́-kʰə́t/ 'NPF-hand' and /tə̄-nū/ 'NPF-younger sibling', and it can be replaced by a possessive pronominal prefix to signal possession, as in /kə̄-nū/ '1SG.POSS.younger sibling'. The existence of the above-mentioned nominal prefix /a-/ on noun roots can similarly be demonstrated by the attempted prefixation of possessive pronominal morphemes or by the formation of nominal compounds, as in /á-uk/ 'NPF-pig' + /á-sáʔ/ 'NPF-meat' → /á-uk-sàʔ/ 'pork' (NPF-pig-meat), or /á-ŋáʔ/ 'NPF-fish' + /tə̄-kə̄p/ 'NPF-skin' → /ŋā-kə̄p/ 'fish scale'. Compound noun formation minimally results in the deletion of a nominal prefix from the head of the compound and is often accompanied by tone sandhi changes, depending upon the tonal environments created by compounding.

The majority of nouns and derived nominals in WM are observed to be disyllabic, often consisting of a prefix plus a root. An iambic pattern of disyllable was originally recognised in Mon-Khmer and in Thai loan words of Mon-Khmer provenance by Henderson (1951) and a similar pattern has since been reported in the disyllabic words of a number of Tibeto-Burman languages spoken in neighbouring regions; for example, in Tiddim Chin (Henderson 1965), in Lahu, but limited to loan words of Tai or Burmese origin (Matisoff 1973), and in Burmese (Bradley 1980). The gradual and non-systematic development of bimorphemic compounds with unstressed, tonally undifferentiated initial syllables in Burmese leads Bradley (1980) to conclude that this pattern has resulted from phonological convergence with Mon.

A prefixal syllable is found in many Tibeto-Burman languages, particularly those of north-eastern India and adjacent areas to the east; Marrison (Vol. I, 1967:110) alone provides a list of 31 languages with examples of their nominal prefixes. Such widespread attestation suggests that a shared genetic source might explain this apparent ubiquity. Yet in spite of this syllable type being recognised in so many Tibeto-Burman languages of the region, in discussing the syllable structure of Jinghpaw (Kachin) Matisoff (1989:164–165) proposes that they have no plausible etymology and are therefore likely to be a relatively recent development in that language. Given the pivotal position accorded to Kachin in the classification of Tibeto-Burman languages by Benedict (1972), this would suggest a similarly recent development of the prefixal syllable in other Tibeto-Burman languages. This is not the case for the reduced pre-syllables of Austro-Asiatic and Mon-Khmer, however, for which Matisoff suggests a venerable antiquity. Regardless of the origin of this phonotactic pattern, he proposes that thousands of words in South-East Asian languages are neither monosyllabic nor disyllabic, but are what he refers to as 'sesquisyllabic'; that is, one-and-a-half syllables long. Authors surveyed are in general agreement that the initial syllable of these disyllabic words is typically unstressed in relation to the following syllable; it carries fewer tonal contrasts than stressed syllables or is not tonally differentiated by contrastive pitch; it is never closed by a consonant; and it has a short vowel with limited oppositions in quality.

Morse (1963:33; 26fn.) makes the observation that many daughter languages of the Naga, Bodo-Naga, Kuki-Chin, Kachin and some Burmic branches of Tibeto-Burman demonstrate verbal prefixes of Proto-Tibetan provenance. These prefixes are formed in the Kachinic language Rawang by the insertion of an anaptyptic schwa between two initial consonants, thus creating disyllabic words with the phonotactic structure of CəC(C)V. The tones of such syllables are non-contrastive, since they are determined by the tone of the following nuclear syllable (Morse 1963:18). The occurrence of disyllabic words in WM having a similar phonotactic pattern to that reported in related languages raises the question of whether the syllable canon posited at (2.1) provides an adequate account of all the phonotactic manifestations of possible words in the language, or whether an additional pattern should be recognised.

Firstly, the observation that these prefixes are not limited to disyllabic words in WM (for example, /tə̄-mə̄kūk/ 'NPF-knee' and /tə̄-mə̄ɹə̄m/ 'NPF-red') refutes any presumed necessity of recognising a specific sesquisyllabic pattern to supplement the phonotactic statement of (2.1). Nevertheless, the nominal prefixes that occur in these words deserve careful examination to determine their phonological and suprasegmental status.

Acoustic measurements reveal that the prefixes /tə-/ and / a-/ have average rhyme durations comparable to polysyllabic words with initial CV syllables that do not constitute morphological prefixes (for example, /wàjàʔ/ 'bird'), so there is no justification for recognising a separate phonotactic class of disyllabic word on the basis of rhyme duration values (cf. those listed in the tables of Appendix A). Furthermore, in contrast to the above-mentioned Tibeto-Burman languages in which the initial syllables of disyllabic words are described as being unstressed, the WM prefixes cannot be differentiated in a non-arbitrary manner from other initial syllables in polysyllabic words on the basis of any of the perceptual cues by which stress is normally recognised, be it duration, intensity, higher pitch or loudness, simply because stress is not a phonologically salient feature of WM.

There is the possibility of analysing the pitch of a nominal prefix as non-contrastive, as Morse (1963) does in his treatment of Rawang. Admittedly, the pitch of the nominal prefix in WM is nearly always identical to the pitch of its root in the data of one native speaker. This suggests that for this particular speaker, it is the word and not the syllable that should be analysed as the domain of tone for this type of disyllabic word. Unfortunately the existence of a number of counter-examples, such as the /High-Low/ tones of /tə́-làŋ/ 'NPF-long' and /á-lù/ 'NPF-field' makes this suggestion less appealing. These examples are taken from the corpus of CL, whose word list in Appendix B demonstrates the greatest harmony in pitch level between the nominal prefix and the root of polysyllabic words. Many more counter-examples can be found in MA's somewhat aberrant corpus (also listed in Appendix B), so it remains to be determined how the deviating tone on the prefix of words such as 'long', 'field', and on all the other counter-examples in the corpus is assigned, if tone is not part of the underlying lexical specification of the nominal prefix.

We may speculate that the creation of at least one of the prefixes originates historically from an anaptyptic vowel inserted to avoid a complex onset in Ao, in which case it is reasonable to assume that originally it was tonally undifferentiated. It also seems entirely plausible than an erstwhile tonally undifferentiated syllable might gradually develop contrastive pitch in a tonal language, just as the words of non-tonal languages borrowed by tonal languages such as Thai are inevitably assigned pitch value. However, if we are prepared to overlook the counter-examples, then it would be expedient to assume that the nominal prefix is neutral with regard to underlying tonal specification, and that its tone is assigned post-lexically by the initial syllable of the root in the process of word formation.

A conclusive analysis of the status of tone on nominal prefixes is not determinable at this stage and requires further investigation. For the present, surface output tones are marked on the nominal prefixes of polysyllabic words.

2.7 The prosodic glottal stop

A glottal stop frequently occurs in WM and demonstrates a number of peculiarities that sets it apart from the consonantal segments. It has a relatively high frequency of occurrence compared to the oral stops, yet is restricted to occurring as the final element of a word. In a sample of 200 words based on the Swadesh word list and comprising a lexical root ± a nominal prefix, 14% had a final glottal stop, whereas only 7% had a final velar stop, 4% had a final dental stop and 5% had a final bilabial stop.

The realisation of the glottal stop is non-predictable and it contrasts with other phonemes and with zero (Ø). In a classical phonemic analysis it could therefore be identified as a segmental phoneme on the basis of its contrastive function. An alternative available in some South-East Asian tonal languages is to treat the glottal stop as one of the exponents of a particular tone, as Okell (1969) does in his description of Burmese. However this is not an option in WM because the glottal stop occurs with all three tonemes, as the following set of minimal and sub-minimal pairs demonstrate:

(2.17)	/jàk/	'beat.PAST'
	/jàʔ/	'hear.PAST'
	/ā-mī/	'NPF-spear'
	/ā-mīʔ/	'NPF-person'
	/tə̄ɹ/	'intestines'
	/tə́ɹʔ/	'height'

In addition to being constrained to occurring word-finally, the glottal stop is the only consonant that can occur after a rhotic in the coda, as demonstrated by the last example of (2.17). Were it to be treated as a segmental phoneme, we would need to acknowledge another optional coda consonant (C_3 = /ʔ/) and revise the phonotactic statement provided at (2.1) above as follows:

(2.18) $(C_1)\ V_1\ (V_2)\ (C_2)\ (C_3)\ T$

Furthermore, it would be necessary to stipulate that if both the C_2 and C_3 positions are filled, then C_2 can only be a rhotic. The additional complexities entailed in accounting for the phonotactic distribution of the glottal stop suggest that despite its contrastive function demonstrated in the examples of (2.17), a phonemic analysis identifying it solely as a segmental phoneme fails to fully explain its phonological function.

This becomes even more apparent when its behaviour under suffixation is also taken into account. A substantial number of lexical verb roots have an underlying final glottal stop. In contrast to the oral stop phonemes, the glottal stop of these verb roots is consistently deleted from its root-final position if modified by suffixing morphology. Deletion of the glottal stop in the environment of a morpheme boundary is demonstrated by the following examples:

(2.19)	/ɳùʔ/	'bite'	+	/-āŋ/	IMPER	→	/ɳūāŋ/
	/jàʔ/	'hear'	+	/-ūkù/	ANT	→	/jāukù/
	/tsʰəʔ/	'pull'	+	/-ʉ̄-líuʔ/	IRR-MOD	→	/tsʰə̄ʉlíuʔ/

A glottal stop also occurs as the final element of some tense/aspect/modality morphemes, but if morpheme concatenation results in the glottal stop of a grammatical morpheme occurring word-internally, then it will similarly undergo the same process of word-internal deletion as the glottal stop of verb roots. This is demonstrated by the deletion of the glottal stop from the causative suffix /-ʉ̄ʔ/ in the example of (2.20) below when it occurs in a word-internal position, versus its realisation when it occurs word-finally. Note that in addition to the deletion of the glottal stop, it appears that the vowel of the causative suffix is unrounded and fronted in the environment of a back rounded vowel. A discussion of this will be taken up in §3.8 of the following chapter.

(2.20)	/tʃàʔ/	'eat'	+	/-ʉ̄ʔ-úkū/	CAUS-ANT	→	/tʃāiūkù/
	/tʃàʔ/	'eat'	+	/-ʉ̄ʔ/	CAUS	→	/tʃāʉʔ/

Trubetzkoy (1977:241–246, first published 1939) posited the notion of *Grenzsignale* [boundary signals] that serve a delimitative function in a language and occur at the boundaries of units of meaning, such as morphemes. The glottal stop appears to perform a similar function in delimiting words in WM when it is realised, somewhat analogously to the word-final unreleased stops of many tonal languages of South-East Asia. From an articulatory point of view nothing could be more indicative of a boundary than the complete cessation of phonation that results from glottal closure, particularly given the glottal stop's consistently word-final distribution. It may therefore be tempting to think of the restricted position of occurrence as being teleologically motivated, or at least as embodying some degree of iconicity.

The phonological status of the glottal stop in WM appears to be somewhat ambiguous because of its dual functions. In some respects its unique distributional characteristics suggest that it does not have the same status as the oral stop segmental phonemes, yet in a phonemic analysis it clearly demonstrates a contrastive function. On the other hand, no other consonant phonemes are deleted word-internally or are restricted to occurring word-finally, so an adequate analysis must therefore be able to account for both its contrastive distribution *and* its word boundary marking function.

An alternative approach to the phonemic analysis is to recognise the glottal stop as a prosodic element or autosegment (depending upon the terminology of one's phonological theory) instead of treating it solely as a segmental phoneme. A prosodic analysis is not necessarily incompatible with the glottal stop potentially demonstrating a simultaneous contrastive function when it is realised in the word-final position. Indeed, the fact that it would be analysed as a segmental phoneme in a classical phonemic analysis is merely contingent upon this word-final distribution, although the evidence presented above suggests that the phonemic approach falls short of accounting for the glottal stop's other peculiarities *vis-a-vis* the oral stop phonemes, particularly its word-internal deletion. Furthermore, a prosodic analysis is not only able to account for the glottal stop's boundary marking function (thereby providing the rationale for word-internal deletion), but also recognises that its contrastive distribution with zero (Ø) and the oral stop phonemes is simply an associated but nevertheless ancillary consequence. In the prosodic analysis suggested here, the boundary marking function is taken as prior and the contrastive function is viewed as a corollary of the word-final realisation.

The glottal stop in WM shares some similarities with the glottal stop segment that occurs in some Australian languages of Arnhem Land (cf. Harvey 1991). While the specific details may differ from language to language, Harvey (1991) observes that the glottal stop of these languages generally demonstrates a contrastive segmental function, has a highly restrictive syllable-final distribution, has specific rules of insertion and deletion in reduplicative processes, can cross morphological boundaries under suffixation and demonstrates a preference for only occurring at reduplication and morpheme boundaries. Harvey proposes that the glottal stop is underlyingly the maximally unspecified segment in the phoneme inventories of these languages and thus can be distinguished by radical underspecification. He also posits that the glottal stop lies on an independent phonological plane to other segmental phonemes in some languages, with rule-ordered cyclical syllabification and affixation rules accounting for its association with a skeletal position.

Issues pertaining to esoteric phonological theory are beyond the scope of this description, but from the perspective of articulatory phonetic reality, radical underspecification *à la* Harvey is somewhat problematic in WM, and indeed generally. If the glottal stop of WM is similarly held to be a completely unspecified segment in terms of articulatory phonetics, then it is difficult to account for the glottal fricative /h/ that also occurs in WM, a consonant that Ladefoged (1982:62) describes as 'simply the voiceless counterpart of the following sound'. In terms of articulatory effort the glottal fricative is even less specified than the glottal stop, which minimally entails glottal closure.

The glottal stop of WM presents something of a dilemma for phonological representation because of its simultaneous segmental and prosodic/autosegmental functions, and because its domain of application is interpretable in different ways, depending upon one's theoretical stance. To illustrate, phonemic theory views the role of the glottal stop from the perspective of the end product, namely its surface contrastive function and word-final distribution (somewhat reminiscent of the simultaneous roles of 'long components' proposed by Harris (1944)), whereas the generative model views it from the morphophonemic perspective of its behaviour in word formation processes and its underlying morpheme-final distribution. On the one hand, it is a property of the word; on the other, an underlying property of the morpheme. A choice must therefore be made between describing phonotactics in terms of surface forms or in terms of underlying representations, with different representations the outcome of each approach.

We must dispense with any phonemic analysis that would identify the glottal stop purely as a segmental phoneme because this is demonstrated to be inadequate in accounting for its distributional and functional characteristics. However there is also the problem of theoretical incompatibility between the underlying morphophonemic representation of the generative model, and the morpheme not being recognised as part of the prosodic hierarchy. In order to be compatible with the tenets of prosodic phonology, either the syllable or the word must be identified as the prosodic domain of the glottal stop, and a deletion rule (formally stated in (2.21) below) might conveniently serve to account for its word-final distribution.

For the purposes of this description the glottal stop is analysed as a non-predictable word prosody that simultaneously functions as a contrastive unit at the segmental tier when it is realised word-finally; it is otherwise subject to a deletion rule in the environment of a word-internal syllable boundary.

(2.21) $\omega \quad = \quad \sigma \ (\sigma) \ (ʔ)$

Deletion rule : $ʔ \rightarrow \emptyset / \_ {}_{\sigma}[$

Alternatively, we might posit a morphemic tier within the autosegmental framework and locate the glottal stop at this level. The domain of the glottal stop could then be posited as being the morpheme, which would account for its underlying distribution in both lexical roots and affixes, and the deletion rule stated above would account for its absence word-internally. Appealing as this approach may be, it is not known how it could be accommodated in a statement of phonotactic distribution in a way that is useful to this description; prosody and morphology are assumed to constitute independent domains, but the glottal stop of WM appears to straddle both.

2.8 Analysis of syllable-final stops

The analysis of C_2 syllable-closing consonants in WM can be approached from a number of different theoretical directions, of which two will be appraised here. The first we will consider is the American structuralist position; this assumes that the unaspirated stop phonemes may occur at either syllable margin, but that the aspirated stop phonemes have a defective distribution and are restricted to syllable onsets. Aspirated and unaspirated stops are therefore only in contrastive distribution in syllable onset position. Inherent in this analysis is the notion that an unaspirated stop that occurs syllable-initially is phonemically the same segment that occurs syllable-finally.

An alternative analysis of WM stops can be based upon the Praguean concept of neutralisation and the archiphoneme, as expounded principally by Trubetzkoy (1977). This would consider the onset a position of relevance and the coda a position of neutralisation, because the feature [aspiration], responsible for the privative opposition between pairs of stop onsets at the bilabial, dental and velar places of articulation, is neutralised syllable-finally.

To illustrate by example, in this framework a WM bilabial stop occurring in the syllable-final position of neutralisation would not be analysed as an allophone of either an aspirated or an unaspirated bilabial stop phoneme, but rather as the representative of an archiphoneme /P/. This is held to be a phonological unit distinct from the member stop phonemes /p/ and /p^h/ that occur in phonological opposition in the syllable-initial position of relevance. The archiphoneme constitutes the lowest common denominator of shared features of the member phonemes in a neutralisable opposition. Thus in this case, /P/ is reducible to [–nasal], [–continuant], [+anterior] and [–coronal].

In a discussion of neutralisation and the archiphoneme, Lass (1984:49–50) summarises four of five types of neutralisation schematically as shown at Figure 2.4 below (his fifth type deals with gradual oppositions and is therefore irrelevant to the discussion at hand). Lass' four types of neutralisation are: type (i), in which only one member of the neutralised opposition appears; type (ii), in which both members of the neutralised opposition appear, but are conditioned by particular neutralisation environments; type (iii), in which the segment appearing in the position of neutralisation has properties of both members of the opposition, plus some of its own; and type (iv), in which either member of the opposition may appear non-contrastively in the same position of neutralisation.

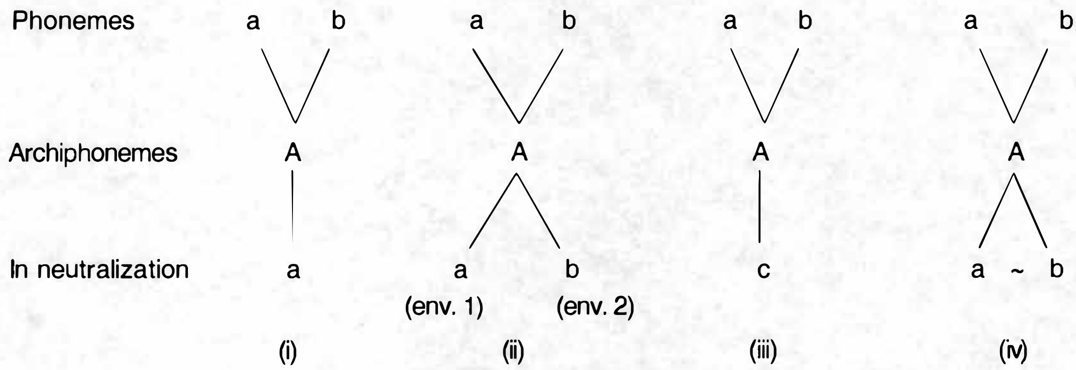


Figure 2.4: Four types of neutralisation as proposed by Lass (1984:49–50)

To the sceptical observer viewing the framework of neutralisation and the archiphoneme from outside the paradigm, in at least one case it might appear that metaphorically, the emperor is wearing no clothes. I refer specifically to Lass' schematised type (i), which he suggests is perhaps the commonest of neutralisation types '... where one member of the neutralized opposition appears to the complete exclusion of the other' (Lass 1984:49). Overlooking the differences in terminology, type (i) does not appear to diverge at all from the American structuralist analysis described above, because the so-called representative of the archiphoneme in the position of neutralisation is indistinguishable from one of the phonemes in the position of relevance. Trubetzkoy is perhaps somewhat misrepresented by Lass here, because rather than one member of the neutralised opposition occurring, what is realised is the sum of relevant features common to the phonemes in neutralisable opposition, i.e. the lowest common denominator, which may or may not necessarily include all the features of one of the members of the neutralisable opposition. The distinction is a fine one, but should nevertheless be acknowledged.

If the putative archiphoneme representatives in WM were always realised as voiceless unaspirated stops in the position of neutralisation, as represented by type (i) above, then one may be led to question the benefit of introducing an additional complexity to the analysis for relatively little gain over the more descriptively economical approach of American structuralism, which assumes that the unaspirated stop phonemes /p, t, k/ may appear at either syllable margin.

The alternative view is that the theory of neutralisation and the archiphoneme is seen to provide a more comprehensive description of the phonotactic distribution of stops than the American structuralist approach because it is able to capture generalisations that would otherwise be obscured by statements of allophonic realisations. The inevitable disadvantage is that it increases the size of the phoneme inventory by three segments, viz. /P, T, K/. It also introduces an added level of complexity to the description at the sacrifice of descriptive simplicity and economy. It is questionable that one approach can be claimed to be superior to the other, as there are no really compelling arguments either way. This book therefore adopts the American structuralist approach, with greater convenience of typing phonemic transcriptions being the trivial motivation.

This concludes the description of the syllable and the phonotactics of WM. In the following chapter I describe the segmental phonology, discuss phonation types and sociolinguistic variation, and propose an orthography for writing the Mongsen dialect of Ao.

3 *Segmental phonology*

3.1 Introduction

This chapter describes the segmental phonology of WM. It is structured as follows. Sections 3.2–3.3 describe the consonant phonemes and their allophonic realisations and discuss the consonant system typology. Sections 3.4–3.5 give a description of the vowel phonemes and their allophonic realisations. Approaches to the analysis of phonation contrasts are discussed in §3.6. Types of linguistic variation in WM are dealt with in §3.7. A tentative distinctive features matrix is suggested in §3.8, and the chapter concludes with a discussion of a practical orthography in §3.9.

3.2 Consonant phonemes

There are twenty consonant phonemes in WM. These are listed in Table 3.1 according to place and manner of articulation.

Table 3.1: Consonant phonemes

	Bilabial	Dental	Post-alveolar	Palatal/ Pal-alv	Velar	Glottal
Stop						
unaspirated	p	t			k	
aspirated	pʰ	tʰ			kʰ	
Affricate						
unaspirated		ts		tʃ		
aspirated		tsʰ		tʃʰ		
Fricative						
voiceless		s				h
voiced		z				
Nasal	m	n			ŋ	
Approximant						
lateral		l				
central	w		ɹ	j		

The following (sub-)minimal pairs of words demonstrate aspiration contrasts between stops occurring word-initially and word-medially at the bilabial, dental and velar places of articulation:

/p/ versus /pʰ/

/pā/	[pa^{33}]	's/he'
/pʰāŋā/	[pʰa^{33}ŋa^{33}]	'five'
/tā-pāʔ/	[tə^{33}paʔ˺^{33}]	'NPF-father'
/tā-pʰā/	[tə^{33}pʰa^{33}]	'NPF-tooth'

/t/ versus /tʰ/

/tā-āp/	[təp^{33}]	'NPF-rotten'
/tʰə̀p/	[tʰəp^{11}]	'throw.at.PAST'
/mātət/	[mə^{33}tət^{33}]	'know.PAST'
/tā-tʰəm/	[tə^{33}tʰəm^{33}]	'NPF-end'

/k/ versus /kʰ/

/kə̀p/	[kəp^{11}]	'shoot.PAST'
/kʰə̄p/	[kʰəp^{33}]	'depart.PAST'
/tā-kūŋ/	[tə^{33}koŋ^{33}]	'NPF-dry'
/tā-kʰūŋ/	[tə^{33}kʰoŋ^{33}]	'NPF-neck'

The following sub-minimal pairs of words demonstrate aspiration contrasts for dental and palato-alveolar affricates occurring word-initially and word-medially:

/ts/ versus /tsʰ/

/tsə̀ŋ/	[tsəŋ^{11}]	'bark.PAST'
/tsʰə́ŋ/	[ts:əŋ^{55}]	'punch.PAST'
/tā-mātsā/	[tə^{33}mə^{33}tsə^{33}]	'NPF-wet'
/tā-tsʰā/	[tə^{33}ts:ə^{33}]	'NPF-short'

/tʃ/ versus /tʃʰ/

/tʃə̀p/	[tʃəp^{11}]	'weep.PAST'
/tʃʰə̀t/	[tʃʰət^{11}]	'rub.PAST'
/ātʃī/	[a^{33}tʃi^{33}]	'fall.PAST'
/átʃʰì/	[a^{55}tʃ:i^{11}]	'sneeze.PAST'

The following sub-minimal pairs of words demonstrate phonemic contrasts in place of articulation between dental and palato-alveolar affricates occurring word-initially and word-medially:

/ts/ versus /tʃ/

/tsə̀làʔ/	[tsə^{11}la̠ʔ˺^{11}]	'earthworm'
/tʃə̀lì/	[tʃə^{11}li^{11}]	'walk.PAST'
/ātsā/	[a^{33}tsə^{33}]	'look.PAST'
/ātʃāt/	[a^{33}tʃət^{33}]	'squeeze.PAST'

/tsʰ/ versus /tʃʰ/

/tsʰə̀k/	[ts:ək^{11}]	'collide.PAST'
/tʃʰə̀t/	[tʃʰət^{11}]	'rub.PAST'

/mə̀-tsʰə̀-là/	[mə^{11}tsːə^{11}la^{11}]	'NEG-pull-NEG.PAST'
/mə̀-tʃʰə̀t-là/	[mə^{11}tʃʰət^{11}la^{11}]	'NEG-rub-NEG.PAST'

The following sub-minimal pairs demonstrate phonemic voicing contrasts between voiced and voiceless dental fricatives occurring word-initially and word-medially:

/s/ versus /z/

/sə̄ŋ/	[səŋ^{33}]	'fill.PAST'
/zə̀ŋ/	[zəŋ^{11}]	'count.PAST'
/tə̄-mə̄sə̄ŋ/	[tə^{33}mə^{33}səŋ^{33}]	'NPF-white'
/tə̄-mə̄zə̄ŋ/	[tə^{33}mə^{33}zəŋ^{33}]	'NPF-claw'; 'NPF-fingernail'

The following sub-minimal pairs demonstrate a phonemic contrast between the dental and velar nasals word-initially, word-medially and word-finally:

/n/ versus /ŋ/

/nūklàŋ/	[nuk^{33}laŋ^{11}]	'hundred'
/ŋùʔ/	[ŋūʔ˺^{11}]	'bite.PAST'
/mə̄nī/	[mə^{33}ni^{33}]	'laugh.PAST'; 'smile.PAST'
/tsə̀ŋī/	[tsə^{11}ŋi^{33}]	'sun'
/tə̄-sə̄n/	[tə^{33}sən^{33}]	'NPF-new'
/sə̄ŋ/	[səŋ^{33}]	'fill.PAST'

The following sub-minimal pairs demonstrate a phonemic contrast between the post-alveolar and lateral approximants word-initially and word-medially:

/ɹ/ versus /l/

/ɹà/	[ɻa^{11}]	'come.PAST'
/lā/	[la^{33}]	'she'[1]
/tə̄-ɹə̄m/	[tə^{33}ɹəm^{33}]	'NPF-middle'
/tə̄-lə̄m/	[tə^{33}ləm^{33}]	'NPF-head'; 'NPF-warm'

3.3 Allophonic realisations of consonant phonemes

This section gives a description of the consonant phonemes and their allophonic realisations, with examples of the phonemes in words.

3.3.1 Stops

The stop series demonstrates three distinctive places of articulation with aspiration contrasts at each place. An instrumental investigation of voice onset time, which is the

1 This pronoun is rarely encountered in spoken language, as /pā/ serves for both masculine and feminine third person singular. It was reported to me that the third person feminine pronoun /lā/ is an innovation introduced by a bible translator to disambiguate pronominal reference in the Chungli Bible. Its source is the suffix /-lā/, a semantic gender marker usually occurring on female names.

acoustic correlate of aspiration, was done on utterance-initial stops. The results are presented in §4.2.

Syllable-final stops are unreleased before obstruents word-medially but are generally released word-finally. Substitution tests confirm that informants also accept unreleased stops word-finally, although these are much less frequently encountered in words uttered in isolation. An unaspirated bilabial stop /p/ may be realised intervocalically as a voiced bilabial stop [b] or as a voiced bilabial fricative [β] in allegro speech, but analogous realisations are not observed for the other unaspirated stops. This is explained by the fact that the lips, cheeks and other soft tissue of the oropharyngeal anatomy allow for a substantial amount of expansion behind the bilabial obstruction, thereby accommodating simultaneous glottal vibration to a perceptually greater extent than that possible for the other oral stops.

/p/ unaspirated bilabial stop.

→[b]~[β]	/ V_V free variation in allegro speech
→[p˺]	/ _]$_\sigma$C [+obstr]
→[p]~[p˺]	/ _# in free variation
→[p]	

examples:

/pá-tsə́ kū/	[pa^{55}tsə^{55}ko^{33}]	'there' (SPEC-DIST + LOC)
/tà tʃʰà-pàʔ-nā/	[ta^{11} tʃʰa^{11}ba^{11}na^{33}]	'because' (thus + do-NML-INST)
/tə̀psə̄t/	[təp˺^{11}sət^{33}]	'kill.PAST'
/pāŋkə̄p/	[paŋ^{33}kəp^{33}]	'lip'

/pʰ/ aspirated bilabial stop

→[pʰ]

examples:

/pʰə̄lī/	[pʰə^{33}li^{33}]	'four'
/ūpʰūŋ/	[u^{33}pʰuŋ^{33}]	'dust'
/tsʰə́ŋpʰáŋ/	[tsːəŋ^{55}pʰaŋ^{55}]	'float.PAST'

The dental stops (as well as other sounds involving dental articulation) might be more accurately characterised as being denti-alveolar, because native speakers describe apical or laminal contact on the teeth accompanied by laminal contact on the area between the alveolar ridge and the point where the upper incisors insert into the gingiva during their articulation of these segments.

/t/ unaspirated laminal dental stop

→[t˺]	/ _]$_\sigma$C [+obstr]
→[t]~[t˺]	/ _# in free variation
→[t]	

examples:

/tə̄lə̄m/	[tə^{33}ləm^{33}]	'warm'; 'head'
/mə̄tūŋ/	[mə^{33}tuŋ^{33}]	'straight.PAST'

/tə̀psə̄t-pàʔ/	[təp^{11}sət$^{˺33}$paʔ^{11}]	'kill-NML'
/ə̀mə̄t/	[ə^{11}mət^{33}]	'hold.PAST'

/t^{h}/ aspirated laminal dental stop

→[t^{h}]

examples:

/t^{h}ənī/	[t^{h}ə^{33}ni^{33}]	'seven'
/át^{h}úʔ/	[a^{55}t^{h}u̯ʔ$^{˺55}$]	'vomit.PAST'
/ānt^{h}ān/	[ən^{33}t^{h}ən^{33}]	'gather.PAST'

/k/ unaspirated dorsal velar stop

→[k˺]	/ _]$_{\sigma}$ C [+obstr]
→[k]~[k˺]	/ _# in free variation
→[k]	

examples:

/kə̄nə̄t/	[kə^{33}nət^{33}]	'we.two' (1DU.EXC)
/tə̄-mə̄kūk/	[tə^{33}mə^{33}kuk^{33}]	'NPF-knee'
/pùkts^{h}ə̄ŋ/	[puk$^{˺11}$ts:əŋ^{33}]	'guts'

/k^{h}/ aspirated dorsal velar stop

→[k^{h}]

examples:

/k^{h}ìʔ/	[k^{h}i̯ʔ$^{˺11}$]	'give.PAST'
/mə̄k^{h}ʉ̄lī/	[mə^{33}k^{h}ʉ^{33}li^{33}]	'smoke'
/nə̀ŋk^{h}ə̀lā/	[nəŋ^{11}k^{h}ə^{11}la^{33}]	'you.all' (2PL)

3.3.2 Affricates

The affricates occur at two distinctive places of articulation and demonstrate an aspiration contrast at each place, in common with the stop series. It is observed, however, that the nature of the phonetic realisation of the aspirated affricates can differ substantially from that of the aspirated stops.

Briefly, what is heard as contrastive aspiration occurring after the release of the aspirated dental affricate is consistently demonstrated in spectrograms to be prolonged noise. The phonemic contrast between the voiceless unaspirated dental affricate /ts/ and voiceless aspirated dental affricate /ts^{h}/ is therefore located in the duration of the strident release; that is to say, /ts^{h}/ is realised phonetically as [ts:]. An analogous situation is not consistently demonstrated to be the basis of the contrast between /tʃ/ and /tʃ^{h}/, except when /tʃ^{h}/ occurs before a high front vowel. In this particular environment it is predictable that the aspirated palato-alveolar affricate will be realised as [tʃ:], and elsewhere it will be realised as [tʃ^{h}]. A detailed description of the acoustic characteristics of aspiration in affricates is presented together with a discussion of voice onset time in §4.3.

The affricate series is analysed as a separate sub-system from the stop series, taking into account the following considerations. Firstly, while the passive articulator of the palato-alveolar affricates /tʃ, tʃ^h/ can easily be accommodated as a distinctive place of articulation in the stop series, a problem for the tabulation of phonemes according to traditional place of articulation arises when the dental affricates /ts, ts^h/ are included in the same series because the passive articulator for both the stop onsets of the affricates /ts, ts^h/ and the stops /t, t^h/ is assumed to be the dental place of articulation. In the absence of palatographic proof demonstrating evidence to the contrary, an analysis identifying two putative stops at the same place of articulation has less appeal than one which identifies a separate series for each of the segments. Addressing this problem, Lass (1984:148) puts the question 'if a language has /t/ and /ts/, is the latter a 'position' or a 'manner'?' He resolves this by assuming that if an affricate and a stop share the same place of articulation, then the affricate is taken to be a member of a different sub-system. Lass' suggestion is adopted here.

Secondly, phonotactic constraints on what may constitute the coda of the syllable might also be used to argue against conflating the affricates and stops in a single series. If we adopt the American Structuralist position outlined in Chapter 2, aspiration is analysed as being a distinctive feature for WM stops when they occur in syllable onset position, but a redundant feature when they occur syllable-finally. Since the unaspirated affricates /ts, tʃ/ differ from the unaspirated stop series /p, t, k/ in not being able to occur syllable finally, on distributional grounds we must therefore conclude that the affricates belong to a different sub-system.

Thirdly, whereas the bilabial, dental and velar stops have corresponding nasal phonemes occurring at each of these distinctive places of articulation, a palatal nasal coinciding with the palato-alveolar place of articulation for affricates is lacking. To sum up, the only evidence in favour of a conflation of stops and affricates is the fact that both series demonstrate aspiration contrasts, and both share the feature [–continuant]. However if symmetry of the phonemic system is held to be prior, then weight of numbers argues in favour of establishing an affricate series separate from the stop series.

/ts/ unaspirated laminal dental affricate

→[ts]

examples:

/tsə̀sə̀/	[tsə^{11}sə^{11}]	'fog'
/ə̄ntsə̄/	[ən^{33}tsə^{33}]	'egg'
/tsìjə̄m/	[tsi^{11}jəm^{33}]	'sea'

/ts^h/ aspirated laminal dental affricate

→[tsː]

examples:

/ts^{h}ə̀ʔ/	[tsːə̠ʔ˺^{11}]	'pull.PAST'
/mə̄ts^{h}ə̄/	[mə^{33}tsːə^{33}]	'knot.PAST'

The dental affricates of WM might be more accurately described as having a laminal denti-alveolar articulation, in common with the dental stops described above. For the sake of phonological symmetry, however, these affricate segments will be treated as occurring at the dental place of articulation.

/tʃ/ unaspirated laminal palato-alveolar affricate

→[tʃ]

examples:

/tʃàʔ/	[tʃa̰ʔ˺^{11}]	'eat.PAST'
/tə̄-tʃə̄n/	[tə^{33}tʃən^{33}]	'NPF-old'
/sə̀ŋtʃàŋ/	[səŋ^{11}tʃaŋ^{11}]	'fruit'

/tʃʰ/ aspirated laminal palato-alveolar affricate

→[tʃː] / _ i

→[tʃʰ]

examples:

/tʃʰìpà/	[tʃːi^{11}pa^{11}]	'fear.PAST'
/mə́tʃʰáp/	[mə^{55}tʃʰap^{55}]	'suck.PAST'
/tʃʰə́m/	[tʃʰəm^{55}]	'wear.PAST'

Two pairs of words in the corpus possibly demonstrate age-related variation between the choice of dental and palato-alveolar affricates. A discussion of linguistic variation concerning affricates is presented in §3.7.1 below.

3.3.3 Fricatives

The fricative series demonstrates two distinctive places of articulation, viz. dental and glottal, and has a voicing contrast at the dental place of articulation. Without the benefit of palatographic studies it is impossible to state with confidence whether the dental fricatives have a denti-alveolar passive articulator, or if the dominant place of articulation is confined to the alveolar area. Once again, for the sake of phonological symmetry the segments /s, z/ are recognised in this description as dental fricatives. A similar problem arises with the identification of the active articulator for these segments, and consequently no assertions will be made. Ladefoged and Maddieson (1996:146) cite studies which find that even for a language as well-studied as English there is considerable disagreement as to what is the most common articulation for the sibilant fricative, and that individual anatomical differences may have a bearing on whether an apical or a laminal articulation is used.

/z/ voiced dental fricative

→[z]

examples:

/zūktʃūk/	[zuk˺^{33}tʃuk^{33}]	'wipe.PAST'
/mə̄zə̄ʔ/	[mə^{33}zəʔ^{33}]	'fire'
/lūŋzāk/	[luŋ^{33}zak^{33}]	'whetstone'

/s/ voiceless dental fricative

→[ʃ] / _i
→[s]

The phones [s] and [ʃ] occur in complementary distribution; the voiceless dental fricative /s/ is palatalised to [ʃ] when occurring before a high front vowel and is realised as [s] elsewhere.

examples:

/sə̄/	[sə^{33}]	'die.PAST'
/á-sáʔ/	[a^{55}saʔ˺^{55}]	'NPF-meat'
/sə̄nsī/	[sən^{33}ʃi^{33}]	'breathe.PAST'
/sītāk/	[ʃi^{33}tak^{33}]	'correct'

The traditional but inaccurate term 'glottal fricative' is conveniently used as a label for the segment /h/; its place and manner of articulation values can largely be left unspecified, since the use of the feature [–oral] minimally differentiates it from all other WM segmental phonemes.

/h/ voiceless glottal fricative

→[h]

examples:

/hʉ̄/	[hʉ^{33}]	'blow.PAST'
/āhʉ̄/	[a^{33}hʉ^{33}]	'snake'
/hàiʔ/	[hai̯ʔ˺^{11}]	'OK'
/hūŋ/	[huŋ^{33}]	'leap.PAST'

The glottal fricative carries a relatively low functional load. Marrison (1967, Vol. I:58) claims that a phoneme /h/ does not occur in Ao. Presumably he is referring specifically to the Chungli dialect and not more generally to all varieties of Ao because he does list some orthographic Mongsen entries containing this segment, such as *huntep* 'angry' (Vol II, 1967:9), *ahü* 'mouse' (Vol. II, 1967:169) and *hau* 'yes' (Vol. II, 1967:308), presumably gathered from the Longchang village Mongsen data of Mills (n.d. Typescript).

The voiced dental fricative /z/ and the palatal approximant /j/ demonstrate at first blush what appears to be free variation when occurring medially in a substantial number of words. This is explored thoroughly in §3.7.2, where I use comparative evidence to argue the case for borrowing between dialects as an explanation for what might otherwise appear to be *bona fide* free variation. Borrowing between dialects is also offered as an explanation in §3.7.3 for the rare occurrence of a voiceless labio-dental fricative [f] in place of a glottal fricative in the data of two native speakers of WM.

3.3.4 Nasals

The nasal phonemes occur at the bilabial, dental and velar places of articulation.

/m/ voiced bilabial nasal

→[m]

examples:

/mūŋ/	[mõŋ^{33}]	'wind'
/jìmāŋ/	[ji^{11}maŋ^{33}]	'road'
/ā-sə̄m/	[a^{33}səm^{33}]	'NPF-three'

/n/ voiced laminal dental nasal

→[n]

examples:

/nàŋ/	[naŋ^{11}]	'you' (SG)
/tə́-nák/	[tə^{55}nak^{55}]	'NPF-black'
/tə̄-sə̄n/	[tə^{33}sən^{33}]	'NPF-new'

/ŋ/ voiced dorsal velar nasal

→[ŋ]

examples:

/ŋùʔ/	[ŋuʔ˺^{11}]	'bite.PAST'
/á-ŋáʔ/	[a^{55}ŋaʔ˺^{55}]	'NPF-fish'
/áwátsə́ŋ/	[a^{55}wa^{55}tsəŋ^{55}]	'lake'

3.3.5 Approximants

WM has one lateral approximant phoneme and three central approximant phonemes. The articulation of the lateral approximant is probably apical alveolar, but will be treated phonologically as dental to accord with the phonemes of other series having a dental articulation.

/l/ voiced dental lateral approximant

→[l]

examples:

/làtà/	[la^{11}ta^{11}]	'moon'
/á-lí/	[a^{55}li^{55}]	'NPF-ground'

/ɹ/ voiced apical post-alveolar central approximant

→[ɹ]~[ɻ]~[ɹ̠] in free variation

examples:

/ɹūŋ/	[ɻoŋ^{33}]	'burn.PAST'
/tə̄-ɹə̄t/	[tə^{33}ɹət^{33}]	'NPF-bone'
/sàɹùwāɹ/	[sa^{11}ɻu^{11}waɻ^{33}]	'animal'

The apical post-alveolar approximant [ɹ] occurs in free variation with a sub-apical retroflex approximant [ɻ] for most speakers. Bidialectal speakers of the Chungli and Mongsen dialects of Ao have provided examples to demonstrate that a voiced post-alveolar approximant also occurs in the Chungli dialect. Gowda (1975:14) identifies a voiced retroflex lateral fricative phoneme in Chungli, a sound previously unreported in any of the world's languages. I have not recognised this sound in either of the two main dialects of Ao, and bidialectal speakers uniformly reject Gowda's example words in which I swap a voiced apical post-alveolar approximant for a voiced retroflex fricative with a lateral release. Some speakers do optionally articulate this sound with a fricative component producing a rhotic that can be phonetically realised as [ɹ̝], but the release remains consistently central for all variants of /ɹ/.

/j/ voiced palatal approximant

→[j]

examples:

/jīm/	[jim^{33}]	'flow.PAST'; 'fly.PAST'
/ā-jā/	[a^{33}ja^{33}]	'NPF-night'

/w/ voiced labial-velar approximant

→[w]

examples:

/wàjàʔ/	[wa^{11}ja̰ʔ$^{˺11}$]	'bird'
/sə̀ŋàwà/	[sə^{11}ŋa^{11}wa^{11}]	'leaf'

3.3.6 Consonant system typology

From a typological perspective, the phonological system of consonant segments tends to conform to dominant cross-linguistic patterns of occurrence identified in typological surveys by Nartey (1979), Maddieson (1984) and others.

WM uses aspiration as a contrastive feature for pairs of segments occurring at each distinctive place of articulation in the stop and affricate series. The nasal series matches the stop series for the bilabial, dental and velar places of articulation, the prevalent pattern in the languages of the world. The voiced approximants, which can be divided into liquids and semi-vowels, also conform to typologically dominant patterns. The only series of consonants that is not found to be typologically consistent is the fricative series. The following comments are restricted to the dental fricatives only, as the glottal fricative /h/ was not treated as a true fricative in the above-mentioned surveys.

The survey of Maddieson (1984) finds that a language is highly likely to have some form of voiceless dental or alveolar fricative. Because only three of the sixty-two languages in the survey with just two fricatives had a contrast between /s/ and /z/, Maddieson (1984:52–53) surmises that languages with only two fricatives tend to avoid a voicing contrast at the same place of articulation. The fricative series of WM is thus typologically rare.

3.4 Vowel phonemes

There are six monophthongal phonemes in WM: the five modal voice phonemes /i, ʉ, u, ə, a/, and the creaky voice phoneme /a̰/. The vowel phonemes are displayed below in Table 3.2.

Table 3.2: Vowel phonemes

	Front	Central		Back
	–round	**–round**	**+round**	**+round**
High	i		ʉ	u
Mid		ə		
Low		a a̰		

Perhaps with great prescience, Trubetzkoy (1977:86–114) proposed three types of property to distinguish vowels perceptually — these are *Öffnungsgradeigenschaften* (properties relating to degrees of aperture: this is equivalent to height, the perceptual correlate of F_1); *Lokalisierungseigenschaften* (properties relating to the position of the articulators, or degrees of localisation: this is equivalent to backness and rounding, the perceptual correlates of F_2); and *Resonanzeigenschaften* (properties related to resonance: this differentiates nasalised vowels from oral vowels). The first two properties serve as coordinates for the description of vowel quality and are used to identify three basic vowel systems (linear, quadrangular and triangular) according to their distinctive characteristics.

Using the parameters of Trubetzkoy's approach to the typology of vowel systems, the vowel phoneme inventory of WM demonstrates three degrees of aperture and three degrees of localisation producing a triangular system of five distinctive segments with modal voice phonation. The five-vowel system is augmented by a low central vowel with contrastive creaky voice, thus adding a sixth distinctive segment /a̰/ to the vowel phoneme inventory.

The following (sub-) minimal pairs of words demonstrate phonemic contrasts between phonetically similar vowels:

/u/ versus /ʉ/

/mījū/	[mi^{33}ju^{33}]	'clean.PAST'
/mījʉ̄/	[mi^{33}jʉ^{33}]	'grow.PAST'

/ə/ versus /ʉ/

/nʉ̀/	[nʉ^{11}]	'lead.PAST'
/nə̄/	[nə^{33}]	'your'

/ə/ versus /a/

/ə̄ɹ/	[əɻ^{33}]	'sew.PAST'
/āɹ/	[aɻ^{33}]	'poisonous creeper'

/ə/ versus /i/

/típ/	[tip^{55}]	'edge'; 'side'
/tə̀p/	[təp^{11}]	'hit.PAST'

/a/ versus /a̰/

/wāŋ/	[waŋ^{33}]	'go-IMPER'
/wā̰ŋ/	[wa̰ŋ^{33}]	'slice-IMPER'

3.5 Allophonic realisations of vowel phonemes

This section gives a description of the vowel phonemes and their allophonic realisations, with examples of the phonemes in words.

3.5.1 Modal voice vowels

There are five contrastive modal voice vowels. Generally, all modal voice vowels have creaky voice free variants when occurring before a glottal stop. Allophones of the back rounded vowel /u/ are nasalised when occurring either before or after a velar nasal consonant.

/i/

→[ḭ]~[i]	/ _ ʔ	in free variation
→[e]~[i]	/ l _	in free variation (for one of the consultants)
→[i]		

examples:

/īpī/	[i³³pi³³]	'sleep.PAST'
/tə́-jíʔ/	[tiʔ˺⁵⁵]	'NPF-horn'
/kʰìʔ/	[kʰḭʔ˺¹¹]	'give.PAST'
/lī/	[li³³]~[le³³]	DAT

/ʉ/

→[ʉ]

examples:

/tʉ̄/	[tʉ³³]	'mother' (Aier clan)
/nə̄m-ʉ̄-pāʔ/	[nəm³³ʉ³³paʔ³³]	'compress-IRR-NML'
/tə́-nʉ́ʔ/	[tə⁵⁵nʉʔ˺⁵⁵]	'NPF-wife'
/pʉ̄/	[pʉ³³]	'that'

The high central rounded vowel /ʉ/ is produced with rather pronounced rounding and moderate protrusion of the lips. Perceptually the vowel has a central quality with respect to backness. Preliminary observations suggest that this phoneme is an innovation confined to varieties of Mongsen spoken in just a few villages located on the Changkikong range. It has the lightest functional load of the modal voice vowels and is found to correspond with /i/ in cognate words of other Mongsen varieties, with the greatest frequency of correspondence seen in grammatical morphemes.

/u/

→[õ]~[ũ]	/(ŋ) _ (ŋ)	
→[ṵ]~[u]	/ _ ʔ	in free variation
→[u]~[o]	in free variation	

examples:

/tə̄-kūŋ/	[tə^{33}kõŋ^{33}]	'NPF-dry'
/ūpʰūŋ/	[u^{33}pʰūŋ^{33}]	'dust'
/tùʔ/	[tu̠ʔ˺^{11}]	'dig.PAST'
/ŋùʔ/	[ŋṵ̃ʔ˺^{11}]	'bite.PAST'
/tʃā-ɹ-ūʔ/	[tʃa^{33}ɹoʔ˺^{33}]	'call-PRES-DEC'
/úpú/	[u^{55}pu^{55}]	'ashes'
/tsə̀ŋ̀ìkū/	[tsə^{11}ŋi^{11}ko^{33}]	'day'

The phones [u] and [o] may occur in free variation in the same words in the corpus and are therefore treated as allophones of a single phoneme /u/, in free variation. A vowel formant plot to be presented in Chapter 4 suggests that the phonetic focus of this phoneme is located between Cardinal Vowel Nos. (7) and (8), although its allophonic scope is observed to encroach upon both vowel spaces. The overwhelming majority of occurrences of [o] are in the vicinity of a velar consonant, yet the velar environment is not found to be mutually exclusive to [u]; for example, [ŋṵ̃ʔ˺^{11}] 'bite.PAST', or [tə^{33}puk^{33}] 'belly'. Shafer (1974:302fn.) also observes frequent fluctuation between [u] and [o] in all Ao dialects and concludes that they are probably allophones of a single back vowel constituting one phoneme. The absence of any words in the WM corpus that contrast on the basis of a phonemic distinction between /u/ and putative /o/ supports his observation, and the results of substitution tests done with a native speaker (MA) in which [o] was swapped for [u] in words without a change in meaning provide conclusive evidence that [o] and [u] are indeed non-contrastive.

/ə/

→[ə̠]~[ə]	/ _ ʔ	in free variation
→[ə]		

examples:

/ə̄ɹ/	[əɻ^{33}]	'sew.PAST'
/tə̄-ɹə̄t/	[tə^{33}ɹət^{33}]	'NPF-bone'
/lísə́/	[li^{55}sə^{55}]	'rope'

The schwa is articulated with the lips slightly spread. Perceptually it has a central quality with respect to both height and backness.

/a/

→[a̠]~[a]	/ _ ʔ	in free variation
→[a]		

examples:

/āɹū/	[a^{33}ɹu^{33}]	'good.PAST'
/tʃàʔ/	[tʃa̠ʔ^{11}]	'eat.PAST'
/tə̄-āpā/	[ta^{33}pa^{33}]	'NPF-thin'

3.5.2 Creaky voice vowel

The creaky voice vowel /a̰/ is analysed as a separate vowel phoneme. Creaky voice is not found to occur contrastively on any other vowels, therefore the vowel system cannot be treated as having two distinct phonation settings — modal and creaky voice — applying contrastively to every vowel quality. I will demonstrate with evidence from the corpus that creaky voice cannot be treated as a separate toneme, nor as a separate syllable- or word-based prosody; nor can it be analysed as a segmental realisation of a glottal stop. Thus the only remaining avenue is to analyse /a̰/ as a vowel phoneme in contrastive distribution with the five modal voice vowels.

/a̰/
→[a̰]

examples:

/wà̰pə̀t/	[wa̰^{11}pət^{11}]	'slope'
/səsā̰-pàʔ/	[sə^{33}sa̰^{33}pa̰ʔ^{11}]	'tear-NML'
/wà̰pàʔ/	[wa̰^{11}pa̰ʔ^{11}]	'slice-NML'

3.6 Analysis of the creaky voice vowel

Creaky voice (also known as laryngealisation, vocal fry, glottal fry) is used contrastively on the low central vowel. This produces two vowel phonemes, /a/ and /a̰/, at one articulatory position. Firstly I will briefly review the literature on creaky voice phonation and, where appropriate, make correlations to the creaky voice encountered in WM, after which I will present evidence in support of an analysis of /a̰/ as a contrastive vowel phoneme.

Ladefoged (1971:14–15) describes laryngealisation as a type of phonation in which '... the arytenoid cartilages are pressed inward so that the posterior portions of the vocal cords are held together and only the anterior (ligamental) portions are able to vibrate'. Laver (1980:126) makes finer distinctions within this phonation type, observing that Ladefoged's description of laryngealisation is more in keeping with his (Laver's) definition of 'creak', rather than of creaky voice. In contrast to Laver (1980) and Catford (1964:32), Ladefoged does not find it necessary to make a distinction between 'creak' and 'creaky voice', because a distinction between the two phonatory settings is not motivated for linguistic phonetics (Ladefoged 1971:15).

Laver (1980:122ff.) provides a comprehensive review of the literature on creak; briefly stated, the consensus of opinion is that (1) the fundamental frequency is below 100 Hz, (2) there is strong adductive tension and medial compression of the vocal cords, (3) only the ligamental glottis is involved in periodic vibration, (4) the vocal tract is damped between glottal pulsations, perhaps by the ventricular folds making contact with the surfaces of the vocal folds, and (5) sub-glottal air pressure is lower than for modal voice phonation.

My auditory impression of creaky voice in WM is that it has two components: a modal voice component produced simultaneously with a sound that Catford (1964:32) describes as having an auditory effect akin to 'a rapid series of taps, like a stick being run along a railing'. I have also compared WM creaky voice phonation to a range of demonstration phonation types produced by Laver on the audio cassette 'Voice Quality Performance' (1976), and I find that WM creaky voice is perceptually most similar to his example of creaky voice phonation.

The creaky voice of WM appears to be consistent with a number of the above-mentioned descriptions of features proposed in the literature, with one exception that I am able to quantify — it is not restricted to occurring at fundamental frequencies below 100 Hz. The example words of (3.2) and (3.3), provided by CL, were found to have a fundamental frequency of approximately 116 Hz. CL's usual fundamental frequency range for mid tone on unstopped syllables is calculated to be 110 to 120 Hz.

(3.1)	/wà̰ʔ/	[wa̰ʔ̚^{11}]	'slice.PAST'
(3.2)	/wā̰-ŋ/	[wa̰ŋ^{33}]	'slice-IMPER
(3.3)	/wā̰-ʉ/	[wa̰ʉ̰^{33}]	'slice-IRR'

In (3.1), the verb root /wà̰ʔ/ has a glottal stop root-finally. The glottal stop, identified in Chapter 2 as a prosodic element constrained to the word-final position, might be assumed to provide the conditioning environment responsible for the realisation of creaky voice on the preceding vowel, were it not for the fact that it is deleted as a result of suffixation of verbal suffixes in (3.2) and (3.3). Note that its deletion does not result in a concomitant loss of the creaky voice on the vowels of these two words, as would be expected if creaky voice were simply conditioned by the following glottal stop.

Whether or not creaky voice can occur contrastively with the high tone is not as clear. As only one example occurred in the corpus, I hesitate to assert any claims on the basis of such limited evidence.

(3.4)	/āsá-wá̰ʔ/	[a^{33}sa^{55}wa̰ʔ̚^{55}]	'ADM-slice'

There is also a possibility that the creaky voice observed in (3.4) might be phonetically conditioned by the following glottal stop. Ideally, to rule out any phonetic segmental influence on voice quality we would require at least a handful of examples of similar syllables having a high tone and contrastive creaky voice on their vowels in the absence of a glottal stop environment. Even without such conclusive evidence, the retention of contrastive creaky voice in the words of (3.2) and (3.3) would seem to suggest that it is a parameter independent of both suprasegmental and segmental influence, since it remains in the presence of tone sandhi phenomena and in the absence of the glottal stop. All things being equal, because creaky voice is observed to occur contrastively on the vowel of a syllable with a mid tone after a sandhi effect has taken place, it seems rather unlikely that this contrastive function would be lost in the presence of a high tone.

An important distinction needs to be drawn between phonetic realisations of creaky voice that result from segmental conditioning environments, and creaky voice that is unequivocally contrastive. Creaky voice is often audible on the vowels of some verb roots having a low tone and a final glottal stop, as in the following examples.

(3.5)	/kʰìʔ/	[kʰḭʔ̚^{11}]	'give.PAST'
(3.6)	/tʃàʔ/	[tʃa̰ʔ̚^{11}]	'eat.PAST'
(3.7)	/tsʰə̀ʔ/	[tsʰə̰ʔ̚^{11}]	'pull.PAST'

The creaky voice phonation on the vowels in the words of (3.5–3.7) is phonetically indistinguishable from that which occurs phonemically in the example of (3.1); nevertheless, it can be proven to be part of the realisation of the glottal stop and therefore non-contrastive by the addition of a suffix to the verb root, as demonstrated by (3.8–3.10) below. When a morpheme is suffixed to these verb roots, the word-final glottal stop is deleted and creaky voice can no longer be heard on the vowel. This indicates that the creaky voice phonation of these examples is attributable to the environment created by the word-final glottal stop and is therefore non-contrastive. This is a converse result to that demonstrated by the example words of (3.2) and (3.3) above, which undergo similar tone sandhi changes after suffixation, yet retain creaky voice on their respective vowels.

(3.8)	/kʰī-āɲ/	[kʰi^{33}aŋ^{33}]	'give-IMPER'
(3.9)	/tʃā-ɲ/	[tʃaŋ^{33}]	'eat-IMPER'
(3.10)	/tsʰə̄-āŋ/	[tsʰə^{33}aŋ^{33}]	'pull-IMPER'

Creaky voice also occurs on the vowels of some nouns with a word-final glottal stop, so a similar test to that used for verbs can be applied to nouns in order to determine whether the creaky voice phonation is predictable or contrastive. For test purposes, a type of possessive suffix /-ɹ/ can be suffixed to a noun whose root has a word-final glottal stop. Although this suffix is usually restricted to human possessors, a non-human possessor personified in the context of a fable was acceptable to a native speaker (MA) for the purpose of testing for contrastive creaky voice. If the creaky voice is predictable, deletion of the glottal stop under suffixation results in the concomitant loss of creaky voice phonation. This is proven to be the case in (3.11) below.

(3.11)	/tsə̀là?/	[tsə^{11}la̰ʔ˺^{11}]	'earthworm'
(3.12)	/tsə̀là-ɹ/	[tsə^{11}laɹ^{11}]	'earthworm-POSS'

Phonemic creaky voice carries an extremely light functional load in WM; when it is found to have contrastive function, it is solely on the low central vowel. To reiterate the main findings of this investigation, firstly there is no evidence to suggest that creaky voice should be posited as the realisation of a toneme separate from the three-level system used to signal lexical contrasts; nor can it be treated as part of an existing toneme because of its extremely limited distribution. Secondly, it cannot be treated as a distinct phonation type for the same reason. And lastly, it cannot be treated as a segmental realisation of the glottal stop, because it is attested in the absence of the glottal stop.

The fact that contrastive creaky voice occurs so infrequently in the corpus gives great cause for suspicion. Nevertheless, having explored all other avenues, there does not appear to be any justifiable alternative to the proposition that creaky voice be analysed as the somewhat idiosyncratic realisation of a separate vowel phoneme.

3.7 Linguistic variation

This section discusses linguistic variation encountered in the corpus and discusses how and why it may have arisen. Where it is warranted, I present comparative evidence in support of claims.

3.7.1 Variation between affricates

It is observed that variation occurs between native speakers in the pronunciation of the affricates /ts$^{\alpha asp}$/ and /tʃ$^{\alpha asp}$/ in two words, these being the only examples hitherto encountered.

	Older speakers			Younger speakers		Gloss
(3.13)	/tsʰə̄-āŋ/	[ts:ə^{33}aŋ^{33}]	(3.14)	/tʃʰī-āŋ/	[tʃ:i^{33}aŋ^{33}]	'take-IMPER'
(3.15)	/tsāʉ/	[tsaʉ̯^{33}]	(3.16)	/tʃāʉ/	[tʃaʉ̯^{33}]	'cooking pot'

The words of (3.13) and (3.15) above were uttered in texts narrated by two women in their sixties, while the words of (3.14) and (3.16) were given as the usual pronunciations of two younger informants (male and female, both in their early thirties) who helped in the translation of those texts and drew my attention to the linguistic variation. The shift from a dental place of articulation in (3.13) to a palato-alveolar place of articulation in (3.14) also causes a perturbation in the following vowel; that is, a schwa assimilates in height to a preceding palato-alveolar consonant.

It is the shift from a dental to a palato-alveolar place of articulation that causes the assimilation in height of the following vowel, and not a shift in the height of the vowel that results in the palatalisation of the affricate. The evidence for this sequence of events is provided by the examples of (3.15) and (3.16) above, which demonstrate that the change from a dental to a palato-alveolar place of articulation by the affricate may occur in the absence of a high front vowel. Substitution tests with a native speaker indicate that these segments cannot be freely substituted in any other words tested, and the existence of sub-minimal pairs such as /tsə̀làʔ/ 'earthworm' and /tʃə̀lì/ 'walk.PAST', or /tʃʰə̀t/ 'rub.PAST' and /tsʰə̀k/ 'collide.PAST' presents conclusive evidence that the dental and palato-alveolar affricates are in contrastive distribution.

One possible reading of this anomaly is that free variation exists between the dental and palato-alveolar affricate phonemes, but only in certain words. Lass (1984:21) provides examples of phonemes demonstrating free variation in English, such as /iː/ ~ /ɛ/ in *evolution*, or /iː/ ~ /ai/ in *either*, *neither*, and observes that it is possible for the distinctive function of a phonemic contrast (demonstrated by minimal pairs such as *beet:bet* and *beet:bite*) to be suspended in a marginal, non-systematic manner in some lexical items.

Another possible interpretation is that these two segments represent sociolinguistic variants, with age group being the primary determinant for the choice of a particular pronunciation. This claim is of course tentative and cannot be confirmed without further investigation and many more examples of variation, although there does appear to be a correlation between age group and the choice of variant for the two examples encountered.

Lastly, there is the possibility of an evolving sound change. While it would be rash to posit that a sound change might be under way in English on the basis of free variation encountered in the initial vowels of a few pairs of words such as those mentioned above, severe limitations on the phonetic environment in which a particular sound can be realised is reasonable cause for suspicion. In contrast to the English vowel phonemes /i/ and /ɛ/, which may occur in an unlimited variety of different environments, the WM dental affricates /ts, tsʰ/ are marked by their extremely limited distribution and the overwhelming majority of occurrences are before a schwa — in a corpus of over eight hundred lexical items, only three words demonstrate a dental affricate occurring before a vowel other than a schwa —

whereas the palato-alveolar affricates /tʃ, tʃʰ/ have no such restriction on their distribution. The gradual loss of environments in which a particular sound can occur seems to be an entirely plausible means by which a phonemic merger might spread through the lexicon.

3.7.2 Dental fricative~palatal glide variation

The phonemes /z/ and /j/ are in contrastive distribution word-initially, as demonstrated by /zūktʃūk/ 'wipe.PAST' and /jūŋ/ 'drink.PAST', yet they seem to occur in free variation medially in a number of words. When the variant /j/ occurs medially it is accompanied by a predictable assimilation in height of a preceding or following mid vowel (cf. 3.17 below). But to further muddy the waters, sometimes bewilderingly unpredictable and unmotivated sound changes can be found in a vowel following the palatal variant, such as the rounding observable in the word-final vowel of /mījʉ̄/ in (3.18) below.

	/z/	/j/	Gloss
(3.17)	/məzəm/	/mījīm/	'poison'
(3.18)	/məzə/	/mījʉ̄/	'fire'
(3.19)	/āzā/	/ājā/	'grass'
(3.20)	/wàzàʔ/	/wàjàʔ/	'bird'
(3.21)	/án-zá/	/án-já/	'chicken-DIM'

There may be a number of reasons for this putative free variation: in order of increasing likelihood, there may be a limited suspension of phonemic contrast when the segments involved occur medially; it could be attributed to an evolving sound change; or it could the result of borrowing. Mills (1926:336) also observed free variation between /z/ and /j/ in many words, noting that 'some villages say *yāni* for "the day after tomorrow," and others say *zāni*'. This was the only pair of examples showing free variation between /z/ and /j/ that he documented. Whilst the /z/~/j/ variation appears to be possible in a substantial number of words tested, it cannot occur freely in just any word containing one of these segments. To prove this, I tried exchanging /j/ for /z/ and *vice versa* in words of the corpus, made assimilatory vocalic sound changes if /j/ was involved, and then tested the words on a native speaker (MA). This had mixed results — some words were accepted, while others were rejected. Examples of rejected words are listed below (fabricated words that were rejected are starred; unstarred words are attested in the corpus).

	/z/	/j/	Gloss
(3.22)	/təzə/	*/tījə/~*/tījʉ̄/	'grandmother'
(3.23)	*/āzū/	/ājū/	'word'
(3.24)	*/təzəmtʰəm/	/tíjímtʰəm/	'hardship'
(3.25)	*/āzā/	/ājā/	'night'
(3.26)	*/āzūŋ/	/ājūŋ/	'river'
(3.27)	*/āzəm/	/ājīm/	'village'

A check was made in the comparative vocabulary of Marrison (Vol. II, 1967) against those WM words that demonstrate apparent free variation of /z/ and /j/. The WM words that are seen to allow either segment to occur medially are those and only those for which cognate reflexes exist in both the Mongsen and Chungli dialects, as suggested by the pairs of words collected from Marrison's word lists and provided below in (3.28–3.32). In a number of cases a borrowed lexical item is segmentally similar, if not identical (often with the exception of the medial consonant), to a reflex extant in the borrowing dialect. The segmental similarity is potentially a red herring aiding and abetting the erroneous analysis of /z/~/j/ variation as *bona fide* free variation. In actual fact, the comparative evidence suggests that what we are possibly dealing with here is the result of indiscriminate borrowing of cognate lexical items from Chungli into WM, leading to the synchronic existence of both reflexes in WM. Compare the following pairs of words (in Marrison's orthography, *ü* = [ə]; *ae* = [ai]):

	Mongsen	Chungli	Gloss
(3.28)	*müsem*	*mim*	'poison'
(3.29)	*mizü*	*mi*	'fire'
(3.30)	*aza*	*aei*	'grass'
(3.31)	*waya*	*ozü*	'bird'
(3.32)	*yi*	*azü*	'rice beer'

The words of (3.28–3.32) above were gathered by Marrison from work done between seventy and one hundred years ago by administrators and missionaries lacking a formal training in linguistics; consequently the transcriptions might be somewhat less than rigorous. This aside, what the selection essentially indicates is a general tendency for a medial dental fricative in Mongsen (represented orthographically by either *z* or *s*) to correspond to a palatal approximant in Chungli (which is interpreted as being represented orthographically by *i*, if not by an overt *y*) in words which are otherwise segmentally similar. Without the benefit of the comparative evidence, these WM words might mistakenly be assumed to contain medial segments in free variation, when in fact both reflexes of an etymon are found to occur within the same dialect as a result of borrowing; this is particularly so when pairs of synonymous words are as segmentally mirrored as are some of the WM examples of (3.17–3.21). Admittedly the correspondence of /z/ forms to Mongsen and /j/ forms to Chungli is more of a tendency than a comprehensive statement of fact, because some examples demonstrate the opposite distribution to that stated above (cf. the example words of (3.31) and (3.32)); however, this might reasonably be expected to occur as a result of borrowing in both directions between varieties in close contact.

The WM words in which free variation was rejected by MA were also checked against the comparative vocabulary of Marrison (Vol. II, 1967). The examples available for comparison suggest that the reason my fabricated words with a medial /z/ or a /j/ (the starred words of 3.22–3.27 above) were rejected was because they do not exist in either the Mongsen *or* the Chungli dialects, as the real data of Marrison (3.33–3.37 below) clearly demonstrate.

	Mongsen	Chungli	Gloss
(3.33)	*tüzü*	*otzü*	'grandmother'
(3.34)	*ayu*	*o*	'word'
(3.35)	*aya, ayanang*	*aonang, amang*	'night'
(3.36)	*ayung, yungba*	*tsü*	'river'
(3.37)	*ayim, yimkung*	*im*	'village'

The evidence presented here overwhelmingly points to borrowing as an explanation for the kind of variation demonstrated in the WM words of (3.17–3.21). It is small wonder that there is such extensive lexical interchange between the two main dialects of Ao, given that Chungli and Mongsen have most probably co-existed for a considerable period of time. Borrowing also must have been facilitated by the high incidence of Chungli/Mongsen bidialectalism, and consequently the boundary between the two lexicons has become blurred over the passage of time.

Cultural and sociolinguistic factors are held to be responsible for creating the conditions under which this type of lexical borrowing between varieties has taken place. For instance, Ao villages often have two or more geographically determined administrative wards known as *khel* in Nagamese, the *lingua franca* of Nagaland. It has been reported that in some villages, one *khel* speaks a variety of Mongsen and the other speaks a variety of Chungli, a linguistic situation still found to exist in Longkhum village in 1996, and perhaps also in a few other villages that I did not have the opportunity of visiting. Mills (1926:3) provides a fascinating account of bidialectalism in one village.

> Of the two "khels" of Sangratsu one consists of Mongsen clans speaking the Mongsen dialect, and the other of Chongli [*sic*] clans speaking the Chongli dialect — the two not twenty yards apart. Each "khel" knows the other's language but speaks its own, and a Mongsen woman married to a man of the Chongli "khel" will speak Mongsen to her husband but Chongli to her baby, for the child is Chongli like its father and must be brought up to speak Chongli. But in Mokongtsü [Mokokchung] village, while there is a Chongli "khel" and a Mongsen "khel", the whole village speaks Mongsen. It must be very inconvenient to speak two languages in the same village, and the tendency to adopt a common language is a natural one.

Waromung village itself has two *khel*; one is called the Mongsen *khel* and the other is called the Chungli *khel*. Nowadays the whole village speaks Mongsen, yet the traditional names suggest that at an earlier period each *khel* spoke a distinct dialect.

The Ao are believed to have migrated from the east. Because Mongsen villages predominate on the western ridges and Chungli villages predominate on the eastern ridges, it is conceivable that Mongsen speakers founded Waromung village and were later followed by a Chungli-speaking group, who established their own *khel* within the village. The original language of the Chungli *khel* is most likely to be the source of the Chungli loan words that have been borrowed into WM long after Chungli has ceased to be spoken as the first language of the Chungli *khel* inhabitants. Perhaps it is significant that all my data comes from native speakers who live in or originate from the Chungli *khel* of Waromung village. There may also have been some superstratum influence exerted on WM by Chungli over the past century because Chungli has become the prestige dialect of the Ao people and the medium of education and proselytism; consequently it is spoken throughout the Mokokchung area.

An alternative explanation for the /z~j/ variation is that intermarriage facilitated the borrowing not of whole lexical items *per se*, but of just the pronunciation of the medial consonant in these particular words.[2] Members of the Chungli and Mongsen moieties freely married in the past and continue to so nowadays, provided that their marriage partners belong to exogamous clans. It is plausible to imagine a situation in which a Chungli woman speaking her native dialect would pronounce the medial consonant of these words with /j/

2 I am grateful to Carol Genetti (pers. comm. 2002) for this suggestion.

while her Mongsen husband would pronounce them with /z/, thereby creating the conditions in a bidialectal household under which their children might interpret /j/ and /z/ as being in free variation word-medially. This would appear to be a somewhat different process to that of the lexical borrowing of whole items suggested above.

Whatever the underlying reasons for the /j~z/ variation may be, we can probably assume that two types of sociolinguistic factors have contributed to the presence of Chungli loan words (or pronunciation) in WM — co-existence and widespread bidialectalism of two erstwhile socially equal dialects, and the influences of a prestige variety.

3.7.3 The labio-dental fricative in loan words

Two WM speakers of a family who had lived in the district town of Mokokchung for many years pronounced the word 'snake', pronounced as [a^{33}hʉ^{33}] by the majority of WM speakers, as [a^{33}fʉ^{33}] in elicitation, but retained the glottal fricative /h/ for other words, such as [hau̯ʔ˥^{11}] 'yes', or [hʉ^{33}] 'blow.PAST'. One of them (MA) also pronounced 'rat' as [a^{33}fə^{11}]; the common WM pronunciation for this word is [aʉ̯^{55}].

The glottal fricative cannot be claimed to have a labiodental realisation before a high central vowel for these speakers, viz. /h/ → [f] /_ ʉ, because the glottal fricative in [hʉ^{33}] 'blow.PAST' does not demonstrate such an allophonic realisation in this environment. As [f] is never encountered in the texts gathered from older WM speakers living in Waromung village, I have treated the aberrations as loan words specific to the idiolects of my two consultants.

Preliminary observations of other Mongsen varieties spoken in the immediate vicinity of Mokokchung (for example, those spoken in Khensa and Mekuli villages) suggest that a voiceless labio-dental fricative has phonemic status in these dialects and these varieties are therefore likely to be the source of the loans containing [f]. Some examples of Khensa Mongsen and Mekuli Mongsen words containing a voiceless labio-dental fricative are [u^{33}fuŋ^{11}] 'dust', [fu^{33}] 'blow.PAST', [a^{33}fi^{33}] 'snake' and [a^{55}fiʔ^{11}] 'rat', with the latter two examples segmentally very similar to MA's pronunciation of words having the same meaning. Since the labio-dental fricative is only observed to occur in a few words, there is no justification for analysing it as a loan phoneme.

3.8 Distinctive features of phonemes

The distinctive features matrix of Table 3.3 (p.51) is for the most part based on distinctive features originally proposed in *The Sound Pattern of English* (Chomsky & Halle 1968; hereafter SPE) and represents the minimum requirements for the identification of the distinctive sound units of WM. In the following appraisal some revisions and additions are proposed to deal with inadequacies encountered in the application of SPE features to a description of WM phonology.

Distinctive features are able to resolve the problem encountered with the traditional classification of stops and affricates discussed in §3.3.2. Both stops and affricates have the feature [–continuant] in common, but the affricates are further differentiated by the acoustic feature [+strident].

Table 3.3: Tentative distinctive features matrix for phonemes. Redundant features are omitted (* the glottal stop is a prosody, not a segmental phoneme)

	p	pʰ	t	tʰ	k	kʰ	m	n	ŋ	s	z	h	ts	tsʰ	tʃ	tʃʰ	l	ɹ	j	w	i	ʉ	u	ə	a	a̰	ʔ*
son	–	–	–	–	–	–	+	+	+	–	–	–	–	–	–	–	+	+	+	+	+	+	+	+	+	+	–
syll	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	–
cons	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	+
cor	–	–	+	+	–	–	–	+	–	+	+		+	+	+	+			+	–							
ant	+	+	+	+	–	–	+	+	–	+	+		+	+	–	–											
high																					+	+	+	–	–	–	
front																					+	–	–	–	–	–	
back					+	+															–	–	+	–	–	–	
low																					–	–	–	–	+	+	
nasal							+	+	+								–	–	–	–							
strident										+	+		+	+	+	+											
lateral																	+	–	–	–							
cont	–	–	–	–	–	–	–	–	–	+	+	+	–	–	–	–	+	+	+	+							–
oral	+	+	+	+	+	+	+	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–
stiff										+	–																
slack										–	+																
spread	–	+	–	+	–	+							–	+	–	+									–	–	
constr	–	–	–	–	–	–							–	–	–	–									–	+	

The laryngeal features [±spread], [±constricted], [±stiff] and [±slack] are introduced after Halle and Stevens (1971; hereafter H&S) because the original SPE feature [±heightened subglottal pressure] has since been demonstrated to be untenable as an explanation for voice onset time contrasts. The features [±spread] and [±constricted] establish a contrast between voiceless aspirated and voiceless unaspirated segments in the stop and affricate series, and in addition serve to distinguish modal voice vowels from the creaky voice vowel. The two other source features [±stiff] and [±slack], although redundant for distinguishing aspirated from unaspirated stops and affricates, or modal voice vowels from the creaky voice vowel, are however required for capturing the voicing opposition that occurs in the fricative series since the earlier SPE feature [±voice] has been replaced by the H&S source features. Some controversy surrounds the use of the latter two source features. A number of phoneticians are sceptical that the H&S model of laryngeal features is actually consistent with observations of glottal activity; part of the problem is the lack of consensus in the literature with regard to how some phonation types such as creaky voice are produced (Durand 1990:56–57).

If we are prepared to overlook these theoretical difficulties, the features [±stiff] and [±slack] can also be extended to the feature specification of the tonal system (described in Chapter 5). This is straightforward for establishing contrasts between the high, mid and low tonemes (cf. Table 3.4 below). Similarly, the falling contour pitch that functions as a boundary tone and marks the narrative converb suffix and a conjunction can be decomposed into high and low components occurring in a serial relationship. Thus at the onset of the contour it is [+stiff] and [–slack], whereas at the termination it is [–stiff] and [+slack].

This might be represented autosegmentally as follows:

(3.38)

$$\begin{array}{cc} \text{H} & \text{L} \\ \begin{bmatrix} +\text{stiff} \\ -\text{slack} \end{bmatrix} & \begin{bmatrix} -\text{stiff} \\ +\text{slack} \end{bmatrix} \\ \multicolumn{2}{c}{\sigma} \end{array}$$

Table 3.4: Distinctive features matrix for tonemes

	/High/	**/Mid/**	**/Low/**
stiff	+	–	–
slack	–	–	+

It was previously mentioned that the segmental phoneme /h/ is not treated as a sonorant, but as a voiceless glottal fricative minimally differentiated from other segmental phonemes by the feature [–oral]. The glottal stop /ʔ/ is also specified in terms of distinctive features but is not recognised as a segmental phoneme because of the prosodic nature of its realisation and its highly restricted distribution. Reasons for this analysis are discussed in detail in §2.7 of the preceding chapter. Although they belong to different domains, if desired the glottal stop and the glottal fricative may additionally be differentiated from each other by the feature [±cont].

SPE binary features cannot account for a vowel system in which a phonemic contrast exists between the [+high, +back, +round] vowels /ʉ/ and /u/. One means of avoiding having to introduce another feature to capture contrasts in the horizontal dimension is to ignore both my auditory impressions and the acoustic evidence of a vowel formant plot in Chapter 4 and treat the high central rounded vowel /ʉ/ phonemically as a front vowel. Here it would contrast with the high front unrounded vowel /i/ on the basis of the feature [±round], as well as contrasting with the high back rounded vowel /u/ on the basis of the feature [±back]. Alternatively, we can adopt a compromise suggested by Lass (1984:86–87). He recommends using [±front] to capture three degrees of localisation with two features in an analogous manner to that used for distinguishing three degrees of aperture. In Table 3.3 the introduced feature [±front] is included in an attempt to capture a phonological contrast that would otherwise not be possible to specify using traditional SPE features.

These compromises will only work if dissimilatory morphophonological changes are not assumed to be responsible for an interesting surface realisation of /ʉ/ as a high front unrounded vowel [i] in certain verbal suffixes when they are employed in word formation. The following discussion explains why this is the case.

The segment /ʉ/ occurs as a constituent of a number of grammatical morphemes, including the irrealis suffix /-ʉ̄/ and the causative suffix /-ʉ̄ʔ/. It is also found to occur with what is functionally an infinitive suffix /-ʉ̄pàʔ/; this appears to consist underlyingly of the irrealis marker /-ʉ̄/ plus the nominalising suffix /-pàʔ/, which together give an infinitival sense to a verb so inflected, e.g. /jūŋ-ʉ̄-pàʔ/ 'to drink'. A process whereby the /ʉ/ of a WM verbal suffix loses its backness and rounding when juxtaposed to a stem-final [+high], [+back], [+round] vowel at first seems plausibly motivated by the need to preserve that suffix's surface segmental representation and the grammatical information it encodes. To illustrate this putative dissimilation, we will look at what happens when the infinitive suffix /-ʉ̄pàʔ/ is juxtaposed to a root-final, [+high], [+back], [+round] vowel /u/.

When the infinitive suffix /-ʉ̄pàʔ/ is suffixed to a root or stem that has /u/ as the final linear segment, such as in the word /jāŋlū/ 'make.PAST', the resulting surface form is [jaŋ^{33}lu^{33}i^{33}paʔ^{11}], not the expected [jaŋ^{33}lu^{33}ʉ^{33}paʔ^{11}]. In effect, this appears to represent a double dissimilation because not only is the initial vowel of the infinitive suffix unrounded, but it is also fronted.

The loss of backness and rounding might be attributed to the phonetic similarity of the juxtaposed segments, and by the fact that if it were not for this apparent dissimilation, important grammatical information encoded in the suffix could be lost as a result of the fusion of features shared by both the high back rounded vowel /u/ of the verb root, and the high central rounded vowel /ʉ/ of the verbal suffix. Confusion could easily arise in the absence of dissimilation because the nominalising suffix is /-pàʔ/; at the speed of normal speech the nominalised verb [jaŋ^{33}lu^{33}paʔ^{11}] 'make-NML', for example, would be rendered almost completely indistinguishable from the infinitive [jaŋ^{33}lu^{33}ʉ^{33}paʔ^{11}] 'make-IRR-NML' if this morphophonological change did not take place.

If we assume that we are indeed dealing with dissimilation, the inadequacies of both compromises suggested above as possible strategies for formalising the phonological contrast between /ʉ/ and /u/ can be demonstrated by application. Firstly we will attempt to treat /ʉ/ phonemically as a high front rounded vowel; this also allows the segment that undergoes dissimilation and the environment responsible for triggering the process to be specified in feature matrices using the original distinctive features of the SPE model, which allows just two degrees of localisation:

(3.39)

$$\begin{array}{ccccc} \text{ʉ} & \rightarrow & \text{i} & / & \text{u} \\ \begin{bmatrix}+\text{high}\\ -\text{back}\\ +\text{round}\end{bmatrix} & & \begin{bmatrix}\\ \\ -\text{round}\end{bmatrix} & & \begin{bmatrix}+\text{high}\\ +\text{back}\\ +\text{round}\end{bmatrix}\end{array}$$

The loss of rounding is accounted for by the fact that both /ʉ/ and its conditioning environment share the feature [+round], therefore this is a plausibly motivated dissimilatory change. However the immediate problem here is that we also have an unmotivated change in that the expected outcome should be just the unrounding of /ʉ/, resulting in the surface form [ɨ]. There should not be an additional change in the degree of localisation to [i], since in our adopted compromise /ʉ/ supposedly already contrasts with /u/ on the basis of the feature [±back]. Taking into consideration this dissimilatory evidence, we might infer that it is untenable to treat /ʉ/ phonemically as a high front rounded vowel in WM because it is impossible to account for the unmotivated dissimilation in the degree of localisation.

Next, if we test the suggestion by Lass (1984:86–87) and use an additional feature [±front] on the basis of the assumption that there really should be a third feature in order to capture three degrees of localisation, then the segments involved in what we are treating as a dissimilation process can be specified by the following feature matrices:

(3.40)

$$\begin{array}{ccccc} \text{ʉ} & \rightarrow & \text{i} & / & \text{u} \\ \begin{bmatrix}+\text{high}\\ -\text{front}\\ -\text{back}\\ +\text{round}\end{bmatrix} & & \begin{bmatrix}\\ +\text{front}\\ \\ -\text{round}\end{bmatrix} & & \begin{bmatrix}+\text{high}\\ \\ +\text{back}\\ +\text{round}\end{bmatrix}\end{array}$$

Once again, while this does explain the loss of the rounding feature as a plausibly motivated dissimilation — both /ʉ/ and the phonetic environment are [+round] — it does not provide the phonetic motivation for fronting, because according to the scheme of things in this compromise, the segment /ʉ/ is supposedly [–front] and [–back], so it already differentiated from /u/, which is [+back]; thus the surface realisation of underlying /ʉ/ as [i] is again held to be a completely unmotivated change.

In summary, both the compromises presented here are proven to be wrong on empirical grounds because they fail to explain an unmotivated change in feature specification, but only if we make the *a priori* assumption that the surface realisation of /ʉ/ as a high front unrounded vowel is a dissimilatory change brought on by morphophonological processes.

To take a completely different perspective on the problem, a more satisfactory explanation for the surface manifestations of /ʉ/ in verbal suffixes might be achieved with reference to morphology and historical sound change, in preference to the fruitless manipulation of phonological constructs. This would shift the analysis out of phonology and into the domain of morphology by positing two allomorphs for each morpheme containing an initial high central rounded vowel. The causative suffix would thus be represented by the allomorphs /-īʔ/ and /-ʉ̄ʔ/, the infinitive suffix by /-īpàʔ/ and /-ʉ̄pàʔ/, and the irrealis suffix by /-ī/ and /-ʉ̄/. A rule could then state that the front unrounded vowel allomorphs are realised in the [+round] environment and the central rounded vowel allomorphs occur elsewhere.

This simple solution is also well supported by comparative evidence. The high central rounded vowel /ʉ/ is unusual in the Mongsen dialect and appears to be confined to varieties spoken in just a few villages. Of the village varieties I have had contact with, only Khar and Waromung have the segment /ʉ/ in their phoneme inventories, and this regularly corresponds with /i/ in the suffixal morphology of all the other Mongsen varieties thus far encountered. The correspondence of /ʉ/ to /i/ is clearly demonstrated in the following list comparing the relevant grammatical morphemes of WM with those of Mangmetong Mongsen, a variety spoken in a village located away to the south of the Mokokchung district (cf. Map 3):

Waromung Mongsen	Mangmetong Mongsen	Gloss
/-ʉ̄ʔ/	/-īʔ/	CAUS
/-ʉ̄pàʔ/	/-īpàʔ/	INFIN
/-ʉ̄/	/-ī/	IRR

The segment /i/ occurs in all Mongsen varieties and generally carries a heavy functional load both in grammatical morphemes and in lexical items. In contrast, /ʉ/ is marked by its limited occurrence in just a few Mongsen varieties, and furthermore, by its very light functional load in the varieties in which it is attested. A perusal of Appendix B will demonstrate just how rarely it occurs in the lexicon of WM. While it overwhelmingly corresponds to /i/ in cognate lexical items, this is not completely consistent, as the lexical cognates of 'woman' in the following list of examples demonstrate. Nevertheless, this is the only correspondence anomaly encountered, so it might therefore be ignored to allow a generalisation to be made.

Waromung Mongsen	Mangmetong Mongsen	Gloss
/nʉ̀/	/nì/	'lead.PAST'
/á-ʉʔ/	/á-híʔ/	'NPF-rat'
/tə́-nʉ́ʔ/	/tə̄-nì/	'NPF-wife'
/á-nʉ́tī/	/á-nə́tī/	'NPF-woman'

While I have not had the opportunity to undertake a comprehensive survey of the vowel phoneme inventories of all Mongsen varieties, native speakers of various Mongsen varieties who have worked with me report that the high central rounded vowel is restricted to a few villages on Changkikong Range, and that it is highly emblematic of the speech of the residents of those villages in the sense that it functions as a shibboleth. The majority of Mongsen varieties otherwise appear to share the vowel phoneme inventory presented in Table 3.5 below.

Table 3.5: Common vowel phoneme inventory of Mongsen varieties

	Front	Central	Back
High	i		u
Mid		ə	
Low		a a̰	

Is it possible that at an earlier stage Mongsen had just three peripheral articulatory positions plus schwa (as represented by the inventory of Table 3.5), then the ancestral language of the Waromung and Khar village varieties innovated an additional central vowel /ʉ/, which was subsequently retained in these daughter languages? If so, this sound change appears to have completely spread through the suffixal morphology, but is still evolving in the lexicon. That /ʉ/ supposedly dissimilates to a high front unrounded vowel in the [+round] environment in WM is highly suspicious when viewed in light of the fact that /ʉ/ is consistently cognate with /i/ in the morphology of other Mongsen varieties. One may infer that this is not actually dissimilation at all; rather, that the archaic verbal suffixes contained /i/, and that this has evolved to /ʉ/ in Waromung verbal suffixes *except* where blocked by the [+round] vocalic environment, thereby causing the retention of the high front unrounded vowel allomorphs. Also, /ʉu/ or /uʉ/ sequences are made conspicuous by their absence in the phonotactic patterns of WM words, so a constraint on [+round] vocalic sequences provides further evidence for the assumption that historical factors may be responsible for these interesting allomorphic manifestations.

Lastly, another advantage of rejecting a morphophonological analysis that forces the assumption of dissimilation in favour of one that posits phonetically conditioned allomorphy is that it allows either of the compromises discussed above to be employed; consequently the phonological contrast between /ʉ/ and /u/ can finally be captured in a distinctive features framework with the addition of the feature [±front].

3.9 A practical orthography

The orthography to be presented in this section is suggested as a means by which WM and other Mongsen varieties may be rendered in writing for general use by native speakers. The proposed orthography attempts to keep digraphs to a minimum and takes into account the established orthography used for writing the Chungli dialect, with which most Ao are familiar.

Some of the suggestions offered here might equally be applied to Chungli as an improvement over the present orthography originally developed by the missionary Clark in the 1880's. This does not represent tone or the glottal stop and is not based on phonemic principles; it therefore falls far short of the ideal. At present, a translation of the Bible (a complete translation of which was first published in 1964) uses an orthography based on the original missionary orthography and serves as the written standard, but there are so many inconsistencies in the orthographic system currently in use that the codification and standardisation of spellings will continue to be an unattainable goal unless major revisions take place. Since 1993 there has been talk by community leaders of producing a new Ao dictionary; however, this will repeat the same mistakes if it is not done in consultation with trained linguists, and of course a major project like this will not come to fruition unless it is supported both financially and in spirit by prominent and respected members of the Ao community, and indeed by the community itself.

Unfortunately speakers of neighbouring languages, particularly those of the Tuensang and Mon districts found to the east and north-east of Mokokchung, have become acquainted with the missionary orthography of Chungli through reading the Ao bible and have subsequently adopted it willy-nilly for writing their own languages. In 1980 new spellings were introduced in Chungli school texts by the State Education Department in an attempt to address some of the shortcomings of the missionary orthography (for example, *h* has been suggested as a symbol for the glottal stop), but not without some controversy in the Ao community (Ngangshikokba Ao, pers. comm. 1996).

Pike (1947:208) addresses some of the problems encountered in devising a writing system for a vernacular language:

> In forming a practical orthography the investigator is constantly disturbed by a dilemma or series of dilemmas. He wishes to make his orthography scientifically adequate in order to get the best and fastest results in the teaching of reading; he wishes his alphabet to reflect the actual linguistic structure of the vernacular spoken by the people. But he wishes also to have an orthography which will not be offensive to the people in the region in which it is spoken or to the national government of the area. He wishes it to be adapted to traditional alphabets of the region and at the same time to be easy to write and print. These two general types of principles, the phonemic and the social ones, do not coincide.

To deal with some of these dilemmas he suggests a number of goals, divided into phonemic and social considerations. The phonemic goals applicable to WM are summarised as follows: (1) there should be a one-to-one correspondence between each phoneme and its symbol; (2) there should be no greater or lesser number of symbols-to-phonemes, so as to avoid the need to remember arbitrary spellings in the former case, and so as to avoid phonemically contrastive words being written identically in the latter; (3) spellings should reflect sounds, so that the native speaker is not required to memorise spellings by an arbitrary set of rules; (4) allophones of a single phoneme in free variation or in complementary distribution should be assigned just one symbol; (5) the orthographic representation of phonemic tone should be based on an adequate analysis of the language and consistently written on every word; (6) assimilated loans should be spelled as they are pronounced in the recipient language, and if possible, spelled using symbols established for representing the sounds of the native language; and (7) if loan words contain sounds that are not native to the recipient language, then the alphabet must be augmented by extra symbols (Pike 1947: 208–210).

Point (6) was reported as being behind the reason for the negative response to the revision proposing the use of *h* for the glottal stop in the missionary orthography (Ngangshikokba Ao, pers. comm. 1996). As both English and the local Indo-Aryan languages have contributed loans containing the glottal fricative /h/, the symbol *h* for the glottal stop was perceived as a poor choice by native speakers. This was in spite of fact that Chungli does not have a glottal fricative in its phoneme inventory, and secondly, that it would occur in complimentary distribution with the glottal stop if it were adopted for representing the glottal fricative in loan words. Perhaps this was not adequately explained to the community prior to the attempted introduction. The failure of this revision underlines just how important it is for any proposed changes to have the acceptance of the community it is intended to benefit.

The applicable social goals proposed by Pike (1947:211–213) can be summarised as follows: (1) use of the orthography should be approved and supported by the people for whom it is intended, so as to warrant the expenditure of effort required to master it; (2) the alphabet symbols should be the same as those used for the national language; (3) diacritics should be avoided where possible; (4) the symbols chosen for the orthography should be easy to print; (5) the symbols proposed for the vernacular language should conform as closely as possible to the trade language of the region to allow for an easy transition from one to the other, and *vice versa*; and (6) the alphabet should as much as possible serve more than one dialect.

The foreseeable problems are mainly with some of the above-mentioned social considerations. It might be argued that the alphabet should at the least be based on an Indic script (such as that used for Assamese or Bengali), particularly since the *lingua franca* of Nagaland is the Assamese-based pidgin known as Nagamese. On the other hand, while Nagamese is widely spoken it is not a written language, and consequently it is highly unlikely that there would be a great deal of support for an orthography based on any form of Indic script, particularly since the majority of Ao have little experience with the scripts of Indo-Aryan languages. In contrast, the Roman script has been used for writing Chungli for over one hundred years now, so Roman symbols may claim to hold a familiarity with native speakers that is altogether absent for Indic scripts. Furthermore, Roman symbols are also used for English, which is one of the official languages of India as well as being the official language of Nagaland.

The proposed orthography is presented below and is followed by discussion justifying the recommendations.

Stops

Phonemic	p	pʰ	t	tʰ	k	kʰ	ʔ
Orthographic	*b*	*p*	*d*	*t*	*g*	*k*	*q*

Affricates

Phonemic	ts	tsʰ	tʃ	tʃʰ
Orthographic	*dz*	*ts*	*j*	*c*

Fricatives

Phonemic	z	s	h
Orthographic	*z*	*s*	*h*

Nasals

Phonemic	m	n	ŋ
Orthographic	*m*	*n*	*ng*

Approximants

Phonemic	ɹ	l	j	w
Orthographic	*r*	*l*	*y*	*w*

Vowels

Phonemic	i	ʉ[3]	u	ə	a	a̠
Orthographic	*i*		*u*	*e*	*a*	*a*

Tones

Phonemic	high	mid	low	falling[4]
Orthographic	*v*	*unmarked*	*x*	:

3.9.1 Segmental representation

The orthography attempts to integrate the established familiarities of the missionary orthography while still adhering to the notion of a phonemic writing system, taking into account Pike's recommendations. As there is already an orthographic connection between the voiceless unaspirated stops and the orthographic voiced symbols (for example, the masculine semantic gender marker /-pāʔ/ that is suffixed to all male names is written as *ba* in the missionary orthography), a decision has been made to extend this distinction analogously to the whole of the stop and affricate series. This allows for an easier transition from the missionary to the phonemic orthography, and also minimises the need for digraphs.

The digraph *ng* suggested for the velar nasal presents a possible ambiguity when it occurs word-medially. For example, given an orthographic form *ninger*, this could be interpreted phonemically to be either /nīŋə̄ɹ/ or /nīnkə̄ɹ/. A simple solution is to insert an apostrophe at the syllable boundary when the first syllable is to be interpreted as having a dental nasal coda and the second syllable a velar stop onset. The absence of an apostrophe therefore indicates a velar nasal onset for the second syllable. Applying these rules, /nīŋə̄ɹ/ is represented orthographically by *ninger*, and /nīnkə̄ɹ/ is represented by *nin'ger*. An alternative approach might assume that this is hardly likely to cause problems of interpretation to native speakers who have the benefit of contextual cues to disambiguate meanings, and therefore the apostrophe is an unnecessary orthographic embellishment. Ultimately it is up to the community of native speakers to decide how precisely representative they want their orthography to be, and ideally they will take into consideration strategies that facilitate the greatest fluency in reading while minimising ambiguity.

There are a number of options for writing the glottal stop. It was mentioned above how important it is for the feelings of the people to be taken into account, and many in the Ao community have already rejected an attempted revision regarding the use of *h* for the

3 The rarity of a vowel phoneme /ʉ/ in Mongsen varieties justifies denying it representation in the proposed orthography. Additional arguments based on sociolinguistic considerations are given in §3.9.1.

4 The colon represents the repetition of a word-final grapheme, whatever its form may be.

orthographic representation of this previously neglected segment in Chungli. The symbol *h* is consequently best reserved for the representation of the glottal fricative phoneme in the Mongsen dialect. A question mark is probably equally out of the question because speakers might want to retain it (although redundantly) for interrogative sentences. The symbol *q* is a possible candidate; it is highly unlikely to be confused with other symbols, it rightfully gives the glottal stop a physical presence that is on a par with other orthographic symbols and it is easily produced by a printing press, typewriter or word processor. A common alternative in devised orthographies is to use an apostrophe. One disincentive for employing this punctuation mark is that it does not carry the same degree of visual weight as the alphabetical symbols do, and its word-final distribution in Mongsen could lead to its inadvertent omission and an ensuing rampant homography of the kind evident in the *Ao-Naga dictionary* of Clark (1990).

Native speakers will need to decide if they want to include representation for the high central rounded vowel /ʉ/ of WM in their orthography. One native speaker consultant (CL) used umlaut *ü* variously to represent both /ʉ/ and /ə/ in notes to help him remember the forms of some WM words in elicitation sessions. I am not sure of the provenance of the umlaut diacritic, as it is not found in the original orthography used for the *Ao Naga dictionary* by Clark (1990), but it is used in *Temeshi Lai*, the 1964 translation of the Bible into Chungli. Grierson (1967:433–435) provides a Chungli and Mongsen word list that uses *ü* in the transcription of some words. This was compiled by A.W. Davis, an Indian Civil Service administrator, and published in *Report on the Census of Assam for 1891* (Grierson 1967:422); the introduction of the umlaut may be attributable to him. Although it could be used to represent a central vowel, umlaut and other diacritics are impractical and likely to be neglected if they are used; they should therefore be avoided.

Personally I think there is little incentive for including representation for /ʉ/ in the proposed orthography, given its light functional load and correspondence with /i/ in other Mongsen varieties. Since /ʉ/ has only been found in two village varieties of Mongsen and the majority of Mongsen varieties appear to have the four-vowel system (plus the creaky voice vowel) presented in Table 3.5, it seems preferable to make the more widespread pattern the standard. This recommendation may also help to achieve some of the social goals of attracting broad support for the orthography, and of being useful for as many varieties as possible.

The vowels [u] and [o] occur in free variation, therefore it is proposed that just *u* is used for the high back rounded vowel /u/. The schwa can be transcribed orthographically with the symbol *e*, and the creaky voice vowel /a̰/ can be safely ignored. Since phonemic creaky voice occurs with such rarity in the language, the omission of a special representation in the orthography for this sound is unlikely to create any difficulties.

A question arises as to whether grammatical morphemes should be represented solely by their underlying forms in the orthography, or whether they should be represented orthographically by their various allomorphs according to the phonetic environment in which each form is realised. For instance, the negative prefix has the allomorphs /mə̀-/, /mù-/, /mì-/, /m-/. How difficult would it be for readers if the underlying form /mə̀-/ were used in all environments? The answer to this will depend upon just how automatic it is for the reader to derive the output from an underlying representation. It would certainly aid sight reading if each grammatical morpheme could be represented by an invariant form. But if testing proves this not to be the case, the orthography could be made less abstract to reduce the cognitive processing demands placed upon its users.

3.9.2 Suprasegmental representation

A fundamental decision needs to be made with respect to the necessity of marking tone in the orthography. It may be the case that native speakers will find tonal representation a burdensome and superfluous task. Once again, empirical testing ideally will guide any final decision that is made regarding the inclusion or omission of tonal representation in the orthography. However if a decision is made to mark tone, then the following considerations might assist in determining an appropriate mode of representation.

Having ruled out the use of diacritics, there does not appear to be any better means of marking tone other than by using easily produced orthographic symbols. The tones could therefore be marked on each syllable by using alphabetical symbols that are not required for representing the phonemes of the language. The /Mid/ tone occurs with the greatest frequency, so this could be represented by a zero grapheme to minimise the density of tone marking without contributing to a concomitant increase in ambiguity. /High/ tone could be signalled by the letter *v* and /Low/ tone by the letter *x*, and these tone markers could either follow the vowel of the syllable whose tone they represent, or else be inserted at the end of the syllable. The choice may be more a matter of aesthetics than function, but this will need to be tested to determine the representation that facilitates the greatest reading fluency. Table 3.6 below provides examples of the two suggested methods for marking lexical tone in segmentally identical words differing only in tonal contrasts.

Table 3.6: Recommended strategies for the orthographic representation of tone

phonemic transcription		tone marker after vowel	tone marker after syllable
/tə̄máŋ/	'all'	*temavng*	*temangv*
/tə̄-māŋ/	'NPF-body'	*temang*	*temang*
/tə̀-māŋ/	'PROH-believe'	*texmang*	*texmang*
/tə́-màŋ/	'NPF-dark'	*tevmaxng*	*tevmangx*

There are two options for dealing with tone sandhi perturbations: either mark citation tones in all cases, or mark output tones. Once again, if tone sandhi perturbations are automatic, it should not create difficulties for native speakers if underlying tones are represented. On the other hand, tone sandhi in WM appears to be quite extensive in the word formation processes examined in Chapter 2. This could create a considerable obstacle to fluent reading if native speakers are forced to process convoluted phonological rules in order to derive the output tones each time a morphologically complex polysyllabic word is encountered in writing. But without the benefit of empirical testing it is difficult to say what would be more appropriate at this stage. One would first need to establish how native speakers actually assign tones to syllables and/or words when reading before choosing the particular level of orthographic representation.

The remaining tone to recognise orthographically is the contour tone that occurs on the conjunction /tə̂ɹ/ 'and' and word-finally on the narrative converb suffix /-ə̂ɹ/; its boundary marking characteristics are discussed at length in §5.4. Although the contour tone is analysed as a sequence of two level tonemes rather than as a fourth toneme in the phonological system, it is important to account for it orthographically because it marks a grammatical contrast between, for example, /jūŋ-ə̄ɹ/ 'drink-PRES' and /jūŋ-ə̂ɹ/ 'drink-NCV'.

Having utilised all the appropriate alphabetical keyboard characters for the representation of phonemes and tones, we must now resort to using either a sequence of the letters marking high and low tones, or in the interest of keeping keystrokes to a minimum, make use of some other easily reproduced character. Since geminated consonants resulting from morpheme concatenation cannot occur word-finally, I propose that the contour tone is most conveniently represented by just repeating the orthographic symbol of the final linear segment of the word it marks. Applying this suggestion, our example word /jūŋ-ə̂ɹ/ 'drink-NCV' is thus represented orthographically as *yungerr*, and the conjunction /tə̂ɹ/ is represented by *terr*. Apart from its simplicity, the repetition of a word-final grapheme has the additional appeal of being somewhat iconic in its representation, because contour tones tend to lengthen the syllables they mark. This is so that the rhyme duration is of sufficient duration to allow the contour to be conveyed acoustically.

Leaving the /Mid/ tone unmarked opens the way for some degree of ambiguity to creep into the orthography, such as in situations where two vowels of independent syllables are juxtaposed and both bear /Mid/ tones (and are therefore unmarked). Whereas the absence of a tone mark indicates the non-syllabic component of a vowel sequence in the phonemic transcription, the second vowel of a sequence without a tone marking letter in the orthographic representation could be interpreted either as the nucleus of a separate syllable bearing a mid tone, *or* as the non-syllabic satellite of the initial vowel. This problem can be easily resolved by the use of the symbols *w* and *y* for /w/ and /j/ respectively when they function as the offglides of phonetic diphthongs, or alternatively by the insertion of an apostrophe to signal syllable boundaries.

3.9.3 Concluding remarks

All of the suggestions offered in the preceding sections are tentative and naturally would require extensive consultation with the Ao community to gauge acceptability, and of course careful testing to ensure the efficacy of the design. Whilst this orthography is believed to be appropriate for quite a number of Mongsen varieties, there remains the challenge of making it useful for other dialects of Ao, if this is desired. To achieve this end additional symbols may be needed if other phonemic distinctions are found to exist. The varieties of Mongsen spoken in Khensa and Mekuli villages, for example, appear to have a phonemic labio-dental fricative and a series of voiceless sonorants. The labio-dental fricative could be accommodated by the addition of the symbol *f*, and the voiceless sonorants could be distinguished from their voiced counterparts by a preceding *h*, for example *hm*, *hn*, *hng* etc. Need should determine further additions.

I close this chapter with the acknowledgement that some of these suggestions may well be considered superfluous by the Mongsen-speaking community. The missionary orthography currently used for the Chungli dialect, for example, does not give representation to tone or the glottal stop, and while these omissions are likely to throw up obstacles to fluent reading, presumably they do not render the written language incomprehensible to Chungli speakers when they have the benefit of contextual setting and native speaker intuition to resolve potential ambiguity. The same may be true for any proposed Mongsen orthography.

On the other hand, the practical experiences of speakers of tonal KCN languages whose opinion I solicited suggest otherwise, because all considered the lack of tonal representation in their respective scripts to be a considerable hindrance to reading and comprehension. One of these consultants, a tertiary-educated native speaker of Khonoma Angami, a language closely related to Ao and having four lexical tones (Blakenship et al. 1993), reports that he has substantial difficulty reading the missionary orthography used for writing the standard Angami dialect Tenyidie because it does not indicate the tones (Visier Sanyu, pers. comm. 2002). Similarly, a tertiary-educated native Mongsen speaker residing in Australia commented on the problems he has in trying to interpret letters written by his mother in the Chungli dialect using the missionary orthography. He reports that he has the same struggle with comprehension of the Chungli Bible, and he also attributes his difficulties to the lack of tonal representation in the missionary script (Temsü Longchari, pers. comm. 2001). All native speakers surveyed also reported a much slower reading speed compared to that for reading English.

Ultimately it is up to the Mongsen-speaking community to determine what level of orthographic complexity is needed to ensure sufficient clarity in the written expression of their language, and this must be balanced against the difficulty of learning and using the chosen orthography, with or without tonal representation.

This ends the description of WM segmental phonology. The next chapter deals with acoustic phonetics and presents the results of investigations into features of phonetic interest identified in the phonological analysis.

4 *Acoustic phonetics*

4.1 Introduction

This chapter commences with an acoustic analysis of the voice onset time of stops and affricates in §4.2 and §4.3 respectively. An acoustic analysis of the vowel formants follows in §4.4, and the chapter concludes with an instrumental investigation of the creaky voice phonation type in §4.5.

4.2 Voice onset time of stops

The (sub-)minimal pairs of §3.3.1 in the preceding chapter present evidence that WM stops have a two-way phonemic distinction in the duration of time between the release of an oral occlusion and the onset of voicing. To further investigate this contrast, an instrumental analysis of the voice onset time of aspirated and unaspirated stops was carried out on utterance-initial segments occurring at each distinctive place of articulation.

4.2.1 Procedure

Recordings of utterances in isolation were digitised with Kay Elemetrics Computerised Speech Laboratory (Kay CSL) using 16 bit quantisation at a sampling rate of 10 KHz. A wide band spectrogram of the sound wave was then made of each utterance using a bandwidth of 293 Hz. The relevant utterance-initial section of the sound wave was expanded and the cursor linked to the spectrogram window to identify the release burst and onset of phonation. This was taken to be the first evidence of sound wave energy, and the peak of the first positive spike without noise excitation respectively. The intervening period was then measured in milliseconds.

4.2.2 Findings

WM demonstrates an average voice onset time (hereafter VOT) of 21 msec for voiceless unaspirated stops; voiceless aspirated stops have an average VOT of 80 msec. Means and standard deviations of voiceless unaspirated versus voiceless aspirated stops are presented in Figure 4.1, and a breakdown of voice onset time means and standard deviations for contrasting stops at each distinctive place of articulation is given in Figure 4.2. The figures of means and standard deviations below are followed by sound waves and temporally aligned wide band spectrograms of utterance-initial voiceless unaspirated and voiceless aspirated bilabial stops (Figures 4.3--4.4) demonstrating segmentation and measurement methods. The speaker is MA.

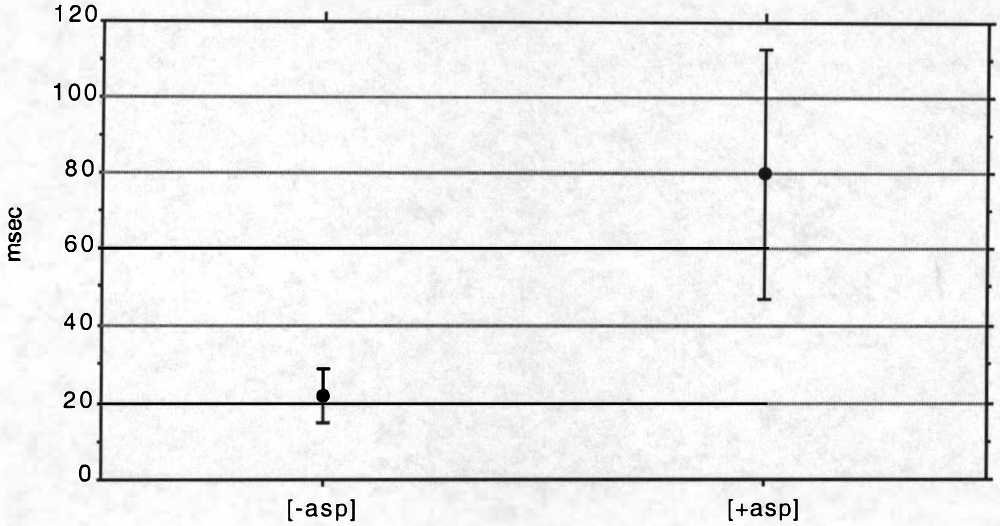


Figure 4.1: Overall means and standard deviations of voice onset time for 24 utterance-initial voiceless unaspirated stops and 22 voiceless aspirated stops in isolated words (in msec).

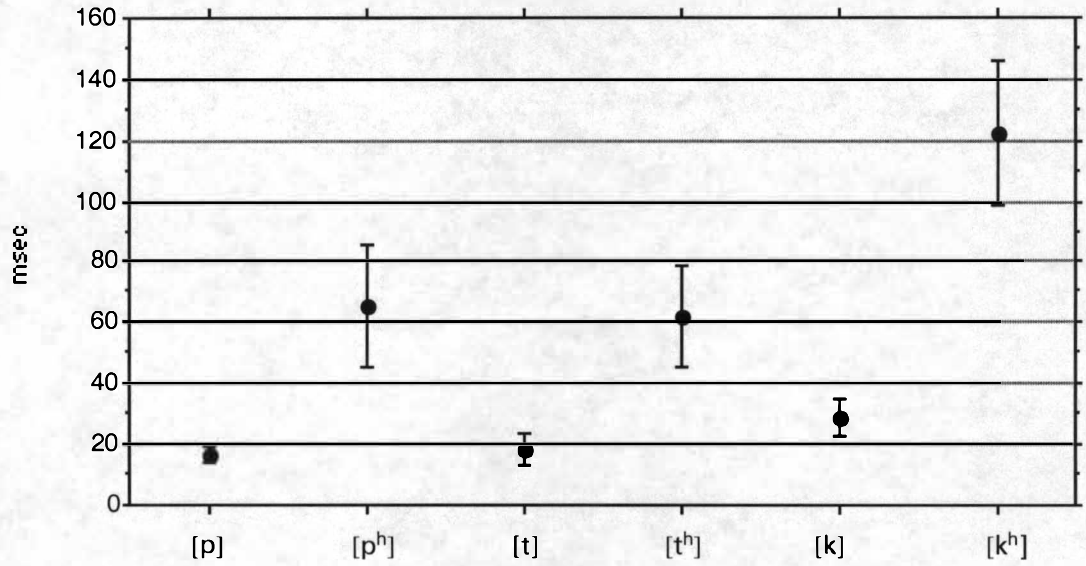


Figure 4.2: Individual means and standard deviations of voice onset time for utterance-initial stops (in msec). Number in samples: [p]=6; [pʰ]=9; [t]=10; [tʰ]=7; [k]=8; [kʰ]=6.

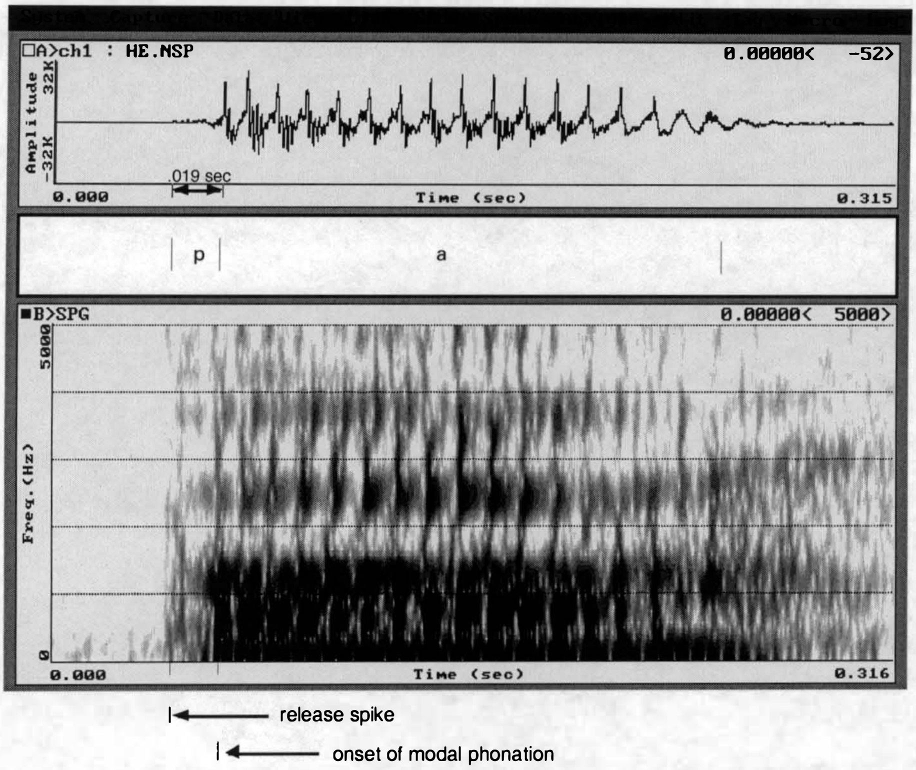


Figure 4.3: Temporally aligned sound wave and wide band spectrogram demonstrating a voice onset time of 19 msec for [p] in the word [pa^{33}] 's/he'.

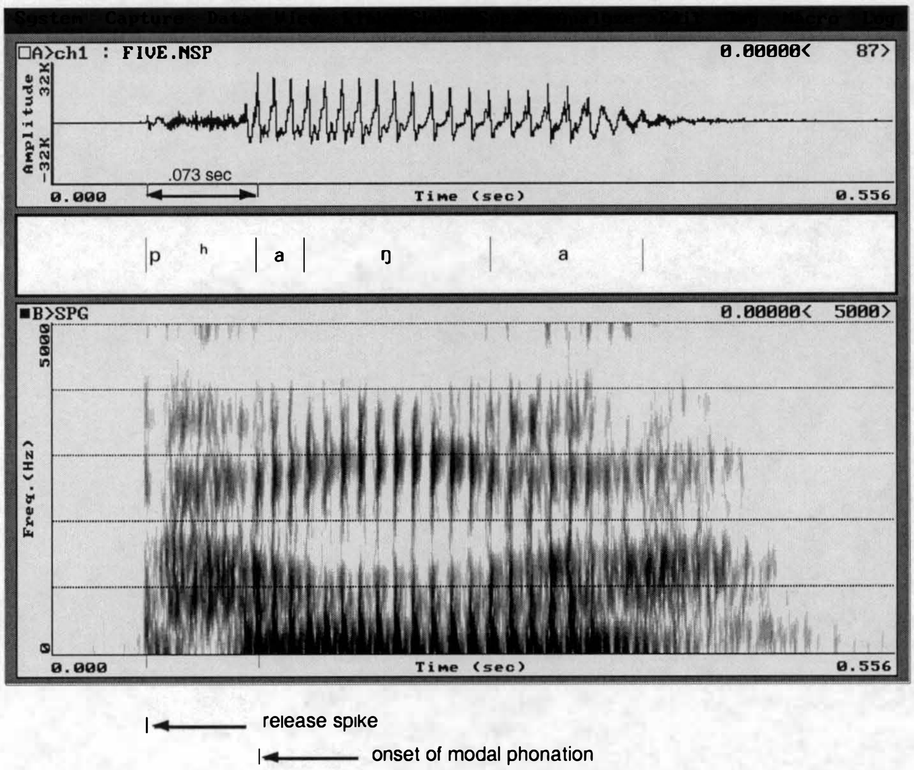


Figure 4.4: Temporally aligned sound wave and wide band spectrogram demonstrating a voice onset time of 73 msec for [p^{h}] in the word [p^{h}a^{33}ŋa^{33}] 'five'.

In a cross-language comparison of the voice onset times of utterance-initial stops, Lisker and Abramson (1964) find that stop categories generally fall into three ranges. These are approximately –125 to –75 msec for voiced stops, zero to +25 msec for voiceless unaspirated stops, and +60 to +100 msec for voiceless aspirated stops, with release of the oral occlusion being the zero reference point. In terms of overall averages, the voice onset times of WM initial voiceless unaspirated and voiceless aspirated stops in isolated words are found to fall within the ranges identified by Lisker and Abramson. It was found that the voice onset time of initial stops in WM are most similar to Cantonese, the only language in the above-mentioned survey that was also observed to demonstrate a two-way opposition between voiceless unaspirated and voiceless aspirated stops.

4.2.3 Voice onset time and place of articulation

Differences in VOT are demonstrated to correlate with place of articulation cross-linguistically (Peterson & Lehiste 1960; Lisker & Abramson 1964; Catford 1977) and there is a universal tendency for VOT lag to become progressively longer as the place of articulation shifts to the rear of the oral cavity. The values given in Figure 4.2 demonstrate the sensitivity of voice onset time to place of articulation in WM. The mean VOT values for unaspirated stops occurring at the bilabial, dental and velar places of articulation are 16 msec, 19 msec and 28 msec respectively. Thus the unaspirated series demonstrates a bilabial < dental < velar gradient in respect to VOT.

Ladefoged (1982:268) attributes this tendency to pressure gradient changes associated with place of articulation. In order for air to flow through the glottis, subglottal air pressure must overcome supralaryngeal air pressure. Airflow through the glottis is less impeded by atmospheric air pressure following the release of a bilabial occlusion than by intra-oral air pressure following the release of an occlusion at the back of the mouth. The impedance caused by the oral cavity is therefore responsible for it taking considerably longer for intra-oral air pressure to diminish to the point where sufficient airflow through the glottis can initiate vocal cord vibration after the release of a velar stop.

For the aspirated WM stops, the mean VOT values are 66 msec, 62 msec and 122 msec for the bilabial, dental and velar places of articulation respectively. Note that in contradistinction to the unaspirated stops, the aspirated series demonstrates a dental < bilabial < velar gradient in VOT. Catford (1977:113) calculated the mean VOT values of stops for the languages surveyed in Lisker and Abramson (1964) and similarly found that languages having an aspirated/unaspirated opposition demonstrated a reversal in position for the mean VOT values of bilabial and alveolar/dental aspirated stops. The velar place of articulation, however, was consistently associated with relatively longer VOT for both the aspirated and unaspirated stops.

4.2.4 Uncontrolled variables

Because of the limited number of examples of each type of stop occurring in utterance-initial position in the WM corpus, one variable that could not be controlled for in the analysis was the influence of vowel quality on VOT. It is conceivable that close vowels would cause a greater degree of impedance to airflow through the vocal tract than open vowels because of the aerodynamic reasons discussed in the preceding section. It might therefore be inferred that the articulation of a close vowel after the release of a stop would result in an increase in intra-oral air pressure, and consequently be reflected in the phonetic realisation of a slightly longer VOT.

Another variable that was not controlled for is the possibility that VOT measurements are susceptible to influence exerted by the syllabic structure of words (i.e. whether mono-, di- or trisyllabic) in the WM corpus. Lisker and Abramson (1964) also may not have controlled for this, as they do not document the particular syllabic structure of words that formed the corpus for their study of utterance-initial stops.

In any event, these uncontrolled potential influences on the production of VOT in WM are assumed to be negligible. The findings of the instrumental analysis correlate extremely well with the both the auditory analysis and the phonemic analysis, bearing in mind that acoustics cannot be used to identify a phonemic contrast between aspirated and unaspirated obstruents in the absence of evidence provided by sub-minimal or minimal pairs. Indeed, Goldsmith (1990:10) sensibly cautions that '... while phonetic reality may motivate a phonological representation, it neither justifies nor ultimately explains it. Phonetic reality provides the stuff of which phonological theory provides the organization'.

4.2.5 Previous descriptions of oppositions in voice onset time

The results of my investigation of voice onset time are rather interesting since it has been claimed by other authors that Ao does not have the two-way distinction of (near-) coincident voice onset time/voice onset time lag that I observe in WM. Gowda (1975:10), for instance, does not recognise any voice onset time contrast whatsoever for stops in a description of the Chungli dialect of Ao and posits a single series /p, t, k/. He states that the aspirated stops [pʰ], [tʰ], [kʰ] occur respectively in free variation with the non-aspirated stops [p], [t], [k] in initial position, are realised intervocalically and after voiced consonants as [b], [d], [g], and are realised elsewhere as the voiceless unaspirated stops [p], [t], [k].

The orthography used for the transcription of Longchang village Mongsen data in the anthropological monograph of Mills (1926) suggests a three-way distinction of VOT lead/ coincident VOT/VOT lag. A brief description of the vowels and consonants (1926:334ff.) lists a voiced series *b*, *d*, *g*, a voiceless series *p, t, k*, and a voiceless aspirated series *ph*, *th*, *kh*. One might surmise from the transcriptions of words in illustrative examples that according to Mills' interpretations, all three contrasts are found word-medially in syllable onset position, but only the voiceless aspirated and voiceless unaspirated stops occur word-initially. Orthographic symbols representative of all three voice onset time contrasts are used to transcribe stops occurring between voiced segments.

Marrison (Vol. I, 1967:58), whose data source for a proposed phoneme inventory of the Mongsen dialect is the work of Mills (1926; n.d. Typescript), states that his analysis does not recognise a phonemic distinction between aspirated and unaspirated plosives in either the Chungli or the Mongsen dialects of Ao. Marrison consequently adjusts the original orthographic transcriptions of the Mongsen data of Mills to accord with his phonological interpretations. In the phoneme tables titled 'The Ao [Chungli] Syllable' (Vol. I, 1967:58; Vol. II, 1967:348), his inventory lists a voiced series /b, d, g/ as well as a voiceless series /p, t, k/, indicating that according to his analysis, a two-way contrast is based upon voicing in stops.

4.3 Voice onset time of affricates

It was reported in Chapter 3 that the affricates demonstrate the same aspiration contrasts as the stops. The nature of the phonetic realisation of the phonemic contrast between voiceless aspirated and voiceless unaspirated affricates, however, can be significantly different to that of the stop consonants. The following interpretations of the phonetic realisation of aspirated affricates in WM are based upon the recorded data of three native speakers (MA, CL and CA).

In §3.2.2 I briefly mentioned that the contrast between the voiceless unaspirated dental affricate /ts/ and voiceless aspirated dental affricate /tsʰ/ is consistently located within the duration of the strident release, rather than in the presence or absence of aspiration *per se.* Consequently all the aspirated dental affricates are transcribed phonetically as [ts:]. The acoustic correlate of aspiration for the dental affricate is random noise distributed throughout the frequencies of the spectrum, the primary source of which is channel and wake turbulence (Catford 1977:37) generated in the oral cavity. Perhaps it is more accurate to describe the nature of this acoustic correlate as quasi-random noise, since there are well-defined concentrations of energy corresponding to the formant patterns of the following vowels in all the spectrograms of WM affricates.

Figure 4.8 (p.72) demonstrates representative spectrographic evidence of the nature of the acoustic phonetic realisation of /tsʰ/. It can be observed that high frequency energy associated with a supraglottal source of friction commences at the release spike and gradually spreads throughout a frequency range of 1–8 KHz, with maximum intensity at 4–8 KHz. The high frequency noise portion of the spectrogram, particularly the noise excited fourth formant, shows little loss of intensity until its termination at the onset of phonation. Between the strident release of the occlusion and the transition to the vowel there is no evidence of a glottal noise source that might be correlated with aspiration.

A similar phonetic realisation of aspiration as sustained strident release is only observed for the aspirated palato-alveolar affricate when it is followed by a high front vowel, as in the spectrogram of Figure 4.10 (p.73); here high frequency noise associated with an oral cavity source can be observed to continue right up to the onset of glottal phonation. This contrasts with the interval of low intensity noise that is observed to occur in Figure 4.11 (p.73) immediately after the strident release of the affricate and prior to the onset of glottal phonation. The phonetic realisation of aspiration in [tʃʰ] thus correlates acoustically with appreciably less distribution of noise through the higher frequencies and is characteristic of a glottal noise source. The palato-alveolar affricate is therefore transcribed as [tʃ:] when occurring in the environment before a high front vowel, and as [tʃʰ] elsewhere.

For the purposes of identifying a phonemic contrast, a prolonged period of stridency may be equated with the parameter of voice onset time lag, as defined by Lisker and Abramson (1964).

4.3.1 Procedure

Recordings of utterances in isolation were digitised at a sampling rate of 16 KHz in order to visualise the nature and distribution of high frequency information up to the Nyquist frequency (half the sampling rate) of 8 KHz. An aligned wide band spectrogram of the sound wave was then made of each utterance using a bandwidth of 469 Hz. The relevant utterance-initial section of the sound wave was expanded and the cursor linked to the spectrogram window to identify the release burst and onset of phonation. As described above for the stops, this was taken to be the first evidence of sound wave energy, and the peak of the first positive spike without noise excitation respectively. The intervening period was measured in milliseconds.

4.3.2 Findings

WM demonstrates an average voice onset time of 86 msec for voiceless unaspirated affricates; voiceless aspirated affricates have an average voice onset time of 148 msec. Means and standard deviations of voiceless unaspirated versus voiceless aspirated affricates are shown in Figure 4.5 below, and a breakdown of voice onset time means and standard deviations for contrasting stops at each distinctive place of articulation is given in Figure 4.6. Interestingly, there is a significant difference in VOT between the dental and palato-alveolar places of articulation for the unaspirated affricates [ts] and [tʃ] ($p < .01$), however a significant difference is not similarly demonstrated for the aspirated affricates [ts:] and [tʃʰ].

The figures of voice onset time means and standard deviations are followed by a selection of sound waves and temporally aligned wide band spectrograms demonstrating the acoustic features of dental and palato-alveolar affricates. Figures 4.10–4.11 in particular show how the articulatory position of a following vowel determines the spectrographic characteristics of quasi-random noise in aspirated palato-alveolar affricates.

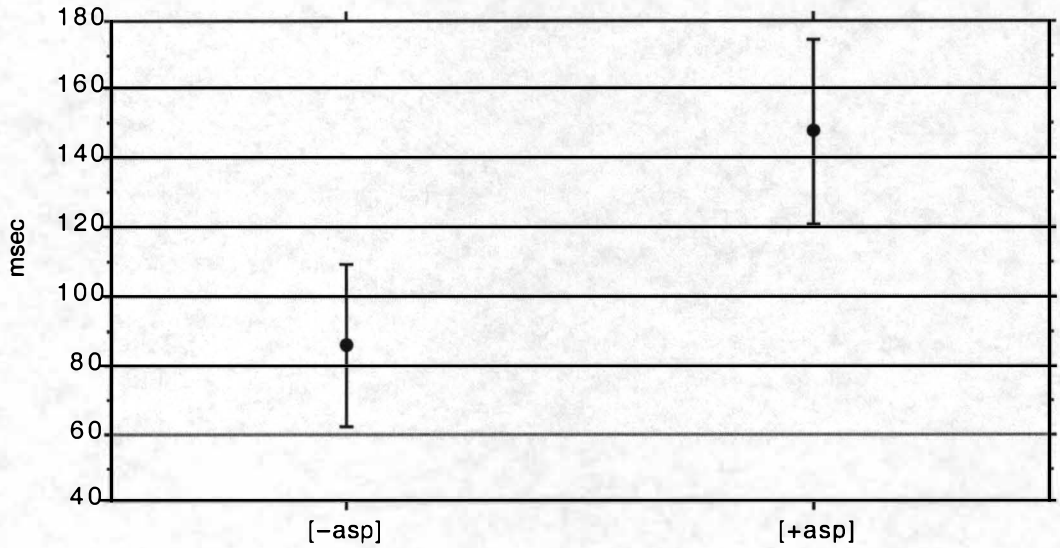


Figure 4.5: Overall means and standard deviations of voice onset time for 19 utterance-initial voiceless unaspirated affricates and 19 voiceless aspirated affricates in isolated words (in msec).

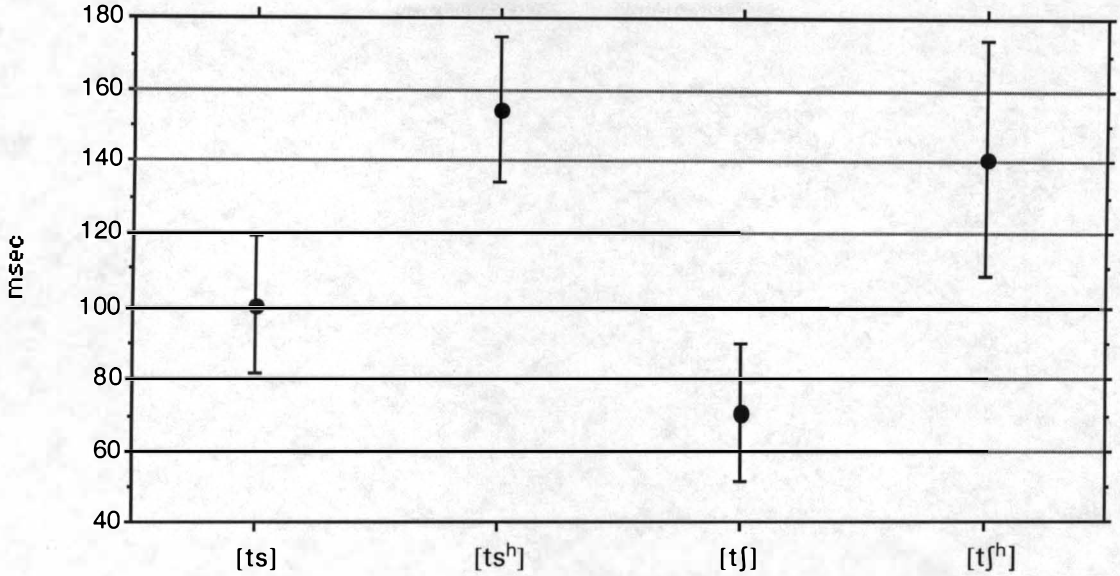


Figure 4.6: Individual means and standard deviations of voice onset time for utterance-initial affricates (in msec). Number in sample: [ts]=10; [tsʰ]=10; [tʃ]=9; [tʃʰ]=9.

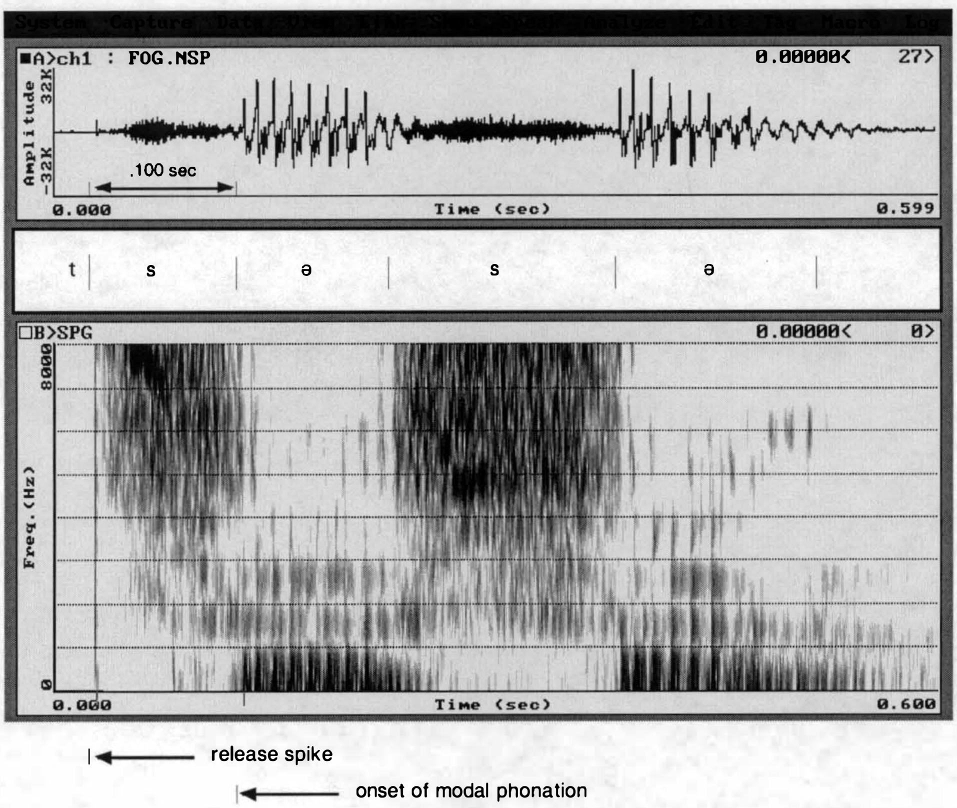


Figure 4.7: Temporally aligned sound wave and wide band spectrogram demonstrating a voice onset time of 100 msec for [ts] in the word [tsə¹¹sə¹¹] 'fog'.

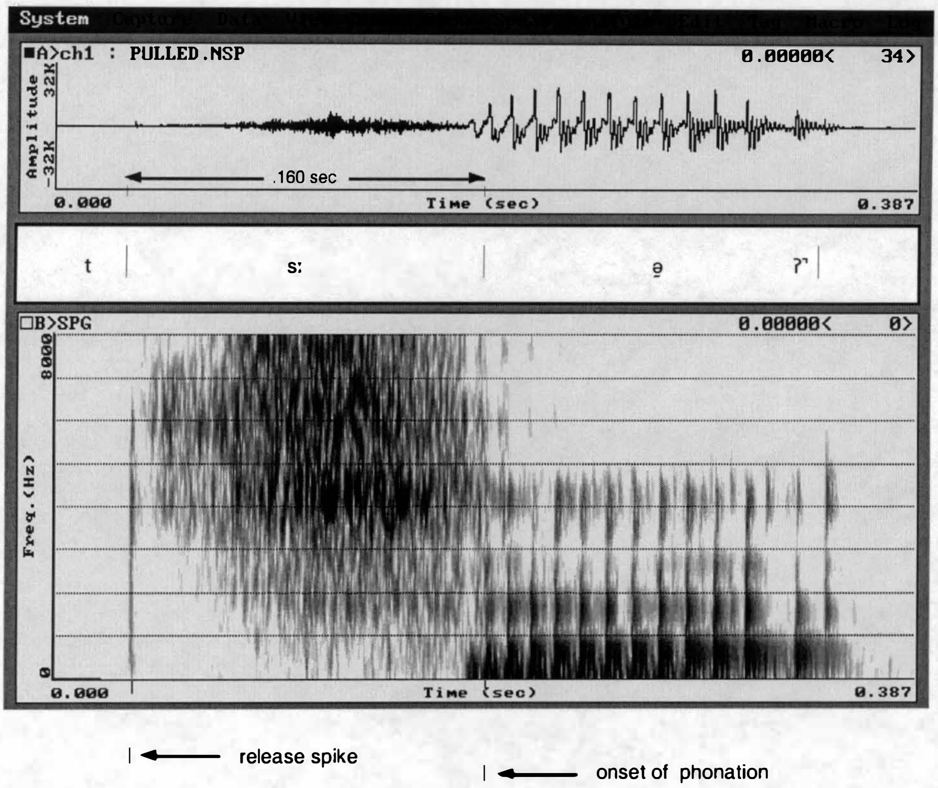


Figure 4.8: Temporally aligned sound wave and wide band spectrogram demonstrating a voice onset time of 160 msec for [ts:] in the word [ts:ə̠ʔ˺^{11}] 'pull.PAST'.

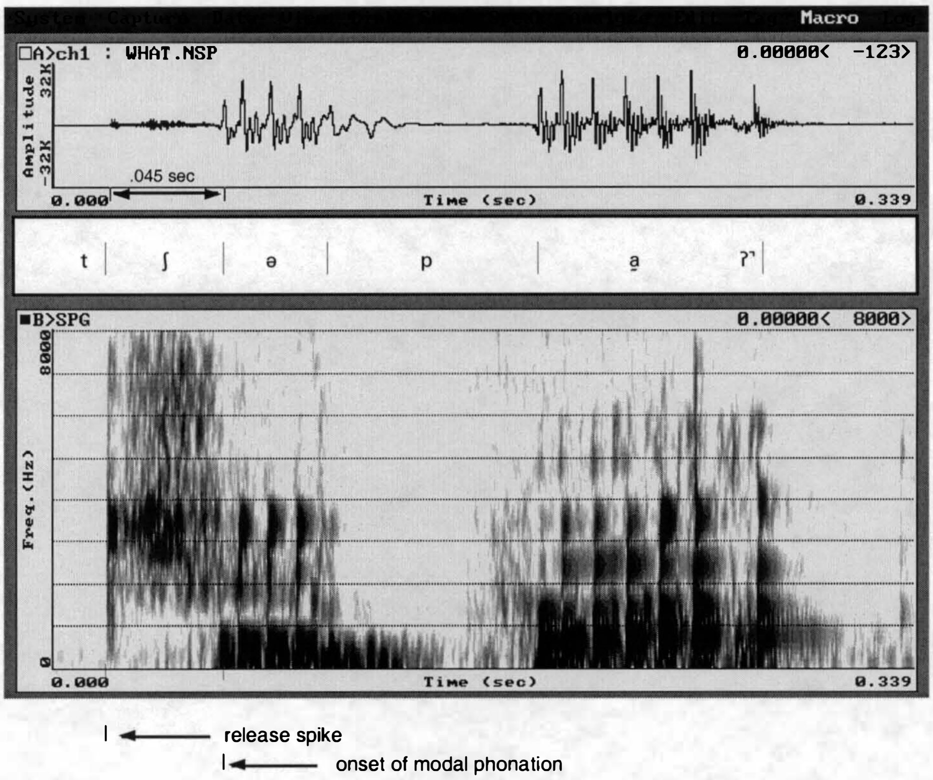


Figure 4.9: Temporally aligned sound wave and wide band spectrogram demonstrating a voice onset time of 45 msec for [tʃ] in the word [tʃə^{55}pa̠ʔ˺^{55}] 'what'.

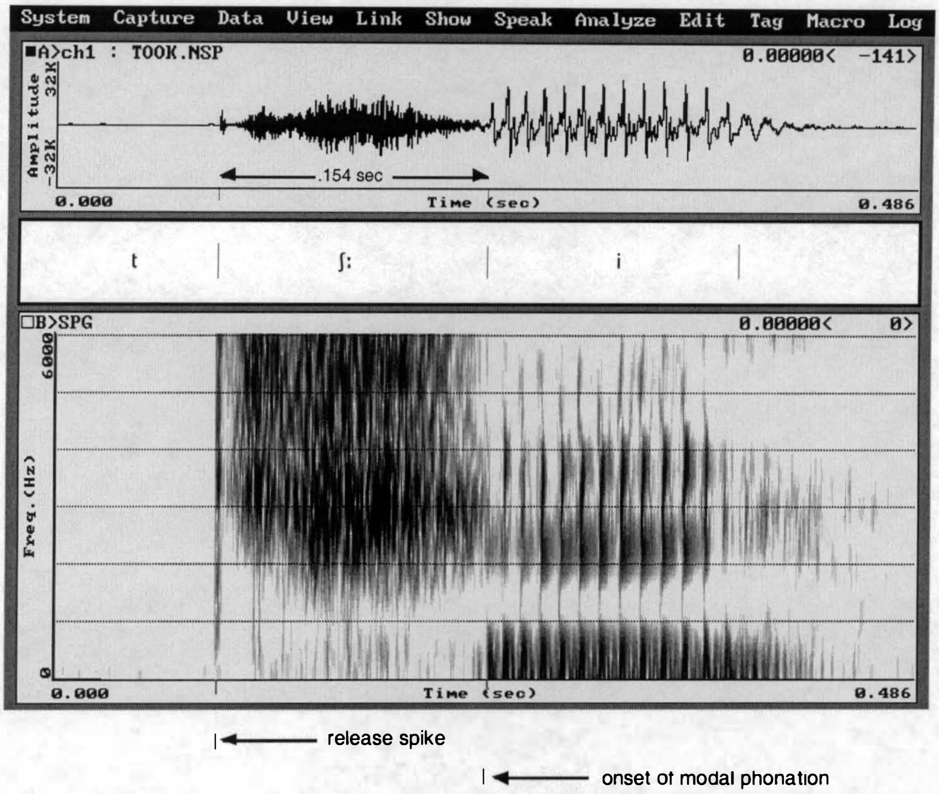


Figure 4.10: Temporally aligned sound wave and wide band spectrogram demonstrating a voice onset time of 154 msec for [tʃː] in the word [tʃːi^{33}] 'take.PAST'.

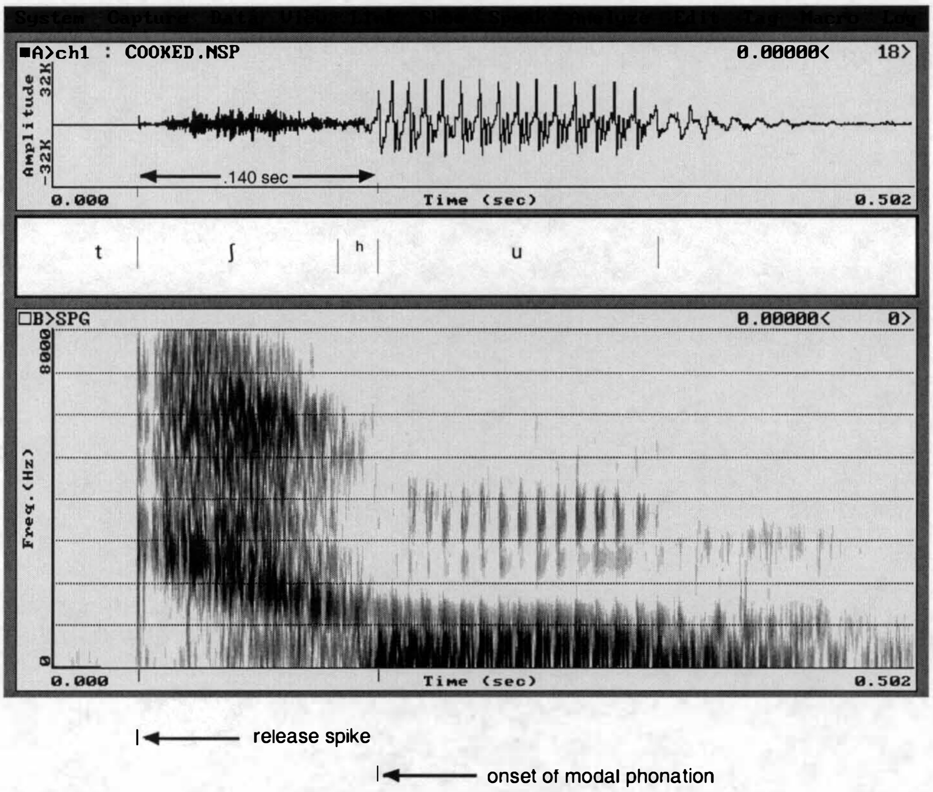


Figure 4.11: Temporally aligned sound wave and wide band spectrogram demonstrating a voice onset time of 154 msec for [tʃʰ] in the word [tʃʰu^{33}] 'cook.PAST'.

4.4 Acoustic analysis of vowel formants

An acoustic analysis of the vowel formants was done on the vowel phones, firstly because it is important to quantify auditory impressions instrumentally, and secondly, to provide the raw data for further investigations involving free variation and phonation contrasts. That the auditory analysis of the vowel system identifies a rather unusual pattern typologically, with three heights centrally but only one peripherally, is in itself enough to justify an acoustic analysis.

A total of forty-two words was digitised to provide six tokens of each vowel investigated. The words were uttered in isolation and recorded in the field using a Sony TCM-5000EV tape recorder and an external Sony ECM-FO1 electret condenser microphone. The speaker is CL.

Inasmuch as the data would allow, tokens were selected only if they were not subject to any perturbation of their formants by adjacent phonetic environments. For example, vowels with dental or palatal frames were avoided if possible. Ideally a less restricted phonetic environment would have been preferred for tokens of [o], but as previously noted, there is an overwhelming tendency for [o] to occur non-exclusively in the environment of a velar; other non-dental environments are not found in the corpus. Similarly, the choice of environment for the creaky voice vowel was limited by its rarity of occurrence. This is offset somewhat by the use of a majority of minimal and sub-minimal pairs for plotting the formant values of [a̰] and [a], so that the resulting differences can be held to be acoustically representative.

4.4.1 Procedure

The data were digitised at a sampling rate of 16 KHz using Kay Elemetrics CSL. Wide band spectrograms (469 Hz) of the sound wave of each word were first made and then the formant histories of the digitised utterance were tracked over time using a frame length of 25 msec and a filter order of 20. A frame length of 25 msec with a filter order of 20 was also used for the LPC frequency response. The CSL instruction manual (1994:385) suggests that if appropriate settings for the sampling rate and filter order have been made, then the formant tracings should be consistently centred on the dominant formants in the spectrogram. This was generally found to be the case for the WM data, and the formant values calculated by the software were consistent with expected outcomes for each vowel quality.

The formant tracings were used to help determine an appropriate position from which to extract formant values. This was judged to be a portion of the spectrogram that demonstrated a steady F-pattern, and which was not affected by transitions to or from adjacent consonants. The cursor was positioned on the steady state interval of the vowel and a Fast Fourier Transform (FFT) power spectrum analysis was made using a frame size of 512 points. This was overlaid with an analysis of the Linear Predictive Coding (LPC) frequency response at the cursor. The centre frequencies of F_1, F_2 and F_3 were then measured.

Representative spectrograms, tracings of LPC formant histories, FFT power spectrum analyses and LPC frequency responses for each vowel are provided below at Figure 4.12 through to Figure 4.18.

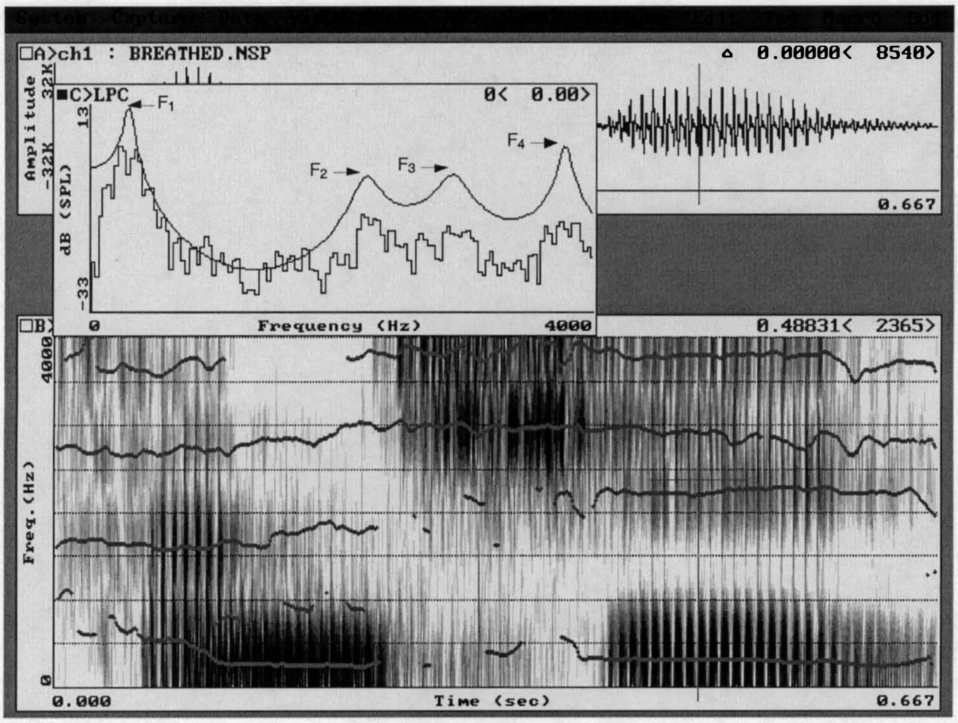


Figure 4.12: Temporally aligned sound wave and wide band spectrogram with superimposed LPC frequency response showing centre frequencies of the first, second, third and fourth formants of the vowel [i] in the word [sən³³ʃi³³] 'breathe.PAST'.

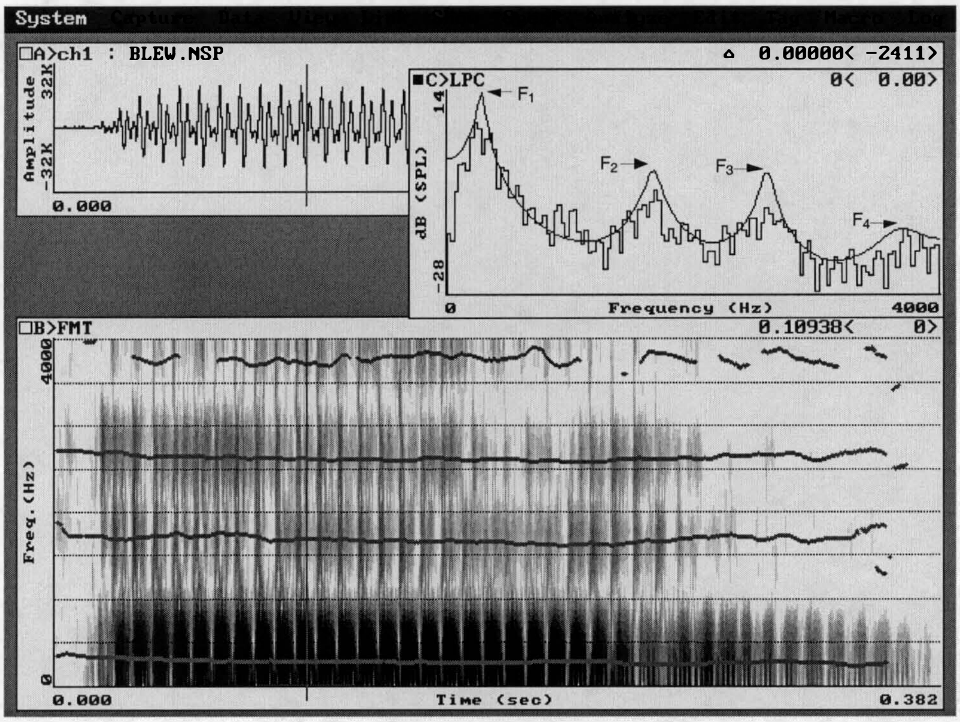


Figure 4.13: Temporally aligned sound wave and wide band spectrogram with superimposed LPC frequency response showing centre frequencies of the first, second, third and fourth formants of the vowel [ʉ] in the word [hʉ³³] 'blow.PAST'.

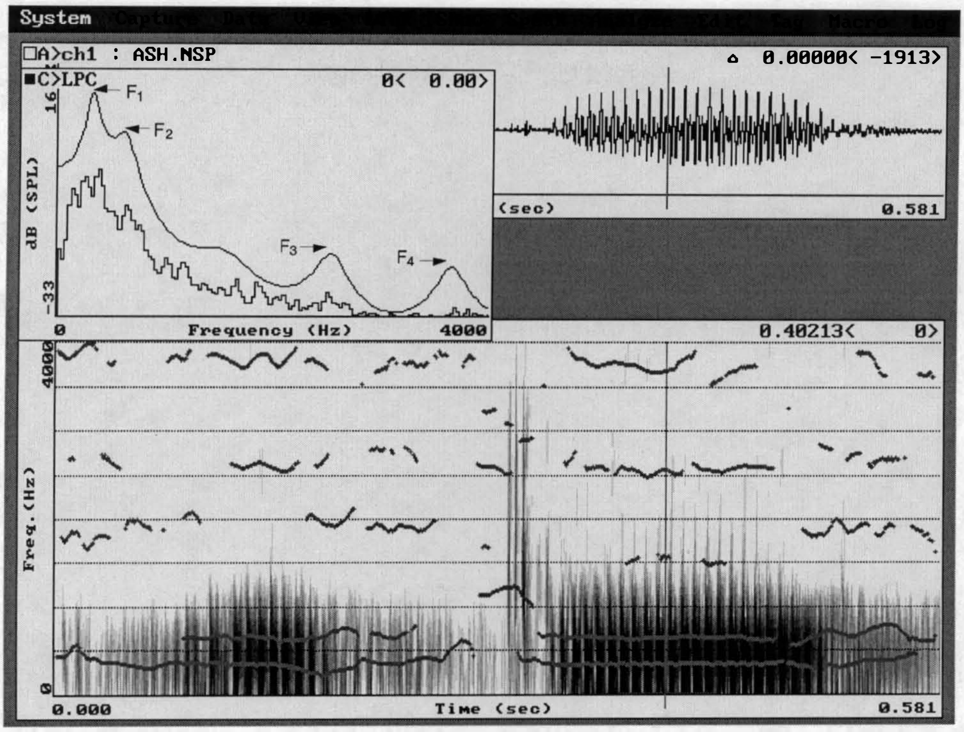


Figure 4.14: Temporally aligned sound wave and wide band spectrogram with superimposed LPC frequency response showing centre frequencies of the first, second, third and fourth formants of the vowel [u] in the second syllable of the word [u^{55}pu^{55}] 'ash'.

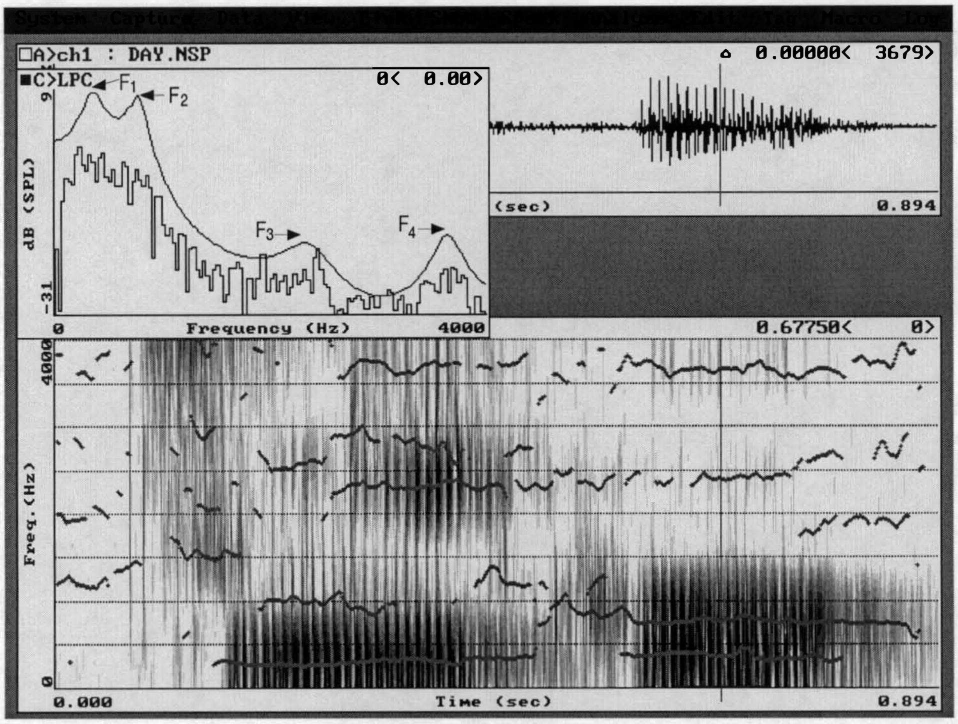


Figure 4.15: Temporally aligned sound wave and wide band spectrogram with superimposed LPC frequency response showing centre frequencies of the first, second, third and fourth formants of the vowel [o] in the word [tsə^{11}ŋi^{11}ko^{33}] 'day'.

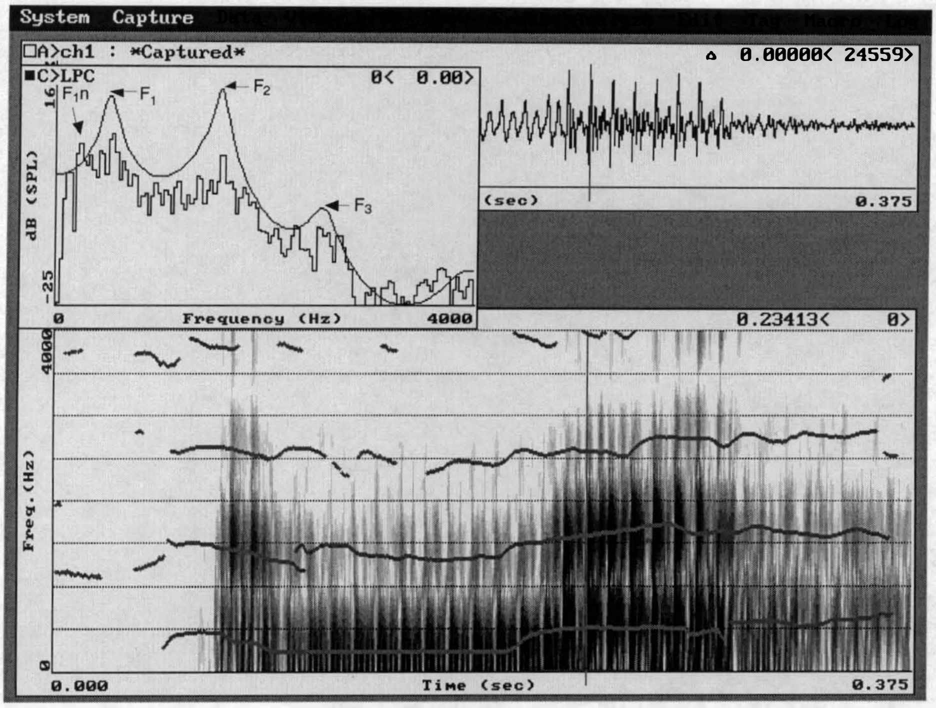


Figure 4.16: Temporally aligned sound wave and wide band spectrogram with superimposed LPC frequency response showing centre frequencies of the first, second and third formants of the vowel [ə] in the second syllable of the word [əm¹¹ət³³] 'hold.PAST'. A nasal formant (F_1n) precedes the first formant in the LPC frame.

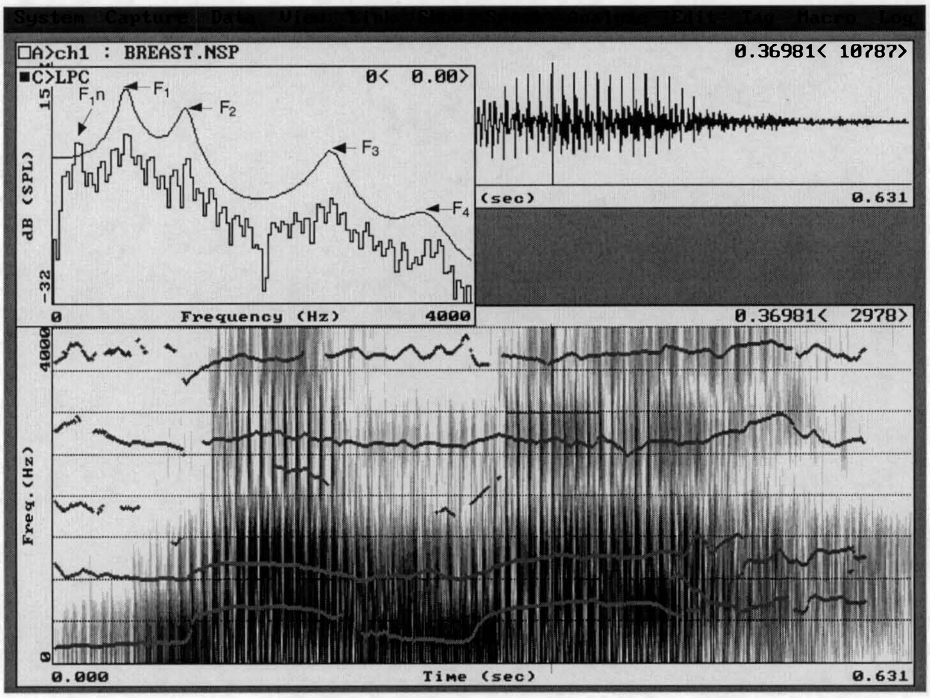


Figure 4.17: Temporally aligned sound wave and wide band spectrogram with superimposed LPC frequency response showing centre frequencies of the first, second, third and fourth formants of the vowel [a] in the second syllable of the word [ma³³ma³³] 'breast'. A nasal formant (F_1n) precedes the first formant in the LPC frame.

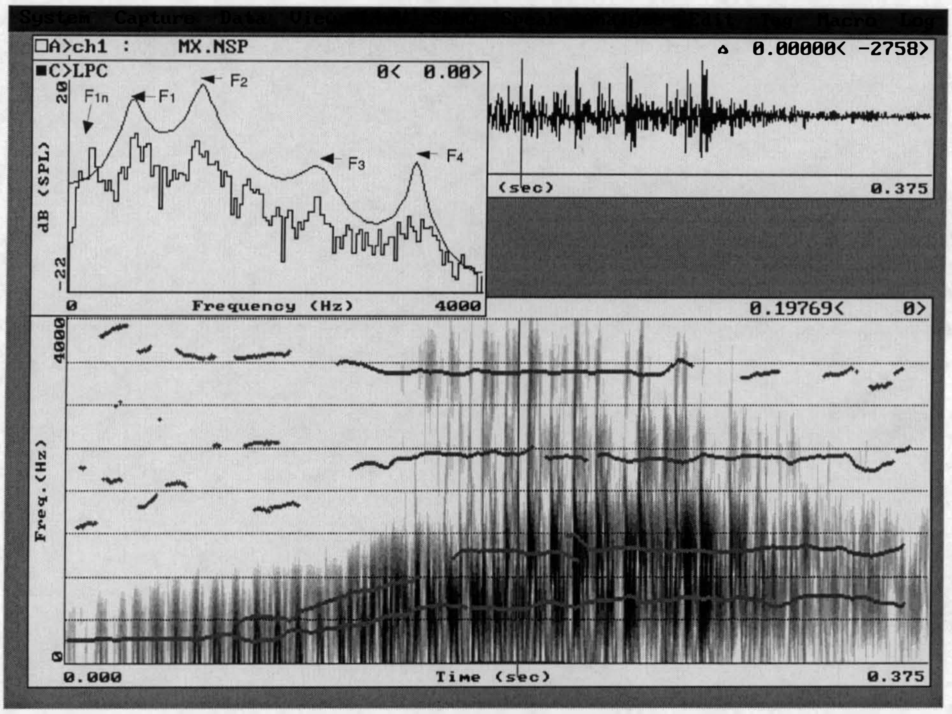


Figure 4.18: Temporally aligned sound wave and wide band spectrogram with superimposed LPC frequency response showing centre frequencies of the first, second, third and fourth formants of the vowel [a̰] in the word [wa̰ʔ[11]] 'slice.PAST'. A nasal formant (F_1n) precedes the first formant in the LPC frame.

Having obtained the values in Hertz for all the tokens, Statview 512+ software was then used to make a scattergram plotting F_2 on the X axis and F_1 on the Y axis. The graph has been rotated 180 degrees so that the vowel chart can be viewed with the customary orientation of the high front vowel in the upper left corner.

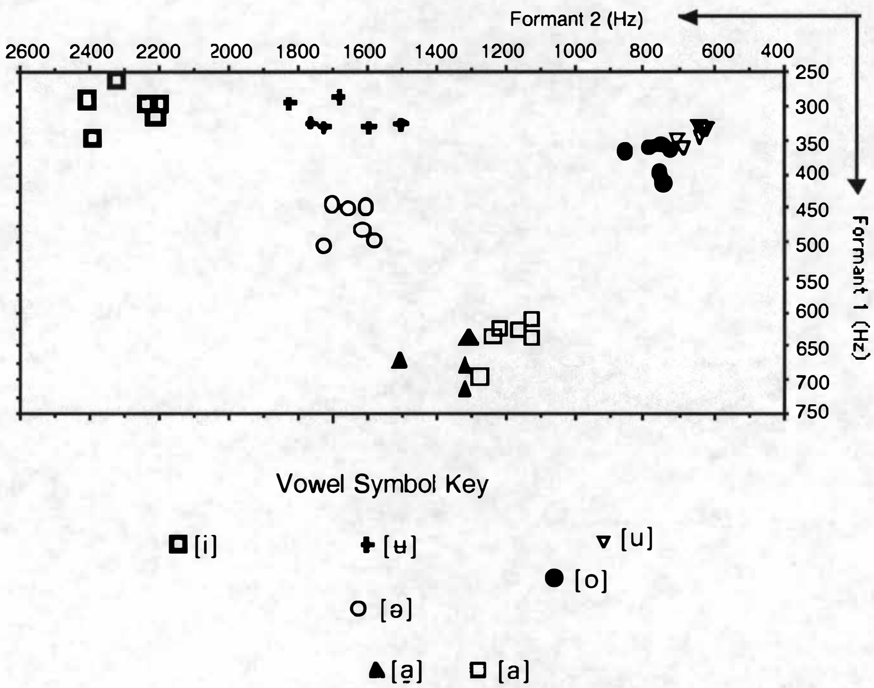


Figure 4.19: Vowel Formant Plot of F_1 against F_2

Table 4.1: Vowel formant values (in Hz). The mean and standard deviation (S.D.) is provided for the first, second and third formants of each vowel. Target tokens are underlined with the exception of the creaky voice vowels, which are identified by a subscript tilde.

Token	**F_1**	**F_2**	**F_3**
[upu̲]	334	624	2553
[u̲pu]	345	644	2648
[təpu̲k	361	689	2523
[u̲pa]	327	648	2595
[u̲pʰuŋ]	350	709	2507
[upu̲]	335	634	2549
Mean	342	658	2512
S.D.	12	33	139

Token	**F_1**	**F_2**	**F_3**
[təməko̲k]	364	724	2594
[tsəŋiko̲]	398	754	2478
[kətseuko̲]	360	744	2415
[atʃo̲k]	415	739	2749
[iko̲]	367	850	2363
[tsəŋiko̲]	361	782	2359
Mean	377	766	2493
S.D.	23	46	153

[əmə̲t]	497	1575	2565
[əmə̲t]	484	1609	2686
[əmə̲t]	452	1656	2707
[mə̲n]	448	1600	2740
[mə̲n]	506	1721	2813
[mə̲n]	446	1697	2791
Mean	472	1643	2717
S.D.	27	58	89

[hʉ̲]	280	1691	2596
[mijəpʉ̲]	325	1607	2493
[anʉ̲ti]	291	1831	2813
[məniʉ̲ʔ˺]	318	1771	2721
[atʃətʉ̲]	323	1734	2781
[əntʃokʉ̲pau̯ʔ˺]	322	1512	2621
Mean	310	1691	2671
S.D.	19	116	122

[təpəti̲]	347	2388	3039
[sənʃi̲]	314	2208	2930
[atʃi̲]	296	2237	2960
[tʃʰi̲pa]	296	2197	2840
[i̲tʃa]	291	2407	3251
[kʰi̲ʔ˺]	261	2323	3201
Mean	301	2293	3037
S.D.	28	92	161

[wa̰ɹ]	668	1508	2238
[wa̰ɹoʔ˺]	639	1309	2369
[wa̰ŋ]	678	1318	2578
[wa̰ʔ˺]	635	1295	2405
[wa̰ʔ˺]	712	1319	2404
[wa̰ʔ˺]	635	1314	2477
Mean	661	1345	2413
S.D.	31	80	113

[wa̲ɹ]	630	1163	2304
[wa̲ɹoʔ˺]	641	1126	2386
[wa̲ŋ]	613	1123	2463
[wa̲]	626	1214	2363
[wa̲]	639	1237	2405
[mama̲]	697	1273	2669
Mean	641	1189	2431
S.D.	29	62	127

4.4.2 Findings

The vowel formant plot of Figure 4.19 demonstrates a pattern of distribution that correlates with the auditory analysis. Firstly, it is observed that there is a slight amount of vowel space overlap of [o] and [u] for their F_1 values but no overlap for their F_2 values, although their respective vowel spaces in the F_2 dimension tightly abut upon each other. A two-tailed t-test was done on the F_1 and F_2 values of [o] and [u] to determine the probability of their constituting significantly different populations. The results are provided at Table 4.2 below.

Table 4.2: Results of a t-test comparing F_1 and F_2 values of [u] and [o]

	F_1 [u]	F_1 [o]	F_2 [u]	F_2 [o]
arithmetical mean	342	378	658	765
standard deviation	12	23	33	46
degrees of freedom	10		10	
unpaired t-value	–3.301		–4.652	
probability	$p < .01$		$p < .01$	

Although the t-test results of Table 4.2 demonstrate that both the F_1 and F_2 values of [u] and [o] are significantly different, no minimal or sub-minimal pairs indicative of a phonemic contrast were encountered in the corpus, and furthermore, tests done with a native speaker, in which [u] was swapped for [o] in words without a change in meaning, demonstrate conclusively that [u] and [o] are in free variation. The significant perceptual cues for the identification of a phoneme /u/ are backness and rounding. These two parameters are given a substantial degree of latitude judging from the broad spread of F_1 and F_2 values, resulting in the auditory perception of the allophone [o] when both values are higher, and [u] when both are lower. The expansive distribution of the formant values of the two allophones is probably attributable to the fact that there are no other back vowels to restrict the vowel space of the phoneme /u/.

As an aid to determining the articulatory position of [ʉ], I recorded myself producing six tokens of the high central rounded vowel in the word [hʉ] 'who' for acoustic comparison with those of WM (I am a native speaker of Australian English). The six tokens were digitised and the formants extracted using the methodology described above in §4.4.1. The mean F_1 value was calculated to be 357 Hz with a standard deviation of 16 Hz, and the mean F_2 value was 1600 Hz with a standard deviation of 25 Hz. If these two mean values are coordinated on the formant plot of Figure 4.20 below, it will be found that the WM high central rounded vowel is marginally higher and more fronted than mine.

A comparison of my production of [a] to that of the WM low central vowel was also made for the same purpose of reference. For this I digitised six tokens of [ha] 'hah' and extracted the formants of the vowel. The mean F_1 value was 777 Hz with a standard deviation of 18 Hz, and the mean F_2 value was 1279 Hz with a standard deviation of 40 Hz. When these two values are coordinated on the scattergram of Figure 4.20, it is found that my low central vowel is a little lower than the WM vowel in the F_1 dimension, but the two are extremely close with regard to backness.

It was previously reported that the phonetic environments of the tokens of the creaky vowel [a̰] are rather restricted due to the low frequency of its occurrence in the corpus. Fortunately a number of minimal and sub-minimal pairs were available in the data so that a comparison of vowels in words that contrast on the basis of phonation type could be made. The mean F_1 value for modal voice [a] is calculated to be 641 Hz with a standard deviation of 29 Hz, as opposed to a mean F_1 of 661 Hz and a standard deviation of 31 Hz for the creaky voice vowel [a̰]. The difference between the mean F_1 values for creaky voice and modal voice low central vowels is therefore 20 Hz.

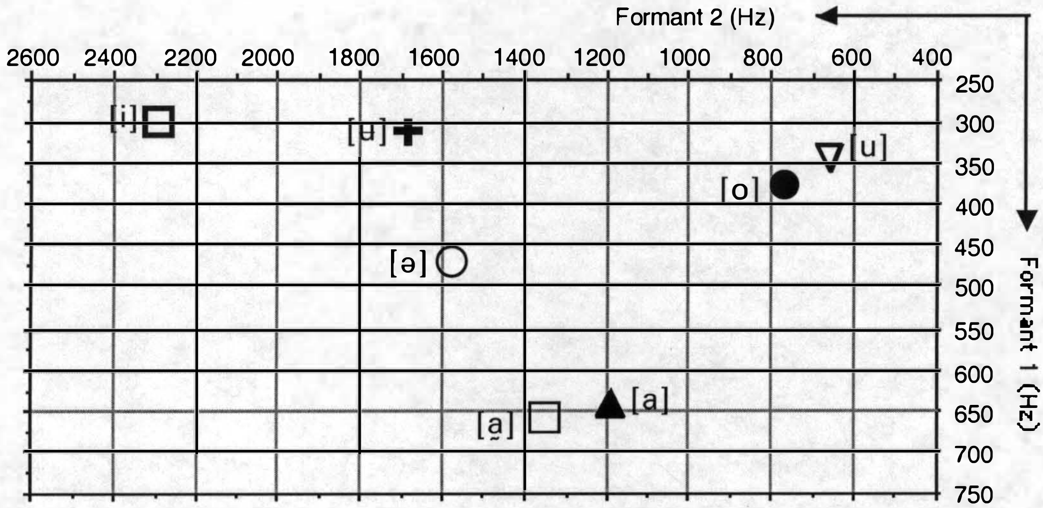


Figure 4.20: Vowel formant plot demonstrating mean formant values of F_1 and F_2

If this small sample based on the data of one WM speaker can be considered to be representative, it would appear that creaky voice phonation slightly raises the frequency of the first formant when it occurs on the low central vowel. In research on phonation types, Ladefoged et al. (1988:298) find that the first formant of the creaky vowels of five Jalapa Mazatec speakers may tend to have a slightly higher frequency than that of modal and breathy phonation types, which they attribute to the raising of the larynx and a consequent shortening of the vocal tract during the production of creaky voice. Formant frequency characteristics, however, were not found to be a reliable measure for the quantification of different phonation types, so the even larger differences between WM modal and creaky voice observed in the F_2 dimension are unlikely to be of significance.

4.5 Acoustic features of creaky voice phonation

The temporally aligned sound waves and wide band spectrograms of Figures 4.21–4.22 below present acoustic data for the comparison of modal versus creaky voice contrasts on the vowels [a] and [a̰] in the minimal pair [waŋ^{33}] 'go-IMPER' and [wa̰ŋ^{33}] 'slice-IMPER'. The speaker is MA; the data was recorded in the recording studio of the Phonetics Laboratory at The Australian National University and digitised using Kay CSL.

The sound wave and temporally aligned wide band spectrogram of [waŋ^{33}] 'go-IMPER' in Figure 4.21 demonstrate regular vibrations of the vocal cords attested by uniform intervals between glottal pulsations consistent with modal voice phonation. This contrasts with the sound wave and temporally aligned spectrogram of [wa̰ŋ^{33}] 'slice-IMPER' in Figure 4.22, in which there are irregular intervals between adjacent glottal pulses, particularly at the onset of the rhyme and continuing for ca. 100 msec. The irregular intervals between glottal pulses are reflected in irregular peak-to-peak distances in the sound wave. There is also more spectral energy present at higher frequencies in the spectrogram of the creaky voice vowel. This is attributable to greater tension of the vocal cords causing the rapid closure of the chords during each glottal cycle, in turn causing sharper pulses and resulting in an increase in excitation of the vocal tract at higher frequencies (Ladefoged et al. 1988:301).

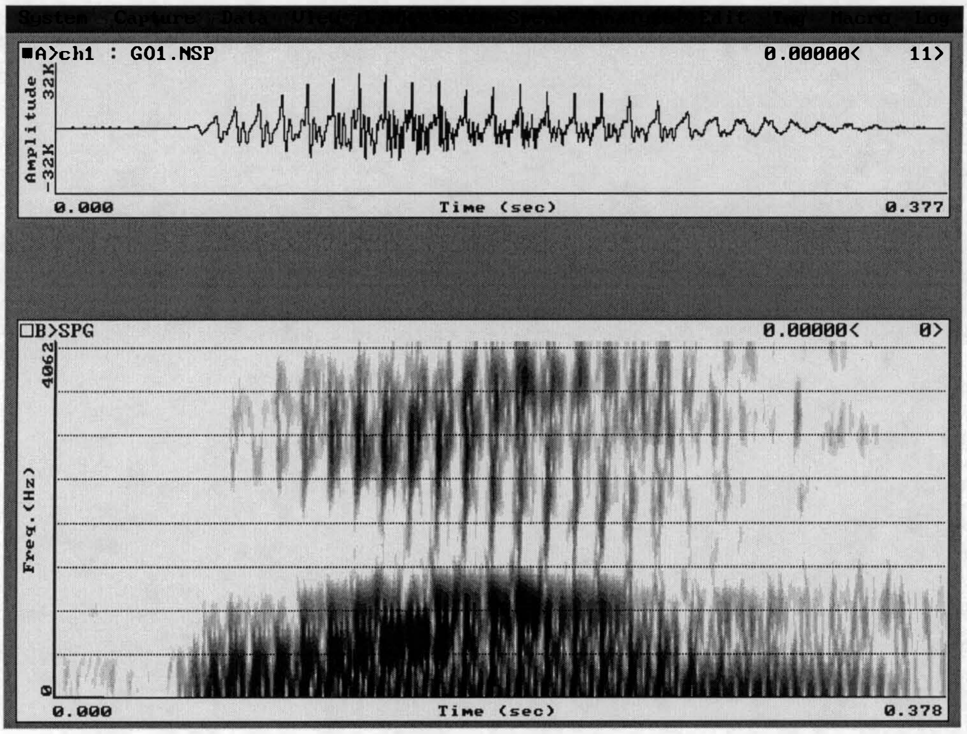


Figure 4.21: Temporally aligned sound wave and wide band spectrogram (293 Hz) demonstrating the acoustic characteristics of modal voice phonation on the vowel [a] in the word [waŋ[33]] 'go-IMPER'.

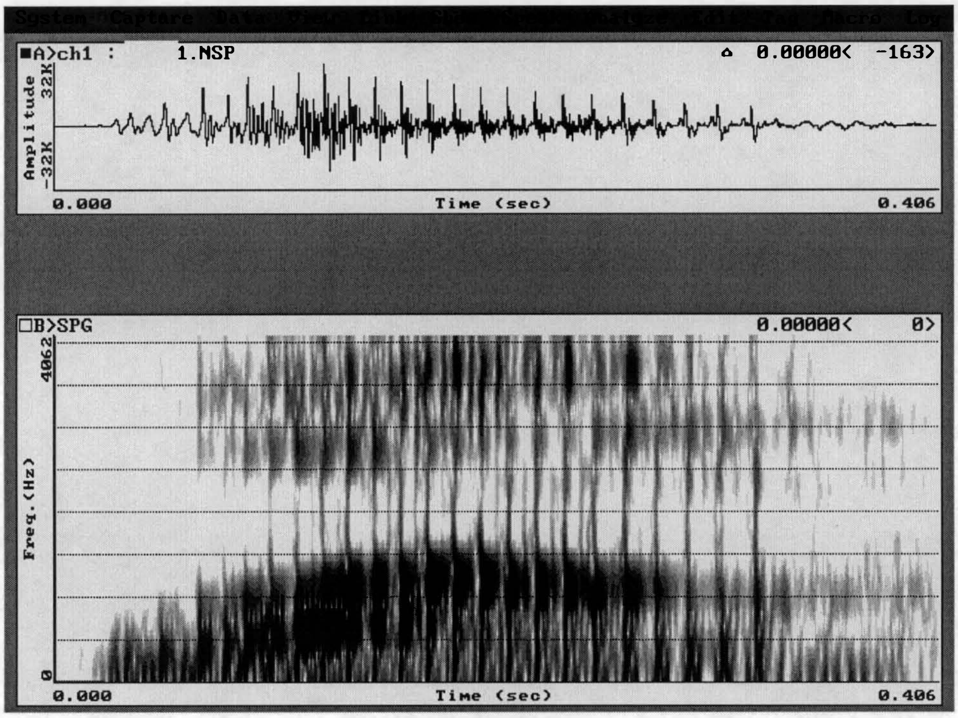


Figure 4.22: Temporally aligned sound wave and wide band spectrogram (293 Hz) demonstrating the acoustic characteristics of creaky voice phonation on the vowel [a̰] in the word [wa̰ŋ[33]] 'slice-IMPER'.

4.5.1 Quantification of jitter

'Jitter' is commonly cited as a salient acoustic feature of creaky voice phonation. Baken (1987:166) defines it as

> ... the variability of the fundamental frequency or, reciprocally, of the fundamental period. When measured during running speech, variability is reflected in pitch sigma [Standard Deviation]. Jitter measurements, however, are concerned with *short-term* variation. That is, jitter is a measurement of how much a given period differs from the period that immediately follows it, and not how much it differs from a cycle at the other end of the utterance. Jitter, then, is a measure of the frequency variability not accounted for by voluntary changes in F_0.

In their work on phonation types in different languages, Ladefoged et al. (1988) report that wide band spectrograms of creaky vowels are marked by uneven vibrations of the vocal cords, producing the irregular striations characteristic of jitter. They also observe jitter in the sound waves of creaky vowels; this is characterised by variations in the intervals between adjacent glottal pulses. In their cross-linguistic investigation, Ladefoged et al. (1988) calculated the mean jitter between glottal pulses for creaky voice versus modal voice vowels. Although there was a considerable degree of variation between the five speakers of the study, the mean value for creaky voice was found to be higher than for modal voice. However, this characterisation of creaky voice is reported to be less useful in distinguishing phonation types cross-linguistically because jitter was not found to be a distinguishing feature of the creaky tone for four speakers of Burmese, one of their languages of investigation.

As the vowels of the words [waŋ³³] 'go-IMPER' and [wa̰ŋ³³] 'slice-IMPER' both carry the same tone and occur in segmentally identical frames, they constitute appropriate data for the illustrative quantification of jitter in modal voice versus creaky voice phonation in WM. The raw data for this comparison was obtained by manually measuring the period in milliseconds between adjacent peaks, with the starting point for measurement being the beginning of the rhyme. This was determined from the sound wave and spectrographic evidence of formant energy consistent with the vowel [a], viz. an F-pattern demonstrating acoustic energy at approximately 700 Hz, 1200 Hz and 2400 Hz. A comparable duration for each vowel was used: 102 msec for the modal voice vowel and 103.2 msec for the creaky voice vowel. This provided a total of nine peak-to-peak measurements for each phonation type.

The simplest method for determining the extent of jitter in different phonation types is to compare the mean and standard deviation of jitter occurring between the glottal pulses of modal voice and creaky voice vowels. Values for all mean and standard deviation calculations of the data are given at the bottom of Table 4.3 below. The significant measure for jitter is the standard deviation, which is found to be 0.3 msec for the modal voice vowel and a considerably greater 1.5 msec for the creaky voice vowel. These findings correlate with differences between creaky and modal phonation reported by Ladefoged et al. (1988) for Jalapa Mazatec, although they differ in scale. The mean and standard deviation differences between phonation types for the data of one speaker of WM are not nearly as great as those observed in Jalapa Mazatec.

Baken (1987:168–188) gives a description of various measurement methods used for the quantification of frequency perturbation. These are actually proposed for the quantification of laryngeal pathology rather than for the phonetic description of non-pathological phonation types. Nevertheless, a number of different methods can be employed to measure the extent of jitter for short durations in the non-pathological setting. As we are not so much concerned with attempting to establish anything like a 'perturbation factor' that might otherwise be useful for distinguishing different manifestations of laryngeal pathologies (Lieberman 1963:353, cited in Baken 1987:173) only those measurement methods judged appropriate for the descriptive linguistic task at hand will be applied to the WM data.

One useful procedure is loosely based on the method of magnitude of durational differences used by Lieberman (1961, 1963) and Iwata and Von Leden (1970) (cited in Baken 1987:169). It plots durational change in milliseconds between successive cycles, as what is most illuminating in determining the extent of difference between modal voice and creaky voice is the individual differences in duration (and therefore F_0) between adjacent cycles. The absolute differences in duration between successive periods were calculated by the following formula:

$$\Delta t = t_n - t_{n-1}$$

where Δt is the absolute difference between the duration of a period n and the duration of a preceding period n–1 (Baken 1987:170).

F_0 was also determined from the duration of the period by using the standard formula for calculating the value in Hertz, being

$$F_0 = 1/t$$

where t is the duration of one period.

Table 4.3: Cyclic duration and calculated F_0 differences between successive cycles in modal and creaky voice phonation

	Modal Voice			Creaky Voice		
Cycle No.	**Period (msec)**	**$t_n - t_{n-1}$ (msec)**	**F_0 (Hz)**	**Period (msec)**	**$t_n - t_{n-1}$ (msec)**	**F_0 (Hz)**
1	11.4		87.7	12		83.3
2	10.8	(–) 0.6	92.6	8.4	(–) 3.6	119
3	11.4	(+) 0.6	87.7	12	(+) 3.6	83.3
4	11.4	0	87.7	11.4	(–) 0.6	87.7
5	11.4	0	87.7	12	(+) 0.6	83.3
6	11.4	0	87.7	11.4	(–) 0.6	87.7
7	11.4	0	87.7	12	(+) 0.6	83.3
8	11.4	0	87.7	12	0	83.3
9	11.4	0	87.7	12	0	83.3
Mean	11.33	0.15	88.24	11.47	1.2	88.24
S.D.	0.2	0.28	1.63	1.18	1.5	11.69

Table 4.3 above compares the period durations, jitter, and F_0 variation of two phonation types for nine successive glottal periods. Glottal periods are numbered from 1 to 9 in the leftmost column, and the algebraic signs (±) indicate the direction of perturbation differences between adjacent periods. The starting point for measurement of durational differences is therefore the second cycle. In terms of F_0, the data are certainly well-matched — the mean is identical — thus excluding one variable that might exert an influence on

peak-to peak duration variation. To give an idea of the cyclic nature of duration perturbation in creaky voice, a perturbation measure proposed by Hecker and Kreul (1971, cited in Baken 1987:173) can be used to determine the 'directional perturbation factor' (DPF). This quantification method is not concerned with the size of the period perturbation. Instead it counts the number of times that successive periods change sign (±); that is, diminish or increase the duration of their period relative to a preceding period. When applied in the speech pathology setting, the DPF sign change count is divided by the number of differences in a sample of data. This derives a percentage value for comparison to a tentative norm.

I have adopted part of this methodology (i.e. the sign changes) because it can be used to show graphically the durational characteristics of WM creaky voice as opposed to modal voice. The values for durational differences ($t_n - t_{n-1}$) provided at Table 4.3 are plotted in Figure 4.23 below.

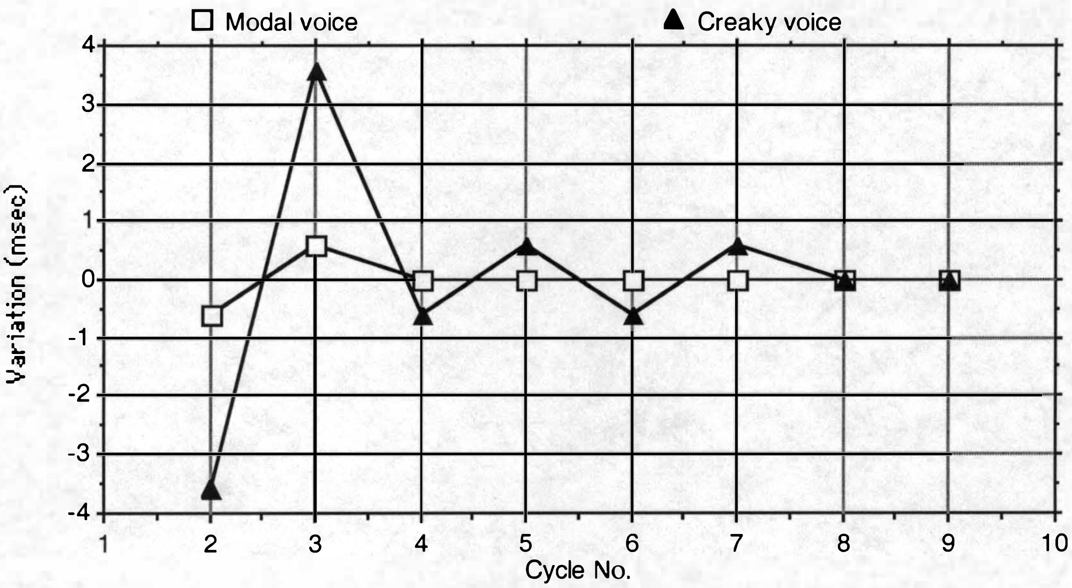


Figure 4.23: Period-to-period differences (in msec) for modal and creaky voice phonation over ca. 100 msec

Cycles 2 and 3 show durational variation in both the modal and creaky voice examples. The magnitude of the change from a negative to positive value for creaky voice, however, is much greater than for modal voice, and these duration differences might be interpreted as being in a relative relationship to the degree of variation seen in successive cycles for both phonation types. After two periods of initial instability, the modal voice durations for successive cycles become consistently regular. This contrasts strongly with the creaky voice example, with its alternating long and short pattern.

Jitter is consistently observed in the spectrograms of all creaky voice vowels occurring in the corpus, which therefore suggests that this feature may be used as a reliable acoustic cue for the identification of creaky voice in WM.

4.5.2 Summary of findings

It would appear that the principal acoustic correlate of creaky voice for one native speaker of WM is jitter. The extent to which this acoustic feature is generally representative of creaky voice phonation in the language cannot be determined until the data of more individual speakers is investigated and the results collated, as was done in the study of Ladefoged et al. (1988). Nevertheless, for one native speaker, two discrete phonation types are demonstrated to form the basis of a phonemic contrast at one articulatory position, and differences between creaky voice and modal voice have been shown to be quantifiable by instrumental means. With the data of more speakers, creaky voice might be further examined to determine the best methods of quantification both within WM and cross-linguistically, and to enhance our understanding of the acoustic characteristics of this phonation setting from a typological perspective.

This concludes an instrumental study of some aspects of acoustic phonetics in WM. In the following chapter I give a description of the tonal system of WM based on an auditory analysis and present the results of a perception test carried out with a native speaker.

5 *Tonal phonology*

5.1 Introduction

This chapter describes the tonal system of WM and has the following organisation. Section 5.1.1 provides an overview of previous research on the tonal systems of the KCN languages of north-east India. Section 5.2 describes the methodology and findings of the auditory analysis of monosyllabic words. Section 5.3 describes the analysis of disyllabic words. Section 5.4 examines the features of the falling pitch. The chapter concludes with a description of two perception tests in §5.5 and an interpretation of the results.

5.1.1 Previous research on tone in Kuki-Chin-Naga languages

Very little is known about the tonal characteristics of Tibeto-Burman languages spoken in north-east India. The first few descriptions, done in the latter part of the nineteenth century and the first half of the twentieth century, by and large tended to ignore tonemic contrasts in languages that have subsequently been reported to have lexical tone. This of course assumes that lexical tone in these languages is not an innovative development of the intervening period. Regrettably, a number of grammatical sketches done more recently on KCN languages of the north-eastern region continue the tradition of giving tonal phenomena short shrift, usually devoting no more than a few cursory lines to its description.

To the best of my knowledge there are only two previous acoustic investigations of KCN languages. Blakenship et al. (1993) present an acoustic description of the Khonoma dialect of Angami; their paper briefly describes the segmental phonemes, voice onset time of stops, and tones, with a main focus on the acoustic characteristics of the voiceless nasal series. Chelliah (1997) gives a short acoustic description of tone in her grammar of Meithei (Manipuri), in which minimal tone pairs were digitised and their fundamental frequencies determined by an automatic pitch extraction algorithm. The pitch shapes of the minimal pairs were then individually plotted on graphs. The results demonstrate a correlation of the acoustic features of tone to an auditory analysis identifying two contrastive pitches.

A number of auditory descriptions of tone have been done on KCN languages. Burling (1960) provides a brief description of the tone system of Angami, and Kapfo (1989) examines tone sandhi phenomena in Khezha, a language closely related to Angami and spoken in southern Nagaland and Manipur. Weidert (1979, 1987) provide the most comprehensive descriptions of KCN tonology to date. These works posit tonal categories arising out of phonation types to account for the development of lexically contrastive tone in subgroups of Tibeto-Burman. Presumably Weidert's account of Tibeto-Burman tonology is largely based upon his own extensive auditory analyses of the tones of Tibeto-Burman languages; unfortunately, detailed descriptions of the individual tonal systems of KCN and other languages, including the methodology by which they were validated, were never made available in print. His publications are concerned with the establishment of tonal correspondences and tonogenetic reconstruction, and consequently give only minimal descriptions of synchronic tonal systems in the many languages examined.

Next to nothing is known about the tone system of Mongsen Ao. This book presents the first detailed auditory and acoustic description of the language.

5.1.2 Data sources

The tonal analysis is based on the citation forms of two speakers (CL and MA). The corpus of CL was recorded under non-laboratory conditions in Nagaland on a portable Sony TCM-5000EV tape recorder with an external Sony ECM-FO1 electret condenser microphone. The corpus of MA was recorded in the recording studio of the phonetics laboratory of the Department of Linguistics (Arts) at The Australian National University. The latter corpus was expanded with recordings done in Sydney using the portable Sony equipment. The elicited data of the two corpora were also compared with texts and additional word lists of native speakers recorded in Nagaland. Different corpora were used exclusively for certain aspects of the tonal analysis. After all the data at hand were transcribed and considered, only CL's corpus was used for the auditory analysis of the disyllables presented in §5.3, and only MA's corpus was used for the perception tests described in §5.5.

5.2 Auditory analysis of monosyllables

5.2.1 Procedure

The main consultants (MA and CL) were given a Swadesh word list of approximately two hundred items and requested to give three repetitions of each word in WM. The original word list that I recorded in Australia with MA revealed that pitch had a lexically contrastive function in the language, as demonstrated by the first minimal pair encountered — [tsə^{55}ŋi^{55}] 'rain' and [tsə^{33}ŋi^{33}] 'sun' — and eventually by other more extensive minimal sets of words that were either encountered in elicitation or offered by consultants.

MA had a naturally low fundamental frequency (hereafter F_0) and a considerably narrow pitch range resulting in very small pitch differences between adjacent tones. He also had a naturally creaky voice. These two factors combined to make it extremely challenging firstly to identify the pitch level of elicited words, and secondly to distinguish phonemic creaky voice from non-phonemic creaky voice. It was reported in Chapter 3 that the creaky voice phonation type only occurs contrastively on the low central vowel, but this could not be established on the basis of MA's data alone. Because of these combined difficulties, the auditory analysis of tone was postponed until I had obtained the elicited data of another native speaker (CL) recorded in Nagaland. Once perceptually significant differences between pitch levels could be established for CL's data, the tonal transcriptions were compared to those of MA's data. The following methodology was then employed to determine the maximum number of contrastive tonemes and the pitch level of each elicited word in the corpora of both speakers.

The data were transcribed phonetically and each syllable was assigned a pitch value according to auditory impressions. Initially, Chao letters (Chao 1930) were used to transcribe tones and each syllable was assigned integers of 1 to 5, with 1 representing the lowest level. As the lexical tones are heard to be quite consistent in pitch height over their durations, generally two integers were found to be more than sufficient to characterise the pitch shape of each syllable. An exception to the level nature of lexical pitch was demonstrated by the only conjunction elicited in the word list of CL, namely [təɹ^{53}] 'and', which has a high-to-mid falling contour. The results of an investigation into the perceptual features of this falling pitch are presented in §5.4 below.

Insofar as it was possible, the data were separated into syllable types to control for rhyme duration and the intrinsic effect of consonant segments on pitch. Ideally, a tonal analysis should also control for the effect of intrinsic pitch of vowels, as F_0 is known to vary with vowel quality (Lehiste 1970:68ff.); however, there was an insufficient number of examples for all three syllable types and tonal groups (described below) to control for this and to still produce representative results. Nevertheless, the corpus listed in Appendix A demonstrates a fairly even distribution throughout each group, therefore it is unlikely that the results have been significantly distorted by the influence of any particular vowel quality's intrinsic pitch.

At the first approximation, four phonetic pitch levels were identified in the data of CL, but no minimal quadruplet could be offered by either CL or MA to demonstrate four *contrastive* pitch levels. Nor was there any distributional evidence of segmentally conditioned allotony that might have indicated less than four contrastive levels. Pike (1948:49) recommends dividing words of the corpus into uniform groups as an aid to tonal analysis, so to check the accuracy of my initial pitch approximations, all examples of each phonetic pitch level were sorted according to the syllabic structure of the word in which they occurred; that is, whether mono, di-, or tri-syllabic, whether stopped or unstopped and, if the latter, whether the coda was filled by a nasal or unfilled (a syllable with a nasal coda will henceforth be referred to as a nasal syllable, after Rose (1992)). Then the members of each tonal group were respectively dubbed from the master tape onto a second tape for comparison.

When there are relatively small differences between pitch levels and no distinctive contours or durations by which to identify a particular toneme as is possible in Thai, for example, with its low-dipping-rising and high-falling contours or short durations on checked syllables (Coupe 2001), hearing all the representatives of each tone group in this manner assists one to identify any inaccuracies in the pitch transcriptions. Aberrations were subsequently re-dubbed with words of the same pitch level and the transcriptions adjusted until consistency between the pitch levels and transcriptions of all members of each tone group was established. Another advantage of this method of sorting according to tone and syllable type is that it can help to identify segmentally conditioned allotony if it exists in the system.

Despite identifying four pitch levels in the data of CL, only three pitch levels were heard in MA's data. No minimal triplet of monosyllabic words differentiated solely on the basis of three lexically contrastive tonemes could be found in the corpus, but sets of disyllabic words constituting minimal tonemic contrasts are attested. Minimal sets of disyllabic words demonstrating three contrasting tonemes were also readily proffered by both native speakers; some examples are provided below in (5.1–5.2). It was noted in Chapter 2 that disyllabic words constitute the dominant pattern in the corpus. This perhaps explains why minimal contrasts are more easily found among disyllabic words than among monosyllabic words.

(5.1)	/tə̄máŋ/	[tə^{33}maŋ^{55}]	'all'
	/tə̄-māŋ/	[tə^{33}maŋ^{33}]	'NPF-body'
	/tə̀-māŋ/	[tə^{11}maŋ^{33}]	'PROH-believe'
	/tə́-màŋ/	[tə^{55}maŋ^{11}]	'NPF-dark'
(5.2)	/tsə́ŋí/	[tsə^{55}ŋi^{55}]	'rain'
	/tsə̄ŋī/	[tsə^{33}ŋi^{33}]	'sun'
	/tsə́ŋì/	[tsə^{55}ŋi^{11}]	'wild dog'

Once the data from both speakers were compared together with the evidence of minimal pairs, it became apparent that the fourth level in the data of CL was phonetically salient but phonemically non-contrastive, perhaps resulting from unconscious shifts in pitch range on the part of the speaker during the elicitation session, or from extreme phonetic variation in pitch height occurring within the range of the high toneme and subsequently heard as a distinctive pitch level.

Since I could not account for CL's extra level and a fourth tonemic contrast could not be demonstrated, a decision was made to treat the two phonetic levels I heard at the upper registers of CL's pitch range as being representative of just one toneme. Syllables transcribed as having pitch levels of [44] and [55] were subsequently all reanalysed as having a pitch level of [55]. Obviously, if these two tonal groups had not been merged into one, I would have had to account for the high toneme having two very close allotonic realisations, i.e. [44] and [55], whose realisations were not seen to be phonetically motivated, and which only occurred in the data of one speaker. Alternatively, positing that pitches [44] and [55] occur in free variation would have been equally as unappealing.

A third word list was also recorded in Nagaland on my behalf by CA, a tertiary educated, female native speaker of WM in her early thirties. This set of data is strikingly divergent from the corpora of the two male speakers because it does not appear to demonstrate lexically contrastive pitch consistently. For many of her citations, CA tended to produce a rising pitch on the first two tokens, followed by a falling pitch on the last token, somewhat in the intonational, list-reading manner of speakers of non-tonal languages. For example, of three tokens of the word 'bark.PAST', the first two tokens had a dipping-rising contour [tsəŋ^{325}], but the third token was completely different again with a mid-falling contour [tsəŋ^{31}]. None of these pitch contours is even remotely similar to the pitch shapes produced by MA and CL for the same citation form uttered in isolation, as both male speakers pronounced the word with a low level pitch, viz. [tsəŋ^{11}].

A number of inferences could be drawn from the peculiarities of CA's corpus: (a) WM does not consistently use pitch to signal lexical contrasts, since in elicitation one speaker did not faithfully pronounce words with a consistent pitch level — WM might therefore be evolving into an intonation language through a process of tonal decay and loss of lexically contrastive tone; or (b) WM is a tonal language but pitch is only used contrastively by some speakers in a stream of speech; that is, in words spoken in isolation, pitch provides redundant information; or (c) WM is a tonal language and pitch is contrastive in elicited words spoken in isolation, but in the data of one speaker, I was unable to determine the parameters by which it was manifested.

While CA's corpus is certainly problematic for this analysis, I assume that the absence of tonal contrasts in her elicited data is more likely to be a feature of her idiolect rather than being representative of the language generally. The assertion that WM is a tonal language for at least some speakers of WM is firmly founded upon the numerous minimal sets of words in which lexical contrast is consistently and unequivocally signalled by pitch, and by the results of perception tests. Furthermore, pitch levels on words uttered in isolation by both MA and CL show a systematic uniformity that would be altogether absent in a non-tonal language. Less convincing evidence but also worthy of mention is the observation that in running speech, all speakers appear to aim for specific pitch targets on each syllable of each word. This contrasts strongly with an intonation language such as English, in which intonations are much more sparsely distributed over phrases and are highly variable.

It would certainly be useful to record sets of minimal contrasts uttered by CA, firstly to see how the tonal contrasts produced by the other speakers in this analysis are manifested in her speech, and secondly, to devise a perception test so as to determine if any of the inferences advanced above can be substantiated. A survey of a sizeable number of native speakers might also be used to determine to what extent contrastive pitch is or is not a characteristic of words uttered in isolation.

Lastly, it is relevant and perhaps revealing to note and that MA and CA are members of the same family and are both fluent speakers of English, so to some extent they have been subject to the same linguistic influences shaping the way they speak Mongsen. A comparison of the word lists in Appendix B demonstrate that while MA's pitch realisations on elicited words show much more concordance with CL's data (as opposed to CA's), there is still considerable deviation in these two speakers' individual pitch realisations for the same words. Subsequent research done with over a dozen native speakers of other varieties of Mongsen clearly demonstrate that MA's pitch realisations, particularly on the word-final syllables of elicited di- and trisyllabic words, deviates substantially from the Mongsen norm.

5.2.2 Results

The auditory analysis identifies three lexical tonemes occurring on all six vowels and on all syllable types. Phonetically, the high toneme is realised at a pitch level of [44] to [55]; the mid toneme has its focus at [33], but may be as low as [22] or as high as [44]; and the low toneme is realised at a pitch level ranging from [11] to [22]. For reasons of descriptive economy argued for in §5.2.1 above, only one allotone is proposed for each toneme.

(5.3)	High:	/´/	→	[55]
	Mid:	/¯/	→	[33]
	Low:	/`/	→	[11]

The following examples of monosyllabic words demonstrate that the three lexical tonemes are not restricted to specific syllable types, nor to initial consonant types, nor to specific word classes.

(5.4)	Stopped syllables:		
	/High/	/áiʔ/	'blood'
	/Mid/	/tʃʰīt/	'eight'
	/Low/	/tʰə̀p/	'throw.PAST'
(5.5)	Nasal syllables:		
	/High/	/tsə́ŋ/	'stab.PAST'
	/Mid/	/sə̄ŋ/	'fill.PAST'

/Low/ /nàŋ/ 'you' (SG)

(5.6) Unstopped syllables:

/High/ /ái/ 'dog'
/Mid/ /sə̄/ 'die.PAST'
/Low/ /nì/ 'I'

The suprasegmental phonology of WM demonstrates minimal pitch height differences between contrasting levels and great homogeneity in pitch shape. Unlike contour tone languages, in which the pitch shape contributes an important cue to the identification of a particular toneme, the level terraces of a register tone language offer precious little in the way of cues to aid in the identification of contrasting tonemes. The only perceptual cue for a toneme might be its pitch height relative to the other tonemes in the system, so that a slight shift upward or downward in key can easily result in confusions with adjacent pitch levels for the transcriber. Morse (1963:35) reports a similar system in Rawang, in which he observes that '[t]one is entirely a matter of contrastive and relative pitch. Thus if a high tone and low tone were pronounced one musical octave apart or only a fraction of a tone apart, it would make no difference as along [sic.] as the difference was contrastive', and in the same passage he proposes that '[a] syllable pronounced in isolation would thus be entirely ambiguous, until another syllable was pronounced to provide a contrastive referent'. Whilst the former statement is a valid observation, the latter claim is shown on empirical grounds to be incorrect for WM, because once one becomes accustomed to a speaker's pitch range, a word uttered in isolation is much more accurately identifiable; this is suggested by the high rate of intelligibility of isolated utterances in the perception tests discussed in §5.5 below.

The pitch levels of WM are not absolute and may therefore overlap; in the absence of phonological cues such as distinctive pitch contours or salient differences in duration, what is held to be phonemically significant is the pitch level relative to that of preceding and/or following syllables. This is because WM has downdrift, a pitch lowering process whereby a phonemically high pitch, for example, tends to be realised at a phonetically lower level towards the end of a sentence relative to its level at the beginning, as the following narrow phonetic and phonemic transcriptions of pitch demonstrate.

(5.7) [a^{55}luŋ^{55} tsə^{55} rak^{11}li^{11}]
á-lúŋ tsə́ ràklì
NPF-stone DET move.PAST
'The stone moved.'

(5.8) [p^{h}ə^{55}niŋ^{55}nok^{55} na^{33} a^{44}luŋ^{44} tsə^{44} rak^{11}li^{11}ʉʔ^{22}]
p^{h}əníŋnúk nā á-lúŋ tsə́ ràklì-ʉ̄ʔ
earthquake AGT NPF-stone DET move-CAUS.PAST
'The earthquake caused the stone to move.'

In (5.7) both syllables of /á-lúŋ/ 'NPF-stone' are heard to have a phonetic pitch level of [55]; in (5.8) /á-lúŋ/ is subject to downdrift due to its position of occurrence in the utterance. While the tonemes of its syllables are still phonemically /High/, they are realised phonetically at the lower pitch level of [44]. Similarly, the causative has an underlying /Mid/ tone giving it a pitch level of [33] usually, but because of its position of occurrence at the end of the utterance it is subject to downdrift and consequently realised at the lower pitch level of [22].

5.3 Auditory analysis of disyllables

CL's corpus of disyllables was transcribed and checked using the procedure outlined for the analysis of monosyllables in §5.2. Once again, the size of the corpus did not allow for control of the effect of intrinsic pitch of vowels, but it did contain a sufficient number of disyllabic words to establish three representative categories on the basis of their segmental and structural uniformity. Words were sorted into two groups according to the structure of their final syllables; this divided the corpus into stopped and unstopped disyllables. The unstopped disyllables were further divided into those with medial sonorants and those with medial voiceless obstruents, thus yielding the following three groups of words having a (C)VCV(C) disyllabic structure:

Category (i): stopped disyllables with a medial sonorant e.g. [to^{33}ŋət^{33}]
Category (ii): unstopped disyllables with a medial sonorant e.g. [tə^{33}ləm^{33}]
Category (iii): unstopped disyllables with a medial voiceless obstruent e.g. [tʃʰi̵^{11}pa^{11}]

In a Swadesh list of approximately 200 lexical items constituting independent words, di- and trisyllabic words were found to comprise 79% of the data (62% and 17% respectively). Of the disyllables, the majority (83% of total disyllables) maintain consistency in pitch height over their two syllables. This allowed each of the above-mentioned categories to be subdivided according to pitch level into three representative tonal groups — /High-High/, /Mid-Mid/ and /Low-Low/. The three tonal groups of each category were later used for the acoustic analysis of disyllables, the results of which are presented in §6.3 of Chapter 6. Other tonal permutations in disyllabic words occurred much less frequently in CL's data; e.g. /Low-Mid/ (10% of total disyllables), /High-Mid/ (3% of total disyllables), /High-Low/ (<2% of total disyllables) and /Mid-High/ (<2% of total disyllables). There were no occurrences of disyllabic words with a /Low-High/ or /Mid-Low/ pattern. The attested sequences demonstrate that there is a strong tendency to avoid differences in adjacent pitch levels in poly-syllabic words, particularly if those differences are extreme, although this cannot be posited as an infallible rule because there are occasional exceptions encountered in disyllabic roots in the corpus, e.g. /tə́-màŋ/ 'NPF-dark', (High-Low) and /á-lù/ 'NPF-field' (High-low). A good number of the di- and trisyllabic words with different pitch levels are composed of a root plus one or more affixes that have become lexicalised with their own specific tones, e.g. /tāmáɹū/ 'bad', transparently formed from the nominal prefix /tə̄-/, the negative prefix /mə̀-/ and the verb root /āɹū/ 'good'. Alternatively, they may be formed from compounds that have subsequently become absorbed into the lexicon e.g. /tə̀psə̄t/ 'kill.PAST', from /tə̀p/ 'hit' plus the fossilised suffix /-sə̄t/ 'do to death', so if the majority of these counter-examples can be accounted for by lexicalisation processes, then WM is found to be a language with an overwhelming preference for all the syllables of its polysyllabic words to share the same pitch level.

5.4 The falling pitch

In addition to the three level tonemes identified in the corpus, a falling pitch is consistently found to occur word-finally on verb stems marked by the narrative converb suffix /-ə̂ɹ/, and it is also found on the one conjunction occurring in citation form in CL's corpus, viz. /tə̂ɹ/ 'and'. The falling pitch is marked separately with a circumflex accent in phonemic transcriptions, although for reasons discussed in §3.7 of Chapter 3 it is more convenient to treat it phonologically as an underlying sequence of /High/ and /Low/ tonemes mapped onto one syllable, rather than to accord it independent tonemic status; another compelling reason for denying it tonemic status is its highly constrained distribution. Since both the narrative converb suffix /-ə̂ɹ/ and the conjunction /tə̂ɹ/ have a similar form and function, it is highly likely that they stem from the same diachronic source, with the clause conjoining function of /tə̂ɹ/ the consequence of a morphological reanalysis of the narrative converb and its clause-final distribution.

The following minimal pairs demonstrate the contrastive function that the falling pitch plays in disambiguating the narrative converb suffix from the segmentally identical present tense suffix.

(5.9)	/jàk/	'beat'	+	/-ə̀ɹ/	PRES	→	/jàk-ə̀ɹ/	'beat-PRES'
	/jàk/	'beat'	+	/-ə̂ɹ/	NCV	→	/jāk-ə̂ɹ/	'beat-NCV'
	/tʃàʔ/	'eat'	+	/-ə̀ɹ/	PRES	→	/tʃà-ɹ/	'eat-PRES'
	/tʃàʔ/	'eat'	+	/-ə̂ɹ/	NCV	→	/tʃâ-ɹ/	'eat-NCV'

The verb roots of the example words in (5.9) above have underlying low tones that are subject to sandhi perturbations when taking certain suffixes. In the examples formed with the verb root /tʃàʔ/ 'eat', suffixation also creates the phonetic conditions under which the glottal stop is deleted word-internally; this in turn results in the deletion of the schwa of the present tense and narrative converb morphemes when juxtaposed with the more stable low central vowel of the verb root. Vowel deletion of this nature involving suffixes was described in §2.5.2 and circumstances resulting in the deletion of the glottal stop were discussed in §2.7.

Recall that features introduced by Halle and Stevens (1971) are used in §3.8 to decompose the falling pitch into /High/ and /Low/ components. The phonological representation of the falling pitch and the distinctive features matrix for tonemes are repeated below for convenience, renumbered as (5.10) and Table 5.1 respectively.

(5.10)

$$\begin{array}{cc} \text{H} & \text{L} \\ \begin{bmatrix} +\text{stiff} \\ -\text{slack} \end{bmatrix} & \begin{bmatrix} -\text{stiff} \\ +\text{slack} \end{bmatrix} \\ \multicolumn{2}{c}{\sigma} \end{array}$$

Table 5.1: Distinctive features matrix for tonemes

	/High/	**/Mid/**	**/Low/**
stiff	+	–	–
slack	–	–	+

From the examples of (5.9) above, we might surmise that the initial /High/ component of the falling pitch marking the narrative converb suffix spreads leftward if it is not blocked by the presence of a root-final consonant, thereby displacing the underlying tone of the verb root (as in /tʃâ-ɹ/ 'eat-NCV'), or else it causes the erstwhile tone of the root to undergo a partial assimilation in pitch height (as in /jāk-ə̂ɹ/ 'beat-NCV'). While acknowledging that many more examples are required to substantiate any interpretation, this does seem to support the decomposition of the falling pitch into a sequence of a /High/ and a /Low/ toneme, as it is only the immediately adjacent /High/ component that exerts an effect on the tone associated with the root. Phonetically, the falling pitch is realised as [53] for the conjunction /tə̂ɹ/ 'and', and as [53]~[31] for the narrative converb suffix. Both demonstrate similar contours when their F_0 values are plotted as a function of duration (cf. the F_0 contours of Figure 6.10), so irrespective of F_0 height at rhyme onset, /High-Low/ serves to represent phonologically all phonetic realisations of the falling pitch.

In addition to marking the contrast between the narrative converb suffix and the segmentally identical present tense marker, arguably the falling pitch has a primary prosodic clause boundary marking function, since both the narrative converb suffix and the falling pitch are constrained to occurring word-finally at the end of subordinate clauses. The following example taken from a text demonstrates this distribution:

(5.11)	*wāɹū*	*tsə̄*	*pʉ̄*	*ɹā-jāʔ*	*ūŋ-əˆɹ*	*wá*	*táŋ*	*tʃʰâ-ɹ*
	crow	DET	PROX	come-PROG	see-NCV	ONOM	just	do-NCV

	jīm-ā	*wā-jūk*
	fly-PCV	go-PERF.PAST

'The crow saw it coming and just cried out "Caw!" and flew off.' (lit. 'The crow seeing it coming, and just crying out "Caw!" flying went away.')

Because of the nature of its distribution and this highly specific boundary marking function, the prosodic domain of the contour tone is best viewed as embracing the entire (subordinate) clause, as opposed to the level pitches of the lexical tones, whose prosodic domain is limited to just their respective syllables. Or to borrow an analogy that Chao (1970:39) offers as an eloquent explanation for the simultaneous existence of tone and intonation in Chinese, the relationship of the level tones of syllables to the contour tone of clauses subordinated by the narrative converb suffix can be equated with 'small ripples riding on large waves'.

5.5 Perceptual analysis

To determine to what extent the tonemes of WM are identifiable to a native speaker, I carried out a perception test with MA using his previously recorded data. A perception test is a valuable means of determining the number of distinctive tonal contrasts that exist in a language. Responses can be used to assess the rigour of the phonological analysis, and confusions can serve to identify commonalities between tonemes that must exist for those confusions to occur. Perceptual testing is therefore an essential component of the auditory analysis of any tone language.

5.5.1 Procedure

The preparation for the perception test was as follows. From MA I elicited tokens of target words constituting sets of minimal tonal contrasts. These were then dubbed in a quasi-random fashion onto a second tape, interspersed with randomly selected words from MA's previously recorded corpus. The end product was a test tape of two hundred and twenty-five words containing nine tokens of each target word. The target words were recorded one month before the perception test was administered, and MA was deliberately kept ignorant as to the purpose and nature of the test. The test tape was played and his response to each test word was noted. He was permitted to replay any words as many times as required, and the entire session was carried out in one sitting at a pace of his choosing. Two sets of minimal contrasts were used for the perception test. These were listed above at (5.1–5.2) and are repeated below for convenience.

Set 1

/tə̄máŋ/	[tə^{33}maŋ^{55}]	'all'
/tə̄-mā̄ŋ/	[tə^{33}maŋ^{33}]	'NPF-body'
/tə̀-mā̄ŋ/	[tə^{11}maŋ^{33}]	'PROH-believe'
/tə́-màŋ/	[tə^{55}maŋ^{11}]	'NPF-dark'

Set 2

/tsə́ŋí/	[tsə^{55}ŋi^{55}]	'rain'
/tsə̄ŋī/	[tsə^{33}ŋi^{33}]	'sun'
/tsə́ŋì/	[tsə^{55}ŋi^{11}]	'wild dog'

I add the caveat that these perception tests were based on the recorded data of just one speaker, and that each of the test words was dubbed onto the test tape nine times. Under ideal conditions it would be more revealing to use at least five or six individuals to record multiple elicitations of each test word, so as to see how well an individual copes with between- and within-speaker variation.

5.5.2 Results of first perception test (Set 1)

The results of the first perception test using the test words of Set 1 are presented in the following confusion matrix of Table 5.2, and their individual F_O shapes are provided in the line chart of Figure 5.1. A description of the methodology by which the F_O shapes and durations of the tonemes were calculated can be found in §§6.1.1–6.1.2.

Table 5.2: Confusion matrix for the four words of Set 1 in which lexical contrast is signalled by pitch

Stimulus → **↓ Response**	/tə̄máŋ/ (MH) 'all'	/tə̄-mā̄ŋ/ (MM) 'NPF-body'	/tə̀-mā̄ŋ/ (LM) 'PROH-believe'	/tə́-màŋ/ (HL) 'NPF-dark'
/tə̄máŋ/ (MH) 'all'	9/9			
/tə̄-mā̄ŋ/ (MM) 'NPF-body'		9/9	4	
/tə̀-mā̄ŋ/ (LM) 'PROH-believe'			5/9	
/tə́-màŋ/ (HL) 'NPF-dark'				9/9

In Table 5.2, stimuli are located on the horizontal axis and responses are located on the vertical axis. Numbers in the shaded cells indicate the number of correct responses to nine tokens of each test word. A number outside of a shaded cell identifies a confusion, with coordinates indicating the stimulus and the response. For example, out of nine responses to the stimulus /tə̀-māŋ/ 'PROH-believe', five were correct and four were incorrect; the coordinates indicate that all four of the incorrect responses were given as /tə̄-māŋ/ 'NPF-body'.

The confusion matrix of Table 5.2 demonstrates a high rate of intelligibility for the words tested (an overall intelligibility of 89%), with confusions only occurring between two stimuli having adjacent pitches; viz. the stimulus /tə̀-māŋ/ 'PROH-believe' (Low-Mid) was confused with /tə̄-māŋ/ 'NPF-body' (Mid-Mid) in four out of nine responses. Interestingly, the confusion was unidirectional, suggesting that this is unlikely to be attributable to chance. If the result *were* dependent upon chance, then we would also expect to find at least some confusions with /tə̀-māŋ/ 'PROH-believe' (Low-Mid) given in response to the stimulus /tə̄-māŋ/ 'NPF-body' (Mid-Mid), yet there was 100% intelligibility for all the stimuli of /tə̄-māŋ/ 'NPF-body' that were tested.

It becomes clear as to why confusions occurred between the /Low-Mid/ tones of /tə̀-māŋ/ 'PROH-believe' and the /Mid-Mid/ tones of /tə̄-māŋ/ 'NPF-body' after their respective F_0 shapes are compared in the line chart of Figure 5.1 below. Firstly, the final syllables of both words have a very similar F_0 height and contour, the main difference being a slightly longer rhyme duration for /tə̀-māŋ/ 'PROH-believe'. As the difference amounts to a mere 60 msec, this is unlikely to contribute a perceptual cue. Also observed in the F_0 shapes of the final syllables is a gradual divergence in F_0 height — /tə̀-māŋ/ 'PROH-believe' maintains quite consistent F_0 height over its duration with a very slight rise-fall mid-way through the rhyme duration, whereas /tə̄-māŋ/ 'NPF-body' shows a consistent decline in F_0 height between rhyme onset and termination. This, however, is also held to be too small a difference to provide a perceptual cue for the correct identification of a stimulus.

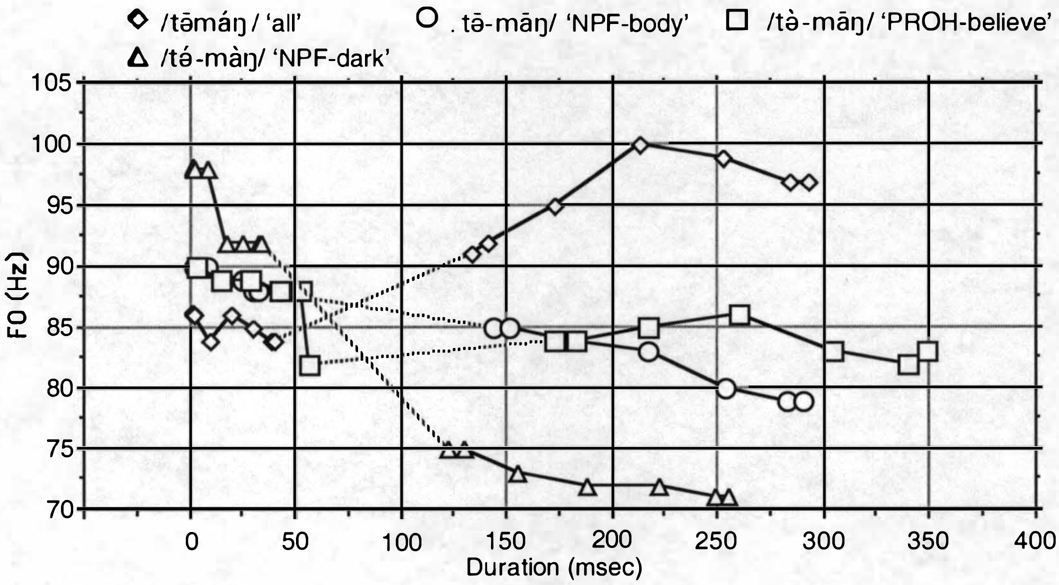


Figure 5.1: Individual F_0 shapes and rhyme durations for the disyllabic words of Set 1

A duration difference of 24 msec in the rhymes of the initial syllables is similarly highly unlikely to signal a contrast. A more probable perceptual cue is the lower F_O of the medial nasal in /tə̀-māŋ/ 'PROH-believe'. This is observed to have an F_O of 82 Hz at onset, whereas the medial nasal of /tə̄-māŋ/ 'NPF-body' has a F_O of 88 Hz at onset. A difference of 6 Hz is audible, and perhaps most importantly, the F_O difference occurs in the salient position of a syllable margin. Although not entered in the line chart of Figure 5.1, the medial nasal segments (whose time courses are represented by the broken line in each F_O contour) were calculated to have an F_O of 82 Hz for /tə̀-māŋ/ 'PROH-believe' and 87 Hz for /tə̄-māŋ/ 'NPF-body' at 50% of their durations, suggesting that the 5-6 Hz difference in F_O between the nasal segments is maintained for a considerable interval of time. The difference is augmented by a drop from 88 Hz to 82 Hz in the F_O contour of the initial syllable of /tə̀-māŋ/ 'PROH-believe', and this occurs at the end of its short duration immediately preceding the nasal onset.

I conclude that the fifty-five percent hit rate for the stimulus /tə̀-māŋ/ 'PROH-believe' is mostly attributable to perceptual cues provided by (a) the drop in F_O occurring in the initial syllable, and (b) the low F_O at the onset of the medial nasal, which is maintained for a substantial interval of its duration. The medial nasal presumably acts as the vehicle by which contrastive /Low/ tone is conveyed.

5.5.3 Results of second perception test (Set 2)

A second perception test was carried out using the same methodology and informant, but with the minimal triplet of Set 2. Once again, stimuli are found on the horizontal axis, and responses on the vertical axis. Shaded cells indicate the number of correct responses to nine tokens of each word. There is little to comment on, since the results demonstrate 100% intelligibility for all tokens tested.

Table 5.3: Confusion matrix for the three words of Set 2 in which lexical contrast is signalled by pitch

Stimulus → **↓ Response**	/tsə́ŋí/ (HH) 'rain'	/tsə̄ŋī/ (MM) 'sun'	/tsə́ŋì/ (HL) 'wild dog'
/tsə́ŋí/ (HH) 'rain'	9/9		
/tsə̄ŋī/ (MM) 'sun'		9/9	
/tsə́ŋì/ (HL) 'wild dog'			9/9

In the line chart of Figure 5.2 below, it is observed that the rhyme durations of the initial syllables in the test words are extremely short and amount to less than 40 msec, which constitutes no more than four or five glottal pulses in their respective digitised sound waves. As in the previous line chart, time courses of the medial nasal segments are represented by the broken lines in each F_O contour. Given the relatively large differences in the F_O height of the final syllables of the test words, it might be inferred that the informant relies on the pitch height of the final syllable to achieve 100% intelligibility in the perception test, and that the pitch of the short initial syllable is tonally irrelevant or perceptually redundant.

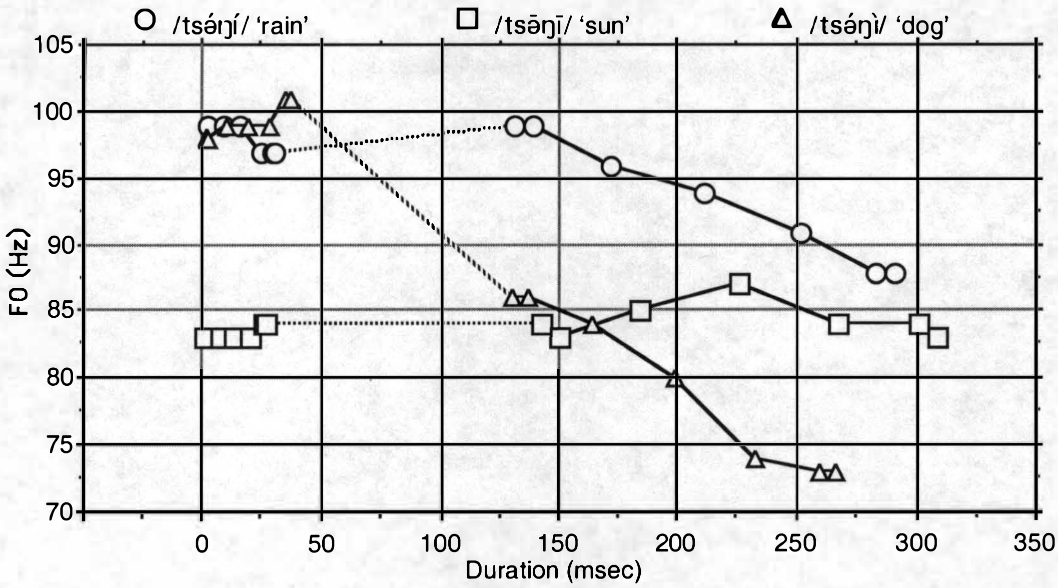


Figure 5.2: Individual F_0 shapes and rhyme durations for the disyllabic words of Set 2

5.5.4 Conclusion

The F_0 shapes of the line charts in Figures 5.1–5.2 demonstrate that as far as one male's tonal F_0 range is concerned, differences in F_0 height between contrasting tonemes are rather small. Nevertheless, the perception tests demonstrate that despite the small differences, tones can be readily identified by a native speaker.

This concludes the description of the tonal phonology of WM. In the next chapter I describe the methodology and results of an acoustic analysis of the monosyllable and disyllable tone groups and the contour tones presented in the auditory analyses of §§5.2–5.4.

6 *Tonal acoustics*

6.1 Introduction

This chapter describes the methodology and results of an acoustic analysis of tone in WM. Section 6.2 describes the procedure used in the instrumental analysis of the corpus and presents the fundamental frequency characteristics of three categories of monosyllables. Section 6.3 presents the fundamental frequency characteristics of three categories of disyllabic words. Lastly, §6.4 describes the procedure and results of an instrumental analysis of the falling pitch.

6.2 Acoustic analysis of monosyllables

6.2.1 Procedure for monosyllables

The acoustic analysis of the monosyllables (as well as the disyllables) is based on data from CL's corpus; these data were also used for the auditory analysis presented in Chapter 5. For the acoustic analysis of monosyllables, citation forms representative of the three contrastive pitch levels identified in the auditory analysis were digitised with Kay Computerised Speech Laboratory (Kay CSL) using 16 bit quantisation at a sampling rate of 10 KHz, after which temporally aligned wide band spectrograms (293 Hz) were made of each digitised utterance.

Next, a comparison of the sound wave to the formants of the wide band spectrogram allowed the rhyme onset and termination point to be identified and marked. The rhyme onset was taken to be the first spectrographic evidence of glottal phonation demonstrating a formant structure consistent with the vowel or with the consonant–vowel transition of the citation form, and termination was judged to be the point at which periodicity was lost in both the sound wave and the temporally aligned spectrogram, usually coinciding with a decrease in peak-to-peak amplitude. The two points were marked and the relevant portion of the spectrogram was expanded to the full width of the frame, in preparation for the extraction of F_o. Because the F_o of syllable onsets is reported to be tonally irrelevant (Howie 1974), F_o was only calculated over the duration of that portion of the syllable identified as the rhyme.

Initially F_o was determined by using the automatic F_o extraction algorithm of the Kay CSL software, but from the outset of the acoustic analysis I found that it sometimes gave spurious results, or was unable to extract F_o over the total duration of the rhyme. The greatest inaccuracy in automatic F_o extraction was observed to be associated with sections of the sound wave manifesting irregular peak-to-peak distances (identified as jitter in Figure

4.22); this is a characteristic of creaky voice phonation, but it may also be associated with non-contrastive glottalisation or the pre-closure phase of the glottal stop (observable in Figure 6.1 below and also in Figures 4.8–4.9 of Chapter 4 respectively). Such sections of the sound wave and temporally aligned spectrogram often correlated with F_0 trace 'drop out'. This can be seen in Figure 6.1, in which there are gaps where the F_0 extraction algorithm has been unable to trace F_0. The F_0 tracking is most erratic over the section of the spectrogram corresponding with the slightly glottalised onset. Furthermore, the overall F_0 trace bears little resemblance to the perceived pitch contour throughout the duration of the rhyme, since the F_0 value is approximately half that determined by a manual extraction method. The last section of the trace may be correct, but it is tonally irrelevant.

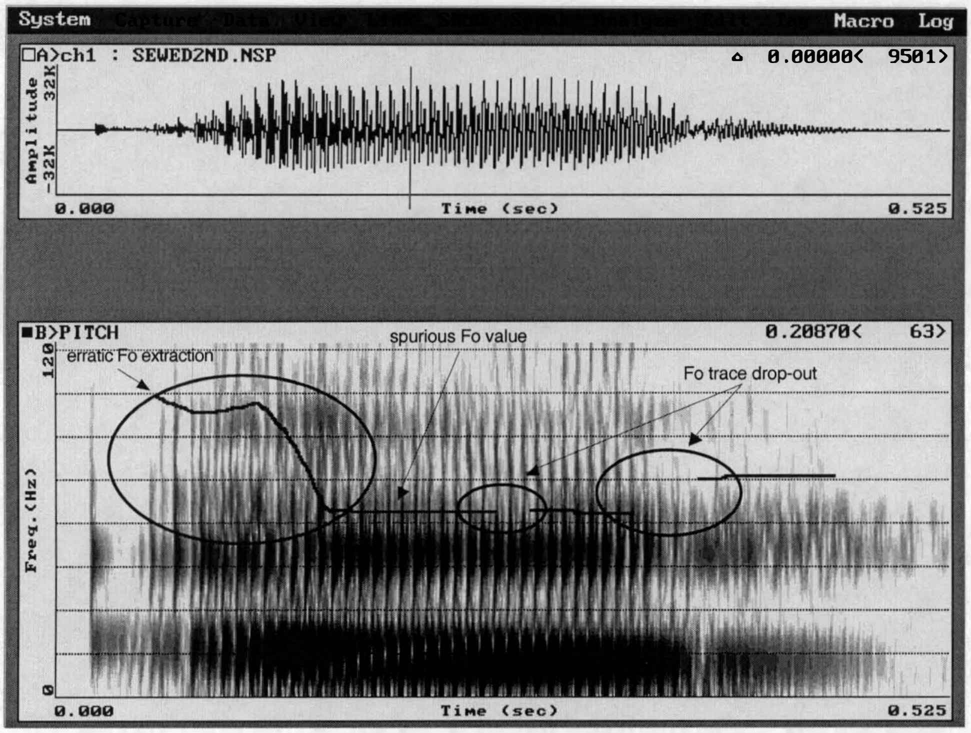


Figure 6.1: Temporally aligned sound wave and wide band spectrogram (293 Hz) of [ʔəɹ^{55}] 'sew.PAST'

Because of the problems encountered with the software's automatic pitch extraction, F_0 was instead determined by the following methodology. Having identified the rhyme by means of the process described above, the view of the wide band spectrogram was deleted and replaced by a narrow band spectrogram (24 Hz) of the rhyme aligned to the marked section of the sound wave. Next, the frequency range was adjusted to 0–1000 Hz. This usually allowed good resolution of the first seven or eight harmonics, but if not, the frequency range was made smaller as required. Values in Hz were then manually measured off the highest harmonic with the best resolution at seven sampling points, viz. 0%, 5%, 25%, 50%, 75%, 95% and 100% of the duration of the rhyme. F_0 is conveniently calculated off any harmonic because of the simple mathematical relationship that the harmonics bear to F_0; it is determined by dividing the frequency of the n^{th} harmonic by n (Baken 1987:140). A comparison of Figure 6.1 to Figure 6.2 below highlights discrepancies between the automatic and manual extraction of F_0 for the same token. In the latter figure, the display range has been adjusted to 0–600 Hz to resolve the first four harmonics. At the onset of the rhyme, the F_0 is calculated off the fourth harmonic to be ca. 120 Hz; this rises to a height of ca. 125 Hz for the duration of the rhyme before dropping very slightly at the termination. The F_0 values manually read off this narrow band spectrogram bear little relationship to those that were extracted by the Kay CSL F_0 tracking algorithm in the previous spectrogram of Figure 6.1.

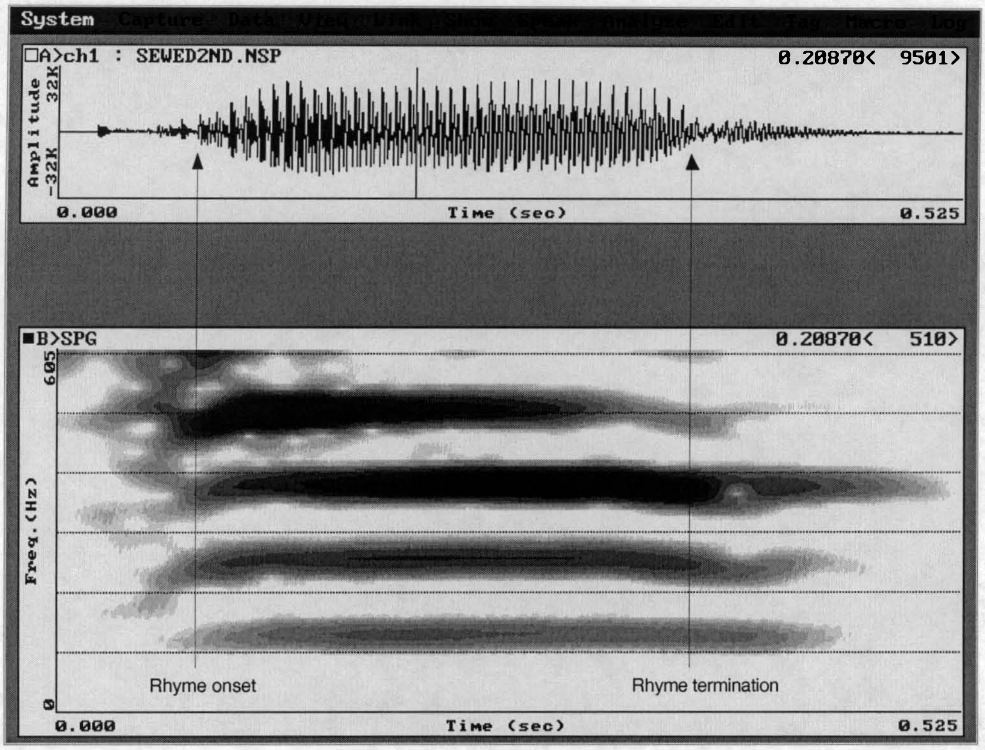


Figure 6.2: Temporally aligned sound wave and narrow band spectrogram (24 Hz) of [ʔəɹ⁵⁵] 'sew.PAST'

The sound wave and aligned narrow band spectrogram of Figure 6.3 on p.103 illustrates the manual method of F_0 extraction used for determining the F_0 shape of a tone as a function of duration. F_0 is measured off the fourth harmonic at seven sampling points, viz. 0%, 5%, 25%, 50%, 75%, 95% and 100% of the duration of the rhyme (275 msec).

It was reported in the auditory analysis of Chapter 5 that words were initially sorted according to syllable structure. Three categories of monosyllabic words were established in preparation for the acoustic analysis: stopped monosyllables (those with a stop coda) e.g. [tiʔ^{55}]; unstopped monosyllables e.g. [ɹa^{11}]; and the latter group were further divided into 'oral' and 'nasal' monosyllables, e.g. [səŋ^{33}]. The nasals were expected to have higher F_0 (Rose 1992), hence the establishment of the separate category of unstopped monosyllables. Each syllable category was subdivided into three groups according to the pitch levels of its members; that is, low [11], mid [33] or high [55], as determined by the auditory analysis. The words of each tonal group were then digitised and temporally aligned wide band spectrograms made of their respective sound waves in preparation for extraction of F_0. Mean F_0 at each sampling point and mean rhyme duration were calculated for each group and the results plotted in line charts using Statview 512+ software.

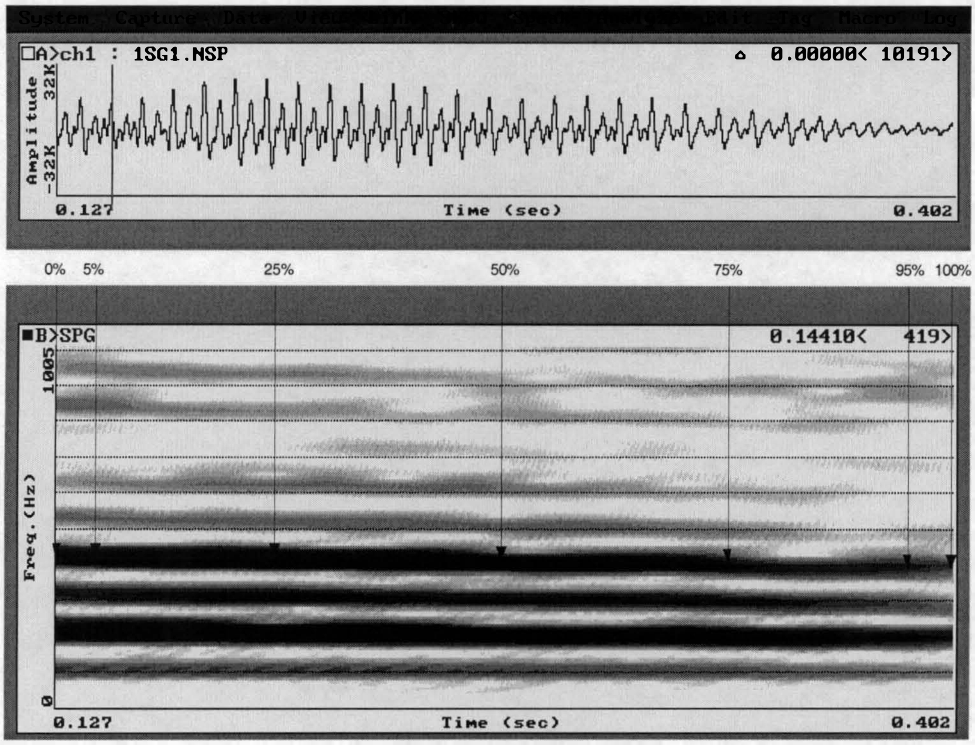


Figure 6.3: Temporally aligned sound wave and narrow band spectrogram (24 Hz) of the rhyme of [ni^{11}] 'I'

The mean F_0 shapes plotted in the line charts of Figures 6.4–6.6 below represent the three categories of monosyllables occurring in CL's corpus. These are based on five to nine tokens of each tone group, with the exception of the /Mid/ toneme stopped monosyllables, for which only three examples occurred in the corpus. Its rarity may be due to defective sampling, since the /Mid/ toneme frequently occurs on stopped disyllables. Individual F_0 values and durations for all tokens together with means and standard deviations for each tonal group are listed in Appendix A according to syllable category.

6.2.2 Results

Figures 6.4–6.6 below present F_0 plotted as a function of absolute mean duration for tones on stopped monosyllables, nasal monosyllables and unstopped monosyllables. The acoustic analysis reveals a hierarchy in terms of F_0 of /High/ > /Mid/ > /Low/ that corresponds closely with an auditory analysis identifying three contrastive tones. F_0 is found to be the only correlate of tone; that is, there are no other correlates reinforced by phonetic features such as vowel quality or phonation type.

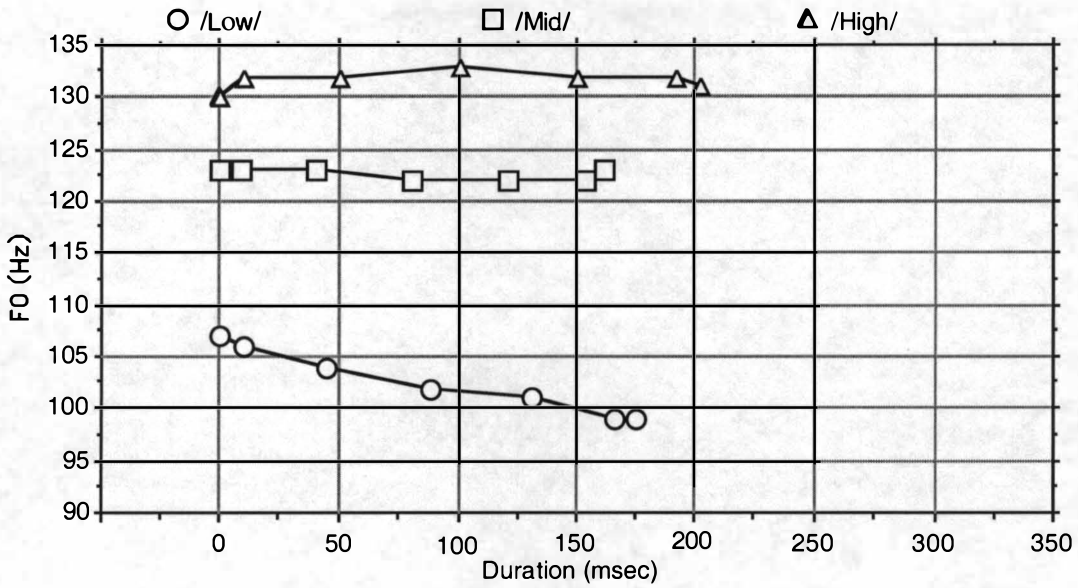


Figure 6.4: Mean /High/, /Mid/ and /Low/ F_o shapes for CL's stopped monosyllables

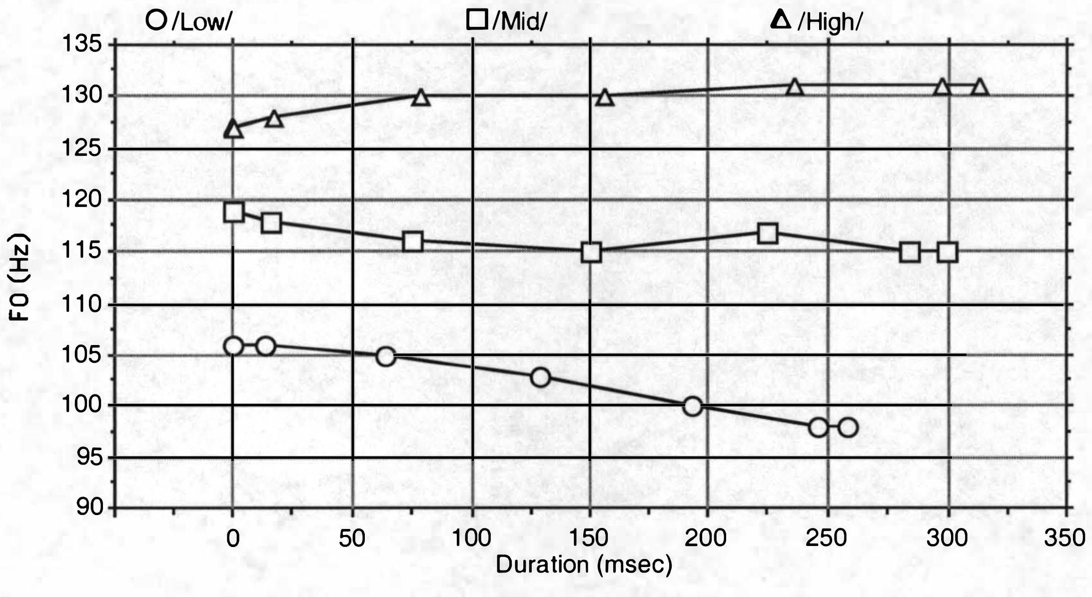


Figure 6.5: Mean /High/, /Mid/ and /Low/ F_o shapes for CL's nasal monosyllables

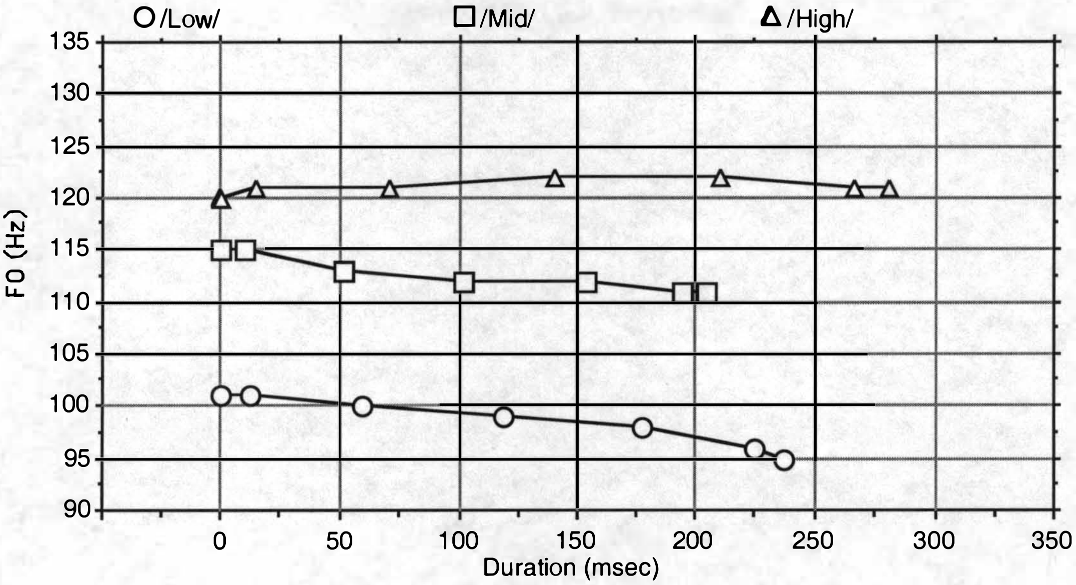


Figure 6.6: Mean /High/, /Mid/ and /Low/ F_0 shapes of CL's unstopped monosyllables

The most remarkable feature of these tones is the very small F_0 differences between contrastive levels — just 10 to 20 Hz. A comparison of F_0 heights for unstopped monosyllables with and without nasal codas demonstrates consistently higher mean F_0 values for syllables with nasal codas. This could be attributable to the fact that the codas of all but two of the nineteen nasal-final monosyllables used in the acoustic analysis had a velar place of articulation. In an acoustic study of a dialect of Chinese, Rose (1992) found that syllables with nasal codas, which happened to be velar, were associated with a higher F_0. There was an insufficient number of nasal syllables in CL's corpus to exclude monosyllabic words with velar nasal codas and still have a sufficient number to establish representative groups for the three tonemes; consequently I cannot ascertain whether the higher F_0 of nasal syllables is attributable to their having the feature [+nasal], or to some other parameter.

As expected, the rhymes of the stopped syllables are significantly shorter in duration than those of the unstopped syllables, while the nasal syllable rhymes are marginally longer than the those of the unstopped syllables.

The tonemes of the stopped syllables are observed to have overall higher mean F_0 values than their unstopped non-nasal counterparts. All F_0 contours demonstrate little variation over time between and within the tonemes of the three syllable types; indeed they are remarkably similar in shape, suggesting that height of F_0 is the sole perceptual cue for the identification of each toneme. The only exception to consistency in F_0 height over time is seen with the /Low/ toneme of each syllable type. In stopped and nasal syllables, it is 8–10 Hz lower in F_0 height at rhyme termination than at rhyme onset; in unstopped monosyllables the diminution in F_0 height between rhyme onset and rhyme termination is slightly less, at approximately 6 Hz. Despite demonstrating uniformity in F_0 shape for all three types of monosyllable, these onset and termination differences in the /Low/ toneme are so acoustically and perceptually minimal that it is unlikely that the sloping contour provides a perceptual cue for its recognition.

6.2.3 Statistics

A one-way ANOVA was done on individual F_o values at 50% of the duration of the rhyme to determine if the three tonemes identified in the auditory analysis differ significantly with respect to their F_o. The results of a Scheffe F-test show a statistically significant difference at 95% for all three tonemes in all three types of syllable. Test statistics are presented in Tables 6.1–6.3 below.

Table 6.1: Results of one-way ANOVA on stopped monosyllables

Toneme	Count	Mean	S.D.	Comparison	Scheffe F-test
/Low/	9	102	3.9	/Low/ v /Mid/	15.8
/Mid/	3	122	105	/Low/ v /High/	53.2
/High/	5	133	8.2	/Mid/ v /High/	3.8
			F = 56.7		
			p = .0001		

Table 6.2: Results of one-way ANOVA on nasal monosyllables

Toneme	Count	Mean	S.D.	Comparison	Scheffe F-test
/Low/	7	103	5.1	/Low/ v /Mid/	12.6
/Mid/	6	115	4.9	/Low/ v /High/	64.0
/High/	6	130	2.6	/Mid/ v /High/	18.4
			F = 64.1		
			p = .0001		

Table 6.3: Results of one-way ANOVA on unstopped monosyllables

Toneme	Count	Mean	S.D.	Comparison	Scheffe F-test
/Low/	6	99	7.1	/Low/ v /Mid/	7.8
/Mid/	5	112	5.7	/Low/ v /High/	26.0
/High/	5	122	4.6	/Mid/ v /High/	3.9
			F = 26.1		
			p = .0001		

6.3 Acoustic analysis of disyllables

6.3.1 Procedure for disyllables

The methodology used for the acoustic analysis of disyllables is the same as that described above for monosyllables, with the following modifications. Firstly, durations of medial consonant segments were measured and their means calculated for inclusion in the line charts.

Secondly, because the initial syllables of some disyllabic words were of such short duration, at times there was insufficient energy present in the sound wave for the Kay CSL software to create a narrow band spectrogram of the rhyme that could be expanded to the full size of the screen. To sidestep this problem, I first made a temporally aligned wide band spectrogram of the whole duration of the utterance to identify the rhyme onset and termination points of the initial syllable using the methodology described above for the analysis of monosyllables. I then expanded the marked rhyme duration to the full width of the screen and noted the duration measurement in milliseconds at 0%, 5%, 25%, 50%, 75%, 95% and 100% of the duration of the rhyme. Next, I deleted the view of the wide band spectrogram, selected a sufficient portion of the sound wave that included the rhyme of the initial syllable, and made a narrow band spectrogram (24 Hz) of the portion. This had to contain enough energy to allow the software to generate a narrow band spectrogram of the selection. Lastly, I manually measured the F_0 off the highest harmonic at my previously noted duration points in milliseconds, representing the above mentioned percentage points of the rhyme. This process was somewhat laborious, but it ensured maximum accuracy of F_0 extraction for the rhymes of the short initial syllables. The F_0 extraction method for the final syllables was the same as described for monosyllables.

It was reported in §5.3 that three categories of disyllabic words were established in preparation for their tonal analysis. To reiterate, these categories were: stopped disyllables with a medial sonorant e.g. [tə^{55}nak^{55}]; unstopped disyllables with a medial sonorant e.g. [tə^{33}ləm^{33}]; and unstopped disyllables with a medial voiceless obstruent e.g. [tʃʰi^{11}pa^{11}]. The three disyllable categories were then sorted according to pitch level. Section 6.3.2 presents the results of an instrumental analysis of disyllabic words bearing sequences of /High-High/, /Mid-Mid/ and /Low-Low/ tonemes. These, it will be recalled, were the most frequently occurring tonal permutations encountered in the disyllabic words of the corpus. With a larger corpus, other tonal combinations in disyllables might also be examined.

6.3.2 Results

Figures 6.7–6.9 below plot mean F_0 values as a function of absolute mean duration for /Low/, /Mid/ and /High/ tonemes on the three categories of disyllables described in §6.3.1: stopped disyllables with a medial sonorant, unstopped disyllables with a medial sonorant, and unstopped disyllables with a medial voiceless obstruent. The F_0 during the medial sonorant is interpolated, represented by the broken line. As reported in the auditory analysis of §5.3, the abundance of disyllables with identical tones on both the initial and final syllable in CL's corpus allowed /High-High/, /Mid-Mid/ and /Low-Low/ tonal groups to be set up. All tokens in the analysis conform to a (C)VCV(C) structure.

Despite their short mean durations, the initial syllables of disyllabic words are heard to have three phonetic pitch levels in the auditory analysis, and these are demonstrated in the line charts of Figures 6.7–6.9. The speaker's F_0 range in the initial syllables is slightly smaller than in the final syllables — initial syllables have a maximum range of approximately 25 Hz, whereas final syllables demonstrate a range of approximately 35 Hz for each category of disyllable. The final syllables compare quite closely with the monosyllables not only in F_0 range, but also in mean differences in F_0 height between the three contrasting tone levels.

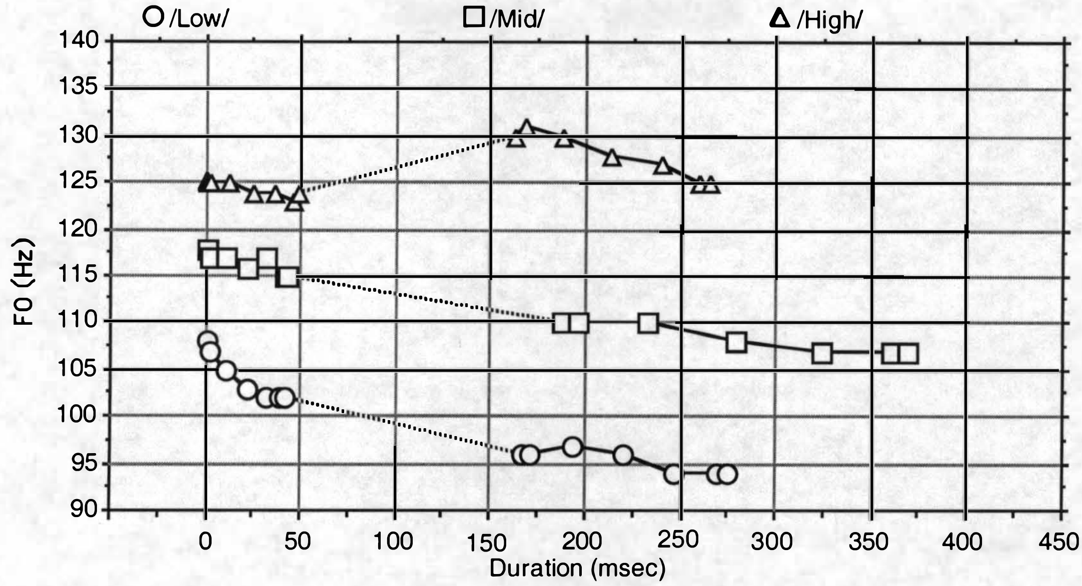


Figure 6.7: Mean /High/, /Mid/ and /Low/ F_o shapes for CL's stopped disyllables with a medial sonorant

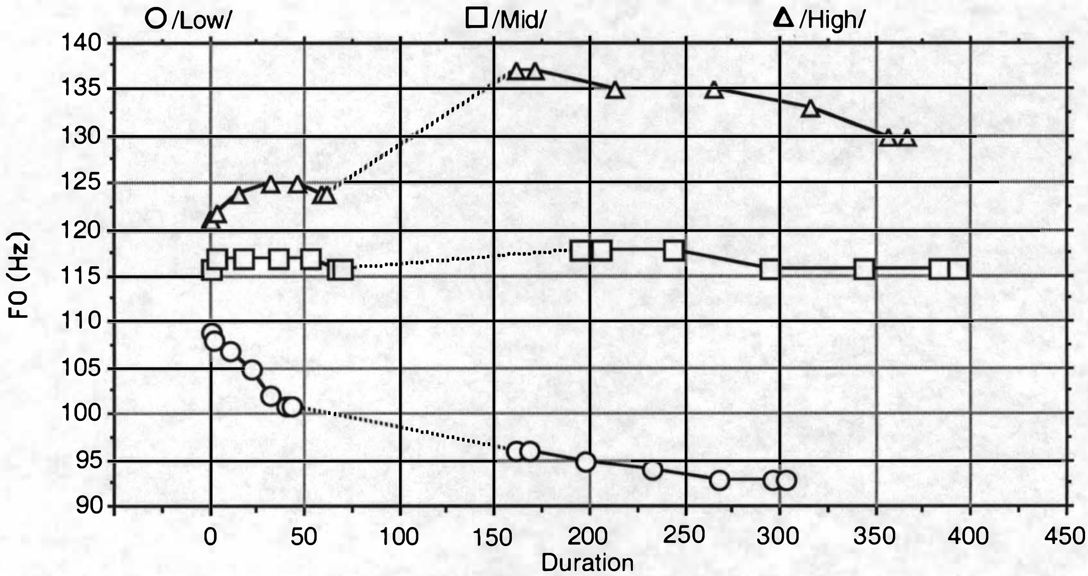


Figure 6.8: Mean /High/, /Mid/ and /Low/ F_o shapes for CL's unstopped disyllables with a medial sonorant

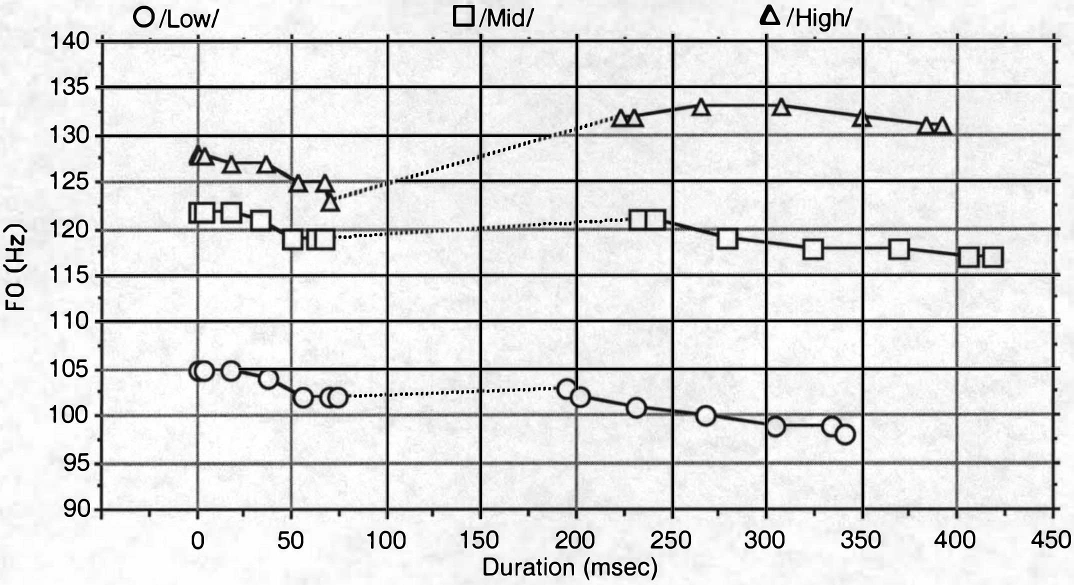


Figure 6.9: Mean /High/, /Mid/ and /Low/ F_0 shapes for CL's unstopped disyllables with a medial obstruent

Another feature of the initial syllables is the very small overall mean differences in F_0 between the /Mid/ and /High/ tonemes, as compared to the much larger overall mean difference in F_0 height between the /Mid/ and /Low/ tonemes for all disyllable categories. A smaller F_0 difference between the /Mid/ and /High/ tonemes relative to the difference between the /Mid/ and /Low/ tonemes was also observed in the stopped and unstopped monosyllables of §6.22. This is unusual, since bigger differences in F_0 usually occur at higher rather than lower frequencies, to ensure acoustic salience.

The mean difference in F_0 between the /Mid/ and /High/ tonemes in the initial syllable F_0 contours of Figure 6.9 is particularly small, constituting no more than 4–6 Hz for the total duration of the rhyme. Nevertheless, the perceptual difference was great enough for a phonetic pitch level distinction to be made in the auditory analysis and this is supported by the acoustic analysis.

Returning briefly to the problem of the phonological status of tone on nominal prefixes, the discussion of the structure of polysyllabic words in §2.6 concluded that the phonological status of tone occurring on the nominal prefixes /a-/ and /tə-/ in polysyllabic words is difficult to determine from the available data. While it is not consistently predictable, a dominant tendency in CL's data is for the tone of a nominal prefix to be the same as that of the stem syllable, hence the high frequency of /High-High/, /Mid-Mid/ and /Low-Low/ tonal combinations in disyllabic words with prefixal syllables. This raises the possibility that the nominal prefixes are lexically unspecified for tone, and that a prefix's tone is assigned by the root only after word formation has taken place. If this is indeed the case, it would be necessary to identify two types of domain over which tone has scope in WM: one domain of tone being the syllable, applicable to mono- and disyllabic roots and grammatical morphemes that are underlyingly specified for tone; the other domain of tone being the word, applicable to polysyllabic words in which the initial syllable is a tonally unspecified nominal prefix. The issue of the nature of tone on prefixal syllables is unresolved for the present but remains a topic worthy of further investigation and may have implications for the typology of tonal systems.

Disyllabic words with nominal prefixes were not separated from those disyllables consisting of just an unsegmentable root prior to the extraction of F_0 and the determination of mean values for each tone group in the acoustic analysis. Consequently no inferences can be drawn from the F_0 pitch shapes of initial syllables in the line charts above, although the contours of the initial syllables of all tone groups are noted to bear a remarkable similarity to those of their respective final syllables. Individual F_0 and rhyme duration values for each token used in the acoustic analysis can be found in the relevant tables of Appendix A.

6.4 Acoustic analysis of the falling pitch

6.4.1 Procedure

To determine the acoustic features of the falling pitch, representative examples identified in the auditory analysis were digitised with Kay CSL software using 16 bit quantisation at a sampling rate of 10 KHz and temporally aligned wide band spectrograms (293 Hz) were then made of the digitised tokens. Rhyme onset and termination points were determined and marked, after which the view of the wide band spectrogram was deleted and replaced by a narrow band spectrogram (24 Hz) of the rhyme. To gain an accurate representation of the time course of each F_0 contour, F_0 was sampled at 10 percentage points over the duration of the rhyme and manually calculated off the highest harmonic with the best resolution. Data for analysis were limited to five tokens for the narrative converb suffix, and three tokens for the conjunction. Raw F_0 values and rhyme durations of individual tokens together with their calculated means and standard deviations are tabled in Appendix A.

6.4.2 Results

Figure 6.10 plots mean F_0 values as a function of absolute mean duration for the falling pitches of the narrative converb suffix /-ə̂ɹ/ and the conjunction /tə̂ɹ/ ‘and’. It is observed that the mean durations of the two F_0 contours are almost identical, as are their F_0 contours over time. The obvious difference between the two is F_0 height at onset and termination; this amounts to a mean difference of 19 Hz at rhyme onset versus 14 Hz at termination.

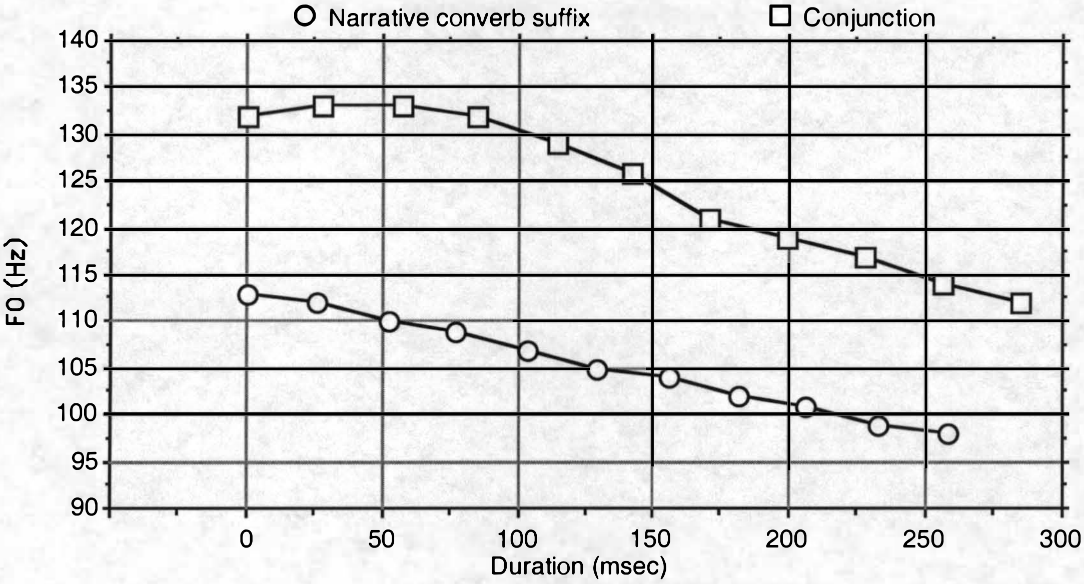


Figure 6.10: Mean F_0 shapes for the falling pitch

The contour pitch is distinctively different in appearance to the level tonemes described above. This is demonstrated most clearly by a comparison of the F_0 contours and rhyme durations of citation tones on the minimal pair /tʃà-ɹ/ 'eat-PRES' and /tʃâ-ɹ/ 'eat-NCV' in Figure 6.11 below. Firstly, the pitch occurring on the verb stem marked by the narrative converb suffix demonstrates a much longer rhyme duration and has a relatively greater fall in its F_0 height over time. The difference between height of F_0 at rhyme onset and termination is a substantial 18 Hz, and is the reason for its perception as a falling contour. In contrast, the lexical tone that occurs on the verb stem marked by the present tense suffix is of much shorter duration and is relatively more stable in terms of F_0 height over its time course, with a difference in F_0 height between rhyme onset and rhyme termination of just 6 Hz. Its relative stability over time is responsible for it being perceived as having a level pitch.

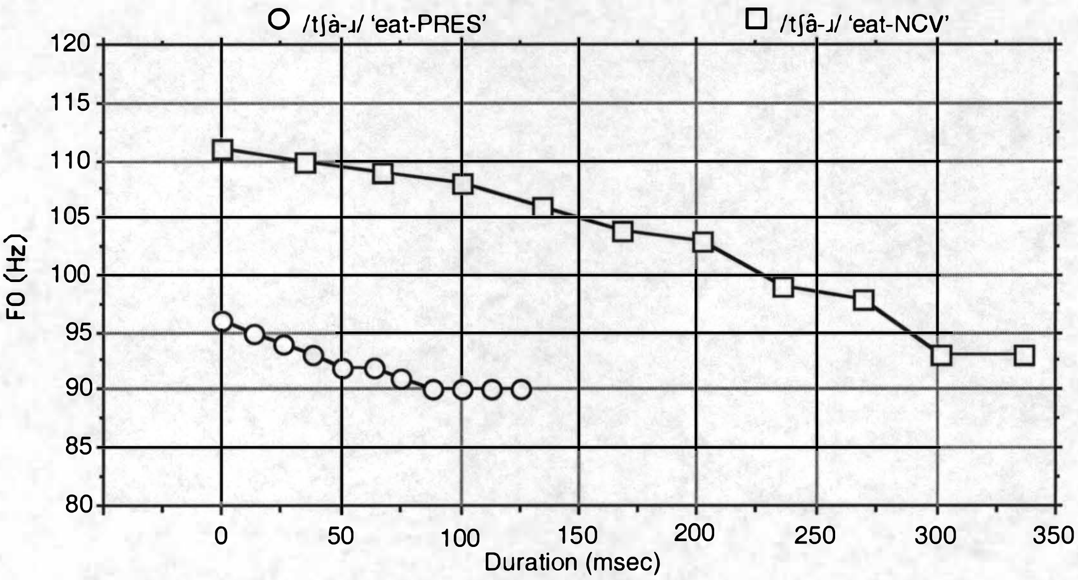


Figure 6.11: Comparison of F_0 shapes in the minimal pair /tʃà-ɹ/ 'eat-PRES' and /tʃâ-ɹ/ 'eat-NCV'

This concludes the chapter on tonal acoustics. The final chapter will provide a synopsis of the main findings of this description and make suggestions for future research on WM and related languages.

7 *Synopsis*

7.1 Introduction

This book has investigated and documented the phonetic and phonological features of a variety of the Mongsen dialect of Ao, a Kuki-Chin-Naga language spoken in Nagaland, north-east India; it has also demonstrated how an acoustic analysis might be used to corroborate and quantify the auditory analysis of the segmental phonology and tone system of an undescribed language. In the following sections I outline the main findings of the research based on the phonetic and phonological analysis, acoustic investigations and perceptual testing, and make suggestions for further research.

7.2 Discussion of main findings

7.2.1 Phonology

Chapter 2 described the linear and metrical structure of the syllable and found that WM allows only one segment to fill the onset slot, but that vowel sequences are permitted in a branching nucleus. It was demonstrated that there is no motivation for treating these sequences as unit diphthongs, nor for treating the rhyme as dominating a branching coda.

Processes of resyllabification and merger resulting from word formation involve assimilation, vowel deletion, diphthongisation and tone sandhi. Whilst some of these processes have been extensively investigated in African and Southeast Asian languages, very little is known of their typological characteristics in KCN languages. The preliminary results of this investigation suggest that WM is quite divergent from the better known tonal languages, particularly with regard to the direction of tone sandhi operating across morpheme boundaries. Some of the interesting tone sandhi manifestations resulting from word formation in WM may be attributable to the fact that verbal tense/aspect/modality affixes are underlyingly specified for tone. Since tone sandhi phenomena were only briefly examined in this description and the results are based on the data of just one native speaker, it would be worthwhile doing further research on this topic in Ao. It would also be useful to look at tone sandhi in other KCN languages; it is highly likely that most if not all are tonal, yet virtually nothing is known about the nature of tone in this branch of Tibeto-Burman.

The glottal stop was identified as having a fundamental prosodic function of marking word boundaries when realised. It can also simultaneously demonstrate a contrastive segmental function word-finally, however distributional evidence suggests that this is best analysed as a superficial consequence of its word-final realisation. Further work on the glottal stop is needed particularly from the point of view of theory, given the difficulty of its phonological representation. A prosodic glottal stop may also occur in related Tibeto-Burman languages of the north-east, but because it is frequently omitted in amateur texts and word lists there is no way of knowing just how widespread an occurrence it has until better descriptions become available. Also of interest is its diachronic origin. Benedict (1972) does not reconstruct a glottal stop for proto Tibeto-Burman, therefore it remains to be determined how it developed in Mongsen if it cannot be accounted for historically with reference to tone, as Burling (1992) posits for the glottal stop of Garo.

The status of tone on the nominal prefixes remains an unresolved issue not just in WM, but in all Tibeto-Burman languages which are reported to have tonally unspecified prefixal syllables with phonologically reduced vowels. If the tones of these syllables are assigned by the tone of the root to which they are prefixed, then it makes sense to identify a system of word tone for those words comprising a nominal prefix plus root; this would be in addition to a system of syllable tone for all the syllables of the language which have lexically specified tone. The diachronic origin of prefixal syllables also needs to be established. They are reported to have a widespread distribution in KCN and in other branches of Tibeto-Burman, yet Matisoff (1989) dismisses genetic inheritance as an explanation for their presence in Jingpaw (Kachin), the linguistic centre of the family both geographically and comparatively, and the hub through which all other Tibeto-Burman divisions are helically related according to the classification of Benedict (1972). Once again, explanations must await additional research in those languages reported to have tonally unspecified prefixal syllables.

Chapter 3 described the phoneme inventory of a fairly unremarkable consonant system. Nevertheless, the analysis does uncover a hitherto unrecognised aspiration contrast in the stop and affricate series. A voicing contrast was found to occur in the fricative series at the dental place of articulation. Cross-linguistically, this is reported to be rather rare in languages with a small number of fricatives (Maddison 1984), and is all the more unusual in WM by virtue of the fact that voicing is otherwise not exploited for signalling phonemic contrasts between segments sharing the same place of articulation.

The vowel phoneme inventory was found to have three heights centrally, but only one peripherally. The existence of two high rounded vowels, viz. /ʉ/ and /u/, is not conducive to establishing a contrast in the framework of traditional distinctive features and requires the introduction of an additional feature [±front]. A potential problem for phonological representation in the distinctive features framework is also presented by an analysis that assumes a morphophonologically conditioned dissimilation of /ʉ/ in word formation involving certain verbal suffixes. Under the morphophonological interpretation, a high central rounded vowel /ʉ/ of some verbal suffixes undergoes a dissimilation to a high front unrounded vowel in the environment of a back rounded vowel, but distinctive feature theory is simply unable to account for the unmotivated fronting in addition to the expected unrounding of /ʉ/. A much more satisfactory explanation is provided by a morphological analysis, and this is supported by comparative evidence. It is observed that the high central rounded vowel /ʉ/ of WM verbal suffixes corresponds uniformly to a high front unrounded vowel /i/ in the same suffixes of other varieties of Mongsen, therefore what *could* be viewed as dissimilation is actually more convincingly explained by an analysis positing phonetically conditioned allomorphy for the verbal suffixes involved.

The high back rounded vowel has two allophones [u] and [o] in free variation, and creaky voice is found to be used contrastively on just the low central vowel /a̰/. As creaky voice does not occur at other articulatory positions, this cannot be treated as a phonation type. The limited distribution therefore forces an analysis identifying /a̰/ as a rather idiosyncratic vowel phoneme. It would be useful to examine phonation in related languages to determine if phonemic creaky voice is peculiar to WM, or a more widespread but hitherto unrecognised phonation type.

The auditory analysis of the tonal system described in Chapter 5 identified three lexically contrastive tones occurring on all syllable types. A perception test carried out with a native speaker found that despite the rather small differences in pitch height between contrastive level pitches and their very similar pitch shapes, they are readily identifiable independent of context.

A falling pitch was observed to have a very limited distribution, only being found on verbs marked by the narrative converb suffix and on the one conjunction that occurred in the corpus. Its limited occurrence at the edges of clauses suggests a primary boundary marking function, as opposed to the lexical tone function of the level pitches occurring on syllables. The morphological structure of the conjunction is suspiciously similar to the converbal suffix, as is its pitch shape, suggesting a shared diachronic source and subsequent morphological reanalysis. As the falling pitch never occurs in other environments it cannot be considered to be part of the lexical tone system.

7.2.2 Acoustic phonetics

The instrumental analysis of voice onset time in Chapter 4 corroborated the auditory analysis. This found that despite claims in the literature that Ao does not have a voice onset time contrast in stops, a coincident VOT versus VOT lag contrast was acoustically demonstrable for both the stop and affricate series of WM. It may be the case that only Chungli lacks a VOT contrast in its stop and affricate series and this has led authors to make assumptions on behalf of all Ao dialects with respect to VOT in stops. The VOT contrasts in the affricates is interesting in that the contrast is not located so much in aspiration *per se*, but in the duration of the noise-generating release. This was consistently demonstrated in spectrograms for the aspirated dental affricate /tsʰ/, but only in the environment before a high front vowel for the aspirated palato-alveolar affricate /tʃʰ/.

The acoustic analysis of vowel formants plotted the formant values of seven vowels and found that all but the back rounded vowels and the low central creaky and modal voice vowels occupied clearly delineated vowel spaces. The back rounded vowels [u] and [o] that were identified as allophones of the same phoneme occurring in free variation in the auditory analysis showed slight overlap for their F_1 values, but no overlap in the F_2 plane. Although the results of a t-test revealed that the formant values of these two vowels are significantly different, tests done with a native speaker ultimately demonstrated that they can be freely substituted in words and thus must be recognised as allophones of the same phoneme.

The creaky voice vowel [a̰] is minimally differentiated from its modal voice counterpart with respect to formant values; the results suggest that formant value characteristics are not useful indicators for the recognition of WM phonation types. In the final section of Chapter 4 I applied modified methods used in speech pathology to quantify the acoustic features of creaky voice and this methodology proved to be a much more convincing diagnostic of creaky voice. Jitter was found to be the principal acoustic correlate.

The acoustic analysis of the tonal system quantified the auditory analysis and revealed that F_0 differences between contrastive tonemes are surprisingly small, amounting to just a 10–20 Hz difference between adjacent tonemes. Despite these small acoustic distinctions, a Scheffe F-test demonstrated a statistically significant difference between all three tonemes occurring on all syllable types.

As previously noted, the nature of tone on the initial syllables of some disyllables is controversial. Nominal prefixes occurring in Ao perhaps do not carry contrastive pitch, instead being assigned tone from the root following word formation in the majority of cases, although counter-examples were found in the corpus. A perception test carried out with the disyllabic words of one native speaker gave inconclusive results. The ambiguities presented by that particular set of data could almost certainly be cleared up by running more perception tests with a number of native speakers.

7.3 Future directions

As for broader suggestions for future research, our knowledge of most Tibeto-Burman languages spoken in north-eastern India and adjacent regions is still so patchy that quality field-based research on virtually any aspect of these languages will make a substantial contribution to Tibeto-Burman studies and help to fill in the considerable gaps in our current knowledge. This need for adequate documentation is acquiring greater urgency as the smaller languages become increasingly threatened with the real possibility of extinction under the pressures of globalisation and the spread of superstratum languages. And just as importantly, the creation of orthographies, grammars and dictionaries will also help to preserve the identities of peoples whose cultures are inseparable from their languages.

Appendix A: F_o values and rhyme durations

The following tables present raw F_O and duration values together with means and standard deviations for data used in the acoustic analysis of WM tones.

Table I: F_O values (in Hz) and durations (in msec) for CL's /Low/ stopped monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[ŋuʔ^{11}]	97	96	96	96	97	90	91	208
[tuʔ^{11}]	105	105	101	96	95	92	91	235
[tʃaʔ^{11}]	104	104	102	104	101	98	99	235
[jak^{11}]	104	104	100	102	95	99	96	109
[ts^{h}əʔ^{11}]	111	111	109	105	105	105	106	353
[nak^{11}]	109	107	106	103	104	102	101	107
[waʔ^{11}]	111	110	110	106	102	103	103	149
[təp^{11}]	110	109	106	106	106	105	105	62
[k^{h}iʔ^{11}]	109	109	105	104	100	95	95	174
mean	107	106	104	102	101	99	99	175
S.D.	4.6	4.6	4.5	3.9	4.2	5.5	5.7	85.7

Table II: F_O values (in Hz) and durations (in msec) for CL's /Mid/ stopped monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[təp^{33}]	124	122	121	121	121	121	122	167
[təp^{33}]	122	123	123	122	122	124	124	172
[təp^{33}]	124	124	124	124	122	122	122	155
mean	123	123	123	122	122	122	123	161
S.D.	1.2	1.0	1.5	1.5	.6	1.5	1.2	9.3

Table III: F_0 values (in Hz) and durations (in msec) for CL's /High/ stopped monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[aiʔ⁵⁵]	105	113	124	129	130	128	127	240
[tiʔ⁵⁵]	126	126	124	123	123	123	123	221
[aʉʔ⁵⁵]	136	139	138	141	138	144	144	205
[kʰiaʔ⁵⁵]	151	151	142	142	138	134	134	173
[tsʰəuk⁵⁵]	131	131	130	130	130	129	129	171
mean	130	132	132	133	132	132	131	202
S.D.	16.7	14.2	8.2	8.2	6.3	8.0	8.1	30.1

Table IV: F_0 values (in Hz) and durations (in msec) for CL's /Low/ nasal monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[tsʰəŋ¹¹]	112	110	108	102	100	97	97	209
[zəŋ¹¹]	99	100	104	101	97	96	97	327
[ɹaŋ¹¹]	97	97	97	96	95	95	95	223
[uŋ¹¹]	113	113	108	105	103	98	98	266
[mən¹¹]	113	113	113	112	108	106	103	272
[naŋ¹¹]	101	100	100	99	94	92	93	277
[tsəŋ¹¹]	107	107	106	103	103	104	104	232
mean	106	106	105	103	100	98	98	258
S.D.	7.0	6.7	5.4	5.1	5.0	5.0	4.0	40.2

Table V: F_0 values (in Hz) and durations (in msec) for CL's /Mid/ nasal monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[ɹoŋ³³]	121	120	116	116	116	115	115	321
[juŋ³³]	116	115	112	113	115	116	116	284
[jim³³]	122	121	117	116	117	117	114	318
[səŋ³³]	114	113	112	106	108	109	109	284
[nuŋ³³]	118	118	118	118	121	115	114	294
[muŋ³³]	121	120	120	120	122	119	119	295
mean	119	118	116	115	117	115	115	299
S.D.	3.2	3.2	3.3	4.9	5.0	3.4	3.3	16.3

Table VI: F_0 values (in Hz) and durations (in msec) for CL's /High/ nasal monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[tsəŋ⁵⁵]	132	132	132	131	131	128	125	325
[tsəŋ⁵⁵]	135	135	136	133	134	135	135	288
[tsəŋ⁵⁵]	135	135	135	133	133	134	134	326
[auŋ⁵⁵]	121	121	127	128	130	130	130	330
[auŋ⁵⁵]	125	125	125	129	131	131	131	296
[auŋ⁵⁵]	115	122	127	127	128	129	128	314
mean	127	128	130	130	131	131	131	313
S.D.	8.2	6.4	4.6	2.6	2.1	2.8	3.7	17.4

Table VII: F_O values (in Hz) and durations (in msec) for CL's /Low/ unstopped monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[ɹa^{11}]	99	98	96	95	93	92	92	221
[ɹa^{11}]	95	95	93	92	92	91	91	224
[ɹa^{11}]	93	93	91	90	90	88	88	174
[ni^{11}]	106	106	105	102	102	98	97	239
[ni^{11}]	108	108	107	107	105	106	104	276
[ni^{11}]	106	105	105	105	103	103	98	279
mean	101	101	100	99	98	96	95	236
S.D.	6.4	6.3	7.0	7.1	6.5	7.2	8.0	39.2

Table VIII: F_O values (in Hz) and durations (in msec) for CL's /Mid/ unstopped monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[hʉ^{33}]	117	117	114	114	114	112	112	215
[sə^{33}]	109	109	107	106	104	105	106	273
[pa^{33}]	112	110	110	109	108	107	107	186
[la^{33}]	115	115	114	112	111	111	111	172
[sa^{33}]	122	122	122	121	121	122	121	173
mean	115	115	113	112	112	111	111	204
S.D.	5.0	5.3	5.6	5.7	6.4	6.6	5.9	42.4

Table IX: F_O values (in Hz) and durations (in msec) for CL's /High/ unstopped monosyllables

	0%	5%	25%	50%	75%	95%	100%	Duration
[ai^{55}]	120	120	120	120	121	121	120	282
[ai^{55}]	119	119	119	119	120	116	116	286
[ai^{55}]	115	115	115	115	115	115	115	244
[tʃai^{55}]	122	122	121	123	124	122	122	237
[əɹ^{55}]	119	120	120	121	122	122	122	327
[əɹ^{55}]	124	126	126	128	128	127	127	291
[əɹ^{55}]	120	123	124	127	126	125	125	293
mean	120	121	121	122	122	121	121	280
S.D.	2.8	3.5	3.55	4.6	4.3	4.4	4.4	30.8

Table X: F_0 values (in Hz) and rhyme durations (in msec) for CL's /Low/ stopped disyllables with a medial sonorant (son = sonorant duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	son	0%	5%	25%	50%	75%	95%	100%	Dur
[wa^{11}jaʔ^{11}]	102	101	99	99	99	99	99	51	161	94	94	95	95	95	95	95	88
[wa^{11}jaʔ^{11}]	105	102	102	101	101	101	101	29	130	94	94	99	100	100	101	100	83
[tsə^{11}laʔ^{11}]	112	112	111	110	107	107	107	55	110	99	99	98	95	96	93	93	127
[tsə^{11}laʔ^{11}]	115	114	111	107	102	102	100	39	116	98	98	97	97	94	94	95	129
[tsə^{11}laʔ^{11}]	105	105	103	102	101	101	101	29	94	96	96	96	93	87	87	86	113
mean	108	107	105	104	102	102	102	41	122	96	96	97	96	94	94	94	108
S.D.	5.4	5.9	5.5	4.6	3.0	3.0	3.1	12.1	25.2	2.3	2.3	1.6	2.6	4.7	5.0	5.0	21.5

Table XI: F_0 values (in Hz) and rhyme durations (in msec) for CL's /Mid/ stopped disyllables with a medial sonorant (son = sonorant duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	son	0%	5%	25%	50%	75%	95%	100%	Dur
[a^{33}miʔ^{33}]	125	122	121	120	118	117	117	63	173	111	111	111	108	109	110	110	218
[a^{33}miʔ^{33}]	111	111	111	113	115	115	115	39	133	111	111	111	110	109	111	111	199
[a^{33}miʔ^{33}]	119	119	120	116	117	114	113	28	126	109	109	109	107	104	102	102	128
mean	118	117	117	116	117	115	115	43	144	110	110	110	108	107	107	107	182
S.D.	7.0	5.7	5.5	3.5	1.5	1.5	2.0	17.9	25.4	1.2	1.2	1.2	1.5	2.9	4.9	4.9	47.4

Table XII: F_0 values (in Hz) and rhyme durations (in msec) for CL's /High/ stopped disyllables with a medial sonorant (son = sonorant duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	son	0%	5%	25%	50%	75%	95%	100%	Dur
[tən⁵⁵ ak⁵⁵]	118	118	115	113	113	110	110	27	115	122	123	120	119	121	121	121	120
[a⁵⁵naʔ⁵⁵]	119	116	121	121	120	119	119	70	131	132	132	131	127	127	129	129	123
[tə⁵⁵nik⁵⁵]	120	120	120	119	119	119	119	47	106	119	120	120	120	120	118	115	101
[ta⁵⁵mak⁵⁵]	117	117	117	116	117	118	118	78	116	120	120	119	117	115	115	118	91
[tu⁵⁵ŋət⁵⁵]	133	133	132	133	132	132	132	50	88	136	135	136	135	132	130	129	70
[nəŋ⁵⁵ət⁵⁵]	121	121	121	121	121	122	122	47	99	128	128	128	127	126	124	124	69
[tə⁵⁵naʔ⁵⁵]	137	137	136	133	131	128	130	29	108	135	135	134	132	131	128	128	123
[maŋ⁵⁵əp⁵⁵]	123	123	123	123	124	124	125	69	107	133	133	132	131	130	130	129	54
[tə⁵⁵nʉʔ⁵⁵]	134	134	134	134	134	135	135	35	108	147	147	147	145	139	135	134	129
[tə⁵⁵miʔ⁵⁵]	130	130	129	127	127	126	126	26	161	132	133	133	130	127	120	120	138
mean	125	125	125	124	124	123	124	48	114	130	131	130	128	127	125	125	102
S.D.	7.5	7.8	7.4	7.5	7.0	7.3	7.5	19.1	20.0	8.5	8.2	8.7	8.4	6.9	6.4	6.0	29.4

Table XIII: F_o values (in Hz) and rhyme durations (in msec) for CL's /Low/ unstopped disyllables with a medial sonorant (son = sonorant duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	son	0%	5%	25%	50%	75%	95%	100%	Dur
[tu^{11}ŋəɹ^{11}]	110	110	110	99	99	96	96	39	115	89	89	89	87	85	87	86	163
[tu^{11}ŋəɹ^{11}]	108	108	104	101	99	99	98	48	84	92	91	89	89	87	89	89	163
[tu^{11}ŋəɹ^{11}]	99	99	99	97	95	92	92	48	97	90	90	90	89	89	89	89	155
[tʃə^{11}li^{11}]	108	108	108	106	103	102	102	34	164	94	94	94	95	94	93	93	98
[tʃə^{11}li^{11}]	106	106	104	102	101	99	99	43	146	97	97	96	95	94	94	94	118
[pʰi^{11}ləm^{11}]	119	118	118	118	111	111	111	35	122	103	103	101	98	100	100	100	139
[pʰi^{11}ləm^{11}]	110	110	109	111	109	106	106	49	104	105	105	104	102	102	102	102	162
mean	109	108	107	105	102	101	101	42	119	96	96	95	94	93	93	93	142
S.D.	5.9	5.7	6.0	7.4	5.7	6.3	6.4	6.4	28.0	6.3	6.4	6.0	5.5	6.4	5.7	5.9	25.7

Table XIV: F_o values (in Hz) and rhyme durations (in msec) for CL's /Mid/ unstopped disyllables with a medial sonorant (son = sonorant duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	son	0%	5%	25%	50%	75%	95%	100%	Dur
[ma^{33}ma^{33}]	113	113	113	113	113	113	114	100	116	117	117	117	114	113	113	113	143
[a^{33}nu^{33}]	112	113	116	115	116	113	111	67	166	114	114	114	114	114	114	114	169
[pʰa^{33}ŋa^{33}]	119	119	119	119	119	119	119	33	117	121	122	122	122	120	120	120	205
[pʰə^{33}li^{33}]	110	110	109	109	109	108	108	95	81	112	112	112	111	110	110	110	222
[tə^{33}ləm^{33}]	119	118	117	116	115	114	114	55	129	112	112	113	112	112	112	112	208
[a^{33}hʉ^{33}]	122	125	127	126	128	128	128	87	159	131	131	131	126	126	124	124	198
[tə^{33}nu^{33}]	118	118	118	119	119	117	117	50	106	120	120	119	118	119	119	119	233
mean	116	117	117	117	117	116	116	70	125	118	118	118	116	116	116	116	197
S.D.	4.5	5.0	5.6	5.4	6.0	6.3	6.5	25.2	29.7	6.7	6.8	6.6	5.6	5.6	5.1	5.1	31.1

Table XV: F_0 values (in Hz) and rhyme durations (in msec) for CL's /High/ unstopped disyllables with a medial sonorant (son = sonorant duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	son	0%	5%	25%	50%	75%	95%	100%	Dur
[tsə^{55}ŋi^{55}]	120	121	121	122	122	120	120	41	116	122	122	122	121	120	118	118	190
[tsə^{55}ŋi^{55}]	121	121	121	121	121	121	121	73	92	119	119	119	119	118	118	118	191
[a^{55}luŋ^{55}]	125	126	132	132	132	132	132	86	82	143	143	143	143	143	144	144	203
[a^{55}niŋ^{55}]	117	117	119	121	122	121	121	61	108	147	.147	141	141	141	136	136	224
[a^{55}niŋ^{55}]	122	122	125	129	128	127	127	64	101	147	147	144	142	137	133	133	239
[a^{55}niŋ^{55}]	123	123	123	122	124	125	125	41	99	142	142	142	141	137	132	132	183
mean	121	122	124	125	125	124	124	61	100	137	137	135	135	133	130	130	205
S.D.	2.7	2.9	4.6	4.8	4.3	4.6	4.6	17.8	11.9	12.7	12.7	11.4	11.3	10.9	10.3	10.3	22.0

Table XVI: F_0 values (in Hz) and rhyme durations (in msec) for CL's /Low/ unstopped disyllables with a medial obstruent (obstr = obstruent duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	obstr	0%	5%	25%	50%	75%	95%	100%	Dur
[tʃi^{11}pa^{11}]	106	106	105	105	104	104	102	57	113	106	106	104	104	104	105	102	174
[tʃi^{11}pa^{11}]	109	109	108	106	104	105	105	57	105	106	106	104	104	103	103	103	170
[tʃi^{11}pa^{11}]	108	108	108	106	104	100	101	55	103	105	105	104	102	102	101	101	141
[la^{11}ta^{11}]	101	101	101	101	100	100	100	82	141	98	96	94	94	94	94	93	138
[la^{11}ta^{11}]	102	101	101	101	99	99	99	101	133	100	100	99	98	95	95	95	134
[la^{11}ta^{11}]	106	106	106	105	106	106	106	84	128	100	100	99	99	96	95	95	124
mean	105	105	105	104	102	102	102	73	121	103	102	101	100	99	99	98	147
S.D.	3.2	3.4	3.2	2.4	2.7	3.0	2.8	19.1	15.7	3.6	4.1	4.1	3.9	4.5	4.8	4.3	20.4

Table XVII: F_o values (in Hz) and rhyme durations (in msec) for CL's /Mid/ unstopped disyllables with a medial obstruent (obstr = obstruent duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	obstr	0%	5%	25%	50%	75%	95%	100%	Dur
[i^{33}pi^{33}]	124	125	127	124	121	121	121	83	145	125	125	124	123	120	117	117	173
[a^{33}tʃi^{33}]	120	120	120	120	118	118	119	94	228	122	122	121	120	118	118	118	168
[ta^{33}pa^{33}]	124	124	124	122	122	122	122	81	136	125	124	122	122	122	120	119	165
[ku^{33}ta^{33}]	124	124	124	122	120	119	119	66	90	121	121	120	118	116	117	117	128
[tə^{33}tʃaŋ^{33}]	122	122	122	120	120	120	119	31	145	119	118	115	112	114	114	114	221
[tə^{33}koŋ^{33}]	116	116	117	116	115	115	114	31	163	118	117	117	114	115	117	117	226
[u^{33}pʰuŋ^{33}]	122	122	122	120	119	119	120	78	181	120	119	116	117	118	116	116	194
mean	122	122	122	121	119	119	119	66	166	121	121	119	118	118	117	117	182
S.D.	2.9	3.1	3.2	2.5	2.3	2.3	2.5	25.5	34.2	2.8	3.0	3.4	4.0	2.8	1.8	1.6	34.4

Table XVIII: F_o values (in Hz) and rhyme durations (in msec) for CL's /High/ unstopped disyllables with a medial obstruent (obstr = obstruent duration)

	0%	5%	25%	50%	75%	95%	100%	Dur	obstr	0%	5%	25%	50%	75%	95%	100%	Dur
[mə^{55}tʃa^{55}]	135	135	135	135	131	132	132	46	141	138	128	140	141	139	137	137	125
[u^{55}pu^{55}]	129	129	130	130	129	128	127	120	157	130	129	131	129	128	128	125	183
[ta^{55}tʃaŋ^{55}]	120	119	117	117	119	119	119	102	173	123	123	124	125	124	125	125	212
[tə^{55}tʃa^{55}]	125	127	127	127	120	120	119	48	156	131	132	133	132	132	131	131	138
[tə^{55}tʃən^{55}]	130	129	128	128	127	124	118	32	134	140	140	137	136	136	135	135	188
mean	128	128	127	127	125	125	123	70	152	132	132	133	133	132	131	131	169
S.D.	5.6	5.8	6.6	6.6	5.4	5.5	6.2	38.8	15.2	6.8	6.9	6.1	6.2	6.0	4.9	5.5	36.4

Table XIX: F_0 values (in Hz) and rhyme durations (in msec) for CL's word-final falling pitch on verbs marked by the narrative converb suffix

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Dur
[jak^{33} əɹ^{31}]	110	110	109	109	107	105	104	102	101	99	98	205
[tʃaɹ^{31}]	111	110	109	109	107	106	105	103	102	101	100	256
[tʃaɹ^{31}]	111	110	109	108	106	104	103	99	98	93	93	336
[jak^{11} təp^{33} əɹ^{31}]	116	115	112	110	108	106	104	104	103	103	100	255
[jak^{11} təp^{33} əɹ^{31}]	115	114	111	107	106	105	102	101	101	100	100	237
mean	113	112	110	109	107	105	104	102	101	99	98	258
S.D.	2.7	2.5	1.4	1.1	.8	.8	1.1	1.9	1.9	3.8	3.0	48.3

Table XX: F_0 values (in Hz) and rhyme durations (in msec) for CL's falling pitch on the conjunction [təɹ^{53}]

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Dur
[təɹ^{53}]	128	130	130	129	128	124	121	119	117	115	113	280
[təɹ^{53}]	130	130	130	129	126	124	120	118	116	113	111	285
[təɹ^{53}]	138	140	140	137	134	129	123	120	117	114	112	288
mean	132	133	133	132	129	126	121	119	117	114	112	284
S.D.	5.3	5.8	5.8	4.6	4.2	2.9	1.5	1.0	.6	1.0	1.0	4.0

Appendix B: Word lists

The following lists represent the WM words elicited from the native speakers CL and MA. I have used CL's forms in the examples of this book where a choice was to be made between differing tonal realisations on the same lexical item; CL's data also was also used exclusively for the acoustic analysis of Chapter 6. All underived verbs are given in their past tense form, the formally unmarked tense in Ao. Prefixes and suffixes on stems are marked by hyphens and glossed accordingly.

CL	MA	Gloss
nì	*nì*	1SG
nàŋ	*nàŋ*	2SG
pā	*pā*	3SG
lā	*lā*	3SG.FEM
kə̄nə̄t		1DU.EXC
īnə̄t		1DU.INC
nə̄ŋə̄t		2DU
tūŋə̄t		3DU
ìlā		1PL.EXC
ìsā	*ìsə̀ŋlā*	1PL.INC
nə̀ŋkʰə̀lā		2PL
túŋkʰə̄lā	*túŋlā*	3PL
tə̄máŋ	*tə̄máŋ*	all
tə̂ɹ		and
sàɹùwāɹ	*sāɹāwàɹ*	animal
	āw-ə̄ɹ	Ao-NS
úpú		ashes
	āʉ	axe
tə̄-sīn	*tə̄-sìn*	NPF-back
tāmáɹū	*tāmáɹū*	bad
tsə̀ŋ	*tsə̄ŋ*	bark.PAST
jàk	*jàk*	beat.PAST
tà tʃʰàpànā	*tà tʃʰàpànā*	because (thus do-NML-INST)
tə̄-pūk	*tə̄-pūk*	NPF-belly
tə̀-pə̀tī	*tə̀-pə̀tí*	NPF-big
ɹàk		bind.PAST

CL	MA	Gloss
wàjàʔ	*wàzàʔ*	bird
ŋùʔ	*ŋùʔ*	bite.PAST
tə́-nák	*tə́-nàk*	NPF-black
áiʔ	*áiʔ*	blood
hʉ̄		blow.PAST
	tə̄-māŋ	NPF-body
tə̄-ɹə̄t	*tə̄-ɹə̄t*	NPF-bone
tə̄-kūlūk		NPF-brain
sə́ŋtáŋ	*sə́ŋtàŋ*	branch/wood
māmā	*māmā*	breast
sə̄nsī	*sə̄nsì*	breathe.PAST
ɹūŋ	*ɹūŋ*	burn.PAST
	tə̄-mūlā	NPF-calf (of leg)
	tə̄-kʰùʔ	NPF-chest
án		chicken
ānū	*ānū*	child
tə̄-mə̄zə̄ŋ	*tə̄-mə̄zə̄ŋ*	NPF-claw/fingernail
mījū		clean.PAST
	sùsī	clothing
tsə̄ŋmī	*tsə̄ŋmīlūŋ*	cloud
tə̄-mə̄kūŋ	*tə̄-mə̄kūŋ*	NPF-cold
	tsʰə̀k	collide.PAST
ɹà	*ɹà*	come.PAST
	tʃʰū	cook.PAST
zə̀ŋ	*zə̀ŋ*	count.PAST
ɹàŋ	*ɹàŋlāk*	cut.PAST
	tsə̀ŋsàŋ	dance.PAST
	tə́-màŋ	NPF-dark
tsèŋìkū	*tsèŋìkʰū*	day
mīn		dense jungle
	kʰə̄p	depart.PAST
	tʰīa	destiny/luck
sə̄	*sə̄*	die.PAST
tùʔ	*tùʔ*	dig.PAST
tə̄-mə̄nə̄n	*tə́-mə́nə̄n*	NPF-dirty
ə̄ntʃūk		discard.PAST
ái	*ái*	dog
jūŋ	*jūŋ*	drink.PAST
tə̄-kūŋ	*tə̄-kūŋ*	dry
ūpʰūŋ		dust
tə̄-nāɹūŋ	*tə̄-nāɹūŋ*	NPF-ear
á-lí-má	*á-lí-mā*	NPF-earth-FACE
phə́níŋnúk		earthquake

CL	MA	Gloss
tsə̀là?	*tsə̀là?*	earthworm
tʃà?	*tʃà?*	eat.PAST
	típ	edge/side
ə̄ntsə̄	*ə́ntsə̄*	egg
	tʃʰīt	eight
lī		exist.PAST
ūkjā		expand.PAST
tə́-ník	*tə́-ník*	NPF-eye
ātʃī	*ātʃī*	fall.PAST
ùpā/tə̄-pā?	*tə̄-pā?*	father/NPF-father
tʃʰìpà	*tə̀-tʃʰìpà*	fear.PAST/NPF-fear.PAST
ə̄ntʃʰáwà		feather
ítʃā	*ítʃā*	few
á-lù		NPF-field
jāk-tə̄p	*jāk-tə̄p*	fight-RECIP.PAST
sə̄ŋ	*sə̄ŋ*	fill.PAST
mə̄zə̄?/mījʉ̄	*mə̄zə̄?*	fire
á-ŋá?	*á-ŋà?*	NPF-fish
pʰāŋā	*pʰāŋā*	five
tsə́ŋpʰáŋ	*pʰūŋə̄t*	float.PAST/flow.PAST
nāɹū	*nāɹū*	flower
jīm	*jīm*	fly.PAST
tsə̀sə̀	*tsə̀sə̀*	fog
tə̄-tʃāŋ	*tə́-tʃáŋ*	NPF-foot/NPF-leg
pʰə̄lī	*pʰə̄lī*	four
sə̀ŋ-tʃàŋ	*sə̀ŋ-tʃàŋ*	fruit (tree-seed)
	ā-tsə̀	NPF-garden
ə̄ntʰə̄n		gather
khì?	*khì?*	give.PAST
wā		go.PAST
t-āɹū	*t-āɹū*	NPF-good
lìə̄ɹ		gooseberry
ā-jā	*ā-zā*	NPF-grass
	tə̄-ɹā	NPF-grass.root
t-ámák	*t-ámàk*	NPF-green
mījʉ̄		grow.PAST
pùk-tsʰə̄ŋ		guts (belly + ?)
kūwā	*kúwá*	hair
tə́-kʰə́t	*tə́-kʰə́t*	NPF-hand
	tə̄-pə̄lā	NPF-happy
tə̄-lə̄m	*tə̄-lə̄m*	NPF-head
jà?	*jà?*	hear.PAST
tə̄-mə̄lūŋtʃāŋ	*tə̄-mə̄lūŋtʃāŋ*	NPF-heart

CL	MA	Gloss
t-átʃáŋ	*t-átʃáŋ*	NPF-heavy
	téɹʔ	height
í-kū	*pí-kù*	here (PROX-LOC)
əmə̄t	*əmə̄t*	hold.PAST
tə́-jíʔ	*tə̄-jīʔ*	NPF-horn
	tə̄-tʃà	NPF-hot
kútá	*kútá*	how
	nūklàŋ	hundred
tə́-pájáʔ	*tə́-pázāʔ*	NPF-husband
ɹə̄tʃə̄p	*ázù*	ice
	mə̄sən	insect lava
	āsājā	jest.PAST
á-uŋ	*ā-uŋ*	NPF-jungle
təpsēt/ānsə̄t	*təpsēt*	kill.PAST
tə̄-mə̄kūk	*tə́-mə́kùk*	NPF-knee
mə̄tə̄t	*mə̄tə̄t*	know.PAST
	mə̄tsʰə̄	knot.PAST
áwátsə́ŋ		lake
mə̄nī	*mə̄nī*	laugh.PAST; smile.PAST
tʰə́ptʃák		lay.PAST
nʉ̄		lead.PAST
səŋàwà	*tə̄-wā*	leaf/NPF-leaf
āitʃā		left side
	pāŋkə̄p	lip
	tə̄-kə̄m	NPF-live
tə̄-mə̄sə̄n	*tə̄-mə̄sə̄n*	NPF-liver
tə́-làŋ	*tə́-làŋ*	NPF-long
ātsə̄	*ātsə̀*	look.PAST
átsə́k	*ātsə̀k*	louse
āpāŋtʃāŋə̄ɹ		man
kúláŋà	*kúlāŋà*	many
á-sáʔ	*á-sàʔ*	NPF-meat
	tə̄-ɹə̄m	NPF-middle; NPF-waist
làtà	*làtà*	moon
ùtʃā	*tə̄-là/t-ʉ̄*	mother/NPF-mother[1]
tə̄-nə̄m	*tə̄-nə̄m*	NPF-mountain
tə̄-pāŋ	*tə̄-pāŋ*	NPF-mouth
	íjàʔ	much
tə̄-nīŋ	*tə̄-nìŋ*	NPF-name
tə́-mə́sə́		NPF-narrow
	tə̄-pʰə̄lā	NPF-navel

1 Some kinship terms differ according to the clan of the speaker.

CL	MA	Gloss
áná	*ánà*	near
tə̄-kʰūŋ	*tə̄-kʰūŋ*	NPF-neck
tə̄-sə̄n	*tə̄-sə̄n*	NPF-new
ā-jā	*á-jā*	NPF-night
tə́-náʔ	*tə́-náʔ*	NPF-nose
mə́tʃʰá	*mə́tʃʰà*	not (NEG+be)
	hàiʔ	OK
tə̄-tʃə̄n	*tə́-tʃə́n*	NPF-old
ā-kʰə̄t à	*ā*	NPF-one + INDEF/INDEF
tùŋə̀ɹ	*tūŋə̀ɹ*	other
ā-mīʔ	*ā-mīʔ*	NPF-person
	á-uk	NPF-pig
tʃʰái		play (sport etc.)
āɹ		poisonous creeper
jʉ̄sī		polish (?)
	tə́-mùʔ	NPF-pork.fat
tsʰə̀ʔ	*tsʰə̀ʔ*	pull.PAST
	tsʰə́ŋ	punch.PAST
nūŋ	*nūŋ*	push.PAST
tsə́ŋí	*tsə́ŋí*	rain
áʉʔ	*āfə̀*	rat
tə̄-mə̄ɹə̄m	*tə̄-mə̄ɹə̄m*	NPF-red
sītāk	*sítàk*	right/correct
	tʃʰáuŋ tāŋī	right side
	tə-niŋ	NPF-ripe
ā-jūŋ	*ā-jūŋ*	NPF-river
jə̀māŋ	*jə̀māŋ*	road
tī-jīŋ/tə̄-zə̄ŋ		NPF-root
lísə́	*līsə̀*	rope
t-āp	*t-āp*	NPF-rotten
tə̄-lūŋlūŋ		NPF-round
	tʃʰə̀t	rub.PAST
mə̄tsə̄ʔ	*mə̄tsə̄ʔ*	salt
lísə́m	*lísə̀m*	sand
sā	*sā*	say.PAST
nàk	*nàk*	scratch.PAST
tsìjə̄m	*tsijə̄m*	sea
ùŋ	*ùŋ*	see.PAST
mə̄tsə̄ʔ	*mə̄tsə̄ʔ*	seed
	tʰə̄nī	seven
ə́ɹ	*ə̄ɹ*	sew.PAST
tə́-tʃʰáʔ	*tə́-tʃʰàʔ*	NPF-sharp
tə̄-tsʰə̄	*tə́-tsʰə̀*	NPF-short
	kə̀p	shoot.PAST

CL	MA	Gloss
	ásà	shout.PAST
	tā-tī	NPF-older.sibling
	tā-nū	NPF-younger.sibling
tən	*tən*	sing.PAST
mān	*mān*	sit.PAST
tā-kāp	*tā-kāp*	NPF-skin
á-níŋ	*á-nìŋ*	NPF-sky
īpī	*jīplī*	sleep.PAST
	wa̰ʔ	slice.PAST
t-ásúʔ	*t-ásúʔ*	NPF-small
mānām	*mānām*	smell.PAST
mākʰūlī	*mūkʰūlī*	smoke
sāsā/tā-sāsā	*sāsā*	smooth.PAST/NPF-smooth
āhʉ	*āfʉ*	snake
	átʃʰì	sneeze.PAST
kʰíaʔ	*tíaʔ*	some
	mìtʃī	soot
	ā-mī	NPF-spear
mātsə́	*mātsā*	spittle
phāktāŋ		split.PAST
ātʃāt	*ātʃāt*	squeeze.PAST
tsāŋ	*tʃʰùʔ*	stab.PAST
jùŋlī	*jùŋlī*	stand.PAST
phātīnū	*phātīnū*	star
á-tʃúk		NPF-stick
á-lúŋ	*á-lúŋ*	NPF-stone
tā-mātūŋ	*īntʰàŋ*	NPF-straight/straight.PAST
mə́tʃáp	*mə́tʃāp*	suck.PAST
tsàŋī	*tsāŋī*	sun
	wàk	swell.PAST
wa̰ʔ	*átsə́ wa̰ʔ*	swim (water + slice.PAST)
tā-mīʔ	*tə́-mí-làk*	NPF-tail/tail-TERM
	tʃʰī	take.PAST
	tʰāɹā	ten
	sātʃāŋ	testicle
	ípáʔ/pá-tsə́	that (SPEC)/SPEC-DET)
ə́ukū	*pá-tsə́ kū*	there (SPEC-DET + LOC)
tā-mālām	*tā-mālām*	NPF-thick
	āwàɹʔ	thief
	tə́-pʰì	NPF-thigh
t-āpā	*t-āpā*	NPF-thin
	pʰìlàm	think.PAST
ípá-ī	*īpà-i*	this (SPEC-PROX)
ā-sām	*à-sām*	NPF-three

CL	MA	Gloss
tʰəp	*tʰəp*	throw.PAST
tsʰə̀ŋ		tie.PAST
tə̄-mə̄lī	*tə̄-mə̄lī*	NPF-tongue
tə̄-pʰā	*tə̄-pʰā*	NPF-tooth
sə̀ŋtūŋ	*sə̀ŋtūŋ*	tree
máŋə́p	*máŋə̀p*	turn.PAST
	mə̄kī	twenty
ā-nə̄t	*á-nə̀t*	NPF-two
ā-jīm	*ā-jīm*	NPF-village
átʰúʔ	*átʰúʔ*	vomit.PAST
tʃə̀lì	*tʃìlì*	walk.PAST
tə̄-lə̄m	*tə̄-lə̄m*	NPF-warm
tsʰə́uk	*tsʰə́uk*	wash.PAST
ā-tsə̄	*á-tsə̄*	NPF-water
tʃʰə́m		wear.PAST
tʃə̀p	*tʃə́p*	weep.PAST
tə̄-mə̄tsə̄	*tə́-mə́tsə́*	NPF-wet
tʃə́páʔ	*tʃə́páʔ*	what?
kújə́m	*kújə̀m*	when?
kə́tsə́u-kū	*tʃə́pá-kū*	where? (what-LOC)
tə̄-mə̄sə̄ŋ	*tə̄-mə̄sə̄ŋ*	NPF-white
sə́páʔ	*sə́páʔ*	who?
sàtə̄m	*sàtə̄m*	wide.PAST
tə́-nʉ́ʔ	*tə́-nìʔ*	NPF-wife
	tsə́ŋì	wild dog
mūŋ	*mūŋ*	wind
tə́-tʃʰá	*tə́-tʃʰá*	NPF-wing
mījʉ̄	*zūktʃūk*	wipe.PAST
ánʉ́tī	*ánʉ́tī*	woman
	mə̄pā	work
ā-kə̄m	*ā-kə̄m*	NPF-year
	háuʔ	yes

References

Anderson, Steven R., 1984, A metrical interpretation of some traditional claims about quality and stress. In Aronoff and Oehrle, eds 1984:83–106.

Aronoff, Mark and Richard T. Oerle, eds, 1984, *Language sound structure: studies in phonology*. Cambridge, MA.: MIT Press.

Baken, R.J., 1987, *Clinical measurement of speech and voice*. Boston: College-Hill Press.

Benedict, Paul K., 1972, *Sino-Tibetan: a conspectus*. (Contributing editor: James A. Matisoff). Cambridge: Cambridge University Press.

Blakenship, Barbara, Peter Ladefoged, Peri Bhaskararao and Nichumeno Chase, 1993, Phonetic structures of Khonoma Angami. *Linguistics of the Tibeto-Burman Area* 16.2:69–88.

Blevins, Juliette, 1996, The syllable in phonological theory. In Goldsmith, ed. 1996:206–244.

Bradley, David, 1980, Phonological convergence between languages in contact: Mon-Khmer structural borrowing in Burmese. In *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society*, 259–267.

1997, Tibeto-Burman languages and classification. In Bradley, ed. *Papers in Southeast Asian linguistics No.14: Tibeto-Burman languages of the Himalayas*, 1–71. Canberra: Pacific Linguistics.

Burling, Robbins, 1960, Angami Naga phonemics and word list. *Journal of Indian Linguistics* 21:51–60.

1969, Proto-Karen: a reanalysis. In Becker, A.L., ed. *Occasional papers of the Wolfenden society on Tibeto-Burman linguistics* (OPWSTBL). University of Michigan: Department of Linguistics.

1983, The Sal languages. *Linguistics of the Tibeto-Burman Area* 7/2:1–32.

1992, Garo as a minimal tone language. *Linguistics of the Tibeto-Burman Area* 15/2:33–51.

Cairns, Charles E. and Mark H. Feinstein, 1982, Markedness and the theory of syllable structure. *Linguistic Inquiry* 13/2:193–225.

Catford, J.C., 1964, Phonation types: the classification of some laryngeal components of speech production. In David Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott and J.L.M. Trim, eds *In honour of Daniel Jones: papers contributed on the occasion of his eightieth birthday, 12 September 1961*, 26–52. London: Longmans.

1977, *Fundamental problems in phonetics.* Edinburgh: Edinburgh University Press.

Census of India, 1981, Census Department, Government of India.

Chao, Yuen Ren, 1930, A system of tone letters. *Le Maître Phonétique* 30:24–27.

1970, *A grammar of spoken Chinese.* Berkeley and Los Angeles: University of California Press.

Chelliah, Shobhana, 1997, *A grammar of Meithei.* Berlin and New York: Mouton de Gruyter.

Chomsky, Noam and Morris Halle, 1968, *The sound pattern of English.* New York: Harper and Row.

Clark, Mrs Edward Winter [aka Mary Mead Clark], 1981, *The Ao Naga grammar.* Delhi: Gian Publications. (First published in 1893.)

Clark, Rev. Edward Winter, 1990, *Ao-Naga dictionary.* Mokokchung (Nagaland): R. Lisen Ao. (First published in 1911.)

Coupe, Alexander Robertson, 1998, The acoustic and perceptual features of tone in the Tibeto-Burman language Ao Naga. In Robert Mannell and Jordi Robert-Ribes, eds *Proceedings of the 5th international conference on spoken language processing*, Vol. 2, 57–60. Canberra: The Australian Speech Science and Technology Association. (Also produced on CD-ROM ISBN 1 876346 17 5.)

1999, A phonetic and phonological description of Ao. MA thesis, The Australian National University.

2001, Thai. In Jane Garry and Carl Rubino, eds *Facts about the world's languages*, 733–739. New York and Dublin: H.W. Wilson.

CSL computerized speech lab instruction manual, 1994, Model 4300B Version 5.X. Kay Elemetrics Corp, N.J.

Durand, Jacques, 1990, *Generative and non-linear phonology.* London and New York: Longman.

Evans, Nicholas, 1995, Current issues in the phonology of Australian languages. In Goldsmith, ed. 1995:723–761.

Fant, Gunnar, 1960, *Acoustic theory of speech production.* The Hague: Mouton.

1973, *Speech sounds and features.* Cambridge MA. and London: MIT Press.

Goldsmith, John A., 1976, Autosegmental phonology. Doctoral dissertation, MIT.

1990, *Autosegmental and metrical phonology.* Oxford and Cambridge MA.: Blackwell.

Goldsmith, John A., ed., 1995, *Handbook of phonological theory.* Oxford and Cambridge MA.: Blackwell.

Gowda, K.S. Gurubasave, 1972, *Ao-Naga phonetic reader.* Mysore: Central Institute of Indian Languages.

1975, *Ao grammar.* Mysore: Central Institute of Indian Languages.

Grierson, Sir George A., ed., 1967, *Linguistic survey of India* Vol. 3, Part 2. Delhi: Motilal Banarsidass. (First published 1903–1928.)

Greenberg, Joseph, 1978, Some generalizations concerning initial and final consonant clusters. In Joseph Greenberg, ed. *Universals of human language, Vol. 2: Phonology*, 243–280. Stanford: Stanford University Press.

Hale, Austin, 1982, *Research on Tibeto-Burman languages.* Mouton: Berlin.

Halle, Morris and K.N. Stevens, 1971, A note on laryngeal features. *MIT quarterly progress report* No.101, 198–213.

Harris, Zellig S., 1944, Simultaneous components in phonology. *Language* 20/1:181–205.

Harvey, Mark, 1991, Glottal stop, underspecification, and syllable structures among the Top End languages. *Australian Journal of Linguistics* 11:67–105.

Hecker, M.H.L. and Kreul, E.J., 1971, Descriptions of the speech of patients with cancer of the vocal folds. Part I: measures of fundamental frequency. *Journal of the Acoustical Society of America* 49:1275–1282.

Henderson, Eugénie J.A., 1948, Prosodies in Siamese: a study in synthesis. *Asia Major*, 189–215, and reprinted in F.R. Palmer, ed. 1970, *Prosodic analysis*, 27–53. London: Oxford University Press.

1951, The phonology of some loanwords in some South-East Asian languages. *Transactions of the Philological Society 1951*, 131–158.

1952, The main features of Cambodian pronunciation. *Bulletin of SOAS* 14:159–174.

1965, *Tiddim Chin: a descriptive analysis of two texts.* London: Oxford University Press.

Hooper [Bybee], Joan B, 1972, The syllable in phonological theory. *Language* 48/3:525–540.

Howie, John M., 1974, On the domain of tone in Mandarin. *Phonetica* 30:129–148.

Hulst, Harry van der and Norval Smith, eds, 1982, *The structure of phonological representations (Part II).* Dordrecht: Foris.

Hyman, Larry M. and Russell G. Schuh, 1974, Universals of tone rules. *Linguistic Inquiry* 5/1:81–115.

Iwata, S. and von Laden, H., 1970, Pitch perturbations in normal and pathological voices. *Folia Phoniatrica* 22:413–424.

Kahn, Daniel, 1976, Syllable-based generalizations in English phonology. Doctoral Dissertation, MIT.

Kapfo, Kedutso, 1989, Tones in Kheza noun constructions. *Linguistics of the Tibeto-Burman Area* 12/2:67–78.

Kenstowicz, Kenneth and Charles Kisseberth, 1979, *Generative phonology: description and practice.* New York: Academic Press.

Ladefoged, Peter, 1971, *Preliminaries to linguistic phonetics.* Chicago and London: University of Chicago Press.

1982, *A course in phonetics.* New York: Harcourt Brace Jovanovich.

Ladefoged, Peter and Ian Maddieson, 1996, *The sounds of the world's languages.* Oxford and Cambridge MA.: Blackwell.

Ladefoged, Peter, Ian Maddieson and Michel T.T Jackson, 1988, Investigating phonation types in different languages. In O. Fujimura, ed. *Vocal physiology: voice production, mechanisms and functions*, 297–317. New York: Raven.

Lapointe, Steven G. and Mark H. Feinstein, 1982, The role of vowel deletion and epenthesis in the assignment of syllable structure. In van der Hulst and Smith, eds 1982:69–120.

Lass, Roger, 1984, *Phonology: an introduction to basic concepts.* Cambridge: Cambridge University Press.

Laver, John, 1976, *Voice quality performance.* Audio cassette accompanying Laver 1980.

1980, *The phonetic description of voice quality.* Cambridge: Cambridge University Press.

Leben, William R., 1973, Suprasegmental phonology. Doctoral Dissertation, MIT.

Lehiste, Isle, ed., 1970, *Suprasegmentals.* Cambridge MA and London: MIT Press.

Lieberman, Philip, 1961, Perturbations in local pitch. *Journal of the Acoustical Society of America* 33:597–603.

1963, Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *Journal of the Acoustical Society of America* 35:344–353.

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

Maddison, Ian, 1984, *Patterns of sounds.* Cambridge: Cambridge University Press.

Marrison, Geoffrey Edward, 1967, The classification of the Naga languages of north-east India. Volume I (Text) and Volume II (Appendices). Doctoral dissertation. SOAS University of London.

Matisoff, James A., 1973, *The grammar of Lahu.* Berkeley and Los Angeles: University of California Press.

1989, The bulging monosyllable, or the mora the merrier: echo-vowel adverbialization in Lahu. In J.H.C.S Davidson, ed. *South-East Asian linguistics: essays in honour of Eugénie J.A. Henderson*, 163–197. London: School of Oriental and African Studies.

Mazaudon, Martine, 1974, Notes on tone in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 1:27–54.

McCarthy, John, 1986, OCP effects: gemination and antigemination. *Linguistic Inquiry* 17:207–263.

1988, Feature geometry and dependency: a review. *Phonetica* 43:84–108.

Mills, James Philip, 1926, *The Ao Nagas.* London: Macmillan and Co.

n.d. Typescript, *Mongsen Ao word list.* Library of SOAS, University of London.

Morse, Robert H, 1963, The phonology of Rawang. *Anthropological Linguistics* 5/5:17–41.

Nartey, J.N.A, 1979, A study in phonemic universals — especially concerning fricatives and stops. *UCLA Working Papers in Linguistics* 46.

Okell, John, 1969, *A reference grammar of colloquial Burmese.* London: Oxford University Press.

Painter, Colin, 1979, *An introduction to instrumental phonetics.* Baltimore Park Press: Baltimore.

Peterson, G.E. and Isle Lehiste, 1960, Duration of syllable nuclei in English. *Journal of the Acoustic Society of America* 32/6:693–703, and reprinted in Isle Lehiste, ed. 1967, *Readings in acoustic phonetics*, 191–201. Cambridge MA. and London: MIT Press.

Pike, Kenneth, 1947, *Phonemics*. Ann Arbor: The University of Michigan Press.

1948, *Tone languages*. Ann Arbor: The University of Michigan Press.

Rabel, Lili, 1961, *Khasi, a language of Assam*. Baton Rouge: Louisiana State University Press.

Rose, Phillip John, 1992, Bidirectional interaction between tone and syllable coda: acoustic evidence from Chinese. In Jeff Pittam, ed. *Proceedings of the Fourth Australian International Conference on Speech Science and Technology,* 292–297. Canberra: Australian Speech Science and Technology Association.

Selkirk, Elizabeth, 1982, The syllable. In van der Hulst and Smith, eds 1982:337–383.

1984, On the major class features and syllabicity. In Aronoff and Oerle, eds 1984:107–136.

Shafer, Robert, 1955, Classification of the Sino-Tibetan languages. *Word* 11:94–111.

1974, *Introduction to Sino-Tibetan*. Weisbaden: Otto Harrassowitz.

Temeshi Lai [The Holy Bible], 1964 (in the Chungli dialect of Ao). Bangalore: Bible Society of India.

Trubetskoy, N.S., 1977, *Grundzüge der Phonologie (Principles of phonology).* Göttingen: Vandenhoeck & Ruprecht. (First printed 1939.)

Weidert, Alfons, 1979, The Sino-Tibetan tonogenetic laryngeal reconstruction theory. *Linguistics of the Tibeto-Burman Area* 5/1:49–127.

1987, *Tibeto-Burman tonology: a comparative account*. Amsterdam/Philadelphia: John Benjamins.

Wolfenden, Stuart Norris, 1929, *Outlines of Tibeto-Burman linguistic morphology*. London: Royal Asiatic Society.