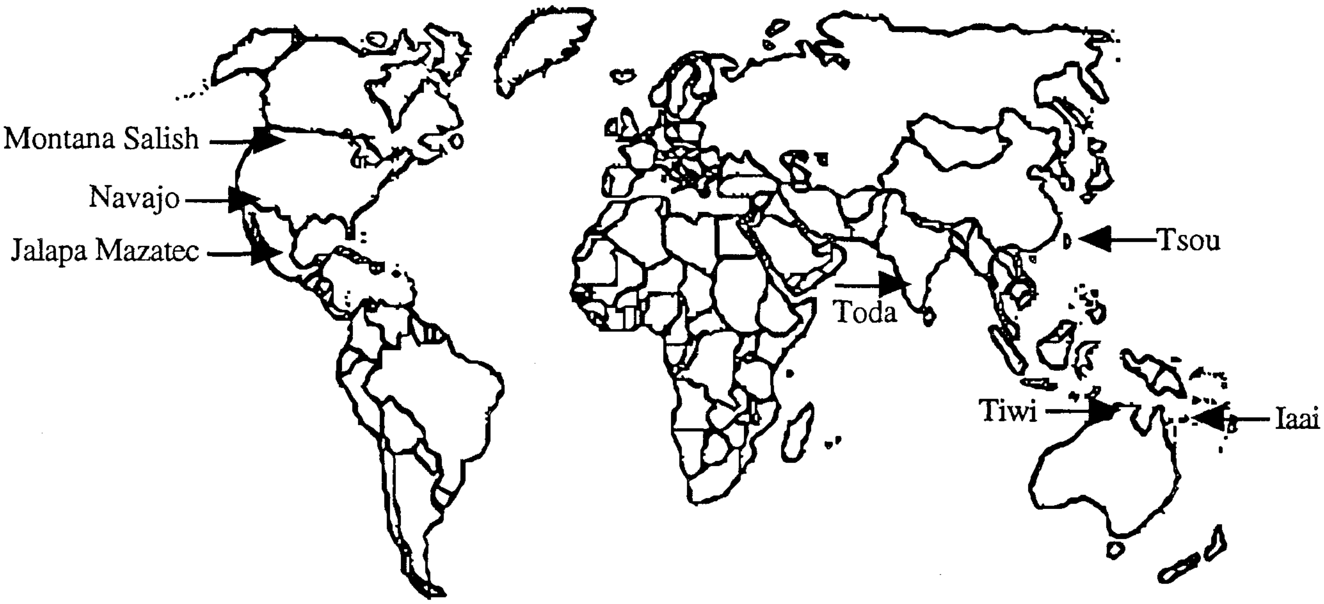
Title
WPP, No. 87: Fieldwork Studies of Targeted Languages II

Publication Date
1994-08-01

Fieldwork Studies of Targeted Languages II



UCLA Working Papers in Phonetics 87

August 1994

UCLA Working Papers In Phonetics 87

August, 1994

Fieldwork Studies Of Targeted Languages II

Table of Contents

Phonetic Structures of Montana Salish

Edward Flemming and Peter Ladefoged
Linguistics Department, UCLA, Los Angeles, CA 90024-1543
and
Sarah Thomason
Linguistics Department, University of Pittsburgh, Pittsburgh, PA 15260.

Montana Salish is an Interior Salishan language spoken on the Flathead reservation in Northwest Montana. The language is most often referred to by linguists as Flathead, but the people themselves refer to it as Salish. As the language family is also often referred to by linguists as Salish, we use the term Montana Salish to make it clear that we are discussing this particular Salishan language. We estimate that there are about 70 speakers of Montana Salish. The two main branches of the Salishan language family, which comprises about twenty-three languages, are the Coast and Interior branches (see Thompson 1979:693 for a complete classification); Montana Salish belongs to the Southern group of the Interior branch. There are few published linguistic studies of Montana Salish, aside from word lists, but there are a number of descriptions of Kalispel and Spokane, which are closely-related dialects of the same (nameless) language (e.g. Vogt 1940, Carlson 1972, 1976, Bates & Carlson 1992, Orser & Carlson 1993).

This paper is based on recordings made in August 1992 at the Flathead Cultural Center on the Flathead Reservation in Montana, thanks to the assistance of the Director of the Flathead Culture Committee, Clarence Woodcock. With the aid of a dictionary which is being compiled by the third author, we had constructed a word list illustrating all the major phonological contrasts, and containing additional material for examining sounds that were of particular phonetic interest. This list was then refined through extensive work with our three main language consultants: Harriet Whitworth, Felicité ("Jim") McDonald and Dorothy Felsman. There was a great deal of discussion. Some forms in the original list were deleted as unknown or not exemplifying the sound sought, others were added, and the spellings were corrected. The complete list of words finally selected is given in the appendix.

Salishan languages have a complex morphological and phonological structure, which makes it impossible to illustrate contrasting sounds using simple minimal sets equivalent to (Californian) English 'pot, tot, cot; bought, dot, got ...' etc. As far as possible, words were selected in which the particular points to be investigated were in the roots. All the words in the original list were attested somewhere,primarily in materials collected in previous fieldwork by the third author or in materials prepared by the Culture Committee. Some of them were somewhat uncommon so that speakers had to be reminded of them before they could recognize them out of context. In some of the materials from which the words were drawn, particles and affixes that would occur with the words in normal, or at least the most common, contexts were omitted. But we have no doubt that all the forms elicited were considered to be proper Montana Salish forms by all our speakers.

When we had agreed on a satisfactory list of 274 words, we made a recording of the three consultants saying it. One of the authors supplied an English gloss as a prompt, Harriet Whitworth said the Salish word, and the other two consultants repeated the word after her. The consultants did have access to a version of the written list, but none of them was reading from it. We also made a recording in which each of these three consultants said the words in a frame sentence. The next day we got together a larger group of ten people consisting of all the Montana Salish speakers available in the Cultural Center at that time. We went through the whole list again, first rehearsing it, and then actually recording it. By this time the three main consultants were very familiar with all the words they were being asked to say. The rest of this larger group consisted of speakers with varying degrees of proficiency in Salish. For the present paper we will restrict our analysis to the three original consultants plus two male speakers, Clarence Woodcock, the Director of the Flathead Culture Committee, and the Associate Director, Antoine (Tony) Incashola. In this way we can be sure that we have a group of three women and two men who are clearly proficient native speakers of Montana Salish.

In this paper words will be cited in a surface phonemic transcription, using IPA symbols and following the majority pronunciation. The only problem raised by using an IPA transcription is that normal IPA practice is to transcribe affricates such as **ʧ** as equivalent to a sequence of the symbols **t** and **ʃ**. But in Montana Salish there is a phonetic (and at least potentially a phonemic) distinction between an affricate **ʧ** as in **ʔiɬtʃtʃé̓n** ‘tender (as, a sore spot)’ and a sequence **t ʃ** as in **stiʕít.ʃən** ‘killdeer’. In this paper **ʧ** will always represent an affricate, unless transcribed with a period between as **t.ʃ**.

Subphonemic features have not been noted except for the vowel **ə**, which has been transcribed, although its placement is largely predictable. This mode of transcription will be printed in bold face, e.g. **q^{w}’ájəlqs** ‘priest (black-robe)’. Where narrower transcriptions are required, they will be enclosed in square brackets, e.g. [q^{w}’ajəɬqs]. Narrow transcriptions will also be used in labeling figures.

Vowels

Vowel Qualities

Montana Salish has five vowels, **i, e, a, o, u.** In addition there is a schwa-like vowel which appears in unstressed syllables only. Vowels do not occur in word-initial position; words that orthographically have an initial vowel are in fact preceded by a glottal stop. Table I contains words illustrating the vowel contrasts in stressed position, in the context **p_l,** except for a, which is in the context **j_l.**

Table 1: Words illustrating vowel contrasts.

i	**pəlpíləlʃ**	‘stagger’
e	**ʧ’upélsi**	‘lonesome’
a	**jál**	‘round’
o	**pólpəlqən**	‘thimbleberry’
u	**púlsəm**	‘he killed something’

Figures 1 and 2 show plots of the vowel formants for the female and male speakers. The formant measurements were taken from the stressed vowels of a word list designed to illustrate the vowel contrasts in similar environments, avoiding vowels adjacent to pharyngeal consonants (1-20 in the word list, excluding 1 and 9 to avoid pharyngeals). The formant values were determined from LPC and FFT spectra on a Kay CSL system.The axes are scaled in proportion to the Bark scale, but are labelled in Hz, and show F1 and F2’-F1. F2’ is a weighted average of F2 and F3 calculated according to the following formula, due to Fant (1973:52):

$$F2' = F2 + \frac{(F3-F2)(F2-F1)}{2(F3-F1)}$$

The ellipses indicate two standard deviations from the mean along the principle components of each vowel distribution.

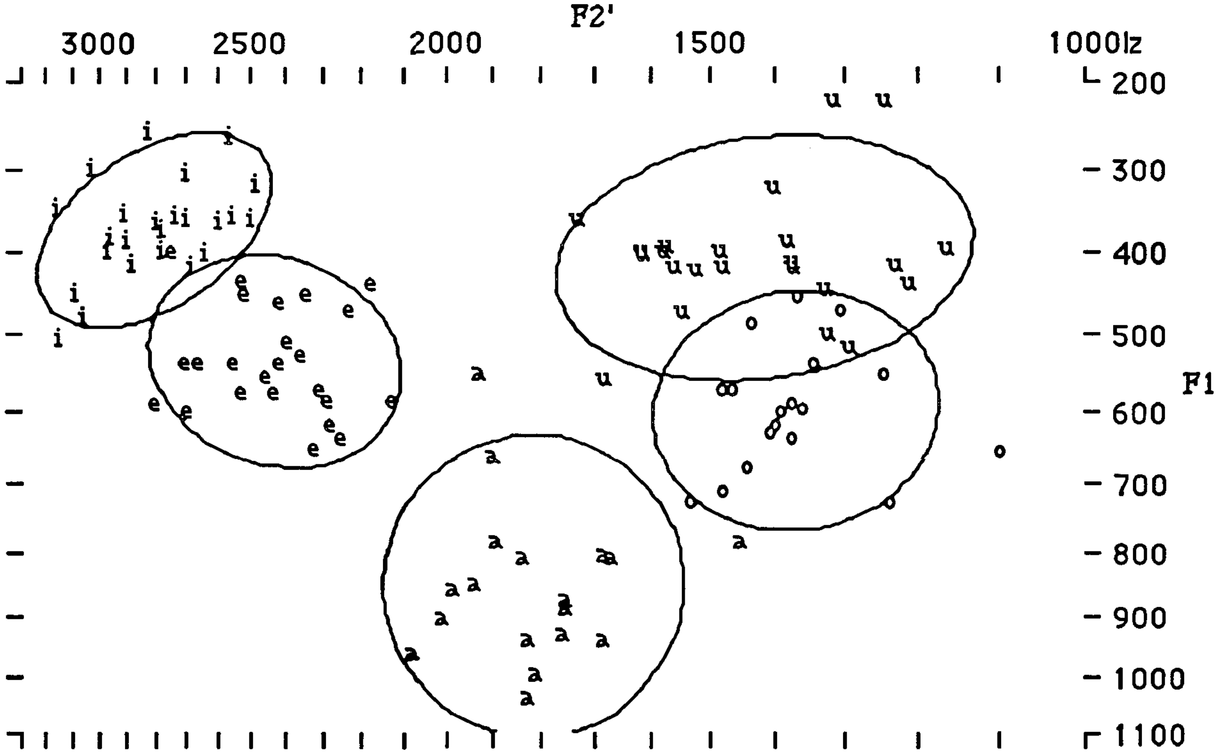


Figure 1. Formants of the Montana Salish vowels from three female speakers (see text for details).

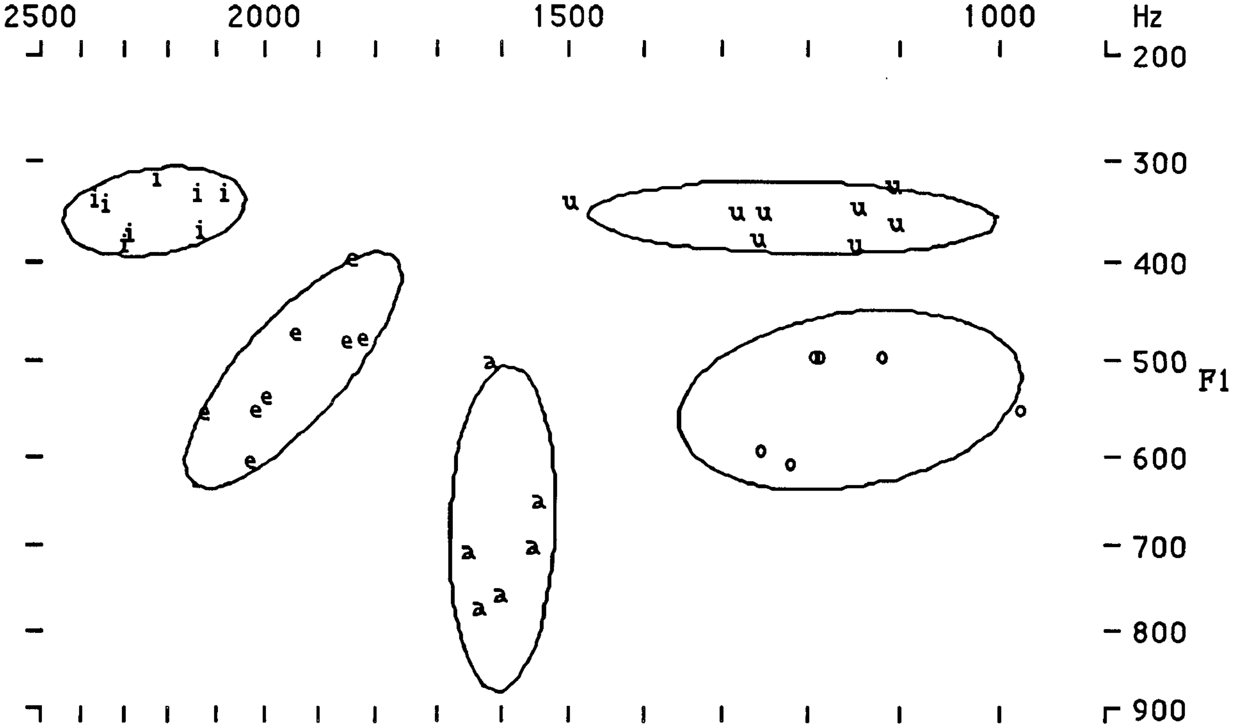


Figure 2. Formants of the Montana Salish vowels from two male speakers (see text for details).

The mean values of the first three formants of each vowel, for the male and female speakers, are shown in Table 2. It may be seen that the vowels are distributed in the vowel space much as in many five vowel languages, with the high back vowel u being not fully back.

Table 2. Mean formant values of vowels in Hz for three female and two male speakers.

VOWEL	FEMALE			MALE		
	F1	F2	F3	F1	F2	F3
i	372	2645	3058	349	2062	2536
e	535	2181	2949	510	1738	2396
a	854	1603	2807	683	1389	2434
o	601	1170	2768	540	994	2222
u	407	1168	2893	355	1011	2265

Intrinsic Pitch

It has been found that the F0 of vowels varies with vowel height: other things being equal, high vowels have higher F0 than low vowels (Ohala and Eukel 1987, and references therein). We tested this generalization against the Montana Salish vowels by testing for a correlation between F0 and F1, as an indicator of vowel height.

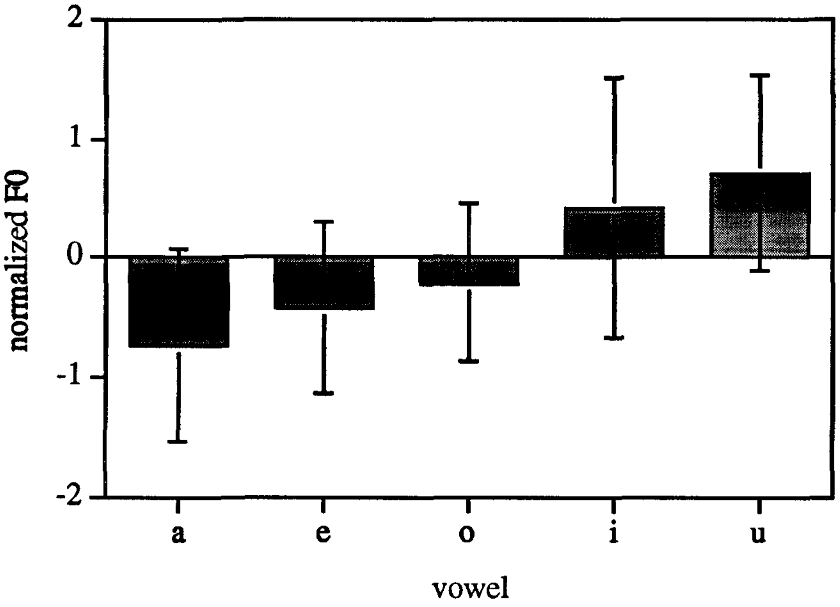


Figure 3. Mean normalized F0 for each vowel, across five speakers.

Montana Salish places a high tone on the accented syllable, so there is an F0 peak over the stressed vowels considered here. F0 measurements were taken at this peak. Then, to reduce the effects of inter-speaker variations in pitch range, the F0 measures were normalized. This was achieved by calculating the mean and standard deviation of the F0 distribution for each speaker individually, pooling all the vowels. Then F0 for each vowel token was converted to a number of standard deviations from the mean F0 for that speaker. F1 measurements were normalized using the same procedure. We found a highly significant, but not especially tight negative correlation (r=0.4, p<.0001) between the normalized F0 and F1 measures, in accordance with the generalization that higher vowels (i.e. vowels with lower F1) tend to have a higher F0. Figure three shows the mean normalized F0 for each vowel, error bars indicate the standard deviations. There is also a tendency for back vowels to have a higher F0 than the corresponding front vowels.

Generalizations concerning intrinsic vowel duration were not tested beacause our dara set does not include a set of words with vowels in sufficiently similar environments for comparison.

Consonants

Overview: The consonant inventory of Montana Salish is shown in Table 3. Words illustrating these sounds are shown in Table 4. We will first present general observations on the realization of these sounds, then turn to durational measurements, and more detailed discussion of the typologically unusual sounds in the Montana Salish consonant inventory: pre-stopped laterals, glottalized sonorants and pharyngeals.

Table 3. Montana Salish consonant phonemes

	bilabial	alveolar	palato-alveolar	palatal	velar	lab. velar	uvular	lab. uvular	pharyngeal	lab. pharyn.	glottal
plosive	p	t			(k)	kʷ	q	qʷ			ʔ
ejective stop	p’	t’				kʷ’	q’	qʷ’			
affricate		ts	tʃ								
ejective affricate		ts’	tʃ’								
lateral ejective affricate		tɬ’									
fricative		s	ʃ			xʷ	χ	χʷ			h
lateral fricative		ɬ									
nasal	m	n									
glottalized nasal	ˀm	ˀn									
approximant				j		w			ʕ	ʕʷ	
glottalized approximant				ˀj		ˀw			ˀʕ	ˀʕʷ	
lateral approximant		l									
glottalized lateral approximant		ˀl									

Stops If we include the palato-alveolar affricates, stops occur at five places of articulation, bilabial, alveolar, palato-alveolar, velar and uvular. At each of these places there is a voiceless unaspirated stop (see below for VOT measurements) and an ejective. The uvular stops may be plain or labialized. The velar stops are nearly always labialized; unlabialized **k** occurs only in loanwords. In addition to these pairs of stops, there is also a pair of alveolar affricates, and an unpaired alveolar lateral ejective affricate. All the ejectives are produced with a considerable lag between the release of the oral closure and the release of the glottal closure (see below for measurements). All the stops are clearly released, even in clusters and in final position.

Laterals: The laterals **l** and **ɬ** are prestopped in most environments by most speakers. Depending on its context it is realized as voiced **ᵈl**, or fricated **ᵗɬ** or **ᵈɮ**. When fricated, it can be phonetically similar to **ɬ**. In addition, as we have already noted, there is an alveolar lateral ejective affricate.

Glottalized sonorants: Nasals, laterals and central approximants all occur in both plain and glottalized forms. Typically, as we shall see below, the glottal constriction precedes the main portion of the sonorant.

Table 4. Words illustrating the Montana Salish consonants before **a**, or, in a few cases, before **e** or **o**.

bilabial			
p	**páʕas**		‘face is pale, grey’
p’	**p’aʕáp**		‘grass fire’
m	**máɬt**		‘mud’
ˀm	**ˀmeˀmstsú**		‘playing cards’
alveolar			
t	**tám**	*tám*	‘it’s not, wrong’
t’	**t’áq’ən**	*t’aq’ən*	‘six’
ts	**tsáqʷəlʃ**		‘western larch’
ts’	**ts’áɬt**		‘it’s cold’
s	**sáχʷ**		‘split wood’
n	**nás**		‘wet’
ˀn	**ˀneˀjxʷéˀ ws**		‘trade’
l	**láq’i**		‘sweatbath’
ˀl	**ʔəˀliáts**		‘red raspberry’
ɬ	**ɬáqʃəlʃ**		‘sit down!’
tɬ’	**tɬ’áq’**		‘hot’
palato-alveolar			
tʃ	**tʃájɬqən**		‘cut hair’
tʃ’	**tʃ’áwən**		‘I prayed (it)’
ʃ	**ʃáll**		‘he got bored’
palatal			
j	**jajáʔ**		‘maternal grandmother’
ˀj	**ˀjeˀjúkʷeʔ**		‘stingy’
velar			
k	**kapí**		‘coffee’
kʷ	**kʷáteʔ**		‘quarter (money)’
kʷ’	**kʷ’áltʃ’qən**		‘lid, cover’
xʷ	**xʷáltʃst**		‘reach (for something)’
w	**wáˀ ləwəˀl**		‘long-billed curlew’

ˀw	**ˀwiˀwá**	‘wild’
uvular		
q	**qáχeʔ**	‘aunt (mother’s sister)’
q’	**q’áq’ɬuʔ**	‘vein’
qʷ	**qʷátsqən**	‘hat’
qʷ’	**qʷ’ájəlqs**	‘priest (black-robe)’
χ	**χáˀ m**	‘dry’
χʷ	**χʷáq’ʷ**	‘to grind or file something’
pharyngeal		
ʕ	**ʕámt**	‘it’s melted’
ʕʷ	**ʕʷóˀl**	‘slippery (oily)’
ˀʕ	**jaˀʕəmím**	‘gathering (as, rocks)’
ˀʕʷ	**ˀʕʷoˀjəntsú**	‘laugh’
glottal		
ʔ	**ʔáwəntxʷ**	‘you said it’
h	**háˀ ʕʷumskʷ**	loosen it!

Duration measurements
VOT: There are no contrasts based primarily on Voice Onset Time (VOT), but VOT for the oral stops varies according to place of articulation as shown in Figure 4. The interval measured is that between the release burst and the onset of periodic voicing in the waveform. These figures accord with the general tendency for dorsal stops to have longer VOTs than labial or coronal stops. Statistical analysis (ANOVA) reveals that the difference between these two groups is significant at $p<.05$).

The VOT for **p** and **k** are close to those reported for French voiceless stops by O’Shaughnessy (1981), but the VOT for Montana Salish **t** is considerably shorter. The VOTs for **p**, **t** and **k** are all long compared to the values reported for voiceless unaspirated stops by Lisker and Abramson (1964), but far short of contrastively aspirated stops. The VOT of 54 ms (s.d. 28) for **q** is very close to the 56ms (s.d. 21) reported for K’ekchi q by Ladefoged and Maddieson (1986:22).

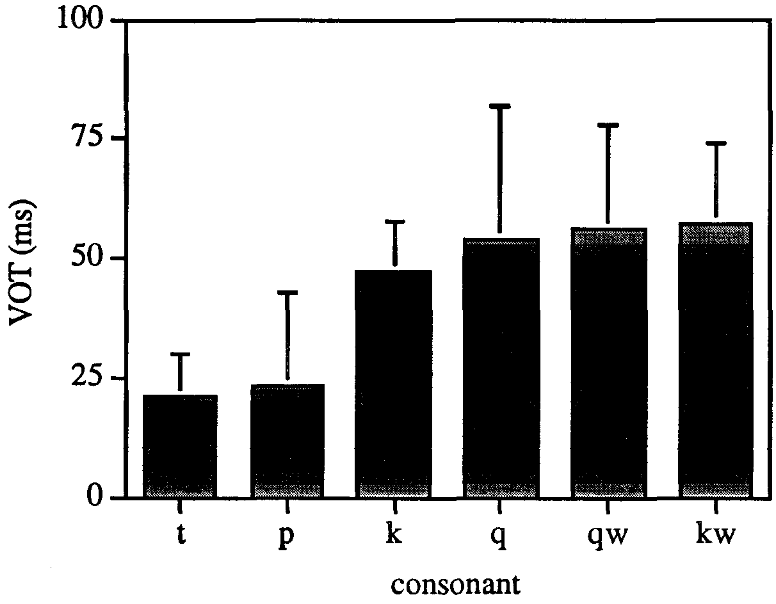


Figure 4. Mean voice onset time (s) by stop category for five speakers.

Glottal lag: In ejectives, the interval between the release of the oral closure and the release of the glottal closure was measured. The mean for each consonant is shown in Figure 5. As may be seen in Table 5, these values are closer to those reported for ejectives in K'ekchi (Ladefoged and Maddieson 1986:22) and Navajo (McDonough and Ladefoged 1993:154-5) than the shorter lags found in Hausa ejectives (Lindau 1984).

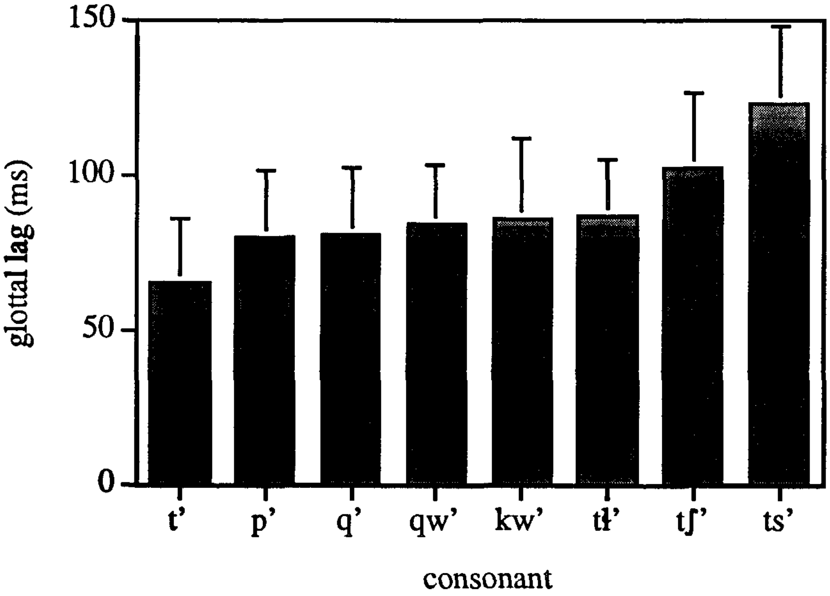


Figure 5. Mean duration of glottal closure after oral release for each ejective (five speakers).

Table 5. Comparison of glottal lags (ms) for ejectives in three languages. Standard deviations are given in parentheses.

CONSONANT	MONTANA SALISH	K’EKCHI	NAVAJO	HAUSA
t’	65 (21)	-	108 (31)	-
ts’	123 (25)	-	142 (41)	-
tʃ’	102 (24)	-	144 (24)	-
tɬ’	87 (18)	-	157 (40)	-
k’ or **kʷ’**	86 (26)	97 (38)	94 (21)	33
q’	81 (21)	92 (38)	-	-

Fricatives: Figure 6 shows the mean durations from onset of frication to the onset of voicing in the following vowel for initial fricatives. The differences in duration are not significant.

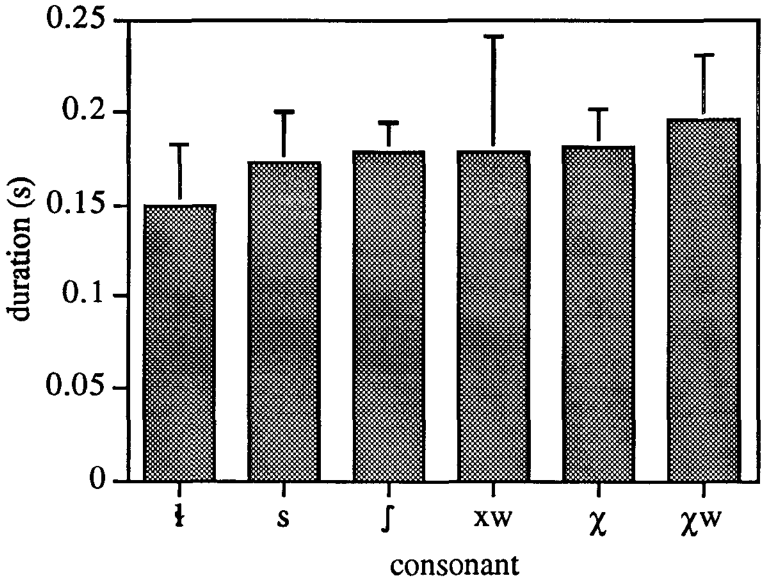


Figure 6. Mean duration by fricative type (five speakers).

Nasals and laterals: The mean oral constriction durations of initial laterals and nasals are shown in Figure 7. In the case of the glottalized nasals, the interval from the onset of voicing to the release of the oral closure was measured. This would exclude any initial glottal closure. The nasal portion of the glottalized nasals is considerably shorter than the plain nasals.

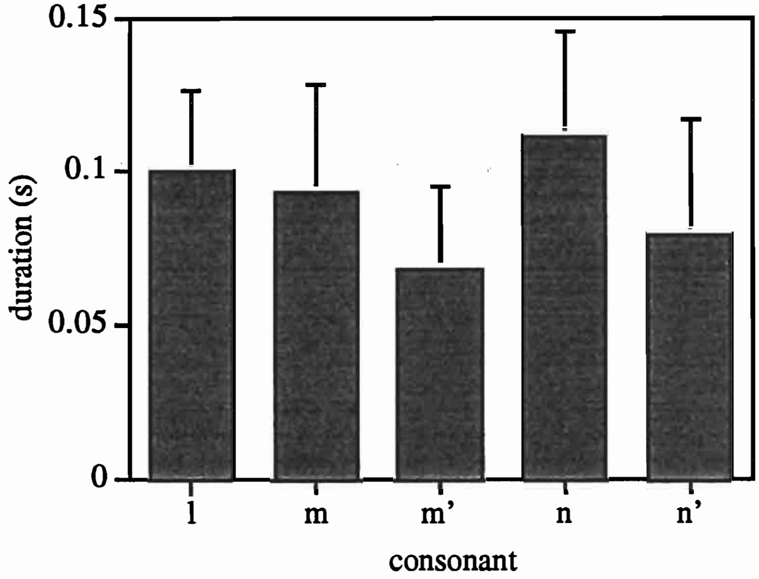


Figure 7. Mean oral constriction durations by consonant type (5 speakers).

Consonant clusters

Extremely complex consonant clusters are one of the most striking features of Salishan languages generally, and Montana Salish is no exception. An initial sequence of five consonants is exemplified in Figure 8, a spectrogram of the word **tʃɬkʷkʷtəˀ néˀws**'a fat little belly'. The initial palatoalveolar affricate is followed immediately by a voiceless alveolar lateral affricate, with no intervening vowel. This in turn is followed by a velar stop, again with no intervening vowel. This labialized velar stop is released, producing what, in a narrow phonetic transcription, could be considered a mid high back rounded voiceless vowel ʊ̥ as the realization of **w**, which is then followed by another similarly released stop. Finally in this cluster there is an aspirated alveolar stop before, for the first time in this word, regular voicing occurs.

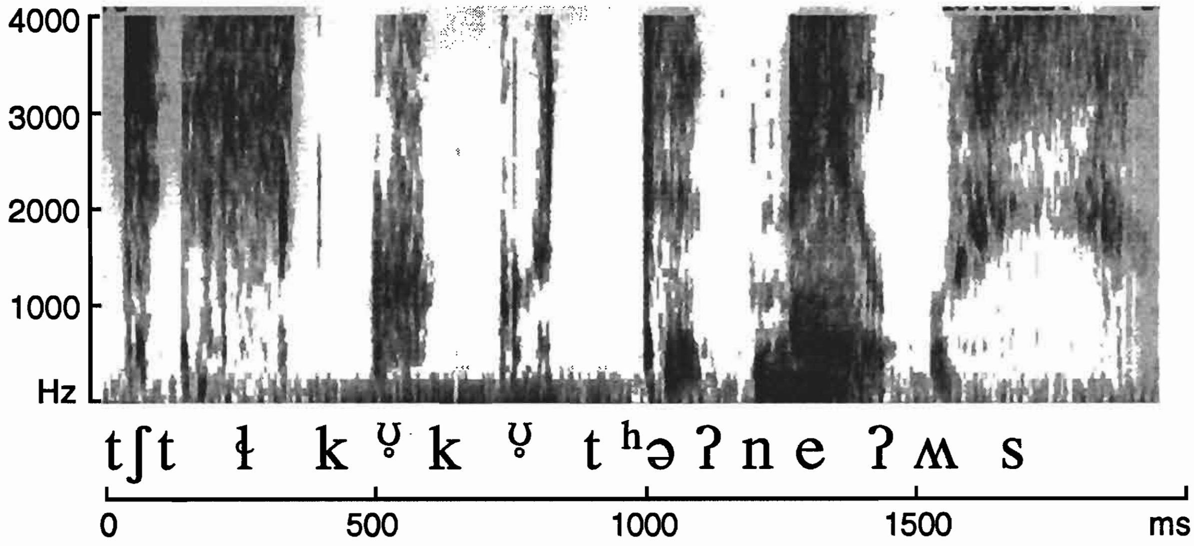


Figure 8. A spectrogram illustrating an initial sequence of five voiceless consonants in **tʃɬkʷkʷtəˀnéˀws**'a fat little belly'.

Figure 9 shows the word **tʃtʃts’éˀlʃtʃən**‘wood tick’, in which there is also a complex sequence of voiceless sounds. In this utterance the two palatoalveolar affricates are followed by an alveolar ejective affriate. After the vowel in this word, there is another complex sequence in which the phonemically preglottalized lateral, which is realized as a glottal stop followed by an alveolar release into a voiceless alveolar lateral fricative, is itself followed by a palatoalveolar fricative, before what is underlyingly a palatoalveolar affricate. In this case the stop part of the affricate does not have a complete closure, and is accordingly symbolized **t̞**, the diacritic indicating a lowered position.

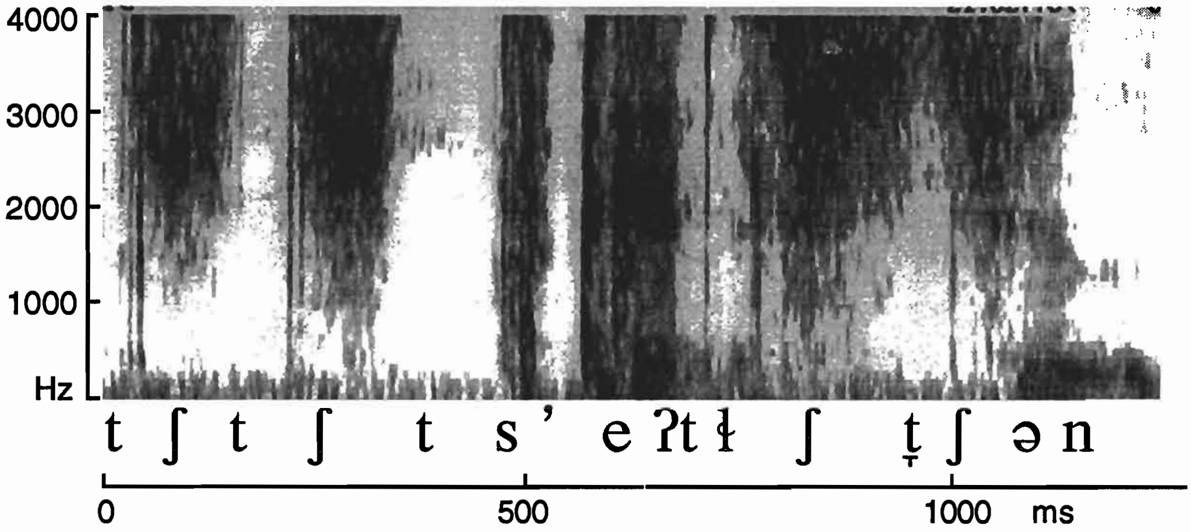


Figure 9. A spectrogram of the word **tʃtʃts’éˀlʃtʃən**‘wood tick’.

As is apparent from these examples, stops and ejectives are always strongly released, even in sequences of identical consonants, as well as in word-final position. Figure 10 provides a further illustration in the word **ppíˀl** ‘pint’. Stops are normally unaspirated (as discussed above, and as can be seen in the case of the second **p**), but in sequences such as this, the first stop is released with considerable aspiration. Since bursts provide crucial cues to the identity of a stop, this pattern of realization aids considerably in maintaining the perceptibility of all the consonants in a cluster.

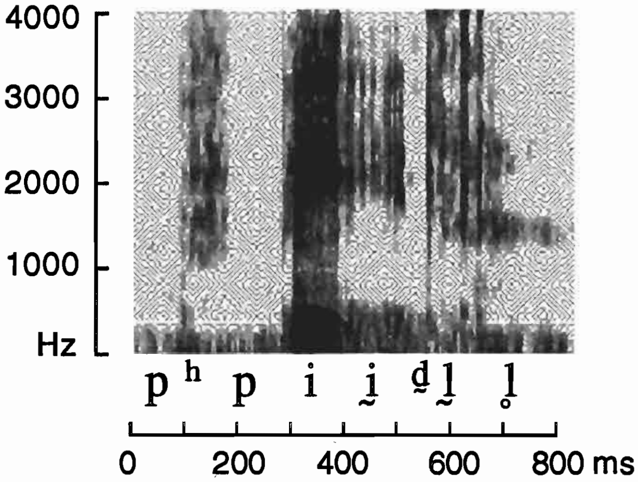


Figure 10. A spectrogram of the word **ppíˀl** (in a narrow transcription **ppíḭᵗɬə̥**) ‘pint’.

While there is considerable freedom in combining consonants in a sequence, there are restrictions. The clearest to emerge from the present study is that sonorants may not follow obstruents in a cluster, they are always separated by a vowel. The vowel transcribed [ə], a transitional vowel of indeterminate quality, only occurs between a sonorant and a preceding consonant. Thus we can regard this epenthetic vowel as inserted simply to break up a sequence of a consonant followed by a sonorant. The class of sonorants for the purposes of this rule consists of **l, m, n, w, j, ʕ, ʕʷ** and their glottalized counterparts. Similar phenomena have been observed in other Salishan languages, e.g. Shuswap (Kuipers 1974) and Thompson (Thompson and Thompson 1992), although in the latter language the schwa sometimes follows the sonorant rather than preceding it. Kuipers regards the schwa vowel as part of the realization of what is essentially a syllabic sonorant; however in Montana Salish it seems more likely that the schwa itself is the syllable nucleus since the following sonorant is often more plausibly assigned to the onset of the following syllable, e.g. in **wáˀləwəˀl** 'long-billed curlew' and **ˀməˀnétʃ** 'excrement', the underlined schwa is found preceding a sonorant, as expected, but the following sonorant is presumably syllabified with the following vowel in each case.

Sonorants are generally separated from other sonorants by a vowel also, but there are two exceptions to this generalization. Sonorants which are identical (without considering glottalization) are not separated by a vowel , as in **sqəllú** 'tale' and **ʔəlˀláts** 'red raspberry' Secondly, in sequences of more than two sonorants, they need not all be separated by vowels; for example, in some utterances of **səˀnəmˀné** 'toilet' there may be two but not three epenthetic vowels. The latter case presumably results from the syllabification of such sequences, but there are insufficient examples of heterogeneous sonorant sequences in the present data set to draw any conclusions about syllable structure in Montana Salish. Sonorants which are nominally in word-initial position preceding consonants are in fact preceded by a glottal stop and a short schwa like vowel, or are apparently syllabic and preceded by a glottal stop, presumably forming a syllable. Nasals are typically realized as syllabic in this environment, whereas laterals are typically preceded by a vowel.

Laterals

Montana Salish contrasts four laterals, a voiced lateral, a voiceless lateral fricative, a glottalized voiced lateral, and an ejective lateral affricate. We will consider the first two in this section; glottalized laterals are described later, together with the other glottalized sonorants. In most environments, the voiced and voiceless laterals are usually produced with a brief stop closure or some other gesture that produces a burst-like transient at the beginning of the lateral. However, this does not always occur. Figures 11 displays two utterances of the same word, **laq'əm** 'he buried', one with and the other without a transient associated with the **l**. In the token on the left, produced by speaker FM, there is not only a transient shortly after the beginning of the lateral, but also evidence of a fricative component in the higher frequencies. We have transcribed this utterence with an initial **t**, but it is difficult to show that a voiceless alveolar stop occurred at the beginning of the word. In the utterance on the right, produced by speaker AI, there is no transient component.

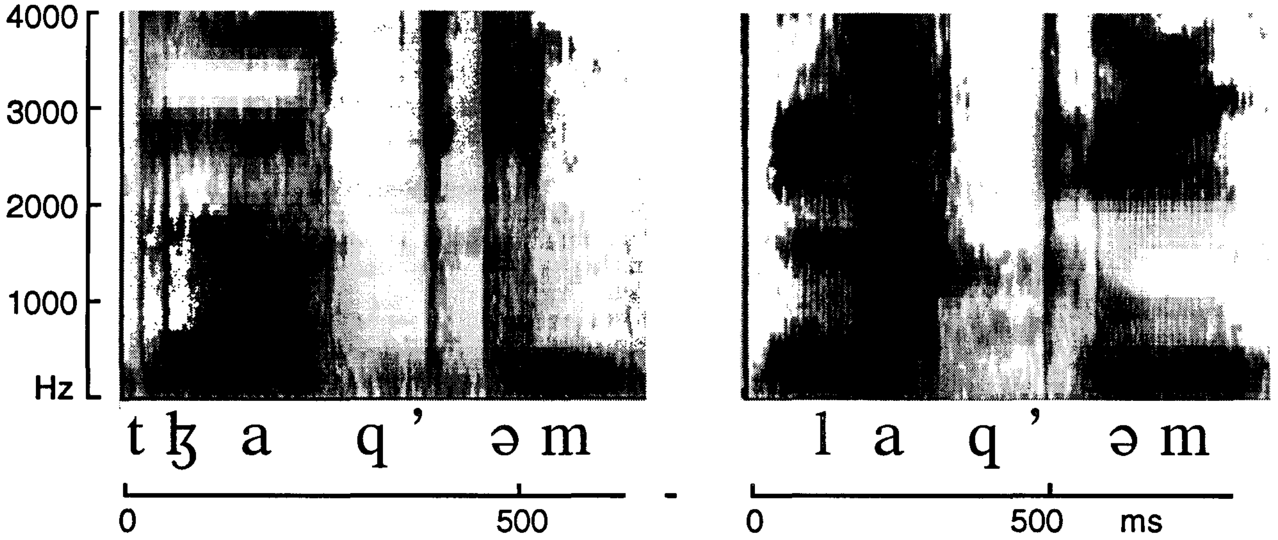


Figure 11. Two pronunciations of **laq’əm** ‘he buried’.

Laterals are preceded by an evident stop closure in most word-internal environments. The context where they are more consistently produced without an intial stop closure is the case of the second lateral in a cluster of laterals. Note that such clusters are possible because similar sonorants are not separated by a schwa. Even here there is sometimes a burst between the two consonants, as illustrated in Figure 12, which shows the sequence of consonants in the middle of the word **p’əlll̆tʃ’tʃ** ‘turned over’. The first of these two laterals has a stop closure preceding it, and a burst as this closure is released. The second lateral has no such closure, but there is a transient in the spectrum, closely resembling that produced by the release of a stop. How this transient is produced is not clear to us at the moment. Such discontinuities indicate that, in these cases at least, the sequences are truly clusters of identical consonants rather than long consonants.

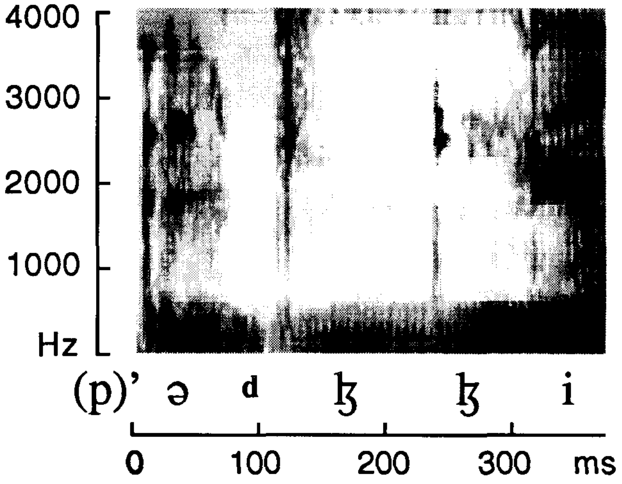


Figure 12. The sequence of lateral consonants in the middle of the word **p’əlll̆tʃ’tʃ** ‘turned over’.

The voiced lateral is often fricated (as can be seen from the noise in the spectrogram of speaker FM’s production of **laq’əm** in Figure 11). It also devoices in word-final position and preceding voiceless consonants. The early portion of the lateral sometimes remains voiced or breathy, but otherwise the result of devoicing is very similar to the voiceless lateral fricative. Figure 13 compares **l** and **ɬ** before voiceless consonants. In the word on the left of the figure, **k^{w}'áltʃ'qən** 'lid, cover', the first 50 ms of the **l** are voiced, but a slightly longer part is voiceless and fricative. In the word on the right, **máɬt** 'mud', there is a short voiceless stop before the entirely voiceless **ɬ**. Maddieson and Emmorey (1984) found that voiceless lateral approximants have lower amplitude than voiceless lateral fricatives, relative to the amplitude of a following vowel. A similar measure did not distinguish Montana Salish devoiced laterals from the voiceless lateral fricatives in similar environments, preceding voiceless consonants, supporting the auditory impression that the devoiced laterals are fully fricated.

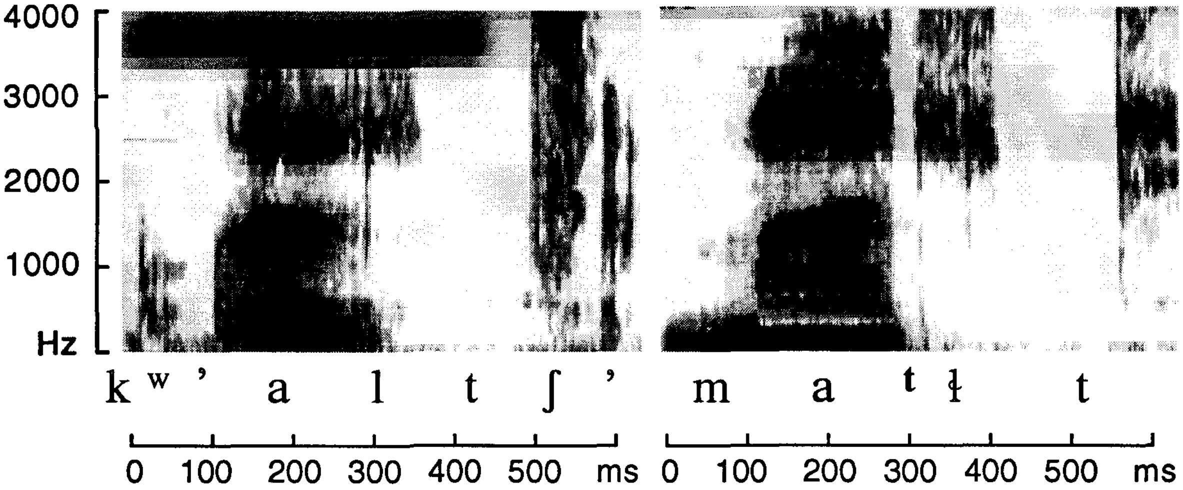


Figure 13. A comparison of voiced **l** becoming becoming partially voiceless before a voiceless consonant in the middle of the word **k^{w}'áltʃ'qən** 'lid, cover', and voiceless **ɬ** also before a voiceless consonant in the word **máɬt** 'mud'.

Steven Egesdal (p.c.) has suggested that vowels are longer preceding the voiced lateral than preceding the voiceless lateral fricative, and that this difference persists before devoiced laterals. A comparison of identical vowels preceding laterals followed by identical voiceless consonants failed to reveal any effect of this kind. However, it was not possible to find word pairs which were matched for number of syllables, so shortening due to word length could have obscured any effect due to the laterals. Given that the voiced lateral is sometimes only partially devoiced, it seems likely that the two laterals are produced differently even in devoicing environments, but the acoustic differences are often subtle.

There is clear evidence that devoiced laterals and voiceless lateral fricatives are phonologically distinct. As discussed above, sonorants are always separated from a preceding consonant by a vowel, while obstruents can form clusters. This rule treats devoiced laterals as sonorants: voiceless lateral fricatives can occur in clusters, whereas devoiced laterals cannot.

Table 6. Contrasts between devoiced laterals and voiceless lateral fricatives

UNDERLYING SONORANT **l**			UNDERLYING FRICATIVE **ɬ**		
q^{w}'ájəlqs	[q^{w}'ajə^{t}ɬqs]	'priest'	**tʃájɬqn**	[tʃajɬqən]	'cut hair'
tsáqwəlʃ	[tsaq^{w}ə^{t}ɬʃ]	'western larch'	**q'áq'ɬu**	[q'aq'ɬu]	'vein'
ɬáqʃəlʃ	[^{t}ɬaqʃə^{t}ɬʃ]	'sit down!'	**tʃ'átənɬq**	[tʃ'atənɬq]	'horsefly'

Glottalized Sonorants

Montana Salish contrasts glottalized and non-glottalized variants of the sonorant consonants **l, m, n, j, w, ʕ**. As discussed above, the glottalized sonorants are also treated as sonorants by the phonology. We will exemplify the general pattern of realization of glottalized sonorants with reference to the nasals, then turn to the specifics of the glottalized lateral and the glottalized glides. Glottalized pharyngeals will be discussed below, with plain pharyngeals.

In almost all positions, glottalized sonorants are typically realized as a glottal constriction followed by the sonorant, i.e. they are preglottalized. As was noted above (see Figure 7), the voiced portion of the glottalized nasal is substantially shorter than that of the plain nasal. The glottal constriction is clearly apparent in an intervocalic environment, as in the word **səˀmú** 'mare' in Figure 14. As the two utterances represented in figure 22 indicate, there is some variation in the degree of glottal constrcition. Often, as in the utterance shown on the left og the figure, there is complete closure of the glottis (the low frequencies evident on the spectrogram are due to the background noise). In other cases, as exemplified by the dpectrogram on the right, the glottal constriction results in a creaky voiced portion of the nasal, without complete closure of the glottis. Glottalized nasals are preglottalized even in final position or before a voiceless consonant, as in **sts'óˀm** 'bone'. In these cases, the nasal portion is devoiced or creaky.

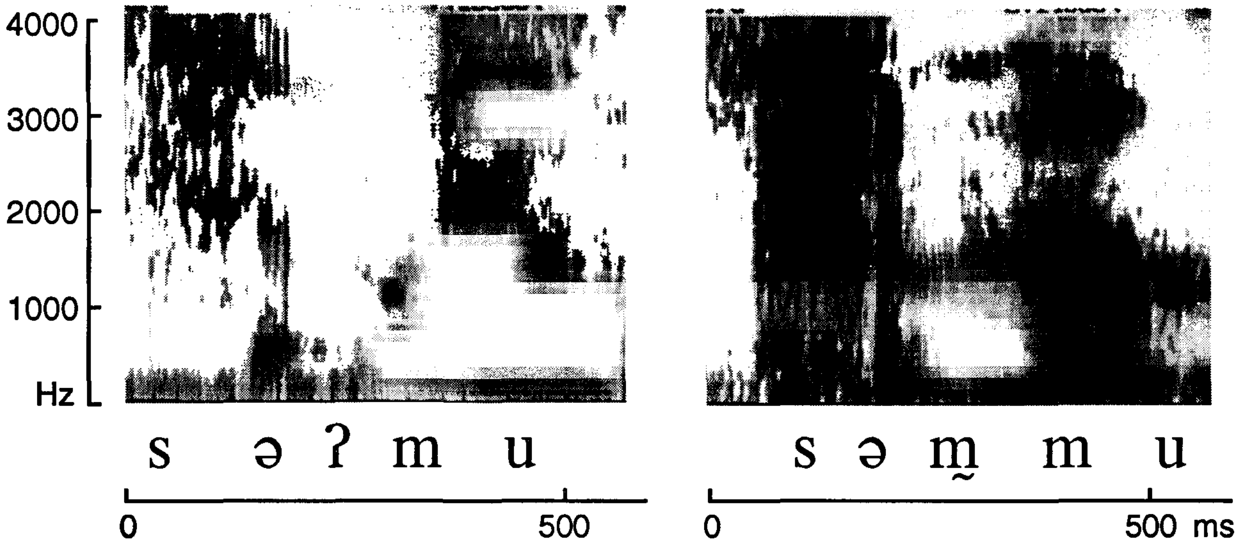


Figure 14. Spectrograms of **səˀmú** 'mare' as pronounced by speaker DF on the left and FM on the right.

The patterns of realization outlined are adhered to consistently except in clusters of glottalized sonorants. It seems that in these cases not all of the sonorants are realized with glottalization, but speakers vary as to which sonorants they glottalize. For example pronunciations of the word for 'soft' include [ɬəmˀmˀmóts], [ɬəm ˀmːóts] and [ɬəˀmːˀmóts], suggesting an underlying representation with three glottalized nasals, **ɬəˀmˀmˀmóts**, not all of which are realized. Thompson and Thompson (1992) note a similar phenomenon in Thompson Salish (p. 45), but state that the initial glottalized sonornat always retains glottalization, which is not the case here.

As described above, plain laterals are typically prestopped. Glottalized laterals are also prestopped. The initial closure of the stop phase is glottal, sometimes, as in Figure 10, extending back to produce creaky voice in the previous vowel. This glottal closure is often not present at the end of the stop, so the release into the lateral is usually coronal. In some cases the glottal constriction results in creaky voicing in the lateral portion. There are no examples of the glottalized lateral in initial position in our data, but figure 10 shows it in final position. In final position and before voiceless consonants, the lateral is devoiced, as with non-glottalized laterals. Deglottalization may occur in clusters of glottalized laterals in the same way described for glottalized nasals.

Glottalized glides follow a similar pattern to the other glottalized sonorants. They are preglottalized in initial, intervocalic and final positions. The degree of glottal constriction varies, as may be seen in Figure 15, which shows the middle section of **ʔesuˀwétʃiʔ** ‘lightning’ by speaker AI on the left and CW on the right. The first speaker has a full glottal stop, whereas the second has creaky voice in the early portions of, or preceding, the glide. In final position, the glide portion is realized as a voiceless, or very creaky, version of the equivalent vowel. While the glottal constriction always occurs mainly during the transition into the glide, some glide transitions may be observed before glottal closure. This phenomenon is particularly marked in final and pre-consonantal glides. As in final position, glottalized sonorants generally take the form **ʔV̥** pre-consonantally, where **V̥** is the vocalic equivalent of the glide. However, **ˀw** is sometimes realized as **ʔw̥ə̥** in these cases. In one word, **ʔajptsín** ‘he talked fast’, the glottalized glide is realized consistently as post-glottalized: [ʔajʔptsín]. It is not clear what conditions this particular realization. Possibly it is the preceding glottal stop.

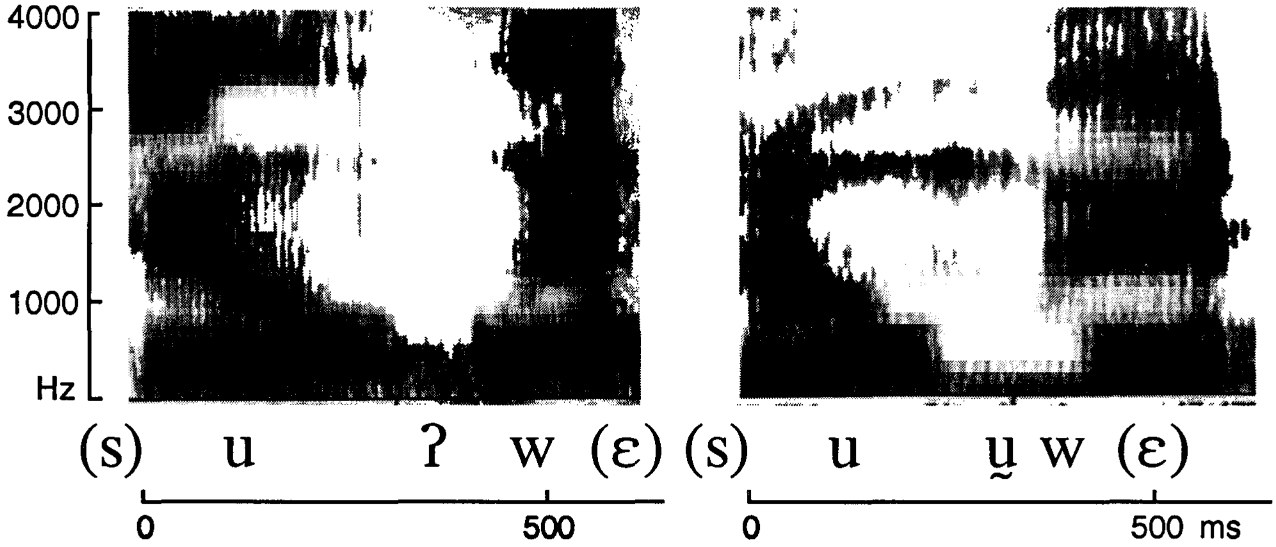


Figure 15. The middle section of **ʔesuˀwétʃiʔ** ‘lightning’ as spoken by speaker AI on the left and CW on the right.

Ejectives

We have aerodynamic records of the three female Montana Salish speakers, as illustrated in figure 16. In making these records, one of the test words was said with the word **tsu** ‘he said...’ before it. The top line is an electroglottographic (EGG) record from the larynx. This type of record cannot be quantified (except in the time domain), but it provides a good indication of the degree of closure of the vocal cords. The middle line is the oral pressure as recorded by a small tube inserted between the lips, with its open end behind the alveolar ridge, and the third line is a record of the oral air flow. The arrows at the top show the moment of release of the bilabial closure. In the case of the ejective on the right of the figure, it may be seen that there is considerable laryngeal activity both at that time and slightly before it. This activity is followed by an interval of about 100 ms before vocal cord vibrations begin. There is far less activity for the plosive on the left of the picture, and vocal cord vibrations begin almost immediately. There is considerably greater oral pressure for the ejective in the second word than for the plosive in the analogous position in the first word. The plosive at the end of the second word also has less oral pressure. For all three speakers, the peak pressure in stem initial ejectives was always greater than in comparable plosives. The mean for 5 utterance (3 from one speaker, and one each from the other two) was 9.0 cm H_2O for the ejectives and 6.1 cm H_2O for the plosives. After the release of the plosive the oral flow rises to above 500 ml/s, whereas in the ejectives there is a comparatively small burst of oral flow, followed by a period in which there is no flow while the glottal closure is maintained. In this sound the vowel begins abruptly as the glottal closure is released.

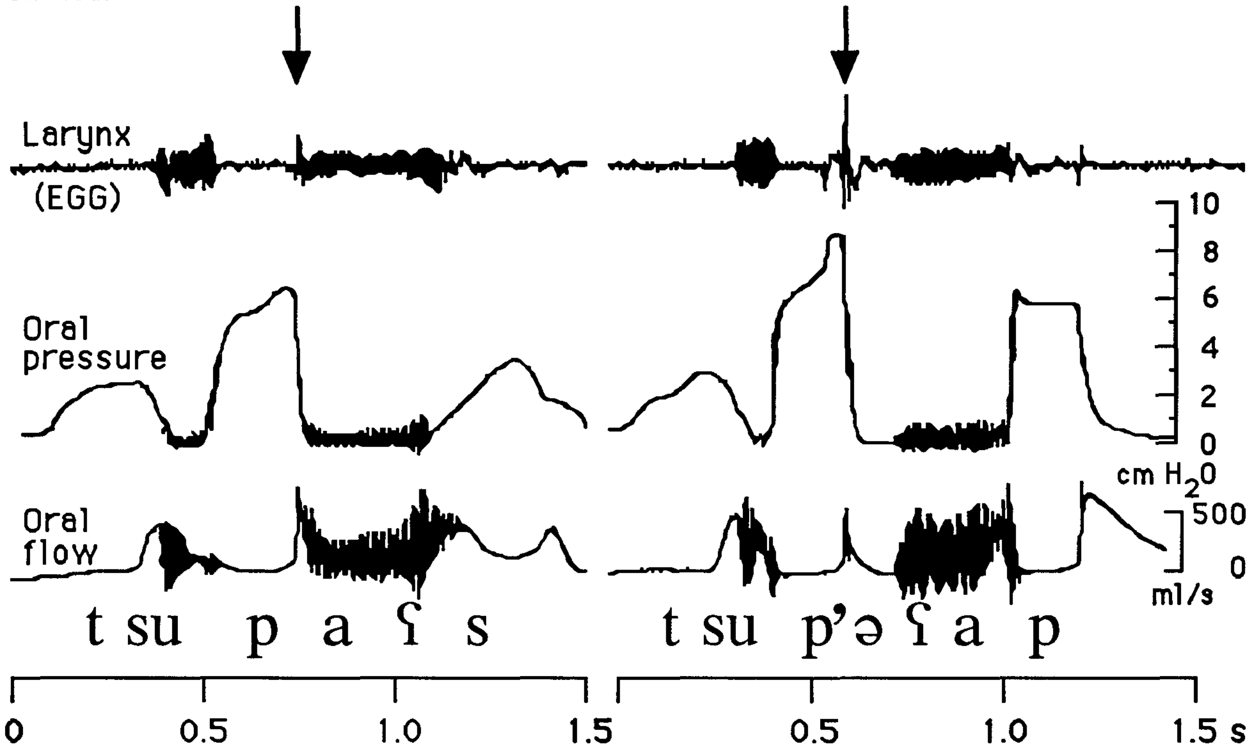


Figure 16. Aerodynamic records of contrasting plosive and ejective in Montana Salish **tsu páʕs** ‘pale face’ vs. **tsu p’əʕáp** ‘the grass/timber caught fire’. (See text for details.)

Figure 17 shows the aerodynamic and laryngeal activity that occurred during the pronunciation of the Montana Salish phrase **tʃ’tʃéˀn** ‘Where to’, enabling us to compare a sequence of two affricates, one with a pulmonic, and the other with a glottalic airstream. We will omit consideration of the top line for the moment. The second line, the larynx record, reflects the laryngeal movements associated with the ejective. It cannot be taken as a direct indication of larynx raising and lowering, both because the gross movements of the larynx do not affect glottal impedance in a way that is directly proportional to larynx movement, and because this record has been band-pass filtered (30-5,000 Hz). If it had not been filtered, the small changes due to the opening and closing of the glottis (which were the major focus of the investigation) would have appeared insignificant in comparison with the large changes associated with the movements of the larynx. Nevertheless, the record clearly shows that there is greater laryngeal activity during the release of the ejective than there is for other sounds. Both the larynx record and the nasal flow record show that in this token there is an epenthetic (non-contrastive) nasalized vowel after the initial affricate, which did not appear in other tokens of this word. The aerodynamic records show that the oral pressure in the ejective was lower than that in the following pulmonic affricate, which is somewhat surprising. The comparatively slow decreases in the pressure and the corresponding increases in the oral flow are typical of affricates.

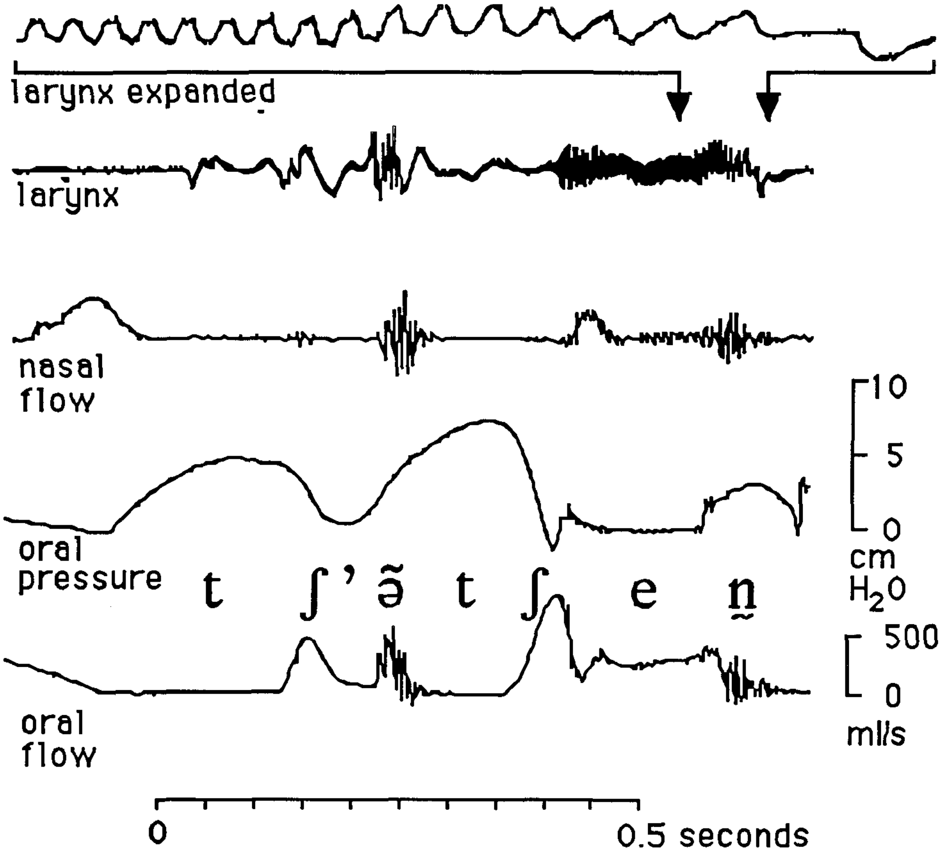


Figure 17. Aerodynamic records of a plosive and an ejective in the Montana Salish word **tʃ'tʃéˀn** 'Where to'.

Pharyngeals

Montana Salish has a voiced pharyngeal approximant which can appear with secondary rounding and/or glottalization. Rounded pharyngeals only appear before round vowels or word-finally, and unrounded pharyngeals only appear adjacent to the low, unrounded vowel /a/. Rounded pharyngeals generally seem to occur before **o**, but there are words in which a rounded glottalized pharyngeal occurs before **u**, e.g. **háʕ̓ʷumskʷ** 'loosen it!' and **səʕ̓ʷúʔ** 'it got low (as of water)'. The pharyngeals are often extremely vowel-like and have often been transcribed as the low back vowels **ɑ** for **ʕ** and **ɔ** for **ʕʷ**. Bessell (1992:92ff.) observes that this is a property of pharyngeals in a number of closely related Interior Salishan languages including Kalispel, Spokane and Colville-Okanagan.

Several features distinguish pharyngeals from vowels, but not all of them are present to any significant degree in any particular utterance. We will exemplify each of these features from clear examples of pharyngeals in the following figures. The first is that pharyngeals are usually characterized by a marked lowering of the fundamental frequency. This effect has been observed with pharyngeals in other languages also (c.f. Maddieson, Spajić, Sands and Ladefoged 1993, on Dahalo, a Cushitic language). In the case of pharyngeals adjacent to an accented vowel, the lowering associated with the pharyngeal results in a dramatic pitch movement to the high tone of the accented vowel, as may be seen in the narrow band spectrogram on the right in Figure 18. (This spectrogram shows, on an expanded frequency scale, the fundamental frequency and the first seven or eight harmonics.)

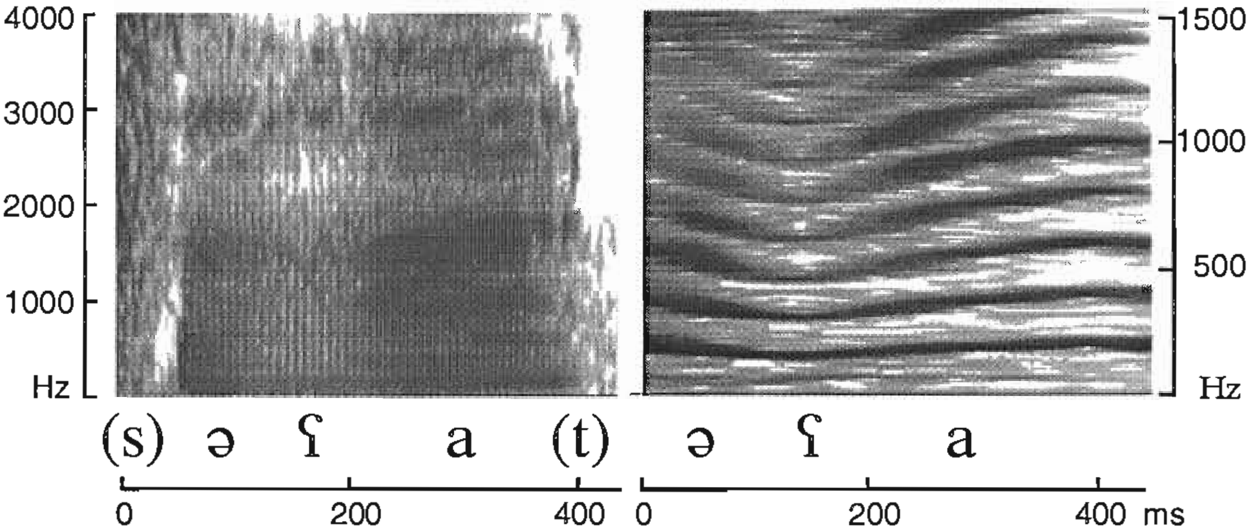


Figure 18. Wide and narrow band spectrograms of the pharyngeal consonant in the middle of the word **ʔesəʕáts** ‘it’s tied, staked’. Note that the narrow band spectrogram on the right has an expanded frequency scale, and the amplitude adjusted so that the harmonics are more clearly visible.

Pharyngeals also generally have lesser intensity than the low vowels which they otherwise resemble, as exemplified in Figure 19, wide and narrow band spectrograms of the word **sχəʕáp** ‘air’ as spoken by speaker HW. This word was spoken on a rising pitch, and there is only a very small drop, or decline in the rate of pitch increase associated with the pharyngeal. In this word, the decrease in amplitude is a much greater mark of the pharyngeal consonant. The drop in intensity may sometimes (as in this figure) be more noticeable at higher frequencies and is thus probably due to a breathy laryngeal setting, which results in a steeper spectral tilt (Ladefoged, Maddieson and Jackson 1988). The breathiness may also be indicated by noise in the spectrum, along with other markers of a change in the mode of vibration of the glottis, such as the sub-harmonic below the fundamental frequency in Figure 19.

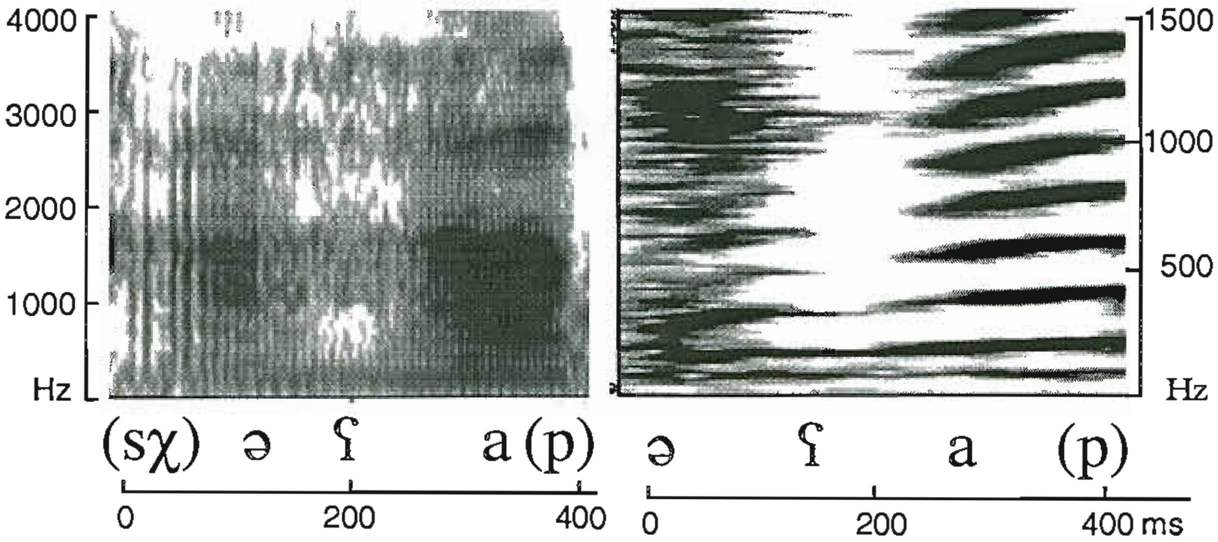


Figure 19. Wide and narrow band spectrograms of the pharyngeal consonant in the middle of the word **sχəʕáp** 'air' as spoken by speaker HW.

Finally, pharyngeal constriction has an effect on formants, raising F1 and lowering F2 (Alwan 1986, Bessell 1992). Montana Salish pharyngeals involve very open approximation, especially in intervocalic position, so the shift in formant frequencies between the pharyngeal and an adjacent low vowel is often slight. It can be seen in Figure 18, but there is very little movement in Figure 19, and virtually no observable movement in Figure 20. A more noticeable case of formant movement is illustrated in Figure 21, a spectrogram of part of the word **ʔinəʕʷól̓** 'empty' in this case the lowering of F2 after the epenthetic **ə** is particularly noticeable, and the high F1 before the **o** is also evident.

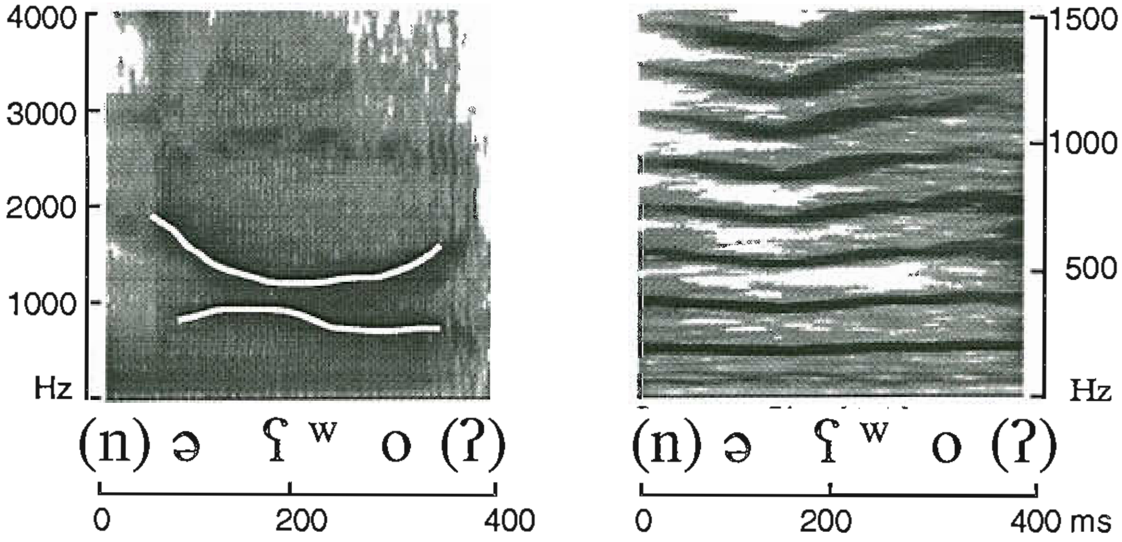


Figure 21. Wide and narrow band spectrograms of the pharyngeal consonant in the middle of the word **ʔinəʕʷól̓** 'empty' .

Even when there are no clear formant movements, the first two formants of both the pharyngeal and the vowel are closer than in the related vowel in other environments. Table 6 shows averages for F1 and F2 in **ʕʷó** and **ʕá** compared to **ó** and **á** in other contexts. The effects are more striking with **o**, because this is otherwise a higher vowel than **a**. There is no clear effect of the pharyngeal on **a** for the males, but there are only two tokens in this case. However, in general it is clear that the presence of a pharyngeal can cause lowering and backing of an adjacent vowel, even in the absence of significant formant movements localized on the pharyngeal itself.

Table 6: mean F1 and F2 in pharyngeal-vowel sequences compared to vowels in other environments.

		F1 (Hz)	F2 (Hz)
females	ʕʷó	832	1016
	ó	601	1170
males	ʕʷó	566	931
	ó	510	994

	F1 (Hz)	F2 (Hz)
ʕá	940	1563
á	854	1603
ʕá	631	1264
á	683	1389

In some utterances none of the characteristics of pharyngeals described above can be identified. The phonemic pharyngeal-vowel sequence then resembles a long vowel, as illustrated in Figure 20, a spectrogram of the same word as in Figure 19, **sχəʕáp** 'air', but spoken by a different speaker. This was produced by a male speaker, AI, so a larger number of harmonics are shown in the narrow band spectrogram on the right. But even considering the absolute difference in fundamental frequency between this male speaker and the female speaker in the previous figure, it is clear that pitch plays a smaller role in distinguishing the pharyngeal in this case. In the utterance shown in Figure 20 the pharyngeal is marked largely by the length of the vowel, with only a small drop in intensity and frequency in the middle. Much of the increased length can be associated with the epenthetic vowel that appears between the voiceless fricative and the pharyngeal. Because of the constant formant frequencies, in the figure legend we have transcribed this epenthetic vowel as **a** rather than **ə**.

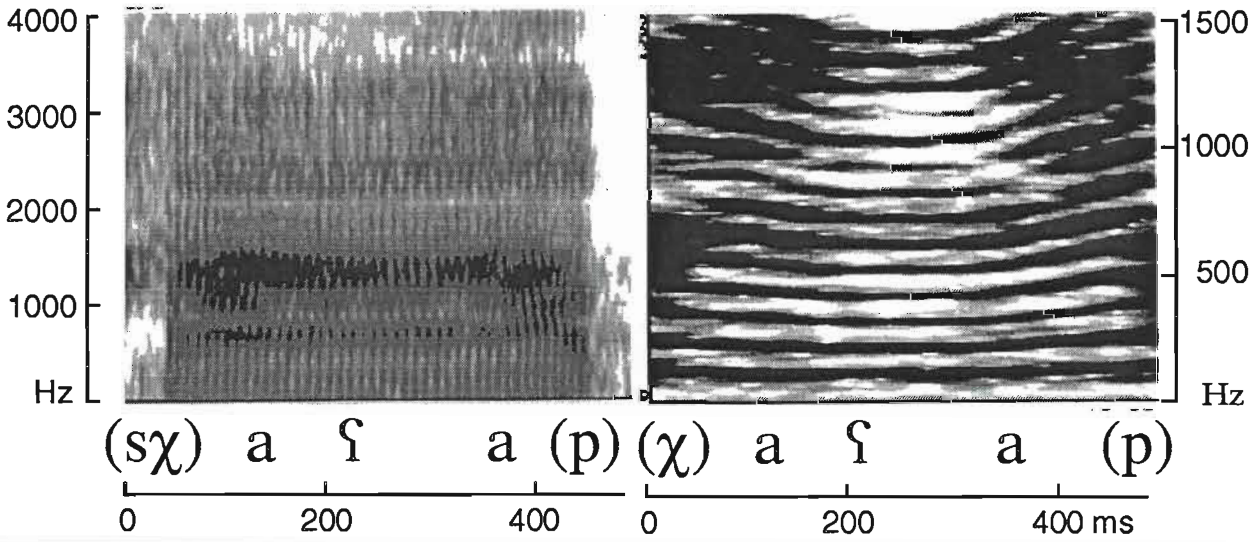


Figure 20. Wide and narrow band spectrograms of the pharyngeal consonant in the middle of the word **sχəʕáp** 'air', as spoken by speaker AI.

In summary, Montana Salish pharyngeals seem to involve three basic elements: a pharyngeal constriction, lowered fundamental frequency and a change in voice quality. However, as noted, not all of these properties are observable in all utterances. Furthermore, a final example suggests some independence between the features associated with pharyngeals. The utterance of **ʔesəʕáts** 'it's tied, staked' shown in Figure 22 contains evidence of F0 lowering, and of F1 raising and F2 lowering, but the fundamental frequency has already been rising for some time when F1 and F2 achieve their closest proximity. This suggests that the pharyngeal constriction and fundamental frequency lowering associated with pharyngeals in Montana Salish are not physiologically connected, but rather are relatively independent correlates of the same phoneme.

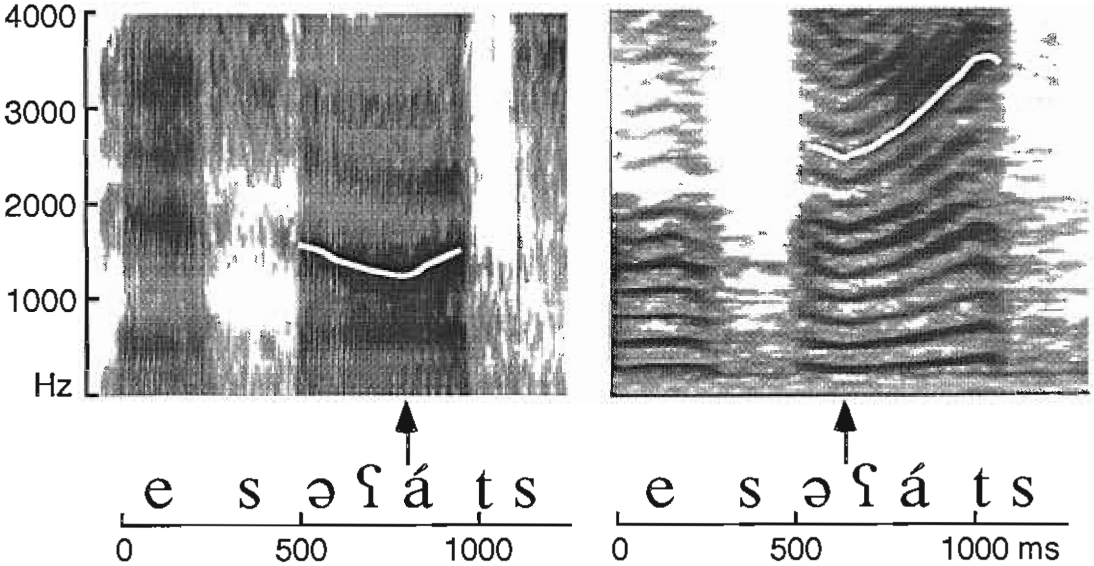


Figure 22. Wide and narrow band spectrograms of the word **ʔesəʕáts** 'it's tied, staked'. Maximum F0 lowering occurs at the time of the arrow below the narrow band spectrogram. Maximum F1 raising and F2 lowering occur at the time of the arrow below the wide band spectrogram.

As Bessell (1992) points out, there is also phonological evidence that, in spite of their vowel-like properties, the pharyngeals behave as sonorant consonants rather than vowels. As discussed above, under most circumstances a schwa vowel is inserted between a sonorant and a preceding consonant. Pharyngeals are treated as sonorants by this rule, so pharyngeals never follow consonants. The preceding schwa can be clearly observed in Figures 21. As we have seen, the realization of **ə** is highly dependent on the adjacent segments, so in some cases preceding a pharyngeal it is close to **a** or **o**, as we transcribed it on the spectrogram in Figure 20.

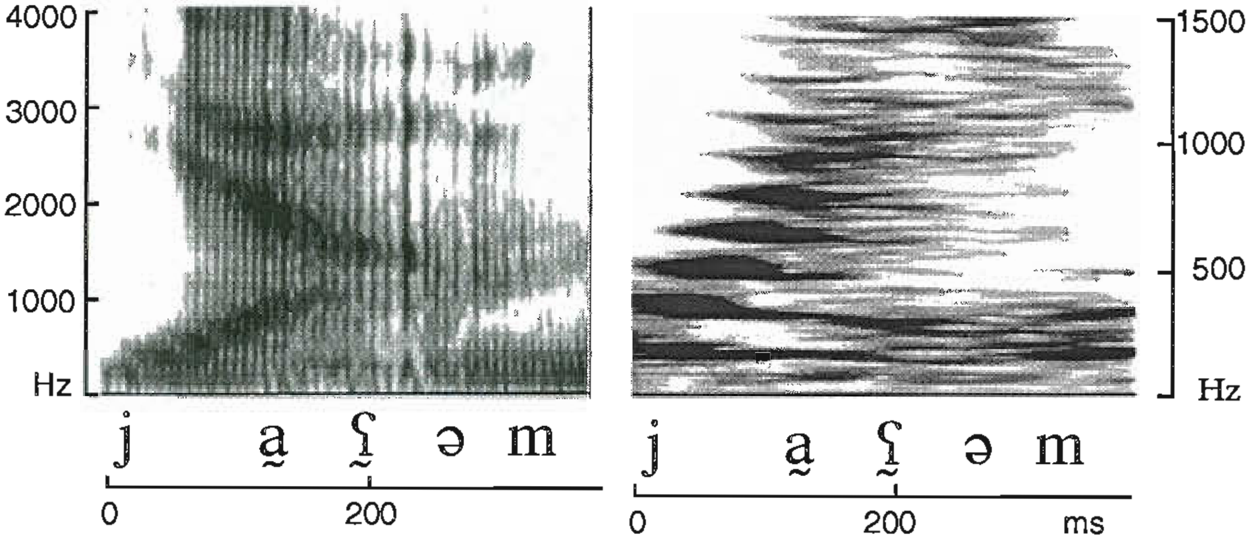


Figure 23. Wide and narrow band spectrograms of the pharyngeal consonant in the middle of the word **jaˀʕəmím** 'gathering (as, rocks)'.

Glottalized pharyngeals show some differences from other pharyngeals. Word-medially, as illustrated in Figure 23, they are generally realized as creaky-voiced pharyngeals , while word-initially they are realized as a glottal stop followed by a pharyngeal. Note that the creaky voicing often makes medial glottalized pharyngeals very hard to identify because they obviously do not share the breathiness characteristic of non-glottalized pharyngeals, and the creakiness can obscure the fundamental frequency of the pharyngeal. **ˀʕʷ** particularly can be hard to distinguish from glottalized **ˀw**. However, cues from formant structure are available, and where it can be determined, F0 is usually low during glottalized pharyngeals. Both these indications are present in Figure 23, where F1 and F2 during the pharyngeal are 990 Hz and 1490 Hz respectively. Note that while the word-initial glottalized pharyngeals are consistent with the general pattern of glottalized sonorants in Montana Salish in being pre-glottalized, the glottalization in medial pharyngeals is generally present throughout. Figure 24 shows a rather different realization of a medial rounded glottalized pharyngeal in the word səˀʕʷuʔ 'it got low (e.g. water)'. In this case, after the epenthetic **ə**, there is a full glottal stop. There is then a further epenthetic vowel, this time more like **u**, between the release of the glottal constriction and the onset of a pharyngeal constriction, which is indicated by the sharp drop in the amplitude of F2, and a slight rise in F1.

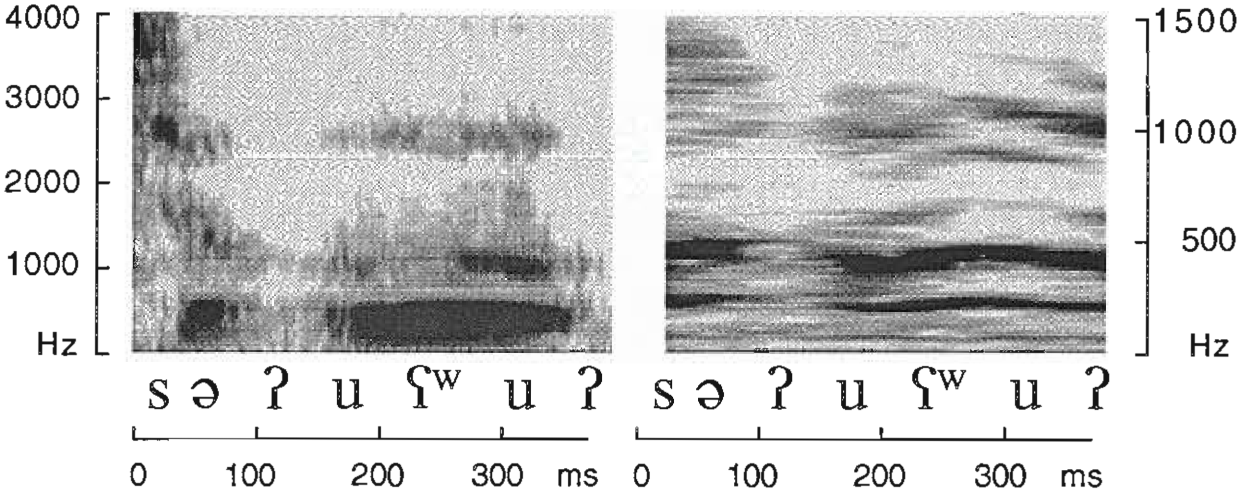


Figure 24. Wide and narrow band spectrograms of the pharyngeal consonant in the middle of the word səʕ̓ʷuʔ'it got low (e.g. water)'.

Coda

This description of Montana Salish has outlined the main phonetic characteristics of the language. But there are many points that are still not clear. For example, we do not know to what extent the variation in the different phones that we have noted are tied to particular individuals. To determine this we need to make further recordings and observe how consistent these speakers are. There are also some sounds for which we cannot relate the acoustic structures we observe with the articulatory gestures that probably produced them. This is particularly so in the case of the transients that appear in some laterals. We hope that future work will clarify these points. Meanwhile, the data presented here, and the recordings which are available to all interested researchers, remain as a first account of Montana Salish phonetics.

Acknowledgements

This paper is for the Salish people. We are especially grateful for the assistance of the Director of the Flathead Culture Committee, Clarence Woodcock, and the Associate Director, Antoine (Tony) Incashola. We are deeply indebted to our three principal consultants: Harriet Whitworth, Felicité ("Jim") McDonald and Dorothy Felsman. Thanks are also due to Joyce McDonough and Steve Egesdal. This research was supported in part by NSF.

References

Alwan, Abeer (1986) *Acoustic and Perceptual Correlates of Uvular and Pharyngeal Consonants*, MA dissertation, MIT.

Bates, Dawn, and Barry F. Carlson. 1992. Simple syllables in Spokane Salish. *Linguistic Inquiry.* 23:653-59.

Bessell, Nicola J. (1992) *Towards a Phonetic and Phonological Typology of Post-Velar Articulation*, Ph.D. Dissertation, University of British Columbia.

Carlson, Barry F. 1972. A grammar of Spokane, a Salish language of eastern Washington. Honolulu: University of Hawaii dissertation. (Univ. of Hawaii Working Papers in Linguistics 4.4.)

Kuipers, Albert H. (1974) *The Shuswap Language*, Mouton, the Hague.

Ladefoged, Peter and Ian Maddieson (1986) *Some of The Sounds of the World's Languages, UCLA Working Papers in Phonetics* 64.

Ladefoged, Peter; Ian Maddieson and Michel Jackson (1988) "Investigating Phonation Types in Different Languages" in Osamu Fujimura (ed) *Vocal Physiology: Voice Production, Mechanisms and Functions*, Raven Press, New York.

Lindau, Mona (1984) "Phonetic differences in glottalic consonants", *Journal of Phonetics* 12, 147-55.

Lisker, Leigh and Arthur Abramson (1964) "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements", *Word* 20, 384-422.

Maddieson, Ian and Karen Emmorey (1984) "Is there a valid distinction between voiceless lateral approximants and fricatives?", *UCLA Working Papers in Phonetics* 59.

Maddieson, Ian; Siniša Spajić, Bonny Sands and Peter Ladefoged (1993) "The Phonetic Structures of Dahalo", *UCLA Working Papers in Phonetics* 84, pp. 25-66.

McDonough, Joyce and Peter Ladefoged (1993) "Navajo stops", *UCLA Working Papers in Phonetics* 84, pp. 151-64.

O'Shaughnessy, D. (1981) "A study of French vowel and consonant durations", *Journal of Phonetics* 9, pp. 385-406.

Orser, Brenda, and Barry F. Carlson. 1993. Sonority in Spokane, Interior Salish pharyngeal resonants. Paper presented at the 1993 Linguistic Society of America Annual Meeting, Los Angeles.

Thompson, Laurence C. "Salishan and the Northwest," *The languages of Native America: A historical and comparative assessment,* ed. by Lyle Campbell and Marianne Mithun, 692-765. Austin: University of Texas Press.

Thompson, Laurence C. and M. Terry Thompson (1992) *The Thompson Language, University of Montana Occasional Papers in Linguistics* 8.

Vogt, Hans. 1940.*The Kalispel language*. Oslo: Det Norsk Videnskaps-Akademi.

Appendix 1: Montana Salish word list

Forms are given in the surface phonemic transcription system used above. The headings in the list have been retained from our original fieldwork notes, so that others can appreciated why particular words are included.

FRAME SENTENCE: **tsu ___ tspists'éʔ** 'He said ___ yesterday.'
(used in one recording session)

Vowels

1	**páʕəlqs**	grey dress
2	**pəlpîllʃ**	stagger
3	**tʃ'upélsì**	lonesome
4	**pólpəlqən**	thimbleberry
5	**púlsəm**	he killed something
6	**jál**	round
7	**séliʃ**	Flathead, Salish
8	**mîlkʷ'**	always
9	**(ʔi)nəʕʷóˀl(kʷ)**	empty
10	**múlʃ**	cottonwood
11	**ʔáw**	ouch!
12	**ʔəlʔéw**	father of a male
13	**ʔesəlí wti**	chapped
14	**ʔestóqʷ**	ravine, draw
15	**tʃ'úw**	it's gone
16	**ˀlˀláts**	red raspberry
17	**qétstʃ**	older brother (of a woman), aunt's or uncle's elder son
18	**tʃitsəntəm**	he met him
19	**łótst**	smashed
20	**ʔn̩łútskʷ**	soak

Consonants

21	**páʕs**	face is pale, grey
22	**p'əʕáp**	grass fire, timber fir
23	**p'úm**	brown, orange
24	**tám**	it's not, wrong
25	**t'áq'ən**	**six**
26	**t'áqʷən**	I licked it
27	**qáχeʔ**	aunt (mother's sister)
28	**qétstʃ**	older brother (of a woman), aunt's or uncle's elder son
29	**q'áq'łuʔ**	vein
30	**q'áˀwχeʔ**	yellow bell (flower)
31	**q'étt**	a hide, pelt
32	**q'éts'**	a sack; weave
33	**kapî**	coffee
34	**kʷáteʔ**	quarter (money)
35	**kʷén**	I took it

36	**kʷ’áltʃ’qən**	lid, cover
37	**kʷ’éˀnəm**	attempt, try
38	**qʷátsqən**	hat
39	**qʷásqʷiˀj**	blue jay
40	**qʷ’ájəlqs**	priest (black-robe)
41	**qʷ’át’χʷ**	bent; banana, cucumber
42	**qʷ’éts’t**	full
43	**tsáqʷəlʃ**	western larch (Larix occidentalis Nutt.)
44	**ts’áɬt**	it's cold
45	**ts’áχ**	fry; it's fried
46	**tʃájɬqən**	cut hair
47	**tʃ’áwən**	I prayed (it)
48	**tʃ’átənɬq**	horsefly
49	**sáχʷ**	split wood
50	**ʃáɬɬ**	he got bored
51	**χáˀm**	dry
52	**χáq’**	pay
53	**xʷáltʃst**	reach (for something)
54	**χʷáq’ʷ**	to grind or file something
55	**máɬt**	mud
56	**máχeʔ**	glacier lily (Erythronium montanum)
57	**ˀmeˀmstsú**	playing cards
58	**ˀməˀnétʃ**	excrement, shit
59	**səˀnəˀmˀné**	toilet
60	**səˀmú**	mare
61	**sts’óˀm**	bone
62	**nás**	wet
63	**náqʷ’**	steal
64	**ˀneˀjxʷéˀws**	trade
65	**ˀneʔótsqe(ʔ)**	when he goes out ...
66	**tʃ’tʃeˀn**	where to?
67	**láq’əm**	he buried
68	**láq’i**	sweatbath
69	**ʔəlˀláts**	red raspberry
70	**ppíˀl**	pint
71	**səˀláχt**	friend
72	**tʃtʃts’éˀlʃtʃən**	wood tick
73	**ɬáq’t**	wide, shovel
74	**ɬáqʃəlʃ**	sit down!
75	**tɬ’áqəneʔ**	pocket
76	**tɬ’áq’**	hot
77	**tɬ’áqʷ’**	beat one's wife
78	**waˀləwəˀl**	long-billed curlew
79	**wénʃ**	dance the war dance
80	**ˀwíˀwá**	wild
81	**ʔesuˀwétʃiʔ**	lightning
82	**ʔiˀwéstən**	death camas
83	**ˀweˀwí**	meadowlark
84	**játəmskʷ**	shake it!

85	**jajáʔ**	maternal grandmother
86	**ˀjeˀjúkʷeʔ**	stingy
87	**ʃeˀj**	that, this

Pharyngeals

88	**ʔáwəntxʷ**	you said it
89	**ʕámt**	it's melted
90	**ʕájmt**	he's angry, mad
91	**ʔesəʕáts**	it's tied, staked
92	**stəʕán**	antelope
93	**sp'əʕás, sp'ás**	nighthawk
94	**səʕáptəni**	Nez Perce
95	**stʃsəʕó**	sunset
96	**sχəʕáp**	air
97	**p'əʕáp**	the grass/timber caught fire
98	**ʔipáʕ**	it's pale
99	**páʕs**	pale face
100	**χʷaχʷáʕ**	fox
101	**qʷ'á(ʕ)ntsú**	to coast, slide
102	**ʔesəjaʕsqé**	shy, reserved
103	**jaˀʕəmîm**	gathering (as, rocks)
104	**ʕʷóst**	lost
105	**ʕʷóˀl**	slippery (oily)
106	**ʕʷóˀllexʷ**	slippery ground
107	**ˀʕʷoˀjintsú**	laugh
108	**ʔiháˀʕʷ**	loose
109	**háˀʕʷumskʷ**	loosen it!
110	**máˀʕʷt**	broken
111	**tsáwəlʃ**	bathe, swim
112	**səˀʕʷúʔ**	it got low (as, water)
113	**ʔnsəˀʕʷóp**	drink up, get dry

Other interesting clusters

114	**p'əˀlp'əˀltʃ'əmím**	she's turning it (the meat) over
115	**p'əllítʃ'tʃ**	turned over
116	**χəˀlχəˀlléxʷ**	teeth
117	**ɬəˀmˀmˀmóts**	soft
118	**sqʷəllú**	tale
119	**səˀməˀlˀlá**	a nosebleed
120	**sɬesəˀlspé**	a little two-year-old
121	**tʃɬkʷkʷtəˀnéˀws**	a fat little belly
122	**kʷ'uˀlˀĺi**	birth, born
123	**ttəˀwít**	youngest, young (boy)
124	**tʃtkʷtɬˀní**	they were walking along a ridge
125	**tq'əˀntʃstá**	six days
126	**qʷéjˀmˀm**	he was in a hurry

"Double vowels"

127	**ləʕáp**	sail; he sailed
128	**ʕátat**	hawk
129	**ʕáj**	bull trout
130	**sts'əʕál, sts'əʔál**	sickness
131	**ʔṇpapáʕ**	cliff, bank
132	**stiʕit.ʃən**	killdeer
133	**sxʷuʔúl**	steam

More ˀm

134	**kʷéˀmt**	and then
135	**ststsəˀmˀmélt**	children
136	**tʃ'ɬtstsəˀmú**	bead, beads
137	**sts'óˀmts'əˀm**	boil (inflamed swelling with pus)
138	**ʔistʃéˀmm**	I'm rejecting it
139	**ʔn̤tʃəˀméɬxʷ**	rob
140	**tʃ'əˀmpstém**	faint
141	**kʷ tʃ'íˀm**	you're crazy, something's dark with your brain!
142	**ʔeˀmt**	feed
143	**heˀmîsheˀm**	mourning dove
144	**hetíˀm**	to tease
145	**ʔiˀmʃ**	he moved (camp)
146	**kʷkʷúsəˀm**	star
147	**skʷkʷ'íˀməlt**	young child
148	**ʔṇɬáɬəˀmqe**	a little black bear
149	**ɬəmɬaˀmá**	frog
150	**ɬəˀmáq**	got a burn
151	**ɬúˀmən**	spoon
152	**estɬ'eˀmstsú**	trouble
153	**ˀməˀlˀməˀlté**	quaking aspen
154	**méstəˀm**	father of a woman
155	**səˀmiˀmîʔ**	news
156	**ˀmiˀméˀjeʔ**	teach
157	**sɬəˀmχé**	a little grizzly bear
158	**sqt'íˀm**	scar
159	**ʔṇqʷqʷ'osə́ˀmî**	dog
160	**súˀmənt**	to smell something
161	**ɬʃəˀméˀn**	a little enemy, a young enemy
162	**sʃutəˀméˀlt**	youngest daughter
163	**téˀm**	what?
164	**itəˀnəˀmú**	nothing, it's nothing
165	**túˀm**	woman's mother
166	**t'əˀmám**	licked (ice-cream cone, sucker)
167	**st'əˀmá**	cow
168	**ɬuˀluˀlíˀm**	wire
169	**ʔi χáˀm**	he's dry
170	**ʔesχáˀmpəmî**	it's getting dry

171	**ˀjaˀjaˀmˀmîˀn**	farm hay rake

More ˀn

172	**ʔaχéˀneʔ**	handbag
173	**ts’sqáˀniʔ, ts’sqáˀneʔ**	chickadee
174	**ʔiɬtʃtʃéˀn**	tender (as, a sore spot)
175	**tʃəˀnétʃs**	handshake
176	**stʃ’îˀn**	pika
177	**tʃ’itʃ’itəˀné**	alder tree
178	**tʃ’əˀnéjən**	snare
179	**ʔeˀnés**	he went, he left
180	**heʔéˀnəm**	eight
181	**skʷəˀnkʷî**	Indian potato
182	**kʷ’îˀnʃ**	how many?
183	**sméˀnxʷ**	tobacco
184	**mîˀnəm**	he painted
185	**tʃ’əɬəˀná**	one person, all alone
186	**ɬəˀnî**	knife
187	**səˀnˀnî**	sheath (for a knife)
188	**ʔespíˀn**	bent
189	**pəˀnîntʃ**	liver
190	**ʔn̩qʷəˀnqʷəˀnˀnéls**	kind person
191	**qʷɬəˀná(lqʷ)**	white birch, paper birch
192	**səˀnsá**	tame
193	**sʃéˀnʃ**	rock
194	**tʃʃəˀnʃəˀnú**	eyeglasses
195	**tkʷtîn**	bullrush
196	**ʔeskʷ’ɬtəwîˀn**	short of, not quite enough
197	**ʔn̩t’oqʷtʃəˀné**	clap
198	**χəˀnút**	nine
199	**ɬχʷχʷəˀnî**	hummingbird

More ˀl

200	**tstsîˀlʃ**	overnight camp, overnight stay
201	**stʃtsəˀlé**	mane
202	**səntʃ’əˀlé**	coyote
203	**ts’aʕəˀlétʃst**	his hand is aching
204	**xʷəˀlstéʕ**	why?
205	**tʃtʃəˀléxʷ**	muskrat
206	**ʔiˀláˀwijeʔ**	great-great-grandparent, great-great-grandchild
207	**ɬkʷəˀlkʷaˀlí**	corn
208	**skʷkʷəˀlíˀl**	sunshine
209	**skʷ’îˀl**	porcupine
210	**lemtəwîˀlʃi**	he's getting happy
211	**ɬˀleˀleputé**	harebell
212	**ɬəˀlˀláq**	thin
213	**kʷ’úˀləm**	he made something

214	**ʔmχʷóˀl**	cradle-board
215	**tʃənîll**	infect
216	**ʔolín**	belly
217	**ɬpəˀlpíˀlˀl**	he's staggering a little
218	**p'aq'əˀləwî**	firefly (it flashes)
219	**sp'əˀlq'é**	large intestine
220	**stʃ'ɬp'əˀlp'əˀlkʷ'î**	cigarette
221	**qəˀləwé**	step
222	**qeʔeseˀlíɬəni**	we were eating
223	**tʃsísp'əˀl**	seven people
224	**ssaˀlú**	hail, hailstones
225	**ʃəˀlʃəˀltʃé**	salamander
226	**stəˀméˀlis**	relative
227	**təlíp**	break, shatter; it broke
228	**təˀlxʷé**	difficult
229	**t'éləm**	she sliced something (meat)
230	**ʔn̩t'pséˀlis**	splice
231	**ʔuttʃ'əˀlé**	safety pin
232	**ʔesəwéˀləm**	it's tilted
233	**wəˀləwéˀllʃ**	he's waddling
234	**ɬχəˀlpú**	dawn
235	**sχʷeʔelí**	camas
236	**jaˀlíˀl**	tangled

More ˀw

237	**səˀwéˀwɬ**	fish
238	**sxʷəˀweˀwéʔ**	the crier
239	**ʔn̩ts'áˀwkʷ**	weak, tasteless liquid (e.g. coffee)
240	**ts'éˀwstən**	washcloth
241	**tʃusəˀwî**	Chinese
242	**ʔn̩ʔemtéˀws**	wait
243	**ʔeˀwéti**	sneak up on
244	**héˀwt**	pack rat
245	**ɬuɬəˀwé**	thorn
246	**ʔm̩t tʃəˀwé**	balsamroot
247	**sənéˀwt**	the wind
248	**sqəlé w**	beaver
249	**q'əˀwî**	round stone with handle, for pounding
250	**qʷ'əléˀw**	pick berries
251	**ti(ʔ)itəˀwî**	horsemint

More ˀj

252	**siˀjéˀnp'**	greasy strips of meat
253	**siˀjeʔəˀjéʔ**	juneberry
254	**χʷχʷəˀjú**	whitefish
255	**sχʷt'ɬéˀj**	mountain goat
256	**χʷχʷəˀjé**	wild rose (entire plant)

257	**xʷéˀjt**	he's in trouble
258	**tqʷ'éˀj**	gum
259	**tamtəméˀj, tamtəmîʔ**	rare, not often
260	**ttʃéˀj**	urine
261	**ʃiˀjú**	pass through
262	**isúˀj**	cold, cool
263	**ʔiq'ʷóˀj, ʔiq'ʷóʕʷˀj**	calm (wind)
264	**q'aˀjé**	pinto
265	**q'eˀjmîn**	paper
266	**qéˀjs**	nightmare
267	**p'iˀjáq**	ripe
268	**tɬ'əméˀj**	few
269	**stʃ'éˀjt**	spider
270	**ts'əˀjú**	yet
271	**tsqaˀjé**	she dried (meat)
272	**tspɬéˀj**	brow (forehead)
273	**ʔaˀjptsín, ʕaˀjptsín**	he talked fast
274	**tʃaˀjuté(ʔ)**	they rode

The rhotics of Toda

Siniša Spajić, Peter Ladefoged
Phonetics Laboratory, Linguistics Department, UCLA
and
P. Bhaskararao
Linguistics Department, Deccan College, Pune

1.0 Introduction

Tongue tip trills occur in about one third of the world's languages (Maddieson 1984). None of the languages in Maddieson's sample have two contrasting apical trills without secondary articulations, although they have been reported in Malayalam (Ladefoged 1971). To the best of our knowledge only one language, Toda, has three contrasting trills; and almost certainly no other language has surface contrasts between palatalized and non-palatalized versions of three lingual trills. This paper examines the phonetic properties of these six Toda rhotics.

Toda is a Dravidian language of Southern India. Other languages in this family include Tamil, Malayalam, Kodaga, Kota, Kannada and Tulu (Subrahmanyam 1983). Toda is spoken by about 1000 speakers in the Nilgiri Hills of Southern India. The most complete description of Toda is Emeneau (1984), which is based on a collection of linguistic data made during a three year period from about mid 1935 to mid 1938. Shalev, Ladefoged and Bhaskararao (1993) describe the collection of a further body of data during fieldwork in 1991. They provide a phonetic analysis of the Toda vowels, stops and fricatives in this data. The present paper analyzes the rhotics in the same body of data.

The Toda rhotics should be considered in the context of the complete phonological inventory of the language. Figures 1 and 2, taken from Shalev et al (1993), illustrate the vowel and consonant inventories. Similar charts can be found in Emeneau (1984). The row and column headings in Figure 2 should be taken as only rough labels. A more detailed account of the stops, fricatives and affricates has been provided by Shalev et al (1993).

	BROAD TRANSCRIPTION			NARROW TRANSCRIPTION		
PRIMARY QUALITIES	**i, iː**		**u, uː**	**i, iː**		**u, uː**
	e, eː		**o, oː**	**e**		**o, oː**
					æː	
		a, aː			**ɑ, ɑː**	
SECONDARY QUALITIES	**y, yː**		**ɯ, ɯː**	**y, yː**	**ɨ, ɨː**	
		ø, øː				**ɵ, ɵː**

Figure 1. The Toda vowel inventory

PLACE ⇒ MANNER ⇓	LABIAL	DENTAL	POST-DENTAL	ALVEOLAR	RETROFLEX	PALATO-ALVEOLAR	VELAR
STOP (AND) AFFRICATE	p b	t̪ d̪	ts dz	t̠ d̠	ʈ ɖ	tʃ dʒ	k g
NASAL	m			n	ɳ		
FRICATIVE	f	θ	s̪	s̠	ʂ	ʃ	x (ɣ)
TRILL			r̪	r̠	ɽ		
APPROXIMANT						j	w
LATERAL			ɬ l		ꞎ ɭ		

Figure 2. The Toda consonant inventory

Previous accounts of the status of voicing in Toda are not completely clear. Emeneau (1984:14) divides the consonants into the following sets (using our symbols):

invariable voiceless consonants:	p t̪ ts t̠ tʃ ʈ k
invariable voiced consonants:	b d̪ dz d̠ dʒ ɖ g l ɭ ʒ ʐ r̪ w
variable voiceless consonants:	f θ x r̠ ɽ ɬ ꞎ s̪ s̠ ʃ ʂ
variable voiced consonants:	m n ɳ j

He claims that the variable sets show allophonic variation, having intervocalic voicing of the voiceless consonants and voiceless allophones of the voiced consonants in final position and after invariable voiceless consonants. The invariable sets are said to show no allophonic variation in voicing. A detailed account of this phenomenon is given in Emeneau (1984:16-17).

All Toda consonants can occur word-medially and word-finally but only the following set, representing a subset of each manner, may appear word-initially:

stops:	p t̪ k
fricatives:	f s̪
nasals:	m n
rhotics:	r̠
laterals:	l
approximants:	j w

What is relevant for the present study is that the various trills contrast with each other only in word-medial or final position. In initial position only r̪ occurs.

1.1 The terms 'flap', 'tap', and 'trill'

All three of the Toda rhotics involve rapid movements of the tongue tip, some of which might be called trills, taps or flaps. The latter two terms have been used very confusingly in the linguistic literature. Ladefoged is responsible for as much of the confusion as anybody, having given different definitions of these terms in three editions of the same textbook (Ladefoged 1975, 1982, 1993). We now think it best not to make a formal distinction between 'tap' and 'flap'. One alternative is to regard these two terms as synonymous (as is implied by Emeneau, 1984:12). Another alternative, which we prefer, is to avoid the use of the term 'tap' altogether.

McGowan (1992) presents an articulatory model of tongue tip trills. In his view of a trill, the tongue tip is a spring loaded trap door separating cavities with different air pressures. The tongue body is braced against the insides of the teeth to allow the tongue tip to swing open and closed. When the oral pressure is great enough it overcomes the torsional tension of the tongue tip trap door, which swings open. As the air rushes out of the oral cavity the pressure drops low enough to be overcome by the torsional tension of the tongue tip trap door, which then shuts and again creates a full closure. As there is continuous pulmonic airflow, air pressure will again build up in the oral cavity and the entire cycle repeats.

We are not sure if we entirely agree with this view of a trill as being the product of a relaxation oscillator — a system like the lid of a boiling kettle bouncing up and down as the steam is let out. It seems to neglect the role of the Bernoulli forces which may suck the tongue up towards the roof of the mouth as air rushes between them, causing the tongue tip to behave as much like a reed in a clarinet as a kettle lid. But in any case the essential factor is that trills are caused by aerodynamic forces, whereas taps and flaps are not. In Ladefoged (1993) both tap and flap refer to a sound 'caused by a single contraction of the muscles so that one articulator is thrown against another'. Even if we do not distinguish between taps and flaps, it is clear that they involve muscular gestures different from those in trills. In this paper we will consider a flap to be produced by a muscular movement involving the tip and blade of the tongue resulting in a brief contact at or near the alveolar ridge. This kind of articulation does not call for lateral bracing of the tongue against the teeth, as does a trill (cf. McGowan 1992). A trill we take to be produced by an initial muscular movement of the tip of the tongue directly upward to somewhere in the general alveolar area, and then aerodynamic forces causing intermittent contact between the tip of the tongue and the roof of the mouth, while the sides of the tongue are braced against the teeth.

In languages such as Spanish, Italian, and, as we shall see, Toda, underlying trills often have single contact allophones. We consider such segments to be single-contact trills and not taps. For this paper a trill is still considered a trill regardless of the number of contacts, as long as it involves the same manner of articulation, i.e. an aerodynamically induced movement of the tongue tip, distinct in nature from the purely gestural contact in a flap or tap of the kind found in American English words such as 'city, better'. The intervocalic consonants in the latter words will always have a single, muscularly produced, contact, and will never have a trill variant.

1.2 Previous accounts of Toda rhotics

The Toda rhotics are usually described as contrasting in place: post-dental, alveolar and retroflex. They are transcribed by Dravidianists as **r**, **r̠**, **ṛ** respectively, with the dental rhotic being unmarked (cf. Emeneau 1984; Subrahmanyam 1983). In accordance with IPA practice, we have opted for **ɽ** rather than **ṛ**, and we have decided to use diacritics for both the dental, **r̪**, and the alveolar **r̠**. In this way we hope to make this paper useful to Dravidianists as well as linguists concerned with more general phonetic problems.

In addition to place of articulation, Emeneau (1984) also suggests that these sounds contrast in the number of contacts within the trills. His general phonetic description of them is as follows (Emeneau 1984:12, our symbols):

> "Of the trills, r̪, is a one-flap with point of contact at the roots of the upper teeth, i.e. post-dental. r̠ is a trill of as many as four or five flaps in about the same position as r̪, ɽ is a trill of the same length as r̠, the initial tap starting with the tongue-tip strongly retroflexed and the remaining taps making contact just back of the alveolar ridge without observable retroflexion."

As we have noted, voicing is contrastive for segments such as stops in Toda, but there are no voicing contrasts among nasals, fricatives, and laterals. Of the rhotics, the post-dental r̪ is classified by Emeneau as being invariably voiced while r̠ and ɽ are both said to be variably voiceless. Summarizing the above information in Table 1, we have the following characterization of these rhotics in Emeneau's terms:

Table 1. Emeneau's general description of Toda Rhotics

	VOICING	PLACE OF ARTICULATION	NUMBER OF CONTACTS
r̪	+ (invariably)	post-dental	1
r̠	- (variably)	alveolar	up to 5
ɽ	- (variably)	retroflex	up to 5

2.0 Procedure

The first goal of this paper is to determine how well our data is described using the general description in Table 1. At this point we can say that it turns out to be appropriate for several instantiations of each of the segments, given a certain amount of latitude for the place of articulation; but as we shall see it is by no means a description which consistently holds true for any of the segments in question and it will have to be qualified in a couple of crucial ways.

2.1 Data collection

The following description of the Toda rhotics is based on analysis of data from six male speakers collected in the field in February 1992. The six male speakers were in two groups of three, one group (speakers 1-3) from the Melgas *mund* (tribal location), and the other (speakers 4-6) from the Kas̠ *mund*, located a few miles away. For both groups, one speaker would say the word in isolation and the other two would repeat it after him. This was done for every token on the elicitation lists used for the two groups. The two lists coincide except for a small subset of tokens elicited for only one or the other of the two groups; members of this subset are marked with an asterisk (*) along with the corresponding speaker numbers. In addition, the Melgas group also produced each of the words on the list in a frame sentence **kad __ essi** 'Kad said __' (*Kad* is a common male Toda name). The entire data set used in this paper is given in Table 2.

Table 2. Entire word list recorded for phonetic study of Toda rhotics.

	TODA	ENGLISH GLOSS	TODA	ENGLISH GLOSS
MANNER ⇒ PLACE ⇓	PLAIN RHOTIC		PALATALIZED RHOTIC	
DENTAL	**y:r̪**	'sweat'	**yr̪ʲ**	'fingernail'
	e:r̪	'to plough'		
	pø: r̪	'name'		
			o:r̪ʲ	'who'(4-6)*
	kar̪	'border of cloth'		
			par̪ʲ	'to ride horse at gallop'
			tu:r̪ʲ	'knife'
ALVEOLAR	**e:r̠**	'male buffalo'(1-3)*	**er̠ʲ**	'to throw'
	er̠	'to scoop up water' (4-6)*		
	ner̠	'to become full' *	**ner̠ʲ**	'sari worn by goddess' (4-6)*
	pø:r̠	'cliff'		
	o:r̠	'six'		
	kar̠	'juice/sap'	**kar̠ʲ**	'to laugh' (4-6)*
	mar̠	'to forget' (4-6)*	**mar̠ʲ**	'young of animals'
	par̠	'vessel'		
	mur̠	'to become angry'	**mur̠ʲ**	'complaint'
	tu:r̠	'bunch of leaves'		
RETRO- FLEX	**me:ɽ**	'to drive buffalo on migration'		
	mø:ɽ	'name of mund' (4-6)*		
	oɽ	'to cook' (4-6)*	**oɽʲ**	'foot' (4-6)*
	poɽ	'to lie down' (4-6)*	**poɽʲ**	'rice given at funeral' (4-6)*
	toɽ	'thigh' *	**toɽʲ**	'pole used at funeral' (4-6)*
	kaɽ	'pen for calves' (4-6)*		
	puɽ	'beating' (4-6)*		
	pɯɽ	'to release' (4-6)*	**pɯɽʲ**	'to open up (as a boil)'(4-6)*

2.2 Acoustic Analysis

The analog recordings were digitized onto a Kay Elemetrics Computerized Speech Lab (CSL) at a sampling rate of 12,500 samples per second. These digitized tokens then served as the basis for the spectrographic and acoustic analysis.

The spectrographic analysis consisted of wide band spectrograms covering a fixed frequency range of 5000 Hz and having a typical window size of 500 ms ± 10 ms. This particular window size was selected because it was found to adequately span the isolated monosyllabic words on our list. As shown in Figure 3, each spectrogram also had an associated waveform which spanned the same window length as the spectrogram.

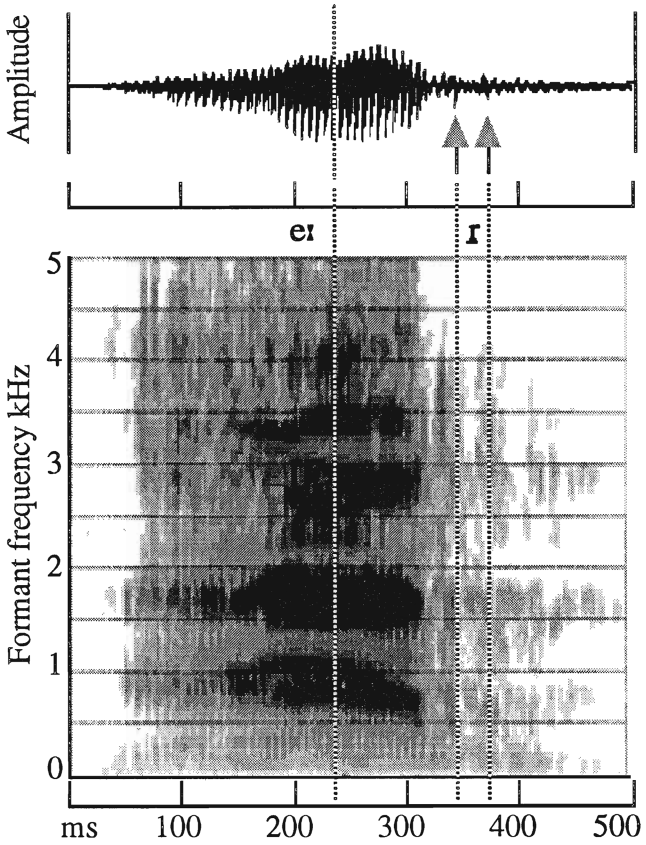


Figure 3. Illustrative spectrogram and associated waveform. The arrows indicate the releases of the aerodynamically induced contacts.

The two dotted lines with associated arrows on the waveform represent two contact releases in the production of this trill. The waveform and the spectrogram combined proved to be very useful in determining how many contacts were present for any given rhotic.

In addition to the spectrograms and waveforms, CSL also allows one to obtain FFT and Linear Predictive Coding (LPC) spectra for any selected point in time in a given spectrogram. The window size chosen for the LPC was 20 ms and for the associated FFT 51.2 ms, with both windows being centered at the same point. Typical FFT and LPC spectra for a selected position are presented in Figure 4.

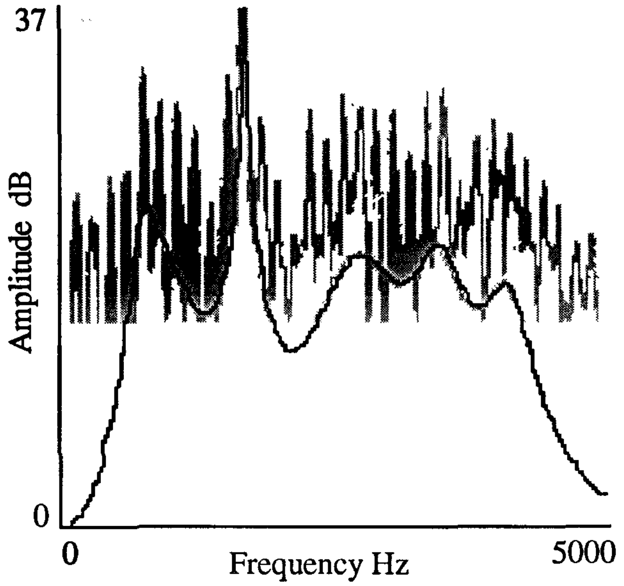


Figure 4. Illustrative LPC (black line) and associated FFT (fainter, gray lines). These spectra are centered on the dotted line at about 230 ms in Figure 3, representing the middle position of the long vowel **eː**.

2.3 Articulatory Analysis

In addition to the acoustic analysis, we obtained video palatographic data for three male Toda speakers of the Melgas *mund* from the following list of words with final rhotics: **kar̪** 'border of cloth', **kaɹ** 'sap, juice', **kaɽ** 'pen for calves'. This list is a subset of the larger one given in Table 2 above.

The palatography procedures were as described in Ladefoged (1993). The speaker had his tongue painted with a solution of 1 part charcoal and 2 parts vegetable oil. The desired token was uttered so that the upper palate will have blackened areas reflecting where the painted portion of the tongue made contact. A mirror was placed in the mouth and rested against the bottom teeth, and the image of the upper palate in the mirror was videotaped with a camera with a fixed focus so as to standardize the size of the video image. We endeavored to videotape the entire period from the time the informant was prompted to produce the desired word to the time that the palate image was recorded in the mirror. This gave the researcher an audio as well as video image of the entire production of each token and offered a cross check as to the production associated with a given palatogram. The tongue and palate were then cleaned off and the tongue repainted for the next word in the list. The video image was transferred onto a computer using Image v.1.7 and reduced to life size. The reductions to life size are made possible from measurements obtained from dental impressions of the upper palate obtained in the field. Casts were made from these dental impressions, and from the casts we were able to make molds of the upper surface of the palate. By cutting these along the mid-sagittal line we obtained the sagittal view which gives us an accurate representation of the shape and depth of the palate itself. by reference to known distances on the palate measured on a dental impression of the mouth. Figure 5 illustrates a typical palatogram which has been digitized and reduced to life size.

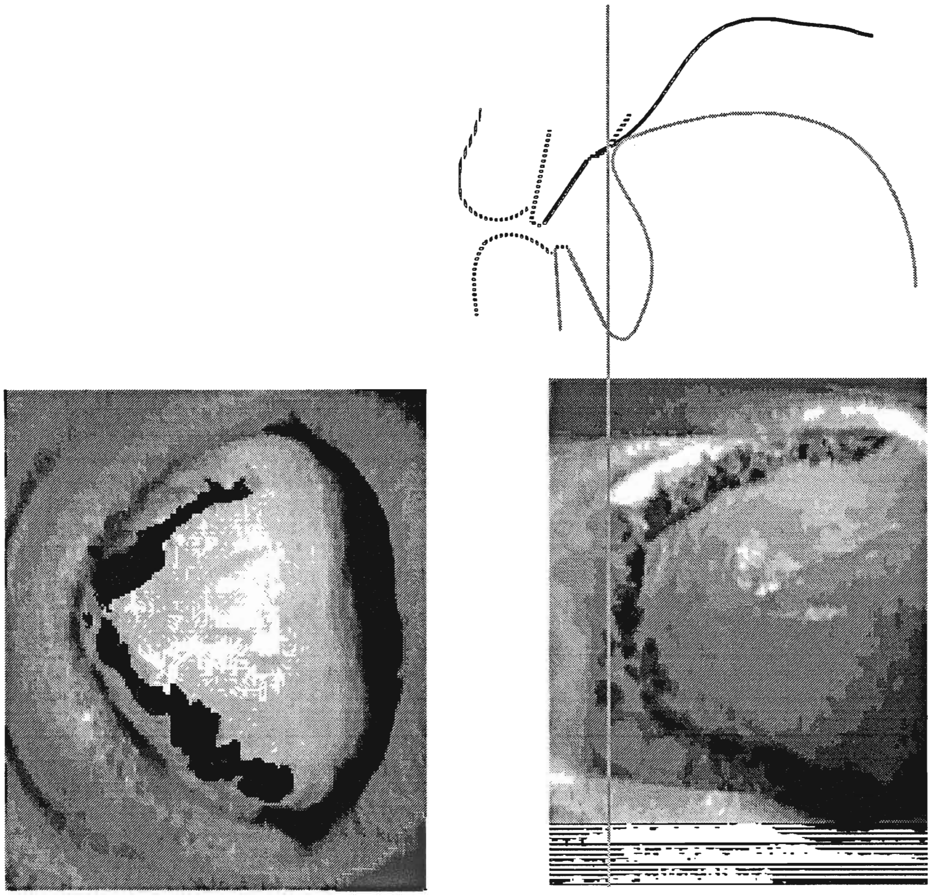


Figure 5. Palatogram, linguagram, sagittal view and inferred tongue position for speaker 2 producing the word **kar̪** ‘border of cloth’.

By lining up a known point such as the point where the teeth meets the gums in both the palatogram and in the sagittal view, we are able to draw guide lines and determine just where along the upper palate the contact was made. By interpolation based on the observed contacts and the known shape of palate we hypothesize the shape and position of the tongue body.

For each speaker we also obtained linguagrams of the three words used for palatography. For each linguagram the palate was painted with the charcoal/oil solution and the tongue cleaned off. The desired token was uttered so that where the tongue came in contact with the painted palate it became blackened. The informant was then asked to protrude his tongue so that it could be video taped. As in the observation of the contacts on the palate, the entire procedure was videotaped so as to provide an acoustic record of that particular production. The left side of Figure 5 illustrates the linguagram of the word **kar̪** ‘border of cloth’ as produced by speaker 2.

Splitting the procedure into two components, palatograms and linguagrams, allows us to obtain much more accurate contact areas on either the tongue or palate. The main drawback is that we are now looking at two different tokens and the informant may have produced the two tokens in slightly different ways. Repeated palatograms of the same word look very much alike from token to token, and it is probable that the tongue contacts also do not vary very much from utterance to utterance The recorded linguagrams show greater variation largely because speakers find it difficult to protrude the tongue for photography in exactly the same way every time.

3.0 Results

3.1 Number of Contacts within trills

As we have seen above, the dental rhotic is described as a single-contact trill and the other two rhotics are described as having 'up to four or five' contacts. Thus we have a potential differentiating feature, at least dividing the rhotics up in terms of dental or not, if the dental is indeed always produced with a single contact and the others, alveolar and retroflex, with multiple contacts.

Table 3. The number of contacts in each trill.

		MELGAS GROUP IN CITATION			MELGAS IN FRAME = INTERVOCALIC			KAS̱ GROUP IN CITATION		
		S1	S2	S3	S1	S2	S3	S4	S5	S6
DENTAL	yːr̪	1	1	1	1	1	1	1	1	1
	yr̪j	1	1	1	1	1	1	1	1	1
	pøːr̪	2	2	3	1	1	1	1	1	1
	kar̪	2	3	2	1	2	1	1	1	2
	oːr̪j	-	-	-	1	1	1	1	1	1
	tuːr̪j	1	1	1	1	1	1	1	1	1
ALVEOLAR	eːr	1	2	3	2	2	2	-	-	-
	er	-	-	-	-	-	-	1	1	2
	erj	2	3	2	2	1	1	1	1	1
	pøːr	3	1	3	2	2	1	1	2	1
	kar	3	3	2	2	2	2	1	1	3
	oːr	3	2	2	1	2	1	1	1	1
	tuːr	1	2	2	1	2	1	1	2	2
RETROFLEX	meːɽ	3	3	2	3	3	2	1	3	2
	møːɽ	-	-	-	-	-	-	1	1	1
	kaɽ	3	3	2	3	3	2	1	3	3
	puɽ	3	3	3	3	3	2	2	4	3

To determine the number of contacts we looked at the waveforms and spectrograms of the respective tokens. Given that a contact represents a moment of full closure of the oral cavity it will usually have a period of 'white space' on the spectrogram reflecting this closure. But given other aspects such as the introduction of other noises in the original recording from the field we don't always obtain this period of pure silence. For these cases contacts then are taken to mean the periods on the spectrogram which do not show clear formant structure and which occur in a cyclic fashion between periods of voiced pulses or some other kind of higher amplitude noise. In addition, not all of the rhotics are consistently produced as trills but sometimes occur as fricatives. For these situations the moment of greatest stricture in a fricative is counted as a contact. These kinds of productions are very rare; in our data we observed only one or two such productions. Table 3 shows the results. A dash represents a token for which we do not have data.

The general kind of impression one gets from Table 3 is that the dental is primarily a single contact trill, the retroflex is usually a two or three contact trill, and the alveolar seems to display the most variation in the number of contacts, occurring as a single contact trill a little less than half the time.

In terms of percentages, across groups and environments the dental rhotic is a single contact in 42 of 51 tokens, which is 82% of the time though there are noticeable differences across groups and environments. Intervocalically it is a single contact 94 % of the time and in citation forms 78%. In citation, the groups differ in the number of contacts present for the dental. This difference might arguably arise from a follow the leader kind of effect, in that in each group the leader of the group said the word first in response to a prompt from one of the authors, and the other two speakers repeated the word after him. Note that in producing words with dental rhotics in citation, speakers 2 and 3 in the Melgas group consistently agree with speaker 1 in having either a single contact or multiple contacts. However, this kind of argumentation is weakened by the disparity in the amount of conformation to the leader's production found across the different rhotics.

The alveolar rhotic, in citation forms, has one contact in 16 of the 36 tokens, which is 44% of the time and 2 or more contacts in 20 out of 36 tokens, which is 56% of the time. This means that one has approximately a fifty / fifty chance of predicting whether there is one or more contacts in an alveolar. There is a significant disparity between the groups (in citation) in that the Kas speakers tend to have a single contact production more of the time than the other group. We see that 13 of their 18 tokens have a single contact production, whereas only 3 out of 18 tokens of the Melgas group show a single contact. This difference is significant by a χ^2 test. Intervocalically the Melgas groups exhibits a single contact production for the alveolar rhotic in 7/18 tokens. The occurrence of such a production in this environment would seem to imply that the number of contacts is not a primary perceptual cue.

The retroflex rhotic exhibits a more consistent pattern, namely it is more than one contact in 16 of 21 tokens, or 76% of the time in isolation and it is never produced with a single contact in a frame. Here again we see a difference in the groups in which the Melgas̱ speakers never have a single-contact trill production while the other group does in 5 of 12 of their tokens in citation.

The overall impression from our data is that there is not an inherent number of contacts to any of these rhotics and that there is great deal of variation across speakers. For example, speaker 4 has a single contact production for all rhotics, except for a multiple contact retroflex trill in the token **puɽ**. For such individuals the number of contacts surely must not be a critical aspect, at least in terms of production. Speakers 1 and 2 come closest to having a consistent pattern in which the dental is always a single contact and the other two rhotics are always multiple contact trills. Given that all of the rhotics may have a single contact production at least in citation form, and that this is observed within speakers as well as across, it seems that we cannot use the number of contacts to consistently differentiate among the rhotics.

This, however, need not mean that the number of contacts within a trill has no role at all. On the contrary, the role of the number of contacts perceptually remains an interesting and open question. Though our data clearly shows that there is no way to consistently predict the number of contacts for any of the rhotics it also shows that there is a general pattern, as mentioned above: the dental is predominately a single contact, the retroflex usually two or more contacts and the alveolar varies greatly in either having one or more contacts. This pattern is certainly reminiscent of Emeneau's (1984) description of the dental being a single tap and the other two rhotics being trills.

In the literature on Toda the dental is described as being, a 'flap' (Sakthivel 1976), a 'one-flap tremulant' (Subrahmanyiam 1983, Zvelebil 1970), and as a one 'one-tap trill' and 'one-flap trill' by Emeneau (1984). In that these sources do not elaborate on how they are using these terms, it is often not clear just what is meant. That is, whether 'flap' is being used simply to mean a single contact irrespective of kind of gesture involved or if it is being used to mean a flap as we have defined above: as found in American English 'bu<u>dd</u>y'. In our data set the dental is predominately produced with a single contact production, as shown in Table 3, but critically we note that there are cases in which all three speakers of the Melgas group produce the dental with more than one contact in both environments, i.e. in citation form and in a phrase where it is intervocalic. Figure 6 shows the production of the words **kar̪** and **kaɽ** in the frame (**kad**)____ **essi**.

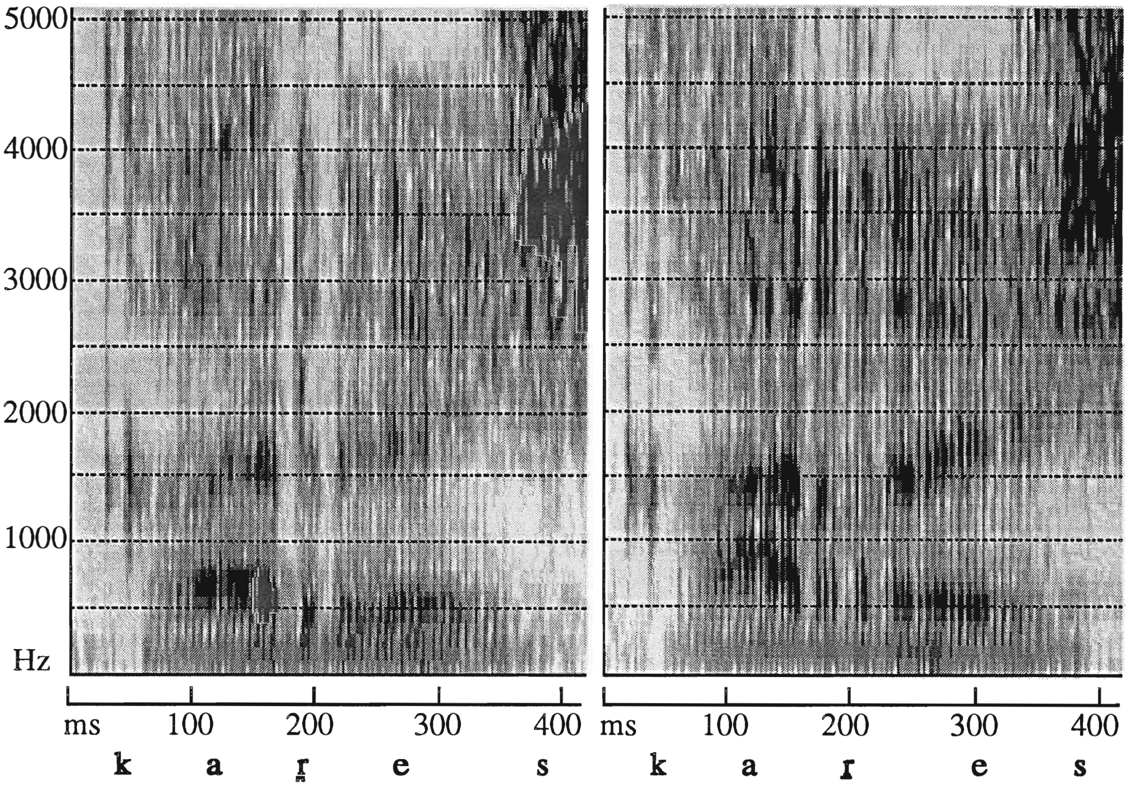


Figure 6. **kar̪** vs. **kar̠**. Speaker 2 : showing both as trills.

The two white spaces which are separated by about 15 ms apart in the spectrogram for the word **kar̪** represent periods of contact of the tongue against the upper palate. It is very improbable, if not impossible, to produce two flaps, as we are using the term, in such a short period of time. The frequency of tongue tip alveolar trills is of the order of 25-30 cycles per second (Ladefoged and Maddieson, in press). This is much higher than the highest rate for taps: 'about five or six per second' (Catford 1977). Therefore both the rhotics in Figure 6 are labeled as ‘trills’ in this production. This in itself does not necessitate that one has to consider all of the dentals, including those which consist of only a single contact, as trills; but if the dental were a ‘flap’ or ‘tap’ underlyingly then one is placed in the awkward position of having a flap become a trill intervocalically. This kind of change is indeed odd when considering the two very different manners of articulation involved here. While it is the case that trills often occur as single-contact productions allophonically it is not been shown anywhere, to our knowledge, that a flap (muscular gesture) becomes a trill (aerodynamic) intervocalically.

The contacts in the dentals display a longer closure period than those of the alveolar for both the first and subsequent contacts. That the first contact in a trill tends to be longer than the others has been pointed out by Ladefoged & Maddieson (in press). But here we are dealing with a length difference across rhotics, that is the first contact of the dental is longer than the first contact of the alveolar, and so forth comparing subsequent contacts. The interesting thing to point out is that even in the one-contact productions of both the dental and alveolar, the contact duration of the dental rhotic tends to be longer. The longer duration of dental closures whether in one-contact or multiple-contact instances suggests a consistent manner of production for dental rhotics regardless of the number of contacts.

For the reasons outlined above, we will consider all three rhotics to be trills, i.e. aerodynamically induced, even when there is only one contact present in the production of the dental rhotic. This is not a necessary conclusion, but it is one which we believe will answer more questions than if we considered the dental a flap or tap. We will return to this issue in the section on rhotic formant frequencies and in the general discussion.

3.2 Voicing

As we have noted, the issue of voicing in Toda is somewhat unclear. Emeneau (1984) suggested that the consonants are divided up into four different voicing paradigms: invariably voiceless and voiced and variably voiceless and voiced. Such a situation is adopted due to a lack of a better solution and Emeneau (1984:14) himself states it is somewhat ‘spurious’. The issue arises because members of the variable sets display allophones but the environment for the allophony has not or perhaps could not be determined. This scenario leads to the issue of which allophone, voiced or voiceless, to consider to be underlying. To determine this Emeneau decided to consider those sounds which were preceded by vowels with more closed allophones as voiced, and those with more open allophones as voiceless (Emeneau 1984:14).

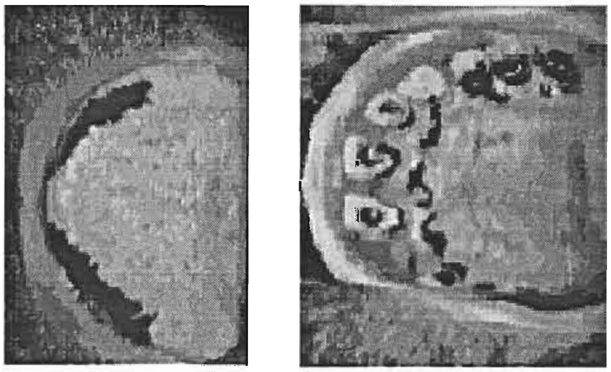
This reasoning is based on the general tendency of Toda vowels preceding voiced segments to be more closed than the same phonemic vowel preceding voiceless segments. Voicing causes the larynx to drop and thus creates a greater back cavity, which in turn causes a lowering of F1 which is interpreted as a raising of the vowel. Using vowel raising as a diagnostic Emeneau classified the dental **r̪** as invariably voiced and the alveolar **r** and retroflex **ɽ** as variably voiceless.

However, this diagnostic has at least three flaws. The first is that Shalev et al (1993) did not find voiced segments to consistently have the described effect on the preceding vowel. That is, though the effect was noticed for **i**, **y**, **yː**, uː, **o**, **oː**, æː and **a** it was not noticed for **iː**, **ɨ**, **u**, **e**, or **ɵː**. The second is that the increase in back cavity area might in fact be due to some other gesture besides a lowered larynx caused by voicing. Such gestures include tongue body and/or tongue root fronting. Since we believe the dental rhotic to have a fronted or perhaps raised tongue body (examined below) it is not clear which gesture is causing the more closed allophone in the preceding vowel. The third problem is that, as Emeneau (1984:14) himself points out and our data confirm, word-finally the first few contacts in the alveolar and retroflex are typically voiced and the remainder are voiceless. Our data also show that the dental rhotic has occasional voiceless allophones word finally. This goes against Emeneau's claim that it is 'invariably voiced'. However, in support of Emeneau's claim, we should note that voiceless allophones of the dental rhotic are unusual, and it is almost invariably voiced. It is also, of course, the only one of these rhotics that appears in initial position, where it is voiced. We have not been able to devise a procedure that gives us reliable measurements of the degree of voicing in the three rhotics in final position, and can report here only that there seems to be a tendency for the dental rhotic to be more consistently voiced.

3.3 Place of articulation

Dravidianists have used the place of articulation as a means of distinguishing between rhotics. This approach has its roots in historical work on Dravidian where these rhotics pattern with the respective natural classes to which they belong, i.e. the dental rhotic follows rules which affect other dentals, the alveolar rhotic groups with other alveolars and so forth. But even in this body of work there are allusions to the fact that this description might not necessarily reflect the situation with the 'dental' and 'alveolar' rhotics. As we have seen, Emeneau (1984:12) insightfully points out that the alveolar, **r̪**, is "produced in about the same place as the **r**" (using our symbols). To examine this issue we used palatography as outlined in section 2.2.

kar̪ 'border of cloth' - Tes̲



kar̲ 'juice / sap' - Tes̲

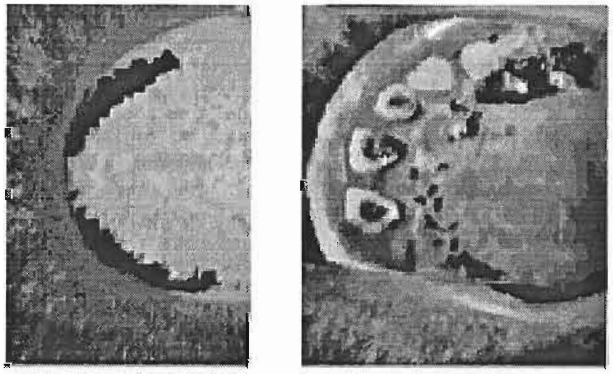


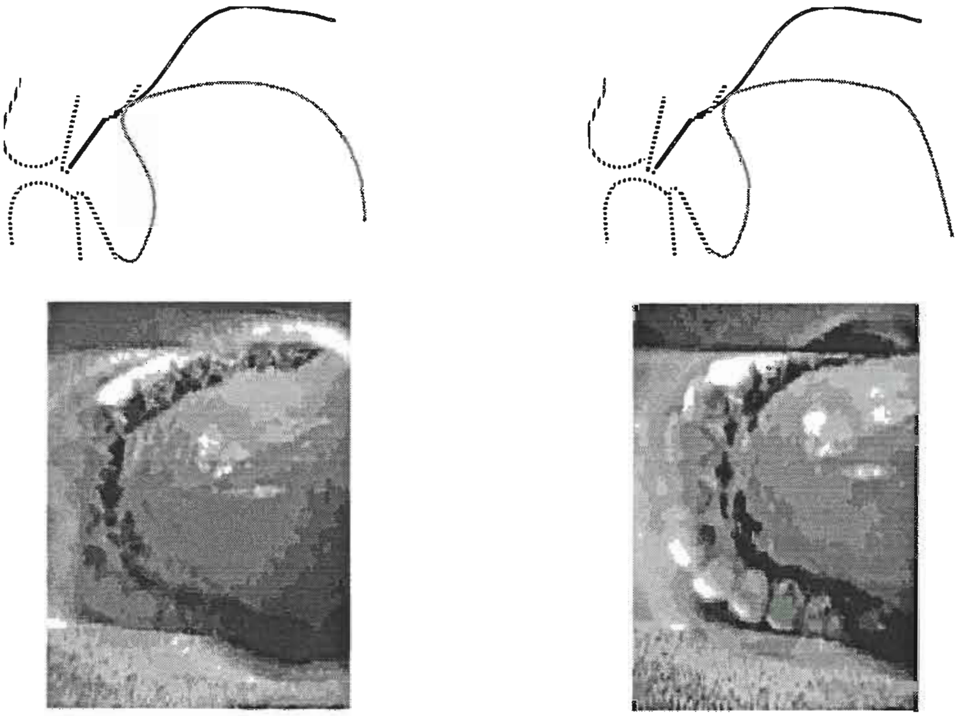
Figure 7. Speaker 1(Tes̲): Palatograms of kar̪ and **kar̲**

3.3.1 Dental vs. Alveolar

The palatograms in Figure 7 illustrate the production of the dental and the alveolar rhotics by speaker 1(Tes̱). The dark areas on the front teeth are stains and not points of contact.

The two palatograms in Figure 7 are very similar. Both have an alveolar contact with the tongue tip and a lateral contact against the side of the mouth. There are some subtle differences in that the 'dental' has more contact on the molars and a narrower contact area on the alveolar ridge which is slightly further forward than the contact of the 'alveolar'. The lateral contact of the alveolar also seems to be more toward the middle (inside of the molars) than the 'dental'. The linguagrams also suggest a very similar contact for both. When combined with the palatograms it suggest that the alveolar has a somewhat more narrowed tongue in its production, at least for this speaker.

The palatography for speaker 3 is unreliable due to secondary contact observed on the video tape (demonstrating the value of video-taping the entire process), but we did obtain reliable palatograms for speaker 2 which are illustrated in Figure 8.



kar̪ . speaker 2

kar̠ . speaker 2

Figure 8. **kar̪** vs **kar̠** for speaker 2.

As in Figure 7, these palatograms also suggest that both segments are articulated at the alveolar ridge at approximately the same distance from the front teeth. The accompanying linguagrams for this speaker are unreliable due to secondary contact observable on the video tape. The palatograms in Figure 8 are different from the data in Figure 7 which showed the 'dental' to be slightly further forward than the 'alveolar'. The 'dental' has a broader contact area at the alveolar ridge than the 'alveolar' rhotic and the lateral contact is more toward the midline than in the alveolar. This is just the reverse of what we saw in Figure 7. These facts suggests that the 'dental' is not consistently produced further forward than the 'alveolar' and that the other subtle differences noted are not part of a consistent difference in articulation between the two rhotics. They presumably fall within an allowable degree of variation for an apical alveolar rhotic in Toda.

Emeneau (1984) described the **r̪** as a post-dental and the **r** as an alveolar. Our palatograms do not support this classification. Shalev et al (1993), in a detailed account of the stops and affricates of Toda based on the same speakers, demonstrated that the dentals and the so called 'post dentals', **t̪** and **ts** respectively, exhibited a very similar contact area which extended from the front teeth (or root of the teeth in some cases) to the alveolar ridge and thus, appropriately, termed both laminal denti-alveolars. But this kind of contact does not reflect the production of either **r̪** or **r**, as seen in Figures 7 and 8, nor by the other speaker for whom we have palatographic data.

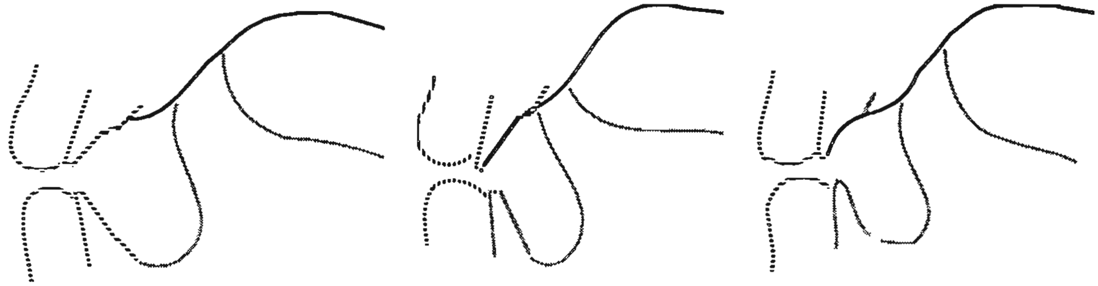
The contact area on the roof of the mouth and on the tongue in the figures above is characteristic of the production of an apical segment as shown in Dart (1991). Shalev et al (1993) also found this same kind apical alveolar contact pattern to be the norm in the production of the alveolar segments in Toda. Given these facts, both of these sounds, **r̪** and **r**, should be considered as apical alveolars as far as shape and point of primary contact is concerned.

In the preceding paragraphs in this section we have been putting the terms 'dental' and 'alveolar' in quotes, as they were obviously not being used to describe places of articulation in the usual way as has been done in other languages (Ladefoged and Maddieson, in press). For the remainder of the paper we will still refer to the rhotic segments as dental, alveolar, and retroflex for ease of reading and for continuity, but keeping in mind that the first two segments are actually produced at essentially the same point of articulation on the roof of the mouth, and with the same part and shape of the tongue at the contact. There is, however, a phonemic difference between these sounds, and we will describe below consistent acoustic differences between the dental and alveolar (and the retroflex for that matter) rhotics. It seems that the dental rhotic has some secondary articulation involving parts of the tongue that are not involved in the contact itself. This secondary articulation is most likely to be some kind of tongue body and/or tongue root fronting. We will support this claim with a detailed account of the formant structure in section 5.2.

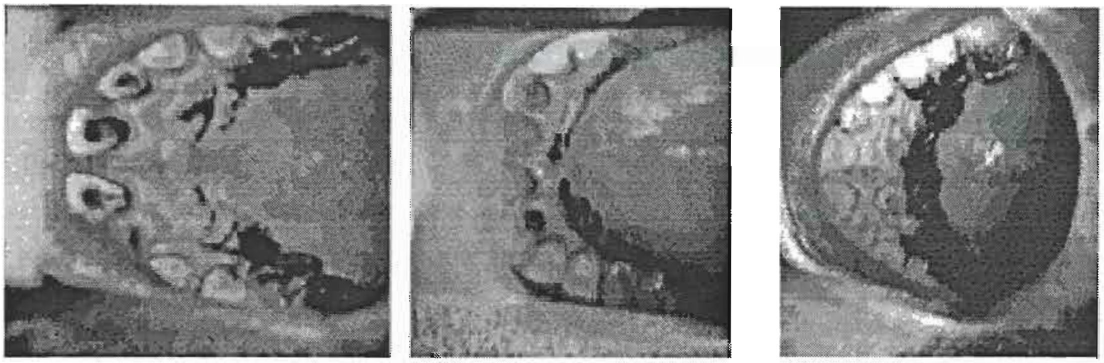
3.3.2 Retroflex vs. Alveolar & Dental

The palatographic data in Figure 9 illustrates that the retroflex is indeed produced further back in the mouth than the other two, which as we have seen are produced in the same place of articulation. However, the palatograms do not reveal all the contact areas as it was not possible to paint the underside of the tongue.

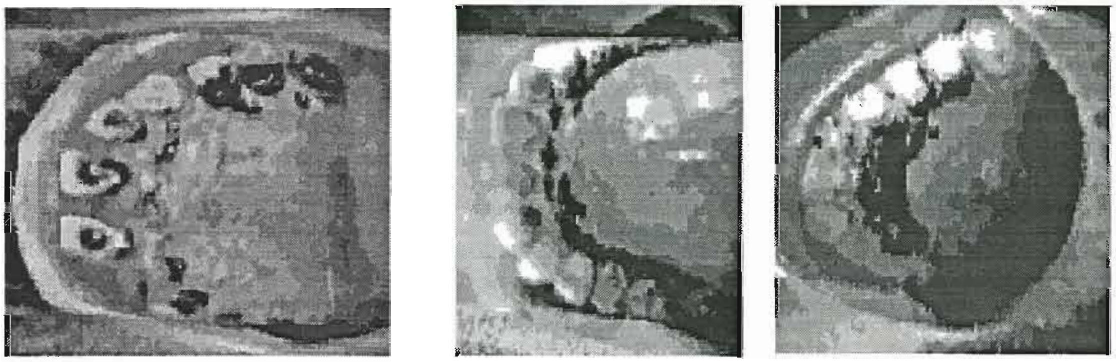
These palatograms demonstrate that the retroflex **ɽ** is distinct from the other two rhotics in place of articulation.



kaɽ **kaɽ** **kaɽ**



speaker 1 (Tes) speaker 2 speaker 3

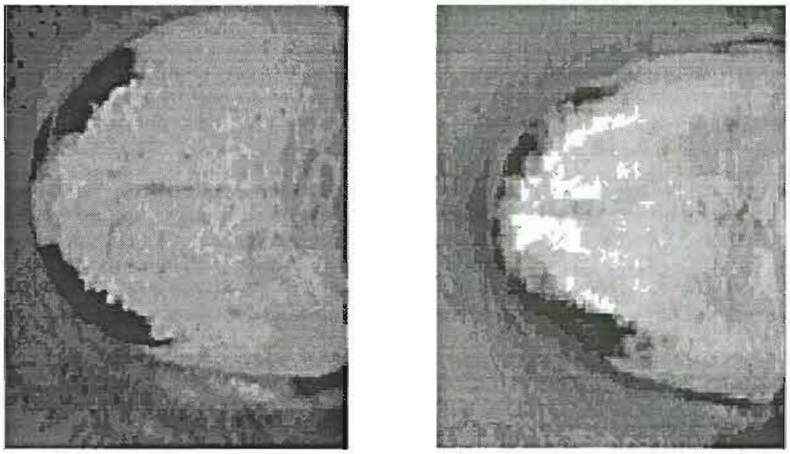


kaɾ̥ **kaɾ** **kaɾ**

Figure 9. Palatograms comparing ɽ with ɾ or ɾ̥ .

We note that the palate contacts for Speaker 2 for ɽ are less distinct than for the other two speakers. This might be explained by one or more factors. Visual inspection of the casts for speaker two shows that he has a palate with a very high vault, higher than either of the other two speakers. This is in terms of both steepness along the mid-sagittal crest and the angle of the lateral walls. This kind of combination, deep and steep, makes it more difficult to place the mirror in the mouth so as to get the optimal 45° reflection of the palate. The end result is that two distinct contact points will not necessarily show up as being all that distinct because of the angles involved. Secondly, in a retroflex trill, as pointed out by Emeneau, and as we shall support with acoustic analysis, the tongue starts out by moving backwards before touching the roof of the mouth; the first contact will be made back of the alveolar ridge and subsequent contacts will be made slightly further forward near or at the alveolar ridge. Thus when we look at the palatogram which represents the production of the retroflex we must bear in mind that we are actually looking at the entire area covered during the numerous contacts which started in the retroflex region and ended up at about the alveolar ridge. As much of the marking medium will have been removed from the tongue on the first contact there may not be any further marks for later contact if the same part of the tongue is used. It is possible that a speaker will simply retroflex the tongue but not make full contact in this position and only make the first contact with the tongue tip at or about the alveolar ridge. Any of these factors might explain this particular production for speaker two.

Our linguagraphic data shows that it is the under side of the tip of the tongue that is used in the retroflex articulation, at least for the primary contact. This is illustrated in the linguagrams for two speakers shown in Figure 10. The linguagram for the other speaker was flawed due to a secondary contact obscuring the contact pattern for the rhotic



Speaker 2: **kaɽ**　　　　Speaker 3: **kaɽ**

Figure 10. linguagrams for the production of the **ɽ** in **kaɽ**

3.4 Summary of the articulatory data

Considering the rhotics in terms of the three basic parameters arrived at from Emeneau's work, we find that the general description needs to be changed in three ways to account for the data which we have examined. The first point is the number of contacts within each trill. There is not an inherent number of contacts associated with any of the plain rhotic trills. But in our data there was a general tendency for the dental to have a single contact, the retroflex to have multiple contacts, and the alveolar to be more variable and to have either a single or multiple contact rhotic. The issue of why this pattern arises remains an interesting open question. The second point to be considered is the voicing. We found no evidence that the three rhotics differed in terms of voicing. The last point is the place of articulation. Our data has clearly shown that, at least for our speakers, the dental is in fact not different in contact from the alveolar and that both are in fact apical alveolars. Given that the two sounds are clearly different acoustically we postulate that the difference must arise from some secondary articulation in the dental involving a part of the tongue not involved in the contact itself. This claim is examined in greater detail below.

4.0 Acoustic analyses

4.1 Previous accounts of rhotic formants

Previous research (Ladefoged 1971, Lindau 1985, Condax and Nathan 1979) has focused on the acoustic properties of rhotics but only in terms of trying to find some property common to the broader set of rhotics which spans manner and place of articulation and which sets them apart from all other consonant segments. Ladefoged (1971) proposed that a lowered F3 close to the second formant was such a feature. Lindau (1980) showed that this is not the case. A lowered F3 was characteristic only for those rhotics produced in the post-alveolar and retroflex regions. Rhotics produced in front of or in back of this mid palate area tended to display a higher, and not lower, F3.

Lindau (1980) also examined the first two formants and concluded that they 'do not seem to correlate well with r-ness. More than anything else they seem to reflect coarticulation with the surrounding vowels'. In this same work she also points out that rhotics generally had weaker formant frequencies than the surrounding vowels. This lower amplitude of the rhotic formants in general acoustic theory is taken to be a characteristic of sonorant consonants, setting them apart from vowels. Although Lindau examined trills, to our knowledge there are no detailed accounts which compare the formant structures of two or more trills in a given language. This next section deals in detail with the formant structure of the different Toda rhotics.

4.2 Procedure

In our analysis below we used the techniques outlined in the general acoustic procedure, section 2.2. We made measurements of each rhotic and the associated preceding vowel in three different positions. These three positions correspond to the onset, middle and offset of each segment.

4.2.1 Vowel formant measurements.

For vowel initial tokens the onset measurement was consistently taken at the point in the vowel which showed the comparatively steady pattern of voicing and formant structure found in the rest of the vowel. This position was chosen to try and avoid introducing other variables into our data set such as the glottalization that sometimes occurred as a speaker started to articulate a vowel (which can affect both the voicing quality and the formant frequency) and the occasional **j** onglide occurring with some tokens with an initial mid or high front vowel. Tokens which began with a voiceless consonant typically had some associated aspiration occur in the release. The onset measurement was taken at 10 ms after the first occurrence of voicing in the vowel. The only other kind of initial segments that we have in our data are nasals. Since nasals tend to affect vowel formants the onset measurement of the vowel was taken at the position in which the effects of the nasal were no longer clearly visible. The vowel middle measurements corresponds to the point in the vowel midway between the onset measurement and the start of the first contact of the rhotic. In longer vowels this reflects a steady state portion, but in short vowels it is typically simply a middle position along a continuum of formant transition. The vowel offset represents a point about 10 ms prior to the first closure associated with the rhotic.

4.2.2 Rhotic formant measurements.

The rhotic onset measurement represents a point approximately 10 ms after the first closure of the rhotic is made. While it is arguable that the actual onset of the rhotic is earlier than this, the contact presents a clear demarcation point which allowed for very consistent measurements in terms of vowel offset and rhotic onset. The rhotic middle position represents a somewhat more varying measurement. For single contact rhotics which then had a voiced vowel like release portion, the release is taken to represent a part of the rhotic and the middle position of this release was viewed as the middle of the rhotic. The middle measurement reflects the middle of this offset which is typically 60 ms after the end of the closure. In rhotics with two or more contacts the middle position is taken to be a point about 10 ms after the second closure. As it turn out this point is also approximately 50 - 60 ms of the first closure. This is a point which is in the open phase of the rhotic where we observe some formant frequencies.

The rhotic offset is a more difficult point to qualify. This measurement essentially represents the furthest right point which still has measurable formant structure. In the case of the palatalized rhotics where there is a clear difference between the formant quality of the rhotic and the associated j̊ release, this measurement was taken before the portion of the spectrogram which was clearly the release.

5.0 Results

5.1 General View

Figure 11 illustrates three spectrograms of the words, **kar̪** 'border of cloth', **kar̠** 'juice/sap', **kaɽ** 'pen for calves' said in citation form. These spectrograms should be viewed as presenting a general exemplar of the production of each plain rhotic but not necessarily the only one possible in that the rhotics tend to vary in the ways which we have outlined above.

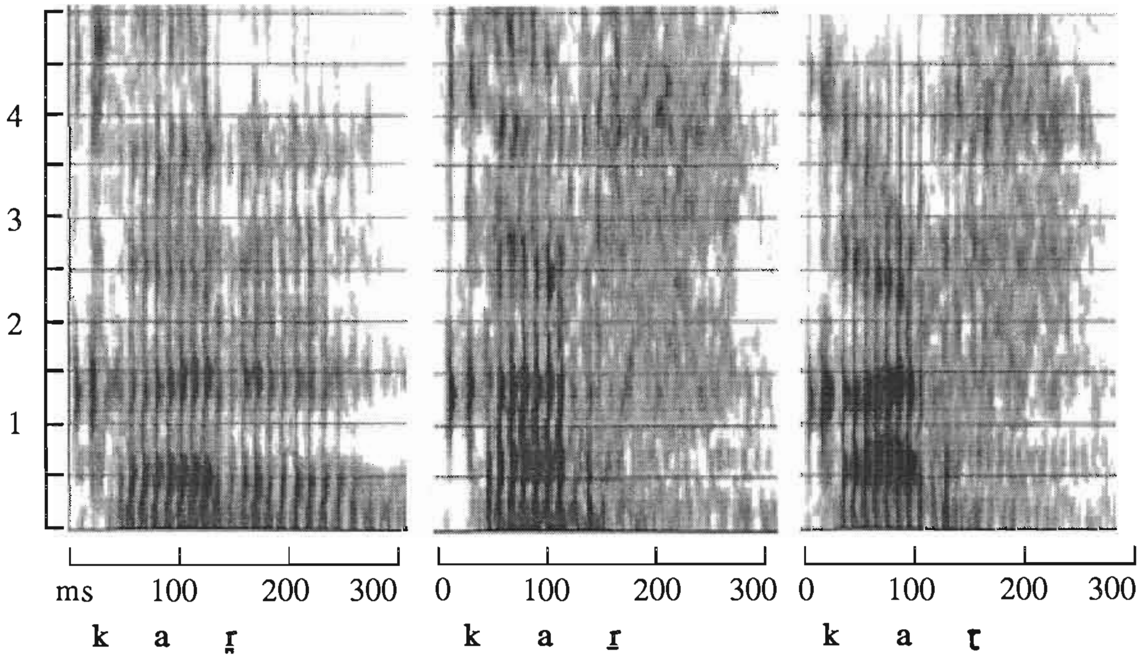


Figure 11. **kar̪**, **kar̠**, **kaɽ** as produced by speaker 4 in citation form.

We see that in this particular production the dental has a fully voiced release which looks very much like a vowel whereas the other words end with a voiceless portion. This might be taken to indicate the different voicing status of the dental (as Emeneau proposed) or as evidence that the dental has some other distinctive characteristic of its release. But this release portion is not always voiced nor is it always present. The alveolar and retroflex rhotics tend to have voicing pulses for the first couple of contacts after which they have a voiceless quality and release. This pattern is typical of their production in words spoken in isolation. (They and the dentals are fully voiced intervocalically.) During the rhotics one can also usually observe formants which are considerably weaker than those found in the preceding vowels.

5.2 Rhotic Formant Structure

Figure 12 shows the averages of measurements of the middle position of the rhotic representing the formant structure of each of the three plain rhotics after different vowel qualities. The measurements are averaged over the number of speakers involved. This was usually six speakers, but in some instances, which are noted in the data set given below, the average is obtained over only three of the speakers (whose identities are indicated by the speaker numbers). The data set used here is:

DENTAL		ALVEOLAR		RETROFLEX	
yːr̪	'finger nail'				
eːr	to plough' (1-3)	**eːɹ**	'male buffalo' (1-3)	**meɽ**	'drive buffalo on migration'
		eɹ	'scoop up water' (4-6)		
pøːr̪	'name'	**pøːɹ**	'cliff'	**møːɽ**	'name of mund' (4-6)
kar̪	'border of cloth'	**kaɹ**	'juice/sap'	**kaɽ**	'pen for calves'
		oːɹ	'six'	**poɽ**	'to lie down' (4-6)
		tuːɹ	'bunch of leaves'	**puɽ**	'beating'.

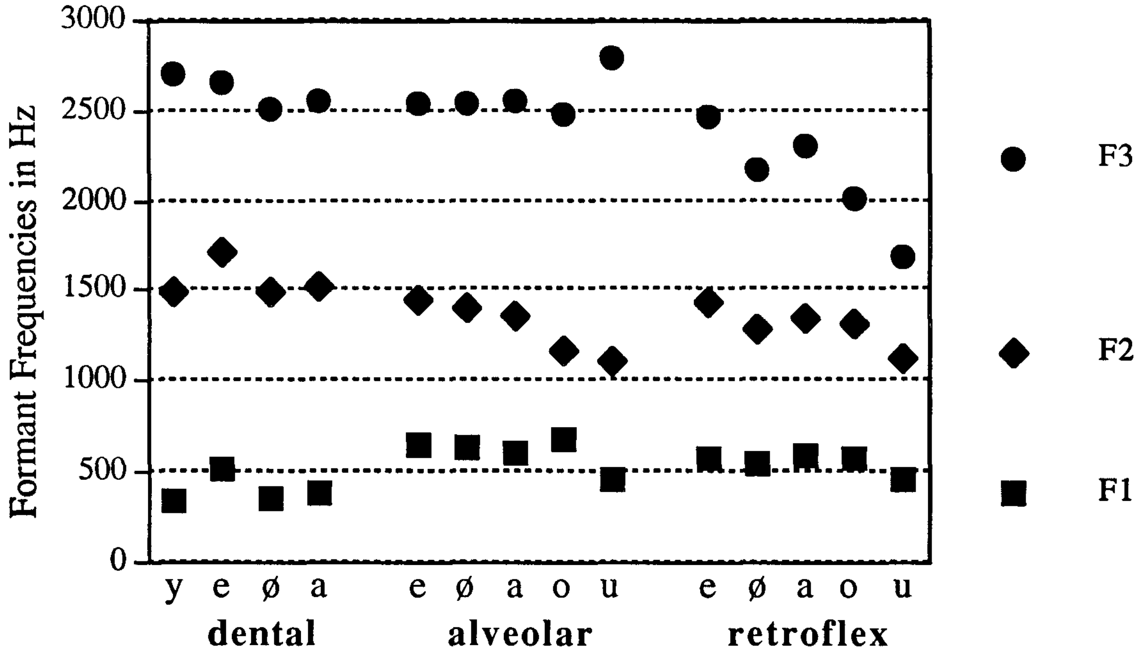


Figure 12. Formant structure of the middle position of the rhotics r̪, ɹ, and ɽ according to vowel context.

This figure illustrates two general points. The first is that these three plain rhotics exhibit distinct formant structures. The second is that vowel quality colors the formant structure of the rhotic and that different rhotics show different degrees of coloring. We can get an idea of the general formant structure exhibited by each plain rhotic by averaging the mean values of the three formants at the middle position for each rhotic across the shared vowel environments **e**, **ø**, and **a.** This is illustrated in the table below. This average across vowels illustrates the point that the different rhotics have different formant frequency patterns.

Table 4. Average of mean values for each rhotic formant across **e**, **ø**, and **a** from the data set used in Figure 12.

		F1	**F2**	**F3**	**F2 – F1**
DENTAL	**ɾ̪**	409	1573	2577	1163
ALVEOLAR	**ɾ̱**	617	1395	2535	777
RETROFLEX	**ɽ**	565	1350	2310	785

The dental rhotic is distinguished from the other two rhotics, most critically from the alveolar, in that it consistently has a considerably greater F2 – F1 value. This difference is maintained in all three positions measured, the onset, middle and offset. Table 5 presents the results of three different analyses of variance, one at each of these positions, comparing F2 – F1 in the dental, alveolar and retroflex rhotics occurring in the words **kaɾ̪**, **kaɾ̱,** and **kaɽ** . It may be seen that the dental rhotic is significantly different from both the other two rhotics throughout its duration. The retroflex and alveolar rhotics, however, do not differ from each other in terms of this F2 – F1 measure.

Table 5. P-values from analysis of variance of F2 – F1 at onset, middle and offset positions in plain rhotics. Comparisons which are significantly different in post hoc tests using Fisher's PLSD are marked with S.

	ONSET	MIDDLE	OFFSET
ALVEOLAR VS. DENTAL	.0004 S	.0006 S	<.0001S
ALVEOLAR VS. RETROFLEX	.8197	.9913	.4816
DENTAL VS. RETROFLEX	.0006 S	.0006 S	<.0001 S

The same pattern of significant differences was found in other vowel contexts. This fact lends support to our earlier conclusions regarding the dental. Though the dental is produced in the same place as the alveolar is does have this different formant structure. This quality leads to positing some kind of secondary articulation associated with the dental. In acoustic theory there are several different factors which can independently lead to a greater F2 – F1, such as a fronted tongue body, raised tongue body, fronted tongue root, or a combination of these factors. Since we have no independent way of determining which of these factors are contributing to the increased F2 – F1 in the dental we will simply conclude that it is something which causes a greater back cavity behind the point of articulation. It is crucial to point out that whatever this gesture is, it manifests itself differently from palatalization, as we shall see in section 6.

The greater F2 – F1 value reported in Table 5 was averaged over both single and multiple contact dentals. When we controlled for number of contacts and then averaged the data we obtained the pattern in Table 6.

Table 6. Formant value averages at the rhotic onset position for the words **kar̪** and **kar̠** controlling for the number of contacts.

		MULTIPLE CONTACT (SPEAKERS 1-3, 6)	SINGLE CONTACT (SPEAKERS 4,5)
ALVEOLAR	F1	622	526
	F2	1262	1458
	F3	2429	2698
	F2 – F1	641	932
DENTAL	F1	397	399
	F2	1489	1630
	F3	2607	2532
	F2 – F1	1092	1232

These results here show several interesting patterns. The dental maintains a greater F2 – F1 value than the alveolar irrespective of how many contacts are present. This can be taken to imply that some aspect of the articulation of the dental is maintained irrespective of the number of contacts. The same is true of the alveolar. Nonetheless, within a given rhotic, there is a difference between multiple and single contact productions, the single contact having a greater F2 – F1. As this pattern holds across rhotics it suggest that single contact alveolars and dentals have something in common. Since these measurements are made on a single token per speaker it is also possible that individual differences unrelated to the number of contacts account for a part of the pattern. For speakers 4 and 5 who have single contact productions of both dental and alveolar rhotics we see that they still maintain a difference by means of formant structure.

The alveolar rhotic is not of great interest to us here in that it really is different only by default. It is in many respects the most basic kind of trill in that it does not have any secondary articulation or the initial tongue tip backing of the retroflex. While the other two differ from each other in two ways, having different F3 and different F2 – F1 values, the alveolar differs from each of the other rhotics in just one way. It is different from the dental by the fact that the dental has a greater F2 – F1 and different from the retroflex by the fact that the retroflex has a lower F3. Note that in Table 5 the F2 – F1 is not significantly different between the alveolar and retroflex at any point in the rhotic.

The retroflex rhotic is usually differentiated from the other two rhotics, and most critically from the alveolar, by a lowered F3. In the chart in Figure 12 we see that the only exception to this generality is the F3 in the middle of the retroflex after **e** and also to a somewhat less extent after **a**. For the other vowel environments, **ø**, **o**, and **u**, the F3 in the middle of the retroflex is lower than the F3 in the dental or alveolar. However, although the average F3 is lower in these latter vowel contexts, it is not found to be a significantly lower F3 value than that of the dental or alveolar by an analysis of variance using Fisher's PLSD post hoc test. We are therefore still left with the task of characterizing the acoustic difference between the retroflex rhotic and the other rhotics.

The acoustic measure we have been discussing has been the value of F3 in the middle of the rhotics. The hypothesized articulation of this sound is that it starts off as being retroflexed and ends up in a position comparable to that of the alveolar. To examine this issue more closely we looked at the formant frequencies of the retroflex at the three positions, onset, middle and offset.

We contrasted the retroflex with the alveolar since the two are so similar. The prediction is that the closer one gets to the beginning of the rhotic the lower the F3 will be.

5.2.1 Alveolar vs. Retroflex

In Figure 13 below we examine the formant history of both the alveolar and retroflex rhotics at onset, middle and offset after the vowel contexts of **e**, **ø**, and **a**. Again the number of speakers averaged over is 6 except where otherwise noted. The data set for this comparison is :

ALVEOLAR		RETROFLEX	
eːr̠	'male buffalo' (1-3)	**meɽ**	'drive buffalo on migration'
er	'scoop up water' (4-6)		
pøːr	'cliff'	**møːɽ**	'name of mund'(4-6)
kar̠	'juice/sap'	**kaɽ**	'pen for calves'
oːr̠	'six'	**poɽ**	'to lie down'(4-6)
tuːr̠	'bunch of leaves'	**puɽ**	'beating'.

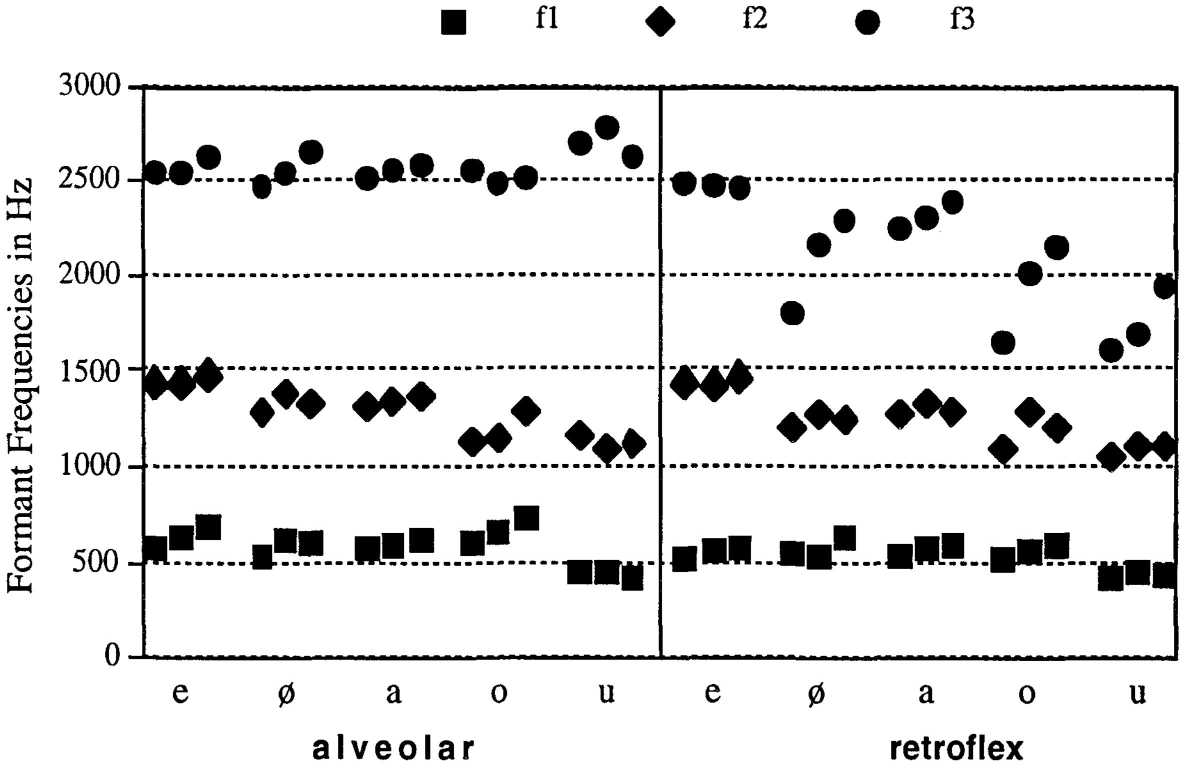


Figure 13. Mean formant values (Hz) of the alveolar and retroflex rhotics at onset, middle and offset.

The results are generally as predicted although the lowering is small with **a** and it is absent with **e**. We have no explanation of the failure of rhotics after **e** to follow the pattern.

The F3 of the retroflex rhotic is noticeably lower in rounded than unrounded vowel contexts, as may be seen by comparing the **ø** context (as well as in the somewhat suspect **e** context). The only other rounded/unrounded pair that we have in our data set is **u** and **ɯ** in the words **puɽ** ‘beating’ and **pɯɽ** ‘to release’. The formant averages of only the first of these words are represented in Figure 13; we did not have an alveolar rhotic to compare with **pɯɽ**. Table 7 below compares the formant averages in the retroflex consonants after comparable rounded and unrounded vowels. As in the comparison of the **ø** context the F3 is lower throughout the rhotic in the rounded context, and there is also a much greater rise from onset to the offset in the rounded context.

Table 7. Mean formant values of retroflex rhotics at onset, middle and offset after **u** and **ɯ**.

	u CONTEXT (SPEAKERS 1-6)			**ɯ** CONTEXT (SPEAKERS 4-6)		
	ONSET	MIDDLE	OFFSET	ONSET	MIDDLE	OFFSET
F3	1596	1677	1933	2231	2420	2392
F2	1070	1120	1121	1379	1365	1417
F1	423	451	570	530	489	450

A parallel effect is noticed in the F3 of the preceding vowel, presented below. A rounded vowel has a substantially lower F3 before a retroflex than it has before the other rhotics. What this suggests is that the combination of lip rounding and retroflexing causes an even lower F3 than just retroflexion alone. Lip rounding alone fails to account for the sharply lower F3 in that the alveolar rhotic does not have lower F3 values after rounded vowels than after unrounded ones. And retroflexion alone fails to account for this lower F3 since no comparable lowering is observed at the onset of the rhotic after an unrounded vowel.

The similarity of the alveolar and retroflex rhotics in the F1 and F2 for any given vowel is demonstrated in Figure 13. Again, using F2 – F1 as a measure of backness we can conclude that the tongue body is in a very similar position in the oral tract for these two sounds.

5.3 Vowel context effects on formant structure.

As seen in Figure 11 above, the vowel context has little effect on the rhotic structure in the case of the dental rhotic. There are slight increases of F1 and F2 for the vowel **e**, and F3 is slightly lower after back vowels. Apart from this there are consistent values of the three formants irrespective of vowel quality. One possible explanation for the small shift associated with **e** is that this segment is averaged over only three speakers, one of which is known to have higher formants than all five other speakers. The fact that for this case his higher values are only averaged with two lower values as opposed to five lower values would account for this upward shift. In conclusion then the dental is viewed as not being affected by the neighboring vowel quality.

The alveolar does show effects of vowel context on the first two formants but not on the F3. The F3 remains at about 2500 Hz irrespective of the context. Surprisingly, the F3 after **u** is higher and not lower as one might expect. It is not clear why this is. The second formant of the alveolar rhotic shows a steady drop in going from front to back vowel contexts, just as one sees in the vowels themselves. The first formant of this rhotic is lowered only after **u**. The difference between the first two formants, F2 – F1, in the alveolar rhotic is greater after front vowels than after back ones. This suggests that the tongue is further forward when this rhotic follows front vowels, taking F2 – F1 as a measure of backness (cf. Ladefoged 1993). The greatest effects on formant structure of the alveolar are observed after the **o** vowel context. The first two formants come close together just as in the vowel **o** itself.

The vowel context affects all three formants of the retroflex rhotic. The F3 is steadily lower in going from front to back vowel contexts. This parallels the F3 difference observed in the vowels themselves. As noted in the previous section the retroflex F3 is lower when the rhotic follows rounded vowels than non rounded ones.

5.4 Effects of rhotic on the preceding vowel.

In the general overview of rhotic formant structure we concluded that the dental rhotic affected the preceding **a** vowel quality, as noted by Emeneau (1984). To see if this holds true of other vowels we examined the formant structure of three different vowel qualities followed by the three rhotics: The results are presented in Figure 14. The data set used here is :

DENTAL		ALVEOLAR		RETROFLEX	
eːr	to plough' (1-3)	**eːr̠**	'male buffalo' (1-3)	**meɽ**	'drive buffalo on migration'
		er̠	'scoop up water' (4-6)		
pøːr̪	'name'	**pøːr̠**	'cliff'	**møːɽ**	'name of mund' (4-6)
kar̪	'border of cloth'	**kar̠**	'juice/sap'	**kaɽ**	'pen for calves'

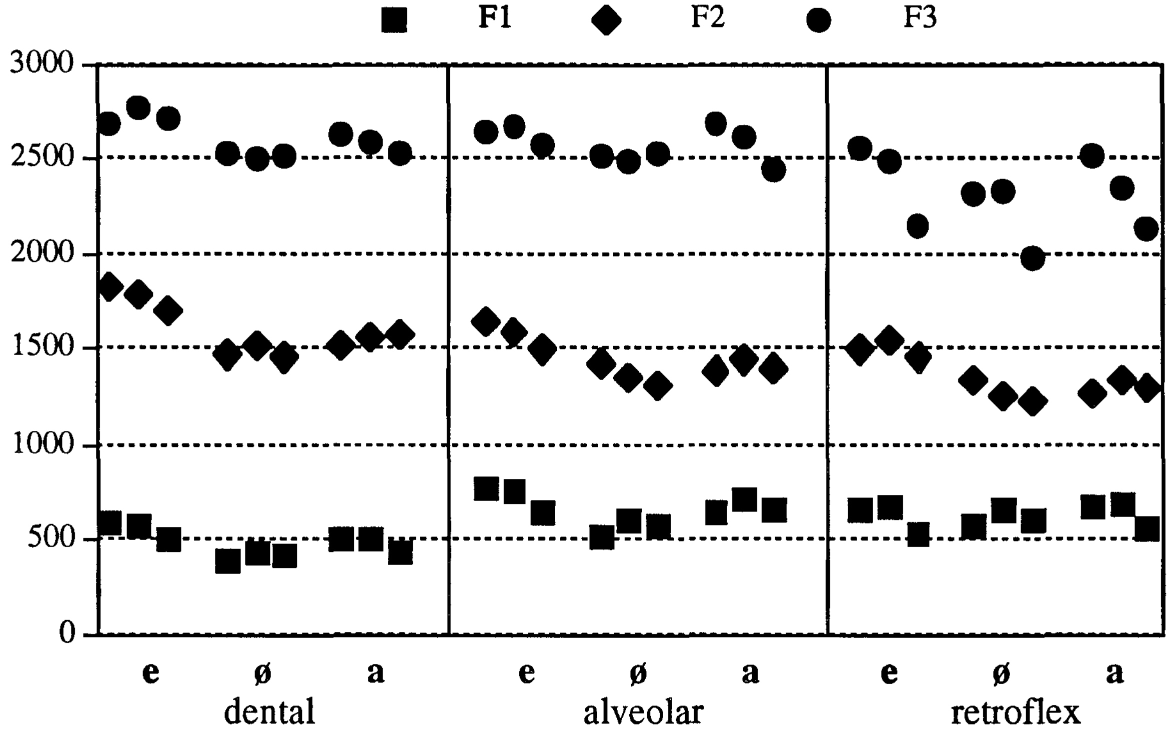


Figure 14. Mean formant frequencies in **e**, **ø**, **a** preceding three plain rhotics

Figure 14 illustrates the fact that the dental affects other vowels in the same way that it affects **a**. The F2 – F1 in a given vowel is greater when that vowel precedes a dental rhotic than when that vowel precedes one of the other two rhotics. That this occurs across vowel quality illustrates that it is due to the dental rhotic itself. Again taking the **a** vowel as the representative of the general pattern, Table 8 illustrates the results of three separate analyses of variance with place of articulation as the independent variable comparing the F2 – F1 of this vowel preceding the dental, alveolar, and retroflex, at the three different measurement positions.

Table 8. P-Values of three separate analyses of variance each comparing F2 – Fl at onset, middle and offset positions of the **a** vowel preceding the three plain rhotics. Comparisons which are found to be significant using Fisher's PLSD post hoc test are marked with S.

	a ONSET	**a** MIDDLE	**a** OFFSET
ALVEOLAR VS. DENTAL	.0119 S	.0003 S	<.0001S
ALVEOLAR VS. RETROFLEX	.1410	.1605	.9498
DENTAL VS. RETROFLEX	.0005 S	<.0001 S	<.0001 S

The dental significantly affects the formant structure of the preceding vowel even at the onset. These results are from a short vowel **a** which in Toda is found to have a mean length of 78 ms when preceding voiceless consonants and 107 ms before voiced ones (Shalev et al 1993). It is more striking to see these effects of the dental as far forward as the onset of a preceding long vowel. Long vowels in Toda are from 1.5 to 2 times as long as the short counterparts (Shalev et al 1993). We compared the mean formant values of the **øː** vowel at the onset, middle and offset in the words **pøːr̪** ' name' and **pøːr̠** 'cliff' and **møːɽ** with measurements on the first two words averaged over six speakers and measurements for the third over three speakers. Here we took a one factor analysis of variance for each measurement position and summarize the results of these three analyses of variance in Table 9. As may be seen, F2 – Fl is very significantly even throughout a long vowel preceding a dental rhotic.

Table 9. p-values for F2 – Fl in **øː** before three different rhotics. Comparisons which are found to be significant using Fisher's PLSD post hoc test are marked with S.

	øː ONSET	**øː** MIDDLE	**øː** OFFSET
ALVEOLAR VS. DENTAL	.0029 S	.00037 S	.0002S
ALVEOLAR VS. RETROFLEX	.1099	.0882	.1764
DENTAL VS. RETROFLEX	.0005 S	.0001 S	.0001 S

The other marked effect on the preceding vowel quality by a rhotic is that of the retroflex. In Figure 13 above, we see that the third formant becomes steadily lower as one goes from onset, to middle, to the offset of a vowel preceding a retroflex rhotic. This pattern holds for **e** and **a** but not for **ø**. The **ø** has a middle F3 which is not lower than the onset. This difference between vowels is most probably due to the length difference between them. In this data set the **e** and **a** vowel qualities represent short vowels, but the **ø** vowel is long. Comparing the levels of F3 at onset for a given vowel across different rhotics, we can see that it is slightly lower when the vowel precedes a retroflex rhotic than one of the other two (again the lowering is enhanced when both lip rounding and retroflexion are combined). This suggests that the effects of the retroflex rhotic can also be seen as early on as the onset of the preceding vowel.

That F3 was lower for the **e** and **a** vowels preceding the retroflex clarifies the issue raised above concerning the lesser degree of F3 lowering in the retroflex rhotic following these two vowel contexts. At the vowel offset position, F3 in all three vowels (**e**, **ø** and **a**)was found to be significantly different according to the place of the following rhotic. For **e** this was a five level analysis of variance with data set **eːr̠**, **er̠ʲ**, **eːr̪**, **meɽ**, for **øː** a three level analysis of variance with same data set as used in Table 8, and for **a** the same data set as in Table 8. The results of the these three different tests are shown in Table 10 below.

Table 10. Summary of p-values of F3 for three separate analyses of variance at the offset position for each vowel before the three plain rhotics.

	e ONSET	**e** MIDDLE	**e** OFFSET
ALVEOLAR VS. DENTAL	.3853	.9173	.5337
ALVEOLAR VS. RETROFLEX	.0015 S	.0037	.0023
DENTAL VS. RETROFLEX	.0010 S	.0043 S	.0076 S

Thus we can conclude that the retroflex rhotic does in fact always cause a lower F3, but for the **e** and **a** vowel contexts this lowering happens primarily in the space occupied by the vowel and not by the rhotic. It might be argued that the position just before the rhotic's first contact is actually the beginning of the rhotic. But regardless of how one decides this issue the fact remains that for the **e** and **a** vowel contexts the lowering of the F3 in the retroflex rhotic primarily occurs prior to the first contact and is only weakly present in the formant structure after that contact. Secondly the fact that the lowering of F3 for these two vowels and their associated retroflexes is not as dramatic as in other vowels is explained by the lack of lip rounding. Finally we should note that the general picture of the retroflex rhotic from the acoustic data supports our notion that this rhotic may begin with a retroflex gesture, but the actual trill is in a similar place to that of the alveolar and dental rhotics.

6.0 Palatalized Rhotics

In addition to the three plain rhotics described in detail in the above sections, Toda also has a set of palatalized rhotics corresponding to the three plain ones. Spectrograms of all six are shown in Figure 15. In the literature on Dravidian these rhotics are transcribed as being followed by **y**. We have opted to denote the palatalization with the IPA raised **ʲ** as is the case in Table 2 in section 2.2.

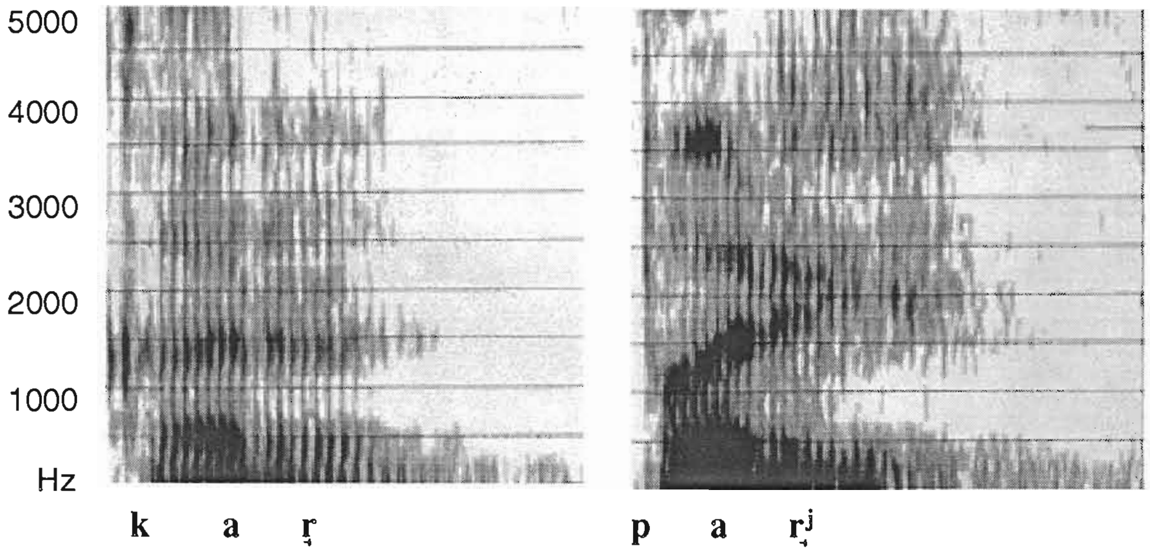


Figure 15a. Spectrograms of two of the six Toda trills produced by speaker 4. The plain form is on the left and the palatalized form is on the right.

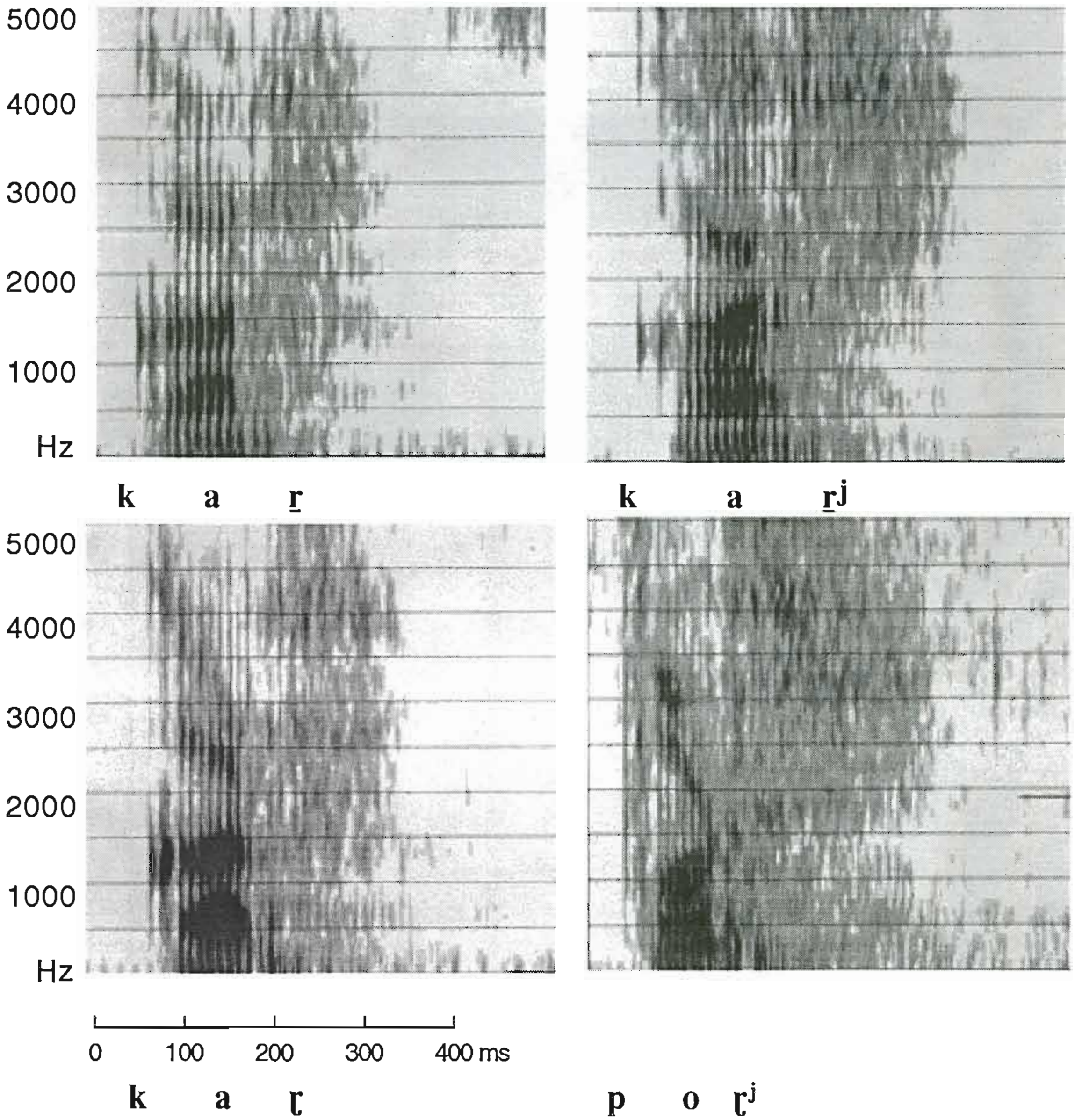


Figure 15b. Spectrograms of the remaining four of the six Toda trills produced by speaker 4. Plain forms are on the left and the palatalized forms are on the right.

The pattern illustrated by these spectrograms is that the palatalized rhotics have the same basic effects on the vowel formant structure as their plain counterparts. Thus both a plain and palatalized dental rhotic will affect the preceding vowel in the same way, increasing F2 – F1. The additional coloring of formant structure that the palatalized segments introduces into the picture is a steadily rising F2 from the onset of the preceding vowel through to the j̊ release of the rhotic. This same kind of effect is seen in the F2 of the palatalized alveolar and retroflex as well as across vowel context. Table 11 illustrates the mean values of the formants and F2 – F1 at each of the six positions measured, three in the **a** vowel and three in the following rhotic.

Table 11. Mean formant values for **a** vowel and following rhotic in words **kar̠**, **kar̪**, **kar̠ʲ**, **par̪ʲ**, **kaɽ**.

			V	**V**	**V**	**R**	**R**	**R**
			onset	**mid**	**offset**	**onset**	**mid**	**offset**
F2 – F1	**kar̠**	alv	**751**	733	739	703	759	713
	kar̪	den	1016	1032	1138	1138	1146	1169
	kar̠ʲ	palv	937	955	1050	1040	1274	1738
	par̪ʲ	pden	705	1091	1336	1358	1581	1705
	kaɽ	ret	642	654	737	758	758	711
F1	**kar̠**	alv	**642**	710	660	602	576	640
	kar̪	den	**506**	526	434	398	371	402
	kar̠ʲ	palv	**652**	719	645	700	589	455
	par̪ʲ	pden	519	547	426	414	369	277
	kaɽ	ret	656	692	559	536	580	597
F2	**kar̠**	alv	1394	1443	1400	1306	1335	1353
	kar̪	den	1523	1558	1572	1536	1518	1571
	kar̠ʲ	palv	1589	1674	1695	1740	1863	2193
	par̪ʲ	pden	1224	1638	1762	1772	1950	1982
	kaɽ	ret	1299	1346	1295	1294	1339	1308
F3	**kar̠**	alv	2681	2610	2451	2485	2523	2552
	kar̪	den	2636	2572	2531	2582	2557	2640
	kar̠ʲ	palv	2500	2448	2457	2461	2423	2634
	par̪ʲ	pden	2638	2574	2536	2458	2553	2569
	kaɽ	ret	2522	2346	2071	2249	2300	2384

The plain dental and the palatalized dental rhotics both cause an increase in the F2 – F1, but the two effects have very different characteristics. Table 11 illustrates that even though the dental and the palatalized rhotics cause an increase in F2 – F1 they achieve this in different ways. The plain dental has a greater F2 – F1 than the other two plain rhotics, which is achieved by causing both the F1 and F2 to shift away from each other. This effect is stable from the onset of the vowel to the offset of the rhotic. The F2 – F1 stays essentially the same from the vowel onset to the rhotic offset, increasing only by 160 Hz. The increase achieved by the end is divided between the two formants, F1 dropping about 100 Hz and F2 rising by about 60 Hz in comparison to the initial position of the vowel.

The palatalized dental on the other hand has a very different pattern. The onset measurements cannot be compared due to the different initial segment, but from the vowel middle to the end we see that the F2 – F1 of the palatalized dental steadily increases from 1091 at the vowel middle to 1705 at the rhotic offset. The total increase in F2 – F1 is 1000 Hz, 785 Hz contributed by the raised F2 and the 215 Hz by the drop in the F1. F1 for both is essentially the same except for the very last position measured where F1 in the palatalized dental rhotic drops further while that in the plain dental rhotic does not.

To tease out the effects shown by palatalization alone we can compare the plain and palatalized alveolars. The F1 is essentially the same for both and is not found to be significantly different at the first 5 positions, the respective p-values are: .8432, .8390, .6551, .0719, .8658. It is only at the last position measured, the rhotic offset, that F1 is found to be significantly different for the plain and palatalized alveolars (p=.0084S). In other words, the effects of palatalization on the F1 are mainly seen in the last part of the rhotic.

The F2 at the vowel onset is higher in the palatalized alveolar, **kar̠ʲ**, than in the plain one, **kar̠**, and it rises steadily to the end of the rhotic. In this minimal pair we can see that even the onset position is affected by palatalization, i.e. an increase in average F2 by 205 Hz. It is only at the end of the vowel that there is a significant difference in the F2 of the plain and palatalized alveolars (p=.0087), but the F2 – F1 is significantly different for these two as early as the vowel middle (p=.0326). Unfortunately we do not have any matched long vowels contexts to determine how far forward the one can see the effects of palatalization.

At one time we had suspected that the difference between the plain dental and the plain alveolar rhotics might be ascribed to some form of palatalization, marked by the increase in F2 throughout the vowel before the plain dental. We now see that each of the rhotics has a genuinely palatalized form. There are two distinct kinds of patterns. One pattern, found in plain dental rhotics, is an increased F2 – F1 which is caused by spreading F1 and F2 equally and sustaining that spread throughout the production of the preceding vowel and through to the rhotic. The other pattern, palatalization, mainly causes an increase in F2 which steadily rises from the onset of the preceding vowel through to the offset of the rhotic. Because of this difference we have opted to view the plain dental, which, as we have seen, actually has an alveolar articulation, as having some sort of tongue body or root fronting resulting in the steady F2 – F1 difference. In contrast, the palatalized forms are associated with a more sharply raised F2 which can be ascribed to a raising of the front or blade of the tongue.

7.0 Conclusion

The places of articulation of the Toda rhotics were not found to be post-dental, alveolar, and retroflex as previous accounts had suggested. Instead we found that the post-dental and alveolar rhotics are in fact both apical alveolars with the 'post-dental' having some secondary articulation. We were not able to precisely pinpoint what this secondary articulation is and conclude that it is some gesture which causes an increased oral cavity behind the point of articulation (most likely tongue body and/or root fronting). The retroflex was found to be produced with at least an initial gesture which was retroflexed but with subsequent contacts occurring at or near the alveolar ridge.

We also examined the number of tongue tip and roof of the mouth contacts involved in a given rhotic. Although there was considerable variation within each of them, a pattern emerged in that the dental was predominantly a single contact, the retroflex usually had multiple contacts, and the alveolar varied greatly, sometimes having a single contact and sometimes being a trill with several contacts.

The most consistent aspect of Toda rhotics was their distinct formant structures and the way in which they colored the preceding vowels. The alveolar is in the middle of this acoustic space and the each of the other two plain rhotics differing from it in one critical way, the dental having an increased F2 – F1 throughout the preceding vowel, and the retroflex being marked by a lowered F3 at the end of the preceding vowel.

The palatalized rhotics exhibit the same basic properties as the plain rhotics with the addition of an increased F2 which steadily rises from the onset of a preceding vowel through to the offset of the palatalized rhotic. There is a clear difference in the way in which a dental and a palatalized rhotic affect a preceding vowel and in their formant structure as well. The dental has an F2 – F1 which is significantly different from the F2 – F1 value in a plain alveolar and plain retroflex as early on as the onset of the preceding vowel. This difference is maintained at a stable level through to the end of the rhotic release, while the palatalized rhotic has an F2 – F1 value which steadily grows larger and larger from the onset of the preceding vowel through to the end of the palatalized rhotics itself.

Acknowledgements

This paper would not have been possible without the great cooperation we had from our Toda consultants, notably Vasamalli and K. Pothili, and Piljin Wiedman. We would also like to thank Murray Emeneau for his pioneering work, and for his continued interest and helpful comments on this research. In addition we would like to acknowledge the significant contributions Ian Maddieson has made to the analyses and the writing of this paper.

References

Burrow, T. and Emeneau, M.B. 1984. *A Dravidian Etymological Dictionary, Second Edition.* Oxford University Press, New York.

Catford, J.C. 1977. *Fundamental Problems in Phonetics..* Indiana University Press, Bloomington.

Condax, I.D. and Nathan, G.S. 1979. It sounds like some kind of ‘r’. Why?. Paper given at the 54th meeting of the Linguistic Society of America.

Dart, S. N. 1991. Articulatory and Acoustic Properties of Apical and Laminal Articulations. *UCLA Working Papers in Phonetics* 79.

Emeneau, M.B. 1984. *Toda Grammar and Texts*. American Philosophical Society, Philadelphia.

Fant, G. 1968. Analysis and synthesis of speech processes. In B. Malmberg (ed.) Manual of Phonetics: 171-277. North-Holland Publishing Co., Amsterdam

Ladefoged, P. 1975. *A Course in Phonetics -First Edition.* Harcourt Brace Jovanovich. New York.

Ladefoged, P. 1982. *A Course in Phonetics -Second Edition.* Harcourt Brace Jovanovich. New York.

Ladefoged, P. 1993. *A Course in Phonetics -Third Edition.* Harcourt Brace Jovanovich. New York.

Ladefoged P. , Cochran A., and Disner S. 1977. Laterals and Trills. *Journal of the International Phonetic Association* , Vol. 7, No. 2: 46-54.

Ladefoged, P. 1993. Linguistic Phonetic Fieldwork: a Practical Guide. *UCLA Working Papers in Phonetics* 84: 1-25.

Ladefoged, P. and Maddieson, I. (in press) *The Sounds of the World's Languages.* Blackwell: Oxford.

Lindau, M. 1980. Phonetic Differences in Nigerian Languages. *UCLA Working Papers in Phonetics* 51: 105-113.

Lindau, M. 1980. The story of /r/. *UCLA Working Papers in Phonetics* 51: 114-119.

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

McGowan, R.S 1992. Tongue-tip trills and vocal-tract wall compliance. *Journal of the Acoustical Society of America* Vol. 91, No. 5: 2903-2910.

Sakthivel, S. 1976. *Phonology of Toda with Vocabulary.* Annamalai University, Annamalainagar.

Shalev M., Ladefoged P., and Bhaskararao P. 1993. The phonetic structures of Toda.. *UCLA Working Papers in Phonetics* 84: 89-113.

Stevens, K. & Keyser, S. 1989. Primary features and their enhancement in consonants. *Language* Vol. 65, No. 1: 81-106.

Subrahmanyam, P.S. 1983. *Dravidian Comparative Phonology.* Annamalai University, Annamalainagar.

Zvelebil, K. 1970. *Comparative Dravidian Phonology.* Mouton, The Hague.

A phonetic study of Tsou

Richard Wright and Peter Ladefoged

Introduction

Tsou is an Austronesian language spoken by approximately 3000 people on Mount Ali in Central Taiwan. Historically it was made up of four dialects, Tapangu, Tfuea, Duhtu and Iimcu, the last of which is extinct (Tung, 1964). In the three remaining dialects there is no significant grammatical and only marginal phonological variation. Aspects of Tsou have been described in various forms in several previous studies: Nevskij (1935), Ogawa & Asai (1935), Lin (1955), Tung (1964), Starosta (1974), Tsuchida (1972, 1976), Ho (1976) and Li (1979). Of these Tung's descriptive study is the most comprehensive, but those by Li (1979) and Tsuchida (1976) provide important additions and corrections to Tung's description of the language.

The present study is based on recordings made in 1993 of 13 speakers, 8 males and 5 females, of the Tfuea, **'tfuja**, dialect as spoken in Punguu village. It reports on the major phonetic characteristics of the Tsou language as determined by both acoustic and auditory analysis, and as such represents the first instrumental study of the phonetic structures of Tsou.

Phonological overview

Tsou has six vowels: **i**, **e**, **i**, **u**, **o**, **a**. According to Tung (1964), there is some dialectal variation associated with **i**, in that Tapangu **i** corresponds to **z** or **iz** in Tfuea and Duhtu in certain environments. Another variation, Duhtu **r** corresponding to Tapangu and Tfuea **j** probably no longer exists, since Li (1979) could find only one Duhtu speaker, over 70, who still preserved it.

Tung describes long vowels as sequences of like vowels as most of the apparent long vowels occur as a result of affixation, reduplication or vowel lengthening. He also describes long vowels as heterosyllabic and cites evidence from stress assignment and reduplication. In light of modern theories of metrical stress (eg Hayes, in press) and morphophonology (McCarthy and Prince, ms), the reduplication and stress assignment in Tsou can be seen as weight dependent processes that count morae rather than syllables. One reason for viewing them this way is that there is no hiatus between ajacent homorganic vowels. The very processes that Tung cites as evidence for treating long vowels as heterosyllabic can also be seen as evidence in favor of long vowels being treated as tautosyllabic. The observation that there is no phonemic length contrast, however, still holds. Stress and reduplication are discussed further below.

Tsou was described by Tung (1964) as having two 'non-syllabic' vowels that are the counterparts of **e** and **o**. However, Ho (1976) and Li (1979) describe the 'non-syllabic' vowels as the glides **j** and **w**, which seems to us a sensible proposal from a phonological point of view, especially given the stress pattern and the process of reduplication seen in the language. It should be noted, however, that **w** has a limited distribution and that in the present study, there is no systematic data available for its analysis. Tung's analysis is reflected in the orthography which, along with a phonetic transcription, will be used in this paper. Thus the orthographic version of the name of the dialect is 'Tfuea' which in our transcription is **tfuja.**

In addition to the two glides, the Tfuea dialect has the following 15 consonant sounds: **ɓ**, **ɗ**, **p**, **t**, **k**, **ʔ**, **ts**, **f**, **s**, **h**, **v**, **z**, **n**, **m**, **ŋ**. Of these, both implosives, **ɓ** and **ɗ**, occur very rarely in the language. Tung (1964) notes that Duhtu has **r** where Tfuea has **j**; however even at the time of his study, Duhtu speakers freely substituted **j** for **r**. Tsuchida, and later Li, noted that only the very elderly speakers maintained the **r**. An additional consonant series, pulmonic ingressive fricatives, has been said to occur in the Tfuea dialect spoken in Punguu (Fuller 1990), but Ladefoged and Zeitoun (1993) note that all of the 14 Punguu speakers they investigated used an egressive airstream to produce the fricatives described by Fuller. One of the most striking features of Tsou Phonology is the combinatorial possibilities of the 15 consonant sounds. Sequences of more than two consonants are forbidden, but a large number of unusual two consonat clusters do occur. These clusters will be discussed below.

Vowels

An acoustic analysis was made of the vowel qualities in Tsou. For five of the six vowels the data consisted of two examples of each vowel in a stressed syllable after a bilabial stop, and two further examples in a stressed syllable after an alveolar stop. This was not possible for the sixth vowel, **ɨ**, for which only two examples from a stressed syllable following an alveolar stop were available in the data which we had recorded. Examples of the vowels are given in Table 1

Table 1: Words illustrating the short vowels of Tsou

VOWEL	ORTHOGRAPHY	IPA	GLOSS
i	*pitu*	**pitu**	'seven'
e	*pepe*	**pepe**	'sky, heaven'
a	*pa'ti*	**paʔti**	'to show'
o	*po'e*	**poʔe**	'to pray in making wine'
u	*tufku*	**tufku**	'to wash clothes'
ɨ	*tʉ'sʉ*	**tɨʔsɨ**	'arrow'

The word list used in making the recordings is given in the appendix. Each word was produced once by each speaker. While minimal sets were difficult to find, the data were selected to avoid consonant environments that would cause allophonic variation. Formant and fundamental frequency measurements were taken in stressed syllables.

Measurements of the frequencies of the first, second and third formants (F1, F2, and F3 respectively) were taken, using spectrographic displays and superimposed FFT and LPC spectra on the KAY CSL system. In order to maximize the accuracy of the spectral analysis, speech sampled at 10 kHz was used in the instrumental analysis of vowels (see Ladefoged, 1993). The LPC calculation used a 20 ms frame with pre-emphasis applied to the signal prior to calculation to make it easier to see the higher formants. In general, a 14th order filter was applied in analyzing the male speech and a 12th order filter was applied in analyzing the female speech. When necessary the filter order was adjusted down to accommodate the higher fundamental frequency of the speech in certain utterances for some speakers. To insure the accuracy of the LPC measurements, a simultaneous 512 point FFT calculation was made and the results were overlaid on the LPC results. These measures were also checked against measures taken from a spectrogram and against auditory impressions. The mean formant frequencies are shown in Table 2.

Table 2: Mean formant frequencies for male and female speakers

	female speakers			**male speakers**		
vowel	**F1**	**F2**	**F3**	**F1**	**F2**	**F3**
i	435	2599	3168	369	2157	2864
e	516	2328	2923	499	2032	2597
a	975	1581	2952	805	1385	2629
o	560	914	2963	539	983	2576
u	468	1118	3158	407	917	2464
ɨ	501	1756	3012	429	1441	2454

Formant frequencies are plotted in Figure 1 (males) and Figure 2 (females). In the figures F1, roughly equivalent to vowel height, is plotted against F2' – F1 which is roughly equivalent to the backness of the vowel. F2' is a weighted average of F1, F2 and F3 calculated with the formula given by Fant (1973: 52): $F2' = F2 + (F3 - F2)(F2 - F1) / 2(F3 - F1)$. The axes are scaled using the Bark scale and labeled in Hz. The ellipses around the points for each vowel have a radius of two standard deviations from the mean.

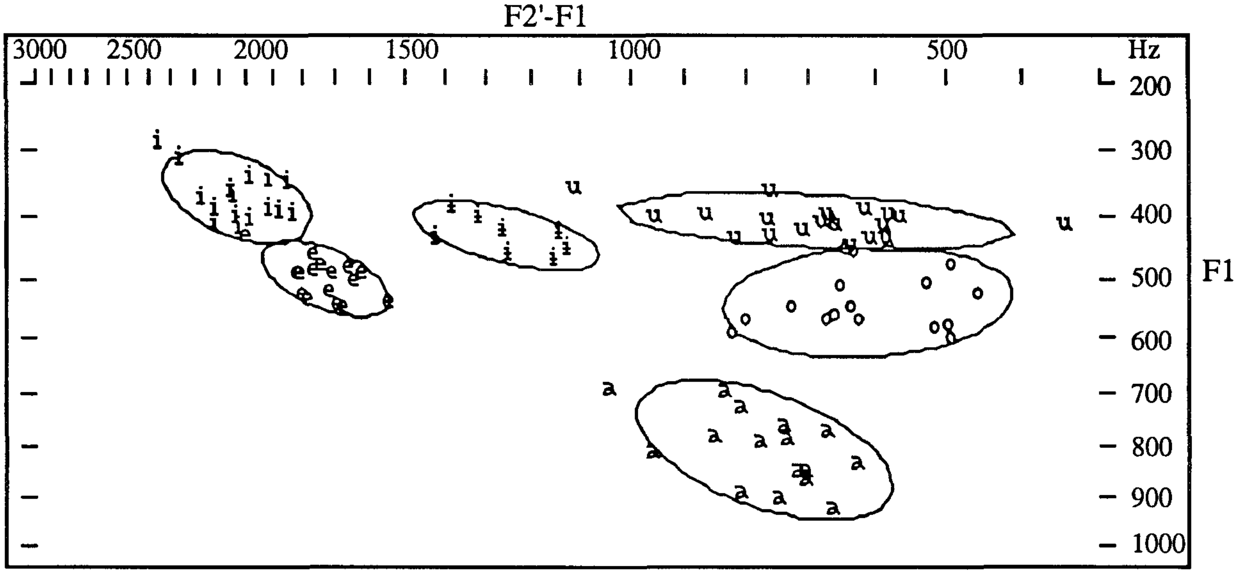


Figure 1: formant plot for male speakers F2' – F1 by F1

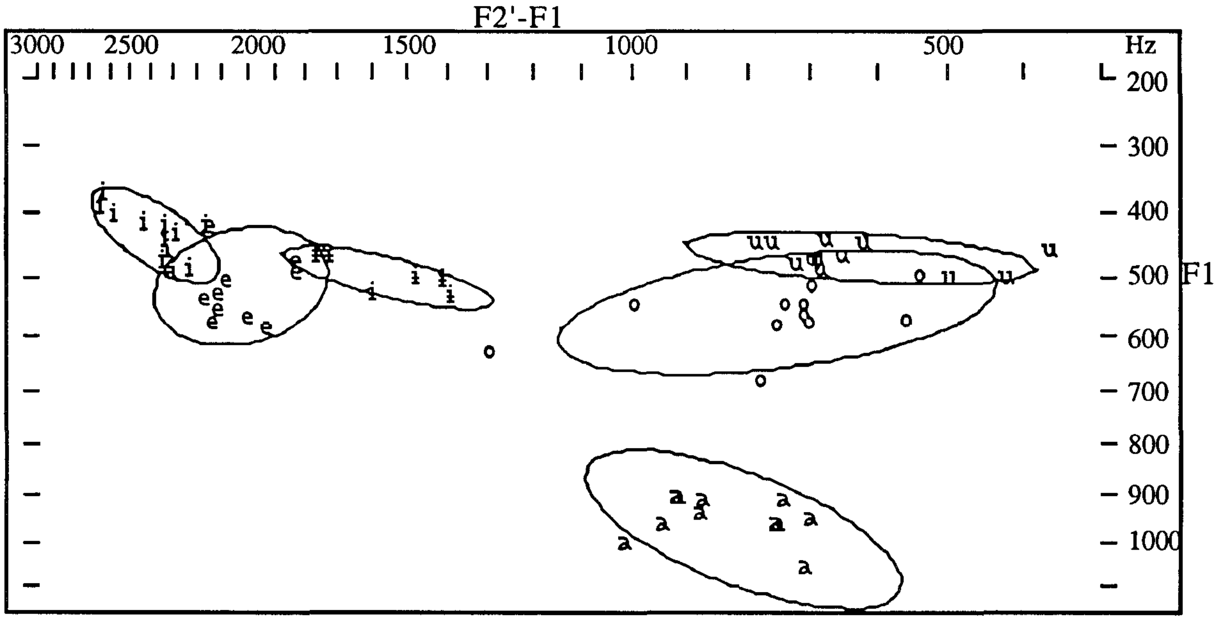


Figure 2: formant plot for female speakers F2' – F1

Based on acoustic data and auditory impressions the following description of Tsou vowels can be made:

i	high front unrounded
e	mid-high front unrounded

ɨ	high central unrounded
u	high back rounded
o	mid back rounded
a	low back unrounded

The vowel **e** has been described by Tsuchida as 'lower mid-front' corresponding to **ɛ**. It is plain from the formant plots that the mid front **e** in Tsou is higher than what is traditionally transcribed as **ɛ**, and is much closer to **i** than to **a.** Tung describes **ɨ** as high central unrounded, but chooses to transcribe it as **ʉ** because his impression is that it is closer to **u** than to **i** and is slightly rounded. It seems to us that it is a high central vowel as described by Li and Tsuchida, and is more accurately transcribed as **ɨ**, in accord with Fuller (1990).

The most notable aspect of the vowel space in Tsou is the centralizing tendency, particularly in the female speakers. The centralization was greatest when the vowel was flanked by voiceless coronals and is most significant in the back vowels. There was a short transition into the vowel and a longer schwa-like transition out of the vowel and into the following consonant. This was particularly pronounced in productions of **o** by one of the female speakers, as is reflected in the wide scatter in the formant plot.

In addition to formant measures, inherent pitch measures were taken for each vowel. Fundamental frequency (F0) was measured from a narrow band spectrographic display that was linked to simultaneous wide band spectrographic, waveform and amplitude contour displays. F0 measurements were taken in a stressed syllable at the point that corresponded to the F0 peak for that syllable. The F0 means are plotted with standard deviations in Figure 3.

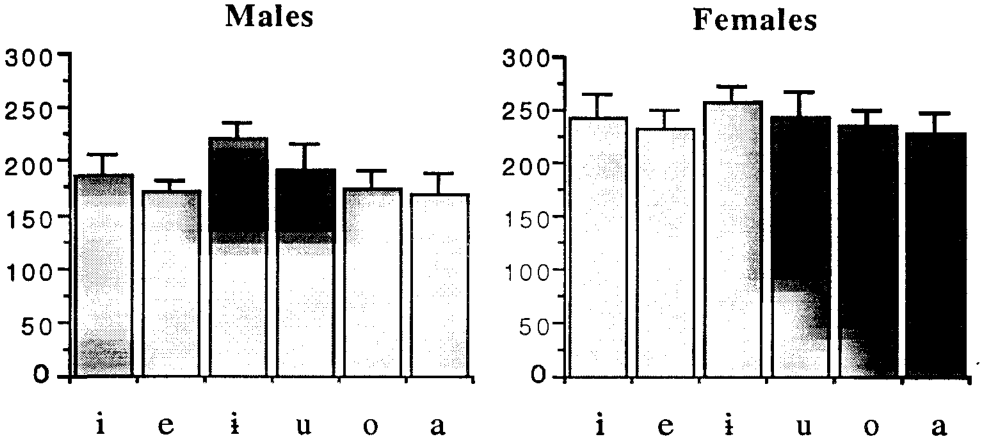


Figure 3. Mean and standard deviations for male and female F0 values for vowels

Surprisingly, markedly higher F0 values were found for the high central vowel than for the other high vowels despite the fact that on the formant plots this vowel is lower than either **i** or **u**. The males had larger differences in F0 between vowels than the females. Normalized F0 and F1 were statistically correlated to determine the reliability of the effect of vowel height on intrinsic pitch. The normalized values in the correlation were the individual deviation from the overall mean for F1 and F0. As is shown in Table 3, a highly significant but weak correlation was found between the deviation from mean of F1 and F0. This result is in agreement with the general finding that vowel height is correlated with fundamental frequency such that high vowels have an intrinsic pitch that is higher than low vowels.

Table 3. correlation of vowel height (F1) and F0

Correlation	P-Value	95% Lower	95% Upper
-.333	<.0001	-.465	-.187

Consonants

As was noted above, the Tfuea dialect of Tsou has the 17 consonants listed in Table 4 below. Both auditory and acoustic analysis were used in arriving at the following description of Tsou consonants. In the recordings there was one repetition of each consonant in onset position before both **i** and **a**, except for the glides **j**, **w**. Words containing the consonants were digitized at either 10 kHz (sonorants) or 20 kHz (obstruents) and analyzed using simultaneous waveform, amplitude and spectrographic displays on the KAY CSL. The wordlist used to record the consonant tokens is found in the appendix.

Table 4. Chart of Tsou consonants

	BILABIAL	LABIO-DENTAL	ALVEOLAR	PALATAL	VELAR	GLOTTAL
PLOSIVE	**p**		**t**		**k**	**ʔ**
IMPLOSIVE	**ɓ**		**ɗ**			
AFFRICATE			**ts**			
FRICATIVE		**f v**	**s z**			**h**
NASAL	**m**		**n**		**ŋ**	
APPROXIMANT	**w**			**j**		

Plosives:

Aspiration is not contrastive in Tsou. While voiceless stops can appear aspirated, Tung cites both distributional and morphophonemic evidence for treating these surface forms as clusters. The strongest distributional evidence is the fact that **h** clusters freely as either the first or the second member of a cluster but blocks further clustering. Tsou permits maximal clusters of two consonants; clusters of **h** and two or more consonants are not permitted. Consonant clusters are discussed below. In certain inflectional processes, the first member of a consonant-**h** cluster can change or in other processes be removed leaving behind the **h**, ex: **mhino** : **phini** 'to trade' (Tung 1964:11).

Description of the voiceless plosives is straightforward: **p** is a voiceless unaspirated bilabial, **t** is a voiceless unaspirated alveolar, **k** is a voiceless unaspirated velar. The status of the glottal stop as a consonant is supported by distributional evidence. Like the glottal fricative, **h**, it is found with the same distribution as the other consonants. Both of these consonants can occur in intervocalic position, in consonant clusters as the first or second member.

The voiced bilabial stop has been described as a preglottalized voiced bilabial plosive by Tung, Li, and Tsuchida, and alternatively as a voiced bilabial implosive by Fuller. Spectrographic analysis of **ɓ** in word onset shows no creakiness or glottalization during voiced closure or in the onset of the vowel. Auditory impressions and visual observations of the speaker's larynx indicate that the voiced bilabial is in fact implosive. Inspection of the waveform and spectrogram reveals that voicing during closure is characteristic of that seen in implosives (Lindau, 1984) in that voicing is strong and there is an amplitude increase in the latter part of the closure. An expanded view of the waveform showing voicing during **ɓ** closure is shown in Figure 4.

Figure 4. Expanded waveform showing increase in amplitude during **ɓ** closure voicing characteristic of implosives.

The voiced alveolar stop has been described variously; Tung describes it as a preglottalized alveolar lateral **l**, Tsuchida describes it as a voiced alveolar stop **d** and Fuller reports hearing it as an implosive **ɗ**. Auditory impressions, acoustic measures and observation of the speakers larynx indicate that it is implosive. Like **ɓ**, it shows an increase in amplitude during the closure voicing, as shown in Figure 5.

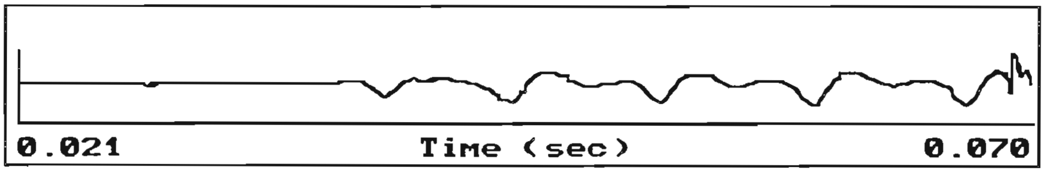


Figure 5. Expanded waveform showing increase in amplitude during **ɗ** closure voicing characteristic of implosives.

Both **ɓ** and**ɗ** may be produced with the preglotttalization that is expected in implosives due to the nature of their articulation. The preglottalization is not evident in word initial position, but in intervocalic position it is clear as can be seen in Figure 6.What is unusual is that one third of the speakers produce the voiced alveolar implosive with a lateral release and a lateral approximant allophone **l** before the low vowel **a**. The lateral approximant is shown in Figure 7 below. The lateral characteristics are probably the source of some of the disagreement seen in past descriptions.

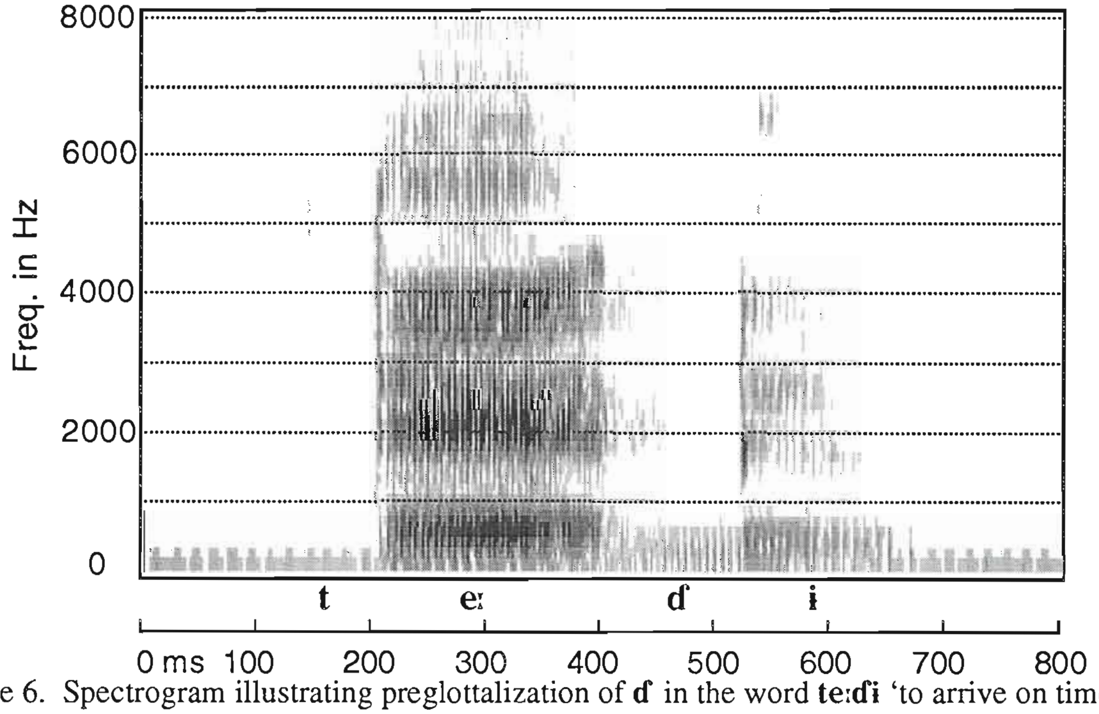


Figure 6. Spectrogram illustrating preglottalization of **ɗ** in the word **teːɗi** 'to arrive on time'

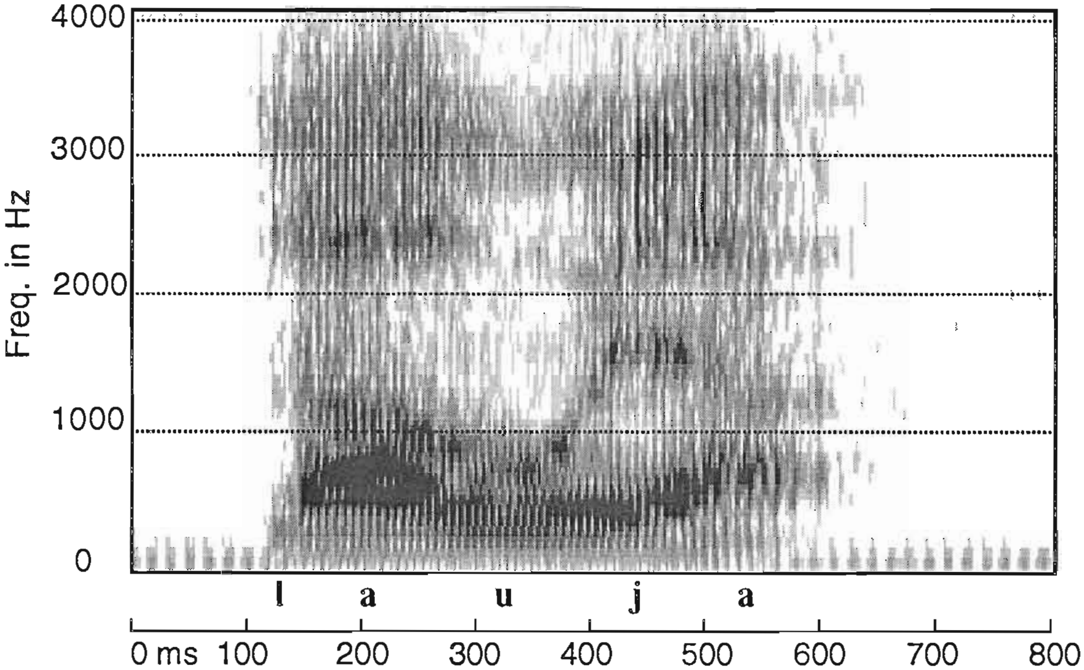


Figure 7. Example spectrogram showing the lateral approximate allophone of **ɗ** in the word **ɗauja [lauja]** 'maple'

In addition to the spectrographic analysis, voice onset time (VOT) measures were taken for the voiceless unaspirated plosives. The interval between the release burst and the first glottal pulse was measured on simultaneous waveform and spectrographic displays. The measurements were submitted to an analysis of variance with VOT as the dependent variable and consonant and speaker as independent variables. The results (F[2,21]=47.15, p<.0001) indicated that there was a significant difference in VOT between the consonants. The mean VOT following the three voiceless plosives is shown in Figure 8 below plotted with error bars. The results agree with the general observation that the further back a consonant closure is made, the longer the duration of the VOT. (Fischer-Jørgensen, 1954).

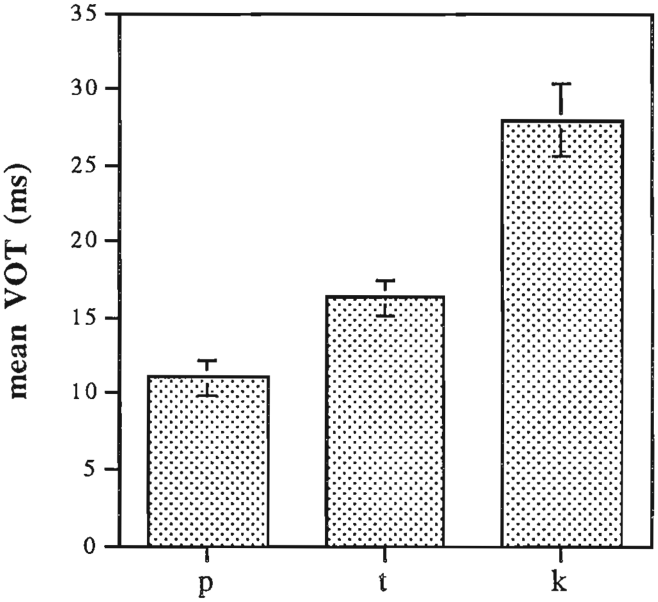


Figure 8. Mean VOT in ms following **p**, **t**, **k**

Affricates:

There is a single affricate in Tsou, the alveolar: **ts**, illustrated by a spectrogram in Figure 9. Its frication is characterized by a concentration of energy between approximately 5 kHz and 6 kHz It is palatalized before the front vowels **i** and **e**, being realized as the palato-alveolar affricate **tʃ**, with a wider concentration of energy that spreads between approximately 3.5 kHz and 5 kHz, illustrated by a spectrogram in Figure 10. **ts** is a single phonological unit rather than a cluster of **t** + **s**, as is shown by its combinatorial possibilities, to be discussed below.

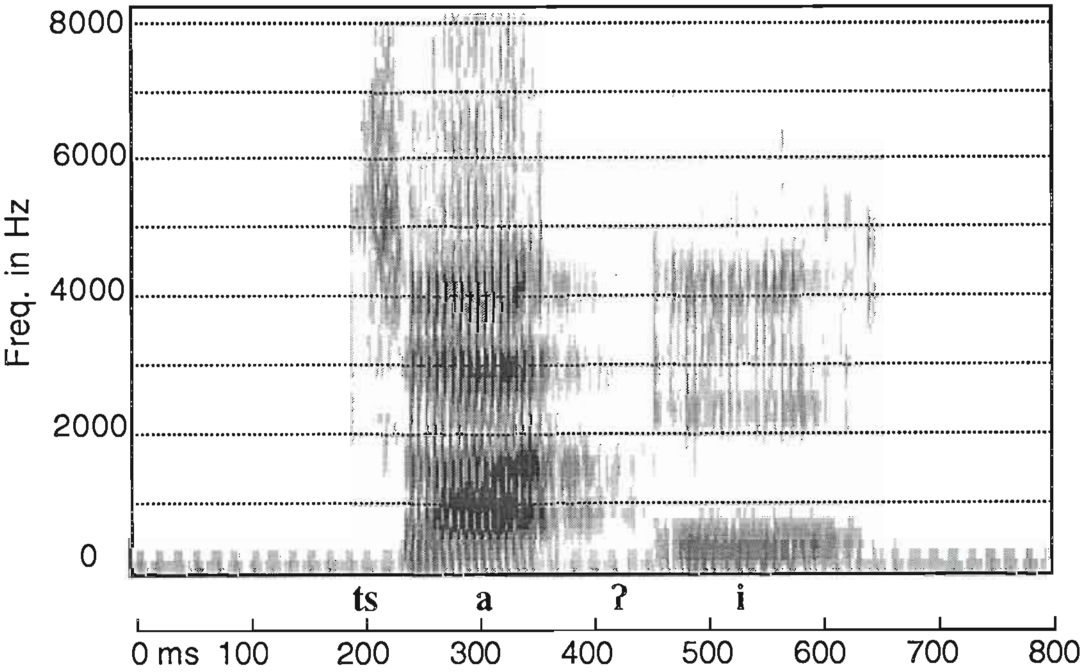


Figure 9. Spectrogram illustrating the alveolar affricate **ts** in the word **tsaʔi** 'dirty'

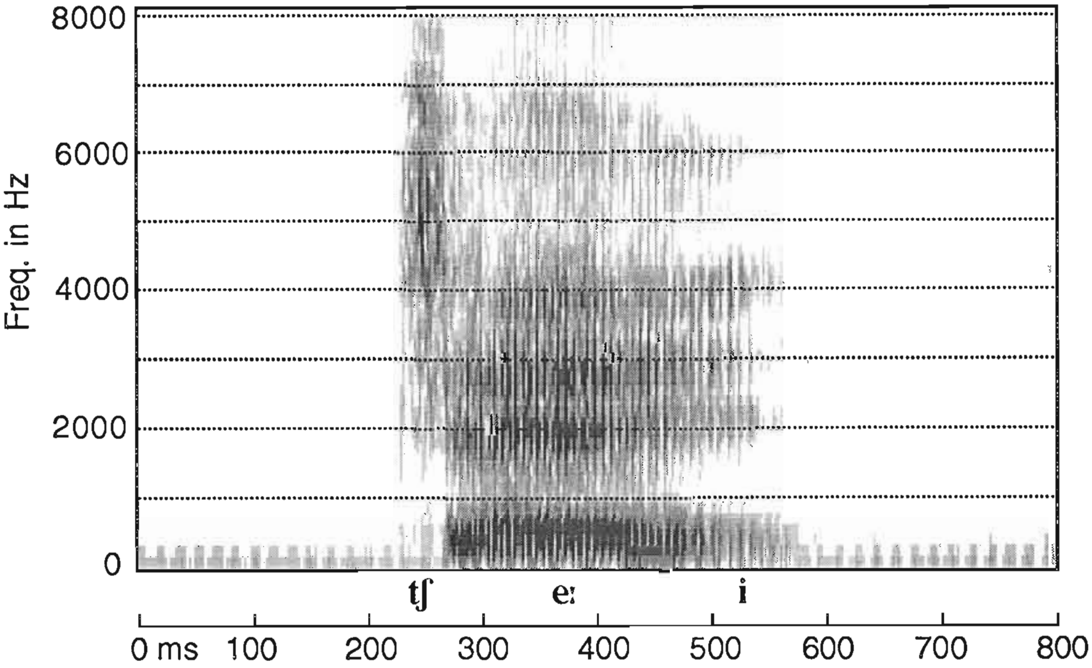


Figure 10. Spectrogram illustrating the palatalized allophone **tʃ** in the word **tʃej** 'dream'

Fricatives:

There are five fricatives in the Tfuea dialect of Tsou: voiceless and voiced labiodental **f, v** and alveolar **s, z**; and the glottal fricative **h**. The frication energy of the voiceless labiodental fricative **f** has low intensity and is widely spread over the spectrum as is typical in labiodentals, illustrated in Figure 11. Some speakers produce the frication with a more concentrated energy particularly in a cluster, such as **ft**, illustrated in Figure 12.

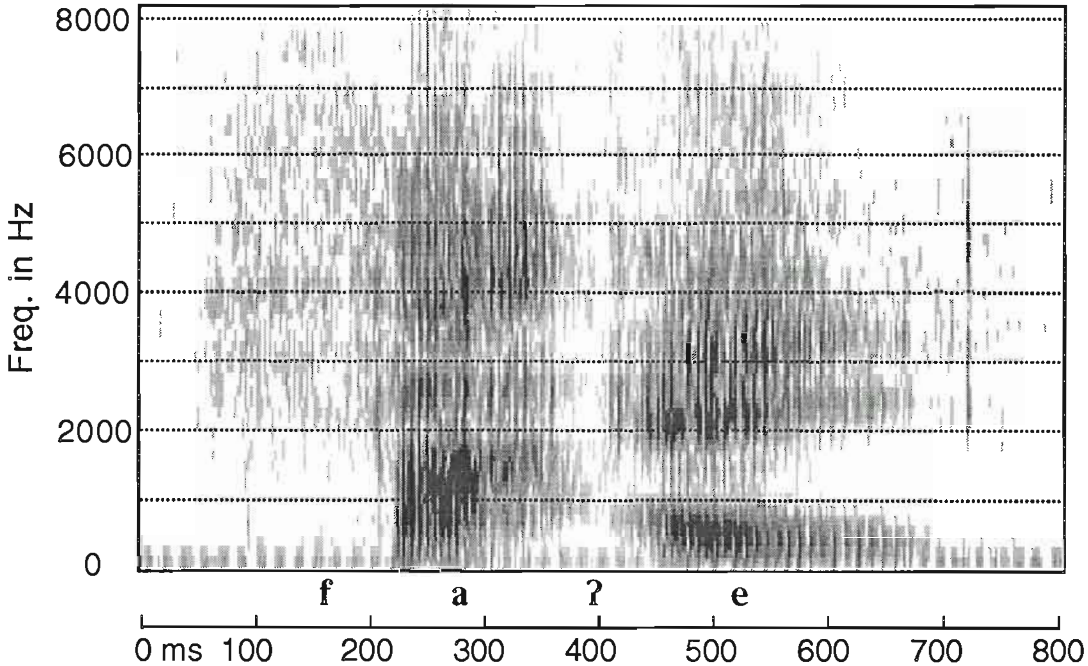


Figure 11. Example spectrogram: **f** with broad spectrum energy during frication in **faʔe** 'egret'

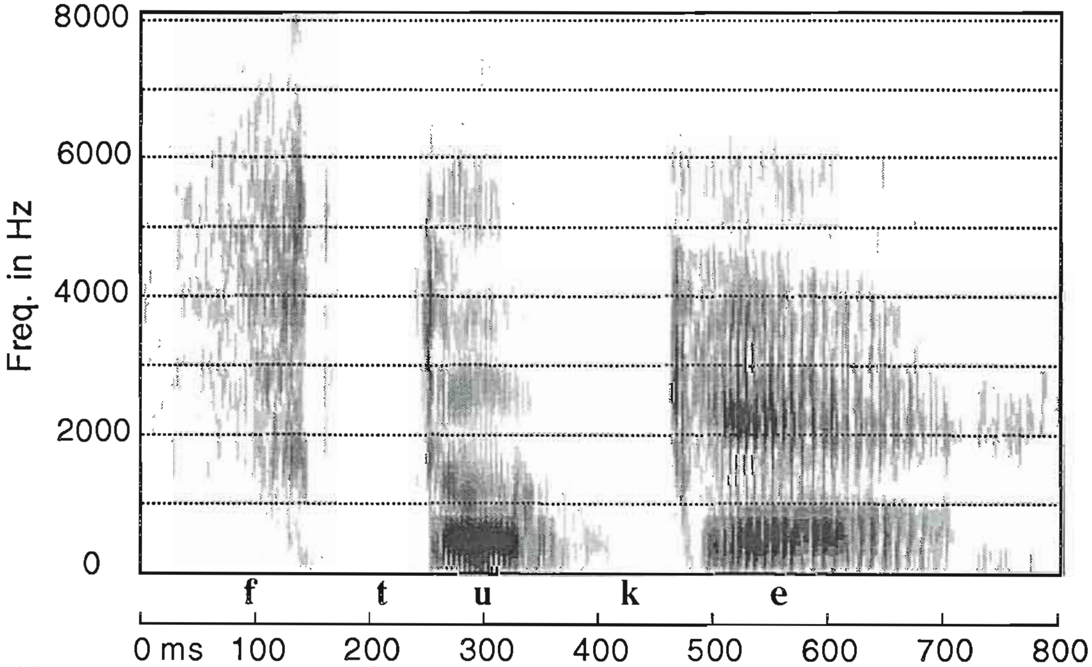


Figure 12. Example spectrogram: **f** in a cluster **ftuke** 'to arch the back'.

The voiceless alveolar fricative, s, is illustrated in Figure 13. The voiceless alveolar fricative is characterized by a high frequency concentration of energy. Like the alveolar affricate, the alveolar fricatives are heavily palatalized before front vowels. The palatalized form, ʃ, is illustrated in the spectrogram in Figure 14.

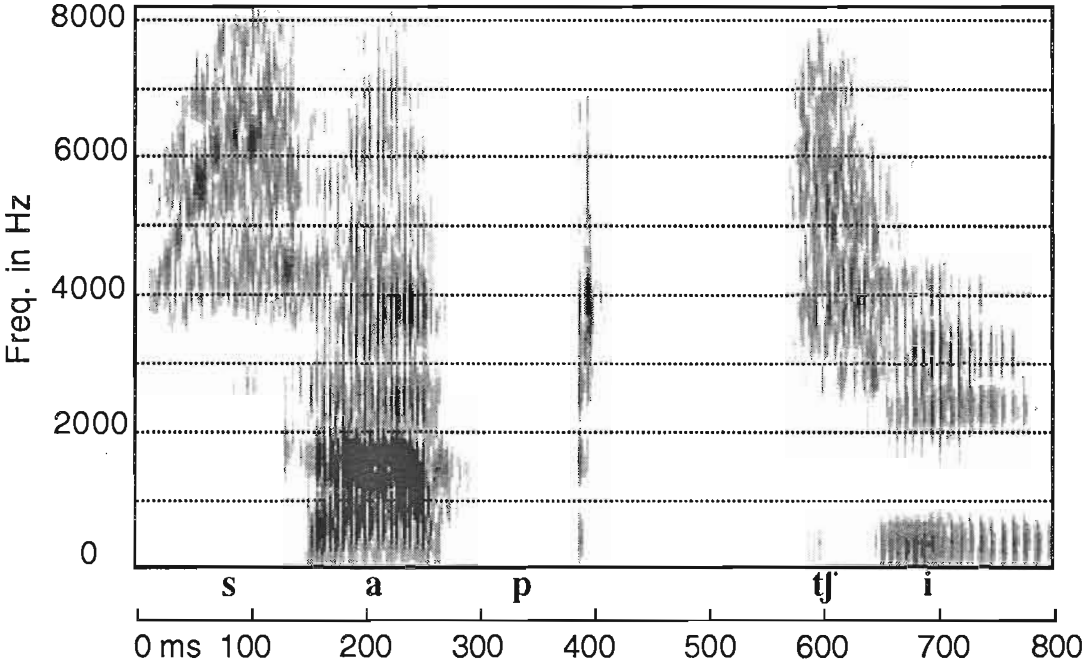


Figure 13. Example spectrogram: s in **saptʃi** 'face'

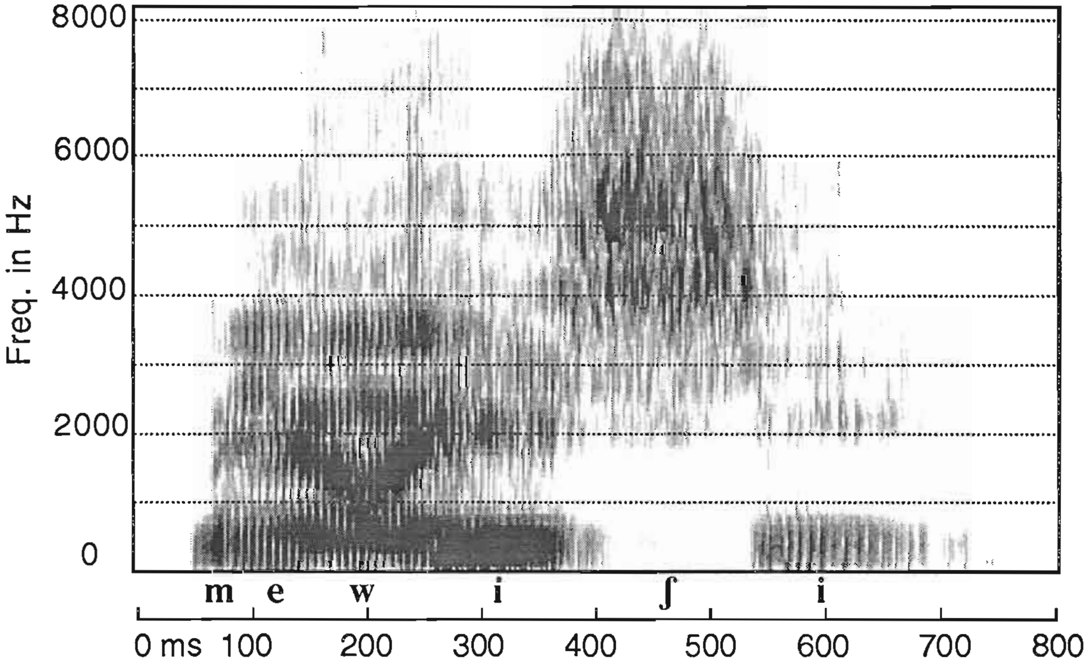


Figure 14. Example spectrogram: ʃ in **mewiʃi** 'large'

Tsou **h** has been described by Tsuchida as a velar fricative and by Tung as a glottal fricative. From a phonetic point of view the data in this study, illustrated in Figure 15, support Tung's description. As is common to **h**, the glottal fricative in Tsou often shows concentration of energy in the region of F2, and sometimes F3, of the following vowel, and takes on the qualities of its vocalic environment. Tsuchida's description may have been influenced by the fact that there is significant coarticulation between **h** and the high central unrounded vowel, **ɨ**. In this environment **h** is often realized with spectral qualities that are characteristic of a velar fricative, an example of which is shown in Figure 16. There is also phonological support for the view that h is a velar fricative, or at least not a glottal one. Tsou has a prohibition against two obstruents at the same place of articulation forming a cluster. But the clusters hʔ and ʔh do occur, as illustrated in Figure 15.

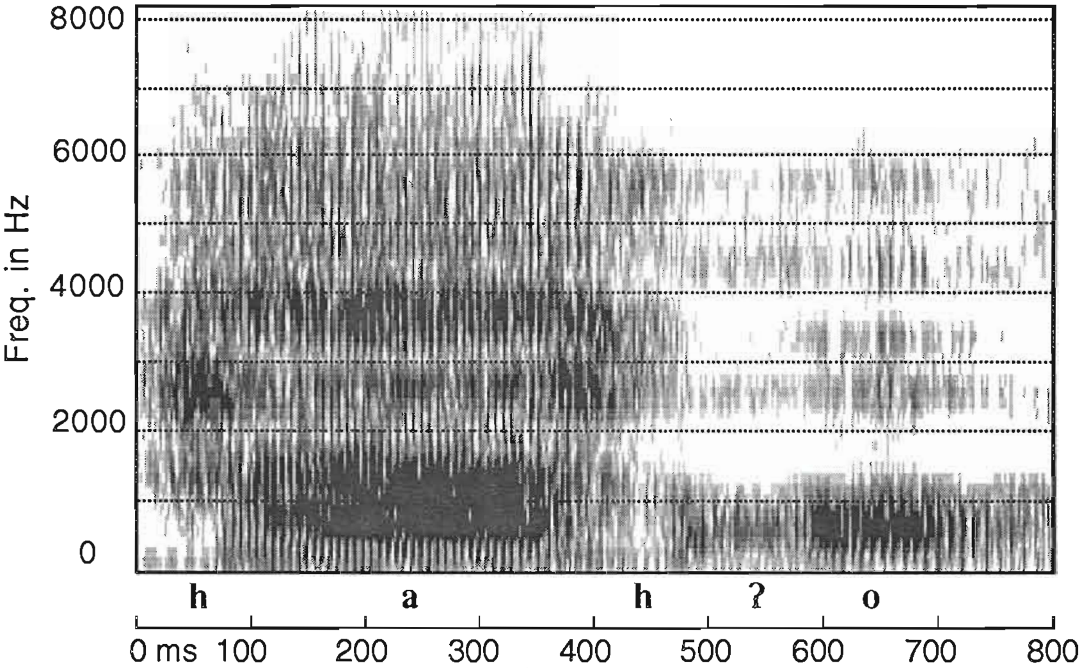


Figure 15. Example spectrogram: **h** in **hahʔo** 'friends'

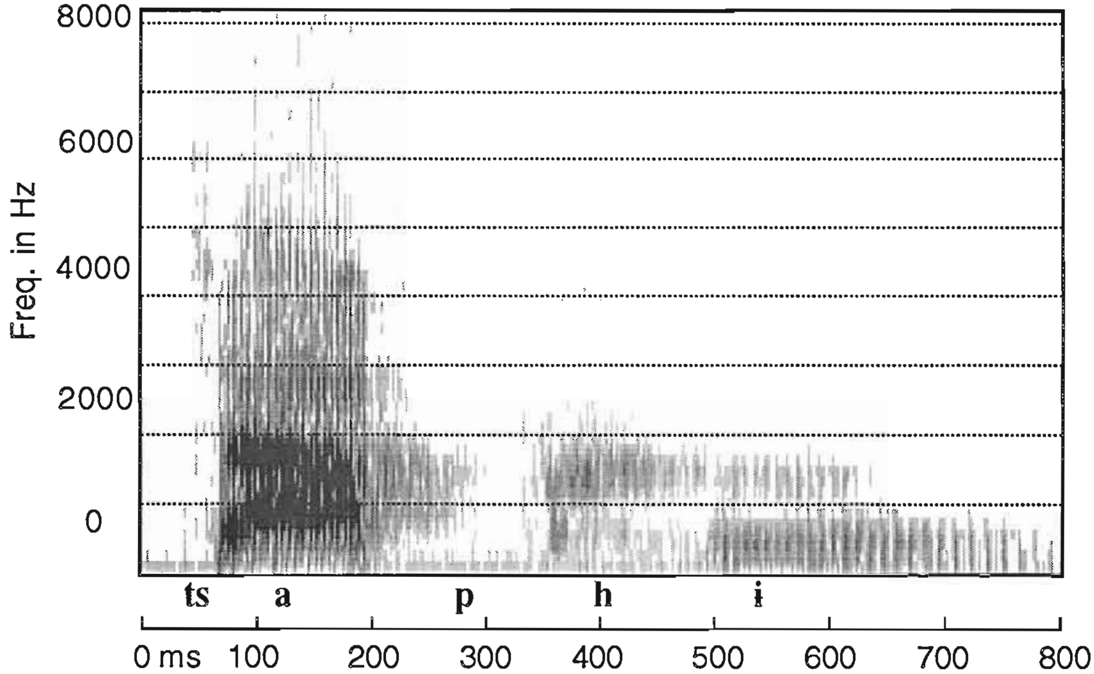


Figure 16. Example spectrogram: **h** in **tsaphɨ** 'palm, sole'

Nasals:

The nasal consonants have three places of articulation paralleling those of the voiceless unaspirated plosives: bilabial **m**, alveolar **n**, and velar **ŋ**. The most remarkable thing about the nasals is that, with the possible exception of **h** and **ʔ**, if these are both considered to be glottal consonants, they are the only segments that are allowed to form homorganic clusters with a following obstruent.

Approximants:

Tung describes two 'consonantal vowels' which he transcribes **ĕ** and **ŏ** in his discussion of vowels but which he fails to differentiate from **e** and **o** elsewhere in his text. Ho and Li have suggested treating the consonantal vowels as **j** and **w**. The rationale is that phonological processes such as stress assignment and reduplication are systematic only if the segments in question are consonantal. For example, certain inflectional processes involve reduplication of the onset and first vowel, or first mora, of the base. The following examples of reduplication and stress assignment are taken from Tung, using his transcription as the orthography.

ex:	*cofkoea*	**tsofkoja**	'clean'
	cocofkoea	**tsotsofkoja**	'very clean'

When **ĕ** is found at the onset of a word, it is reduplicated as if it were a consonant.

ex:	*eoskʉ*	**joskɨ**	'fish'
	eoeoskʉ	**jojoskɨ**	'fish' pl.

When **ĕ** is found as part of a syllable preceded by a consonant and followed by a vowel, it is treated as part of a consonant onset cluster and it is reduplicated together with the preceding consonant and first vowel of the base.

ex:	*beahci*	**bjahtʃi**	'fruit'
	beabeahci	**bjabjahtʃi**	'fruits'

If **ĕ** is treated as a syllabic vowel, the rules for reduplication become very complicated indeed. If, on the other hand it is treated as a glide, rules for reduplication are straightforward: reduplicate the onset and first vowel of the base. Stated in terms of prosodic morphology, this is a process of light syllable reduplication. The syllabic front mid-vowel **e** may itself form a syllable nucleus as is evident in forms such as **pepe** ‘sky’. Furthermore, **e** is reduplicated in words in which it makes up the nucleus of the first syllable:

ex:	*evi*	**e̲vi**	‘tree’
	eevi	**e̲e̲vi**	‘trees’

This type of evidence from reduplication is not available for **ŏ**, but evidence from stress assignment does point to treating both **ŏ** and **ĕ** as glides, distinct from the front and back mid vowels. In Tsou, stress generally falls on the penultimate syllable.

ex:

óko	**'oko**	‘child’
okósi	**ok'osi**	‘his child’
mʔúmʔu	**'mʔumʔu**	‘hair’
eamʔumʔúa	**eamʔu'mʔua**	‘hair’
úa	**'ua**	‘deer’
kaíu	**ka'iu**	‘a large jar’
iunióu	**iuni'ou**	‘to burn’
ciuéi	**tʃiu'ei**	‘to pierce the ears’

The first two pairs of examples show that stress is mobile and occurs on the penultimate syllable; when a suffix is added to the word, the stress moves to the syllable that was final in the non-suffixed form. The fact that stress is mobile in Tsou is evidence in favor of an analysis, such as that by Tung, Li and Ho, in which Tsou stress is quantity sensitive and not assigned underlyingly. The last four examples above show that in a series of vowels, the stress falls on the second to last vowel. As can be seen both **o** and **e** are treated as syllabic as far as these forms are concerned since both receive stress. In other forms, however, what have been transcribed by Tung as **o** and **e** are treated as non-syllabic.

ex:

kóeu	**'koju**	‘ear’
koeúsi	**ko'jusi**	‘his ear’
aúeu	**a'uju**	‘earlier’
aueúhe	**au'juhe**	‘much earlier’
híoa	**'hiwa**	‘work’
eahioáa	**eahi'waː**	‘to have work’
fúeŋu	**'fujŋu**	‘hill’
epúeu	**e'puju**	‘to fall’
céoa	**'tʃewa**	‘earth’
eóu	**'jou**	‘big fish’
éou	**'ewu**	‘thief’

Taken together, the examples above show that certain mid-vowels are treated as syllabic, they receive stress and participate in morphophonemic processes such as reduplication, while at the same time, what have been transcribed as mid-vowels are treated as if they are non-syllabic as far as processes of stress and reduplication are concerned. These facts are what led Tung to conclude that Tsou had both syllabic and non-syllabic mid vowels and led Li and Ho to propose the glides **j** and **w**. Tung's proposal is motivated by two factors. The first is that he finds it more parsimonious not to add glides to the phonemic inventory. The second reason is that, at least in the case of **j**, in intervocalic position and preceding or following a consonant the auditory impression created is that the glide has the qualities of **e**. In word initial and word final position, however, the auditory impression of both glides is similar to that of high vowels. From a phonological point of view, a glide is essentially a non syllabic vowel, and therefore in some sense Tung, Li and Ho are in agreement. The problem arises in Tung's transcription in which he fails to explicitly mark the difference between a syllabic and a non-syllabic vowel.

The wordlist used for making the recordings of consonants did not include examples of **j** and **w** in positions that are equivalent to the other consonants. However, several of the words in the wordlists do illustrate **j** word internally and in word initial position. Therefore, a preliminary acoustic study of **j** can be made. By chance **w** did not appear in word initial position with the exception of a single name. Therefore, it is not possible to perform a comparable analysis of **w.** However, as can be seen in Figure 14, its intervocalic characteristics are typical of a glide. Figures 17 and 18 below show words illustrating the front glide in word initial and intervocalic positions.

When the front glide is in word initial position, it typically shows no steady state. Furthermore its F2 and F1 transitions begin at points that are typical of the high front vowel **i**. However, when the front glide is found in intervocalic position, the maximum frequency for F2 is lower and the minimum frequency for F1 is higher than in onset position. Altogether, the spectral characteristics for the front glide in intervocalic position place it nearer the mid-high front vowel **e** rather than **i**. Further study of the front glide is needed to determine whether or not it is actually a mid-high glide, as Tung's description indicates, that raises in word onset and word final position or a high front glide, as in Li and Ho's description, that laxes in intervocalic position and following or preceding a consonant. But our preliminary investigation suggests that it is similar to the English glide in 'yellow, a yellow ball, unite, reunite'**.** Since the convention is to transcribe a front glide as **j** in the absence of any height contrast, it is transcribed as such here.

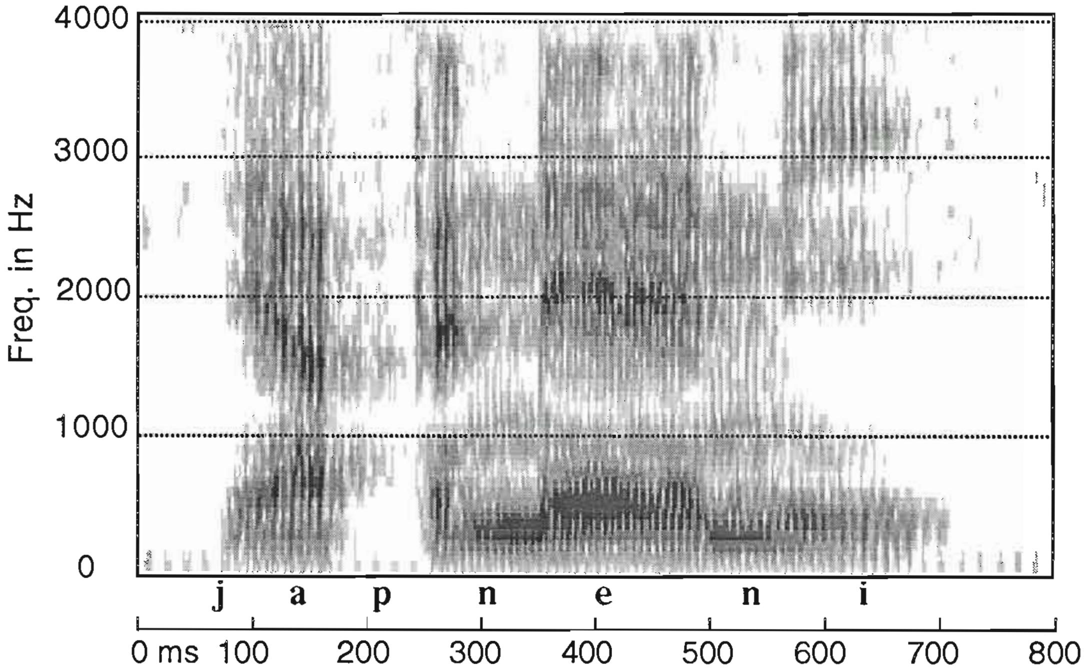


Figure 17. Example spectrogram showing **j** in **japneni** 'to sow'

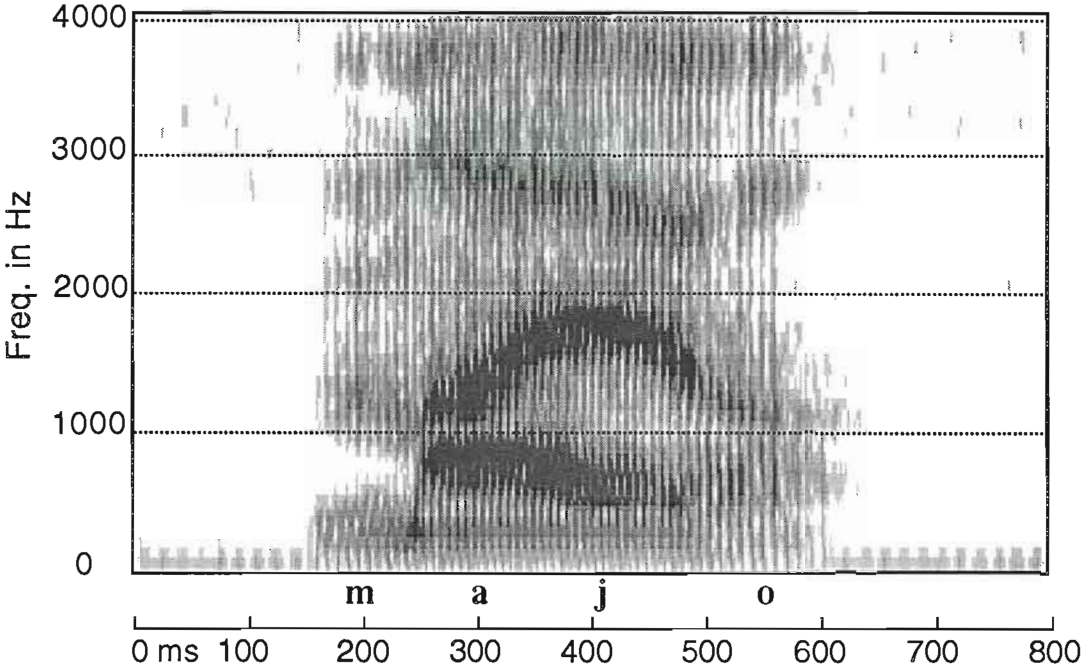


Figure 18. Example spectrogram showing **j** in **majo** 'to take'

Stress

With the exception of a handful of words, stress is predictable in Tsou. As was seen in the examples above, stress tends to fall on the penultimate syllable.

ex: (From Tung)

óko	**'oko**	'child'
okósi	**ok'osi**	'his child'
mamespíŋi	**mames'piŋi**	'woman'

This, however, is not the full story. When a word ends in a long vowel, stress falls on the last syllable.

ex: (From Tung)

emóo	**e'moː**	'house'
eémo	**'eːmo**	'to build a house'

The full generalization, then, is that stress falls on the final syllable if it has a long vowel, else on the penultimate syllable. This pattern is one of Tung's motivations for describing long vowels as disyllabic; he wanted to say that stress was always on the penultimate syllable. Tung, however, was writing before the advent of metrical theories of stress such as that described in detail in Hayes (1993) and thus lacked the tools for analyzing this pattern as moraic stress. In a theory of moraic stress, a syllable containing a long vowel is heavy and bears two morae. A syllable with a single vowel is light and bears a single mora. The stress pattern in Tsou is characteristic of moraic trochees: main stress falls on the final syllable if it is heavy, else on the penultimate syllable. Because the main stress falls at the right edge of a word, stress is assigned from right to left. The moraic trochee foot template is as follows:

(x .)   (x)
∪ ∪ or   –
Where ∪ is a light syllable (μ) and – is a heavy syllable (μμ)

ex:		
	(x .) óko	'child'
	(x .) okósi	'his child'
	(x) emóo	'house'
	(x) eémo	'to build a house'

Evidence for moraic trochees comes from the Duhtu dialect of Tsou which has a productive process of vowel deletion described in Tsuchida. Alternating vowels preceding the main stress are deleted. Vowel deletion does not occur if a cluster of three or more consonants would occur, or if the vowel is in an onsetless word initial syllable or in word final position. Vowel deletion also fails to occur if the vowel in question is long, i.e. if the syllable is heavy. This pattern is also characteristic of moraic trochees. Every heavy syllable receives stress, every other light syllable receives stress, all stressless syllables excluding those mentioned above are deleted. Our data were collected without this point in mind, and do not allow us to fully determine vowel deletion rules in the Tfuea dialect. From what we have observed it is clear that vowel deletion is an active process in Tfuea, and it seems likely that both dialects have the same stress pattern: main stress is on a final heavy syllable else on the penultimate syllable.

ex: from Tsuchida
seʔe- nətəh- a [sʔen'təha]
'to cut off with a bolo' (general focus)

			X.		main stress
	(X	.)	(X	.)	trochaic stress
∪	∪	∪	∪	∪	
σ	σ	σ	σ	σ	syllabic
|	|	|	|	|	
μ	μ	μ	μ	μ	moraic
|	|	|	|	|	
s e	**ʔ** e	n ə	**t** ə	h a	
s	ʔe	n	tə	ha	vowel deletion

Consonant clusters

One of the most interesting aspects of Tsou is the wide array of consonant clusters that are permissible in word initial position. In general a consonant cluster is made up of two and only two consonants, and almost any two of the Tsou single consonants may appear in either order. There are, however, a few exceptions to this generalization. With the exception of **h** and **ʔ**, homorganic fricative-stop or stop-fricative clusters do not occur. If **h** is regarded as not having a glottal place of articulation, **hʔ** and **ʔh** need not be considered exceptions. Strengthening the notion that in the phonological definition of possible clusters h should be regarded as a velar fricative is the fact that ph and th do occur as initial and medial clusters, but kh occurs as neither. Similarly hp and ht occur in both positions, but hk has been observed only medially, in **kuhku** 'fox'. We do not know whether there is something special about this word, or whether the lack of hk in our data is only an accidental gap. Only nasals precede homorganic stops or fricatives. The preglottalized alveolar implosive **ɗ** clusters only with the glottal stop. As most of the data was compiled using Tung's text where the glides are unmarked, no systematic data was collected on the clustering patterns of **j** or **w**. Not all of the consonant clusters that are possible in Tsou have been documented in word initial position. Tung's description is corpus based, and therefore any absence of a particular cluster could be due to chance. The present study expands on Tung's list of initial consonant clusters, yet any absences may too be due to chance. Some time was spent trying to elicit words that begin with the clusters that are listed but not exemplified in Tung's text. When this process failed it is a strong indication that the cluster is at least dispreferred, but it does not rule out the possibility of such a cluster in rarely used words. Other clusters that are not even listed in Tung may also occur. The Appendix contains a table which is an expanded version of that found in Tung. The glosses have been corrected where necessary. The following table illustrates the combinations of consonants found in our data.

Table 4. Consonsant clusters attested in the data
x = clusters that appear in word initial and word internal position
(x) = clusters that appear only in word internal position

	C2 →	p	f	v	ɓ	m	t	ɗ	ts	s	z	n	k	ŋ	ʔ	h
C1	p						x		x	x	(x)	x	x	x	x	x
↓	f						x		x			x	x	x	x	
	v								x			(x)			(x)	x
	ɓ											x	(x)			
	m	x	x		(x)		(x)		x	x	x	x			x	x
	t	x	x	x	(x)	x						x	x	x	x	x
	ɗ															
	ts	x	x	x	(x)	x						x	x	x	x	x
	s	x		x	x	x						x	x	x	x	
	z														(x)	
	n	(x)				x	x		(x)	x	(x)					(x)
	k								(x)	x		x			(x)	
	ŋ			x		(x)	(x)		(x)	(x)	(x)		(x)			x
	ʔ	x	(x)	x		x	x		x	x		(x)	(x)			(x)
	h	x		x		x	x		x	(x)	(x)	x	(x)	(x)	x	

Both members of a stop-stop cluster are fully released in word onset position and word internally. There is an audible burst that varies in amplitude, and almost invariably no voicing between a voiceless stop and a following voiceless segment. (One of the speakers in the study produced voicing forming a short epenthetic vowel in this environment). When a voiceless stop precedes a nasal there is a small amount of voicing, 3 to 5 glottal pulses, between the release of the stop and the closure of a following nasal segment. During the brief period of voicing, what little formant structure is discernible is highly dependent on the consonantal context. Figure 19 illustrates the characteristics of voiceless stop clusters and Figure 20 illustrates those of voiceless stop plus nasal clusters.

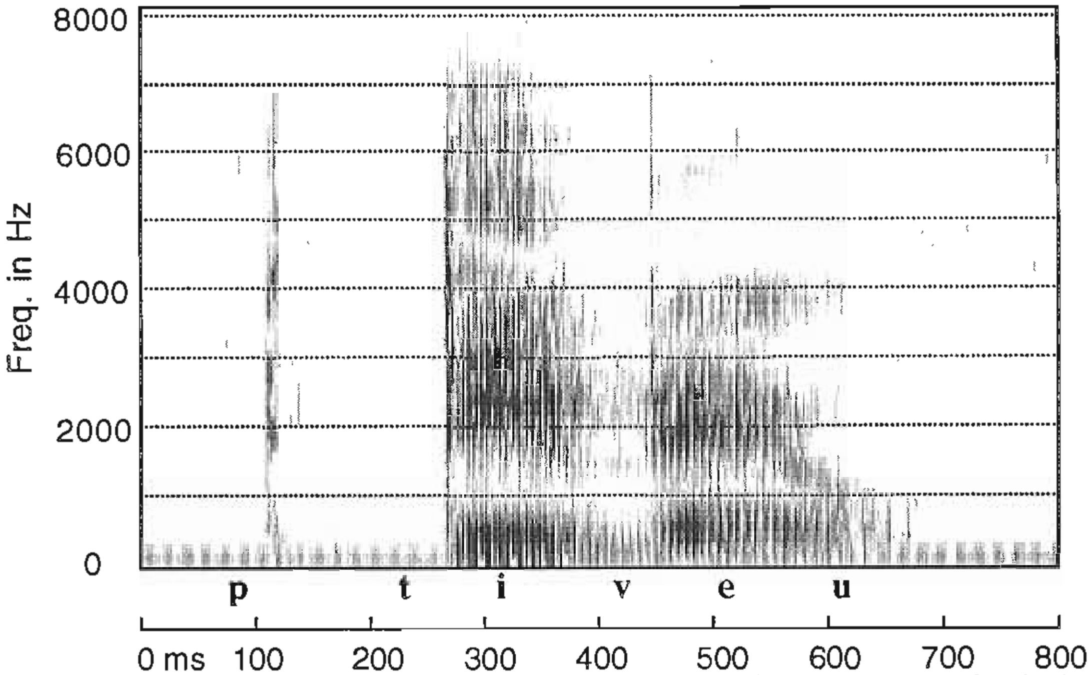


Figure 19. Example spectrogram showing the characteristic voiceless release of voiceless stop clusters in **ptiveu** 'a kind of reed'

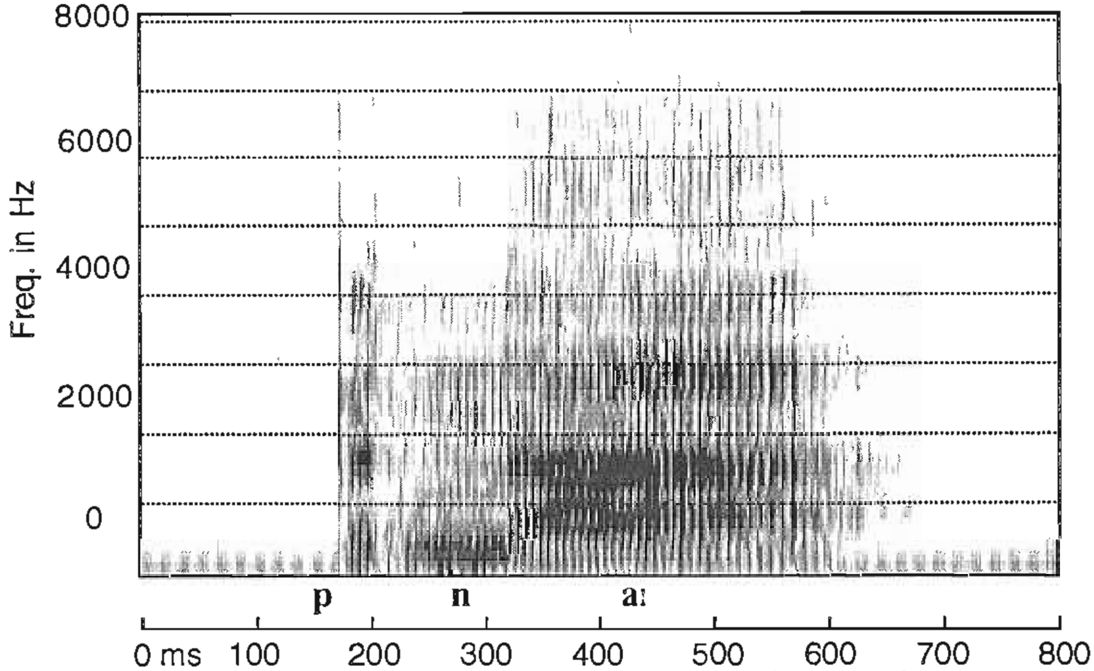


Figure 20. Example spectrogram showing the characteristic vocalic release in voiceless stop + nasal clusters in **pnaː** 'to shoot'

Nasal-stop and nasal-nasal clusters vary between having a vocalic release and having no vocalic release. The variation appears to be speaker dependent. Figures 20 and 21 illustrate the two possibilities.

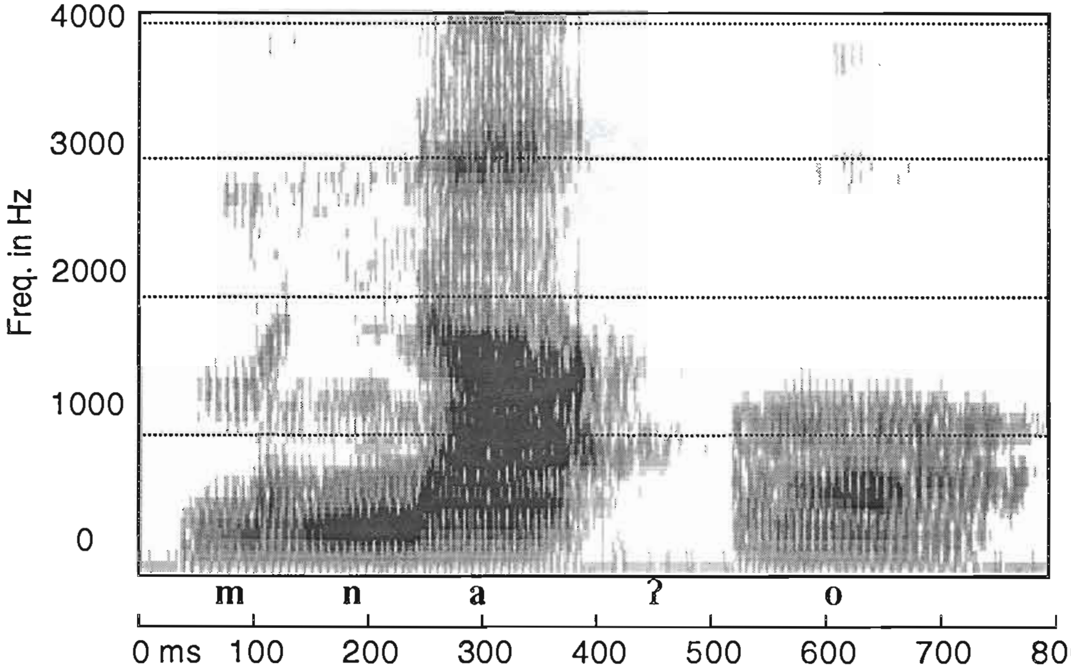


Figure 21. Example spectrogram showing the vocalic release in a nasal-nasal cluster in **mnaʔo** ‘come to do one thing’

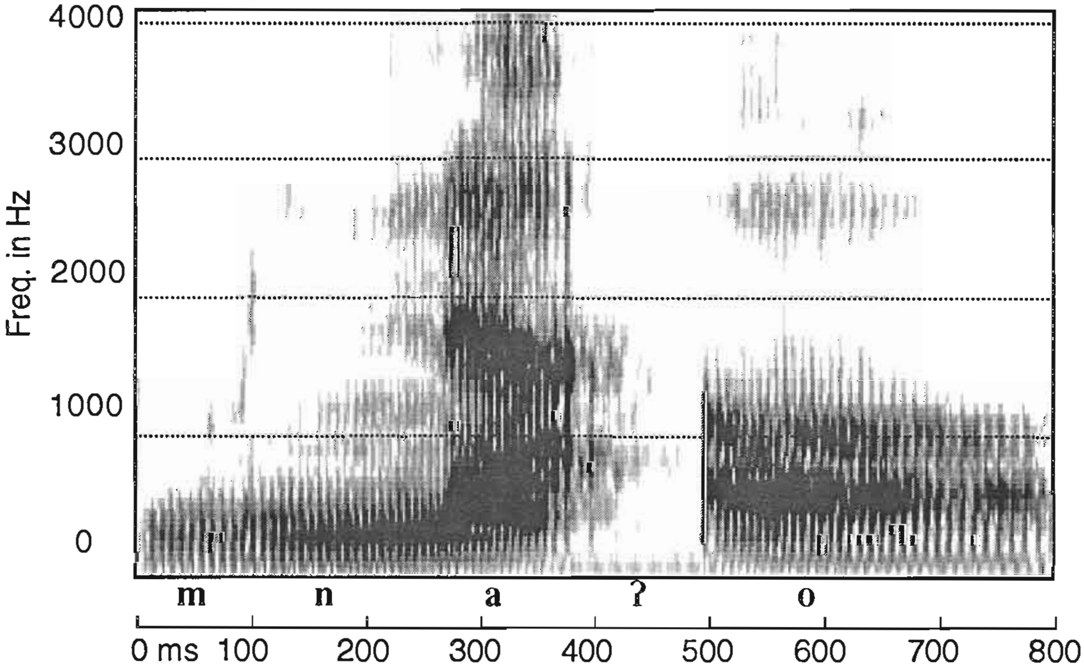


Figure 22. Example spectrogram showing no vocalic release in a nasal-nasal cluster in **mnaʔo** ‘come to do one thing’

Some of the most unusual clusters involve **h** and **ʔ**. Clusters such as **ph** and **th**, (but not, as we have observed, **kh**) occur, and are much like aspirated consonants in other languages. What is, to the best of our knowledge, unknown in other languages are initial clusters with **h** as the first member, such as **ht**, illustrated in Figures 23 and 24. For most of the speakers in the study the frication associated with the **h** has low intensity and spreads widely across the spectrum, as illustrated in Figure 23. In Figure 24 the initial aspiration is somewhat stronger than that produced by the other speakers (chosen so that the acoustic characteristics were clearly visible).

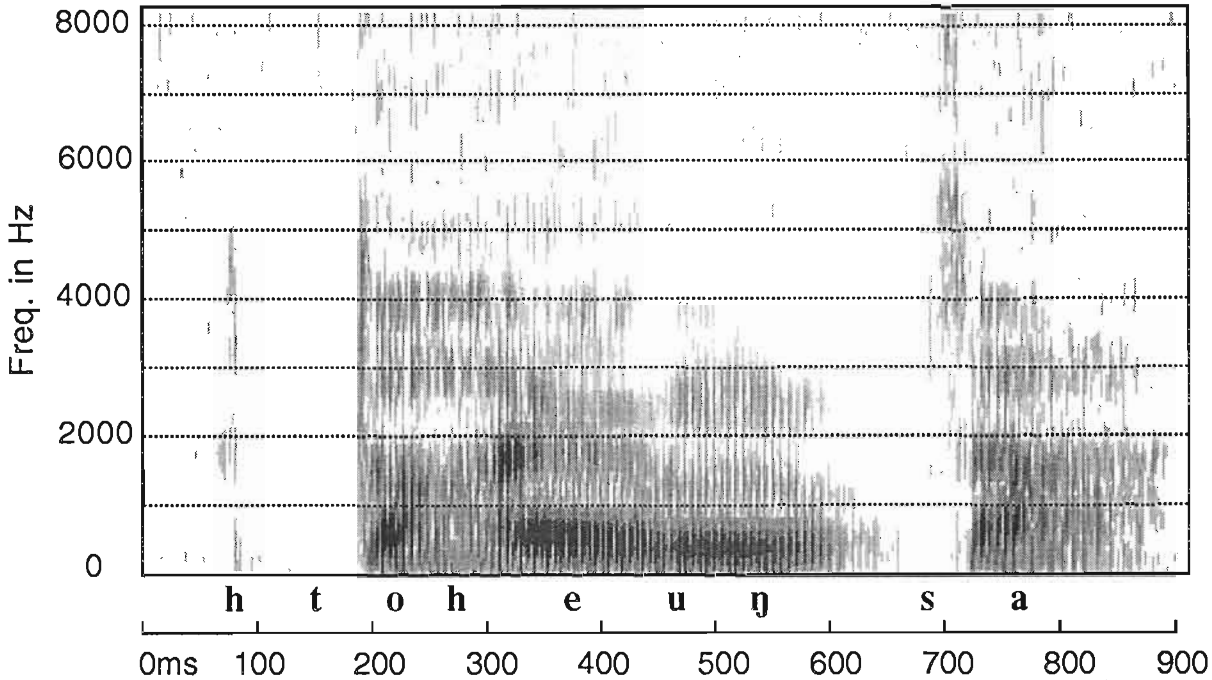


Figure 23. Spectrogram of **htoheuŋsa** 'repeated' illustrating the typical low intensity, broad spectrum frication with concentration of energy at the closure of the following stop.

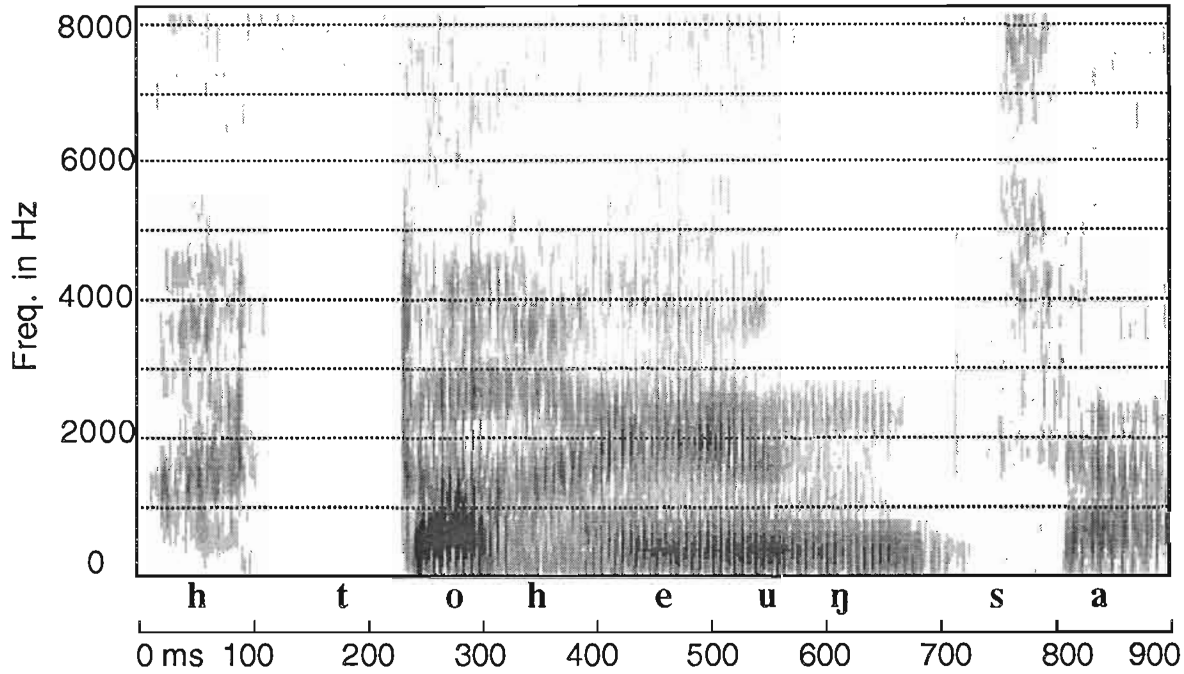


Figure 24. Example spectrogram of **htoheuŋsa** 'repeated' illustrating stronger frication.

Before the alveolar stop there is a movement so that the acoustic structure becomes more like that associated with the alveolar burst: the narrowing of the alveolar constriction causes a concentration of the energy of the frication between 2000 and 5000 Hz. In Figure 24 the movement of the tongue towards the alveolar closure is clear in the change in frequency of the formants within the frication.

Similarly, clusters such as **pʔ** and **tʔ** are sometimes sequences of a plosive and a glottal stop, but are sometimes much like ejectives in other languages; and again what is unknown in other languages are initial clusters such as **ʔp** and **ʔt**, illustrated with **ʔt** in Figure 25. Perhaps the most exceptional of all is the cluster **hʔ** illustrated in Figure 26.

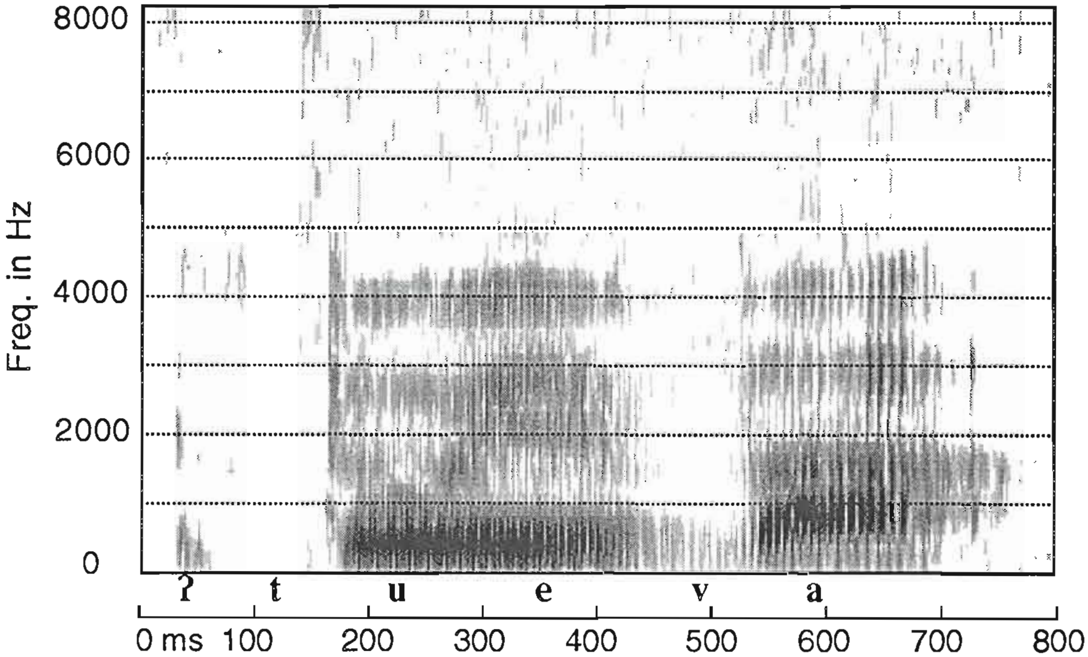


Figure 25. Example spectrogram of **ʔtueva** ‘the third month’ illustrating the characteristic glottal release .

In clusters of glottal stop and voiceless stop the glottal stop is usually released before the closure of the following consonant is complete. Often this is visable on a spectrogram as two or three creaky glottal pulses. Occasionally the creakiness persists into the beginning of the vowel.

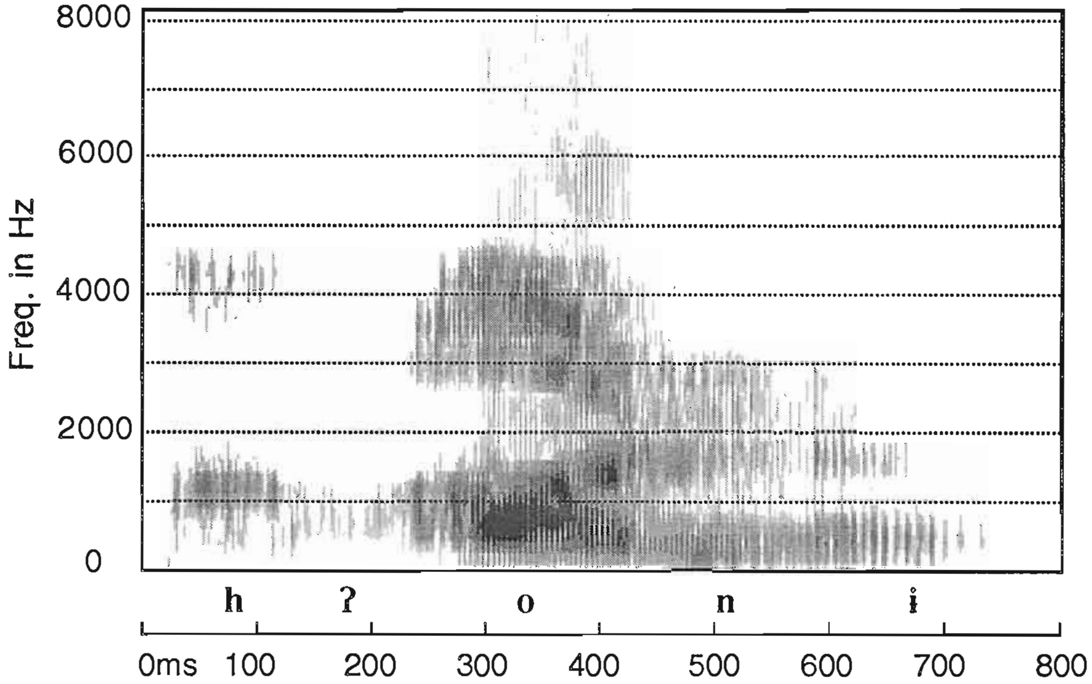


Figure 26. Example spectrogram of **hʔonɨ** 'liver' showing coarticulation of **h** with the vowel following the glottal stop.

In the **hʔ** cluster, the formants visible in the frication show coarticulation with the vowel following the glottal stop. In this case the concentration of energy around 1000 Hz with very little energy in higher frequencies indicates that there is lip rounding throughout the production of the **h**.

Conclusion

This paper has outlined the main phonetic characteristics of Tsou. But there are many points of phonetic interest still to be investigated. In particular, further research on the consonant clusters is needed. We are not completely certain whether there are some consonant combinations that are possible but do not appear in our data. In previous analyses, all syllables are said to have the form (C)(C)V so that there are no codas in the language, regardless of the position of the syllable within the word. These analyses are based in large part on the fact that there are no consonants at the ends of words. Our data show that in medial position the first consonant in a cluster is always released which is a characteristic normally associated with onsets. We need to conduct further investigations to establish the full range of consonant clusters in Tsou. In addition, further investigation of the nature of unusual clusters like those in Tsou will shed light on questions of the nature of the makeup of possible syllable onset clusters and the role of saliency in determining ordering of articulations in speech.

Acknowledgements

This research was supported by the National Science Council of Taiwan and the Academia Sinica, Taiwan, and by NSF. We are grateful to Elizabeth Zeitoun for fieldwork assistance, Paul Li for his encouragement and valuable comments, and to Chiu-yu Tseng for her efficient arrangements. Special thanks are due to Mo'o Muknana who hosted our visit in Punguu and to all the villagers who cooperated so enthusiastically.

References

Fant, Gunnar (1973). *Speech sounds and features.* Cambridge, MA, MIT Press.

Fischer-Jørgensen, Eli (1954). Acoustic analysis of stop consonants. **Miscellanea Phonetica* **2**: 42-59.

Fuller, Michael (1990). Pulmonic ingressive fricatives in Tsou. *Journal of the International Phonetic Association* **20**: 9-14.

Hayes, Bruce (in press). *Metrical Stress Theory: Principles and case studies.* UCLA

Ho, Dah-An (1976). Tsou phonology. *Bulletin of the Institute of History and Philology, Academia Sinica* **47**: 245-74.

Ladefoged, Peter (1993). Linguistic phonetic fieldwork: a practical guide. *UCLA Working Papers in Phonetics* **84**: 1-24.

Ladefoged, Peter and Zeitoun, Elizabeth (1993). Pulmonic ingressives phones do not occur in Tsou. *Journal of the International Phonetic Association* **23**(1): 13-15.

Li, Paul J.-K. (1979). Variations in the Tsou dialects. *Bulletin of the Institute of History and Philology, Academia Sinica.* **50**: 273-96.

Lin, Hen-Li (1955). On the quasi indirect passive voice in the Tsou grammar in Ali Mountain. *Journal of Chinese Ethnography* **1**: 207-216.

Lindau, Mona (1984). Phonetic differences in glottalic consonants. *Journal of Phonetics* **12**: 147-155.

McCarthy, John J. & Prince, Alan S. (ms). *Prosodic Morphology: Constraint interaction and satisfaction.* University of Massachusetts, Amherst and Rutgers University,

Nevskij, N.A. (1935). *Materialy po Govoram Jazyka Cou.* Moscow, Trudy Instituta Vostokovedenija.

Ogawa, N aoyoshi & Asai, Erin (1935). *The myths and traditions of the Formosal native tribes.* Taipei, Taihoku.

Starosta, Stanley (1974). Causative verbs in Formosan languages. *Oceanic Linguistics* **13**: 279-369.

Tsuchida, Shigeru (1972). The origins of the Tsou phonemes /b/ and /d/. *Gengo Kenkyu* **46**: 42-52.

Tsuchida, Shigeru (1976). *Reconstruction of Proto-Tsouic phonology.* Tokyo, Studies of languages and culture of Asia and Africa, Monograph Series.

Tung, T'ung-Ho (1964). *A Descriptive Study of the Tsou Language.* Taipei, Institute of History and Philology, Academia Sinica Special Publication.

APPENDIX

I: WORD LIST USED IN VOWEL RECORDINGS

IPA	GLOSS	IPA	GLOSS
pakʔi	‘bad, wicked’	**paʔti**	‘to show, to let see’
peːla	‘can, be able to’	**pepe**	‘sky, heaven’
pitu	‘seven’	**pio**	‘how many’
pohe	‘maize’	**poʔe**	‘to pray making wine’
puzu	‘fire’	**puːtu**	‘Chinese’
taso	‘strong’	**tafʔu**	‘kind of skirt’
teːlu (teːɗu)	‘to arrive on time’	**teːʃi**	‘rope, cord’
timʒi	‘to hold’	**titha**	‘to use’
tofi	‘gourd’	**toti**	‘kind of fungus’
tufku	‘to wash clothes’	**tutu**	‘seed, grain, ball’
tisʔi	‘clear (of water)’	**tiʔsi**	‘arrow’

II: WORD LIST USED IN CONSONANT RECORDINGS

paʔmiʒi	‘to let pass’	**peʔpiʔi**	‘eyebrow’
baʔi	‘grandmother’	**bjahtʃi**	‘fruit, flesh’
maʔkuvʔo	‘to suprise’	**majo**	‘to take’
meoiʃi	‘large big		
faʔei	‘egret’	**feango**	‘body’
vaʔhi	‘stream’	**vjovjo**	‘kind of deer’
tatsu	‘lamb’	**teʔi**	‘to shit’
lauja	‘maple’	**liŋki (ɗiŋki)**	‘muddy’
navju	‘cooked rice’	**neʔe**	‘there (out of sight)’
tsaʔi	‘dirty’	**tʃei**	‘dream’
saptʃi	‘face’	**ʃeu**	‘intestine’
zomi	‘bird’		
kaʔli (kaʔɗi)	‘knee’	**keːi**	‘eye secretion’
ŋajo	‘mouth’	**ŋeːi**	‘hemp’
hahʔo	‘friends’		

III: CONSONANT CLUSTERS IN THE PRESENT DATA

CLUSTER	INITIAL	GLOSS	MEDIAL	GLOSS
pt	**ptiveu**	‘a kind of reed’	**sipti**	‘four’
pk	**pkaːko**	‘to escape’	------	-------
pʔ	**pʔuŋu**	‘body joint (knee)’	**sopʔo**	‘light (weight)’
pn	**pnaː**	‘to shoot’	**japneni**	‘to sow’
pŋ	**pŋei**	‘pestle’	**apŋu**	‘a basket’
pts	**ptsokni**	‘tall bamboo’	**optsoi**	‘to kill’
ps	**psipŋa**	‘to thrust into’	**jopʃi**	‘dusk’
pz			**ŋipza**	‘mildew’
ph	**phiŋi**	‘door’	**tsaphi**	‘palm, sole, paw’
t				
tp	**tposi**	‘drawing’	**atpita**	‘to win’
tɓ	------	-------	**ketbi**	‘storage’
tk	**tkiti**	‘miser’	**kitkiti**	‘wrist’
tʔ	**tʔaŋo**	‘foot’	**atʔiŋhi**	‘surely’
tm	**tmali**	‘to hear’	**matmomane**	‘sing for a long time’
tn	**tniji**	‘rain’	**nitni**	‘lung’
tŋ	**tŋoː**	‘pot’	**botŋoni**	‘many’
tf	**tfuja**	‘place name’	**fitfi**	‘edge’

	tv	**tvohusa**	‘bamboo tube’	**teakotva**	‘narrow’
	th	**thoepza**	‘to make bigger’	**aothomi**	‘to try’
k					
	kʔ			**akʔi**	‘grandfather’
	kn	**knaju**	‘to cheat’	**konakni**	‘species of plant’
	kts	------	-------	**zotikci**	‘to press’
	ks	**ksiksi**	‘burning charcoal’	**ksiksi**	‘burning charcoal’
ʔ					
	ʔp	**ʔpitva**	‘monkey’	**peʔpiʔi**	‘eyebrow’
	ʔt	**ʔtujva**	‘the 3rd month’	**aopaʔto**	‘to appear’
	ʔɗ	------	-------	**kaʔɗi**	‘knee’
	ʔk	------	-------	**koʔko**	‘therefore’
	ʔm	**ʔmaska**	‘the 10th month’	**maʔmio**	‘to think’
	ʔn			**naʔno**	‘very’
	ʔts	**ʔtsonza**	‘the 1st month’	**smujuʔtsu**	‘to pierce’
	ʔf			**tmauʔfu**	‘to stir’
	ʔv	**ʔvoja**	‘the 8th month’	**iʔvaho**	‘again’
	ʔs	**ʔsia**	‘the 9th month’	**toʔso**	‘to throw’
	ʔh	------	-------	**vaʔhi**	‘to try’
	ɓn	**bnivhi**	‘plum’	------	-------
	ɓk	------	-------	**tiɓkoɓkotsa**	‘therefore’
m					
	mp	**mpiti**	‘to hold’		
	mɓ	------	-------	**seʔimɓa**	‘five hundred’
	mt	------	-------	**aimti**	‘actually’
	mʔ	**mʔumʔu**	‘hair’	**amʔuhu**	‘early, quick’
	mn	**mnaʔo**	‘come to do one thing’	**imni**	‘good’
	mts	**mtsoː**	‘eye’	**mamtsino**	‘to bathe’
	mf	**mfeiʔsi**	‘to cover’		
	ms	**msapie**	‘put on shoe’	**nomso**	‘have enough room’
	mz	**mzaː**	‘our’	**omza**	‘upper side’
	mh	**mhino**	‘to buy’	**mimho**	‘to permit’
n					
	np	------	-------	**ainpinpi**	‘at will’
	nt	**nteʔo**	‘conditional prefix’	**tʔuantihi**	‘star's name’
	nk	------	-------	**kunkunu**	‘wood carving’
	nʔ	------	-------	**tanʔe**	‘here’
	nm	**nmijo**	‘adult animal’	**anma**	‘sixth month’
	nts	------	-------	**maintsa**	‘to say (agent focus)’
	nv	------	-------	**sinvi**	‘to chew’
	ns	**nsoː**	‘pond’	**zonso**	‘lie in wait for an animal’
	nz	------	-------	**ʔtsonza**	‘first month’
	nh	------	-------	**inhe**	‘sow’
ŋ					
	ŋt	------	-------	**hoŋtuŋtu**	‘to burn one end’
	ŋk	------	-------	**oŋko**	‘name’
	ŋʔ	------	-------	**anʔosi**	‘two friends
		------	-------	ganging up on	a third’
	ŋm	------	-------	**joŋma**	‘to be in ambush’
	ŋts	------	-------	**eiŋtsa**	‘mad’
	ŋv	**ŋvoj**	‘to dry in the sun’	**tuŋva**	‘to powder’

	ŋs	------	-------	**moŋsi**	'to cry'
	ŋz	------	-------	**meŋzu**	'lance'
	ŋh	**ŋhou**	'monkey'	**atʔiŋhi**	'only'
ts					
	tsp	**tspiho**	'hip'	**putspuhu**	'inner mountains'
	tsb	------	-------	**etsbihi**	'to feel heavy'
	tsk	**tskosi**	'to level'	**putsku**	'navel'
	tsʔ	**tsʔojha**	'stream'	**jatsʔi**	'to stand'
	tsm	**tsmoi**	'bear'	**hutsma**	'to sing for long'
	tsn	**tsnimi**	'banana'	**aːtsni**	'surely'
	tsŋ	**tsŋuhu**	'chestnut'	**atsŋihi**	'completed'
	tsf	**tsfuo**	'stomach'	**teotsfuŋa**	'convergence of streams'
	tsv	**tsvoʔha**	'pierce'	**tatsvohʔi**	'long'
	tsh	***tshumu***	*'water'*	***botshio***	*'to know'*
f					
	ft	**ftuke**	'to arch the back'	**moftiʔi**	'to flash'
	fk	**fkoi**	'snake'	**tsofkoja**	'clean'
	fʔ	**fʔue**	'sweet potato'	**emufʔo**	'to bury'
	fn	------	-------	**jofna**	'evening'
	fŋ	**fŋuː**	'head'	**emifni**	'to hide'
	fts	**ftsuju**	'gun'	**faftsuja**	'a castrated man'
v					
	vn	------	-------	**ɓivni**	'flower'
	vʔ	------	-------	**avʔu**	'dog'
	vts	**vtsoni**	'spouse'	**tsivtsi**	'tail'
	vh	**vhovhoni**	'lean'	**tsovhi**	'distant'
s					
	sp	**spoja**	'not have'	**mamespiŋi**	'woman'
	sɓ	**sɓuːku**	'bamboo shoot'		
	sk	**skoʔa**	'to hug'	**joski**	'fish'
	sʔ	**sʔofi**	'stick'	**kosʔoza**	'shrimp'
	sm	**smuː**	'dew'	**esmi**	'to pass by'
	sn	**snifi**	'skin'	**esnithi**	'to cut and break'
	sŋ	**sŋisŋijo**	'a place name'	**asŋithi**	'constantly'
	sv	**sviji**	'earring'	**majasvi**	'dance'
z					
	zʔ	------	-------	**mahizʔo**	'poisoning'
h					
	hp	**hpihpiŋi**	'world'	**sohpoi**	'distressed'
	ht	**htoeuŋsa**	'repeated'	**ehti**	'branch'
	hk			**kuhku**	'fox'
	hʔ	**hʔoni**	'liver'	**smohʔo**	'to hatch'
	hm	**hmuju**	'blood'	**tsohmo**	'to pour in'
	hn	**hnou**	'small house'	**smuhnu**	'to dispatch'
	hŋ	------	-------	**oahŋu**	'relatives'
	hts	**htsuju**	'hill'	**bjahtsi**	'fruit'
	hv	**hvosi**	'enlarged genitals'	**etohva**	'smoking pipe'
	hs	------	-------	**ɓohsopi**	'to stick in'
	hz	------	-------	**tsohzona**	'noon'

Vowel enhancement and dispersion in the vowel space of Western Navajo: a study of traditional Navajo speakers

Joyce McDonough, UCLA and OSU
Martha Austin-Garrison, Navajo Community College

Navajo is from the Southern or Apachean branch of the Athabaskan language family, a family known for its especially complex morphology and morphology-phonology interface. Some closely related languages are Chiricahua, Jicarilla and Mescalero Apache (Hoijer 1945a). Navajo is by far the most vital of the Athabaskan languages and in fact one of the most vital American Indian languages. It is spoken by around 160,00 speakers who, for the great part, live on the Navajo reservation in northern New Mexico and Arizona. Only a small percentage of the speakers are monolingual. As language use is rapidly changing, a careful and systematic description of the speech of monolingual or near monolingual speakers is particularly important for the insights it can provide into the characteristics and patterns of language change and dialect within the community. This paper discusses the acoustic characteristics of the vowels of five speakers of Navajo from the Navajo Mountain area of the reservation in central Arizona. We call this dialect Western Navajo. We will refer to the speakers as 'traditional Navajo speakers'.

Issues

There are two main issues that concern us. One is the asymmetric vowel space of Navajo. The high back vowel /u/, which is present in many Athabaskan languages and in the reconstructed proto-Athabaskan, is absent in Navajo. Vowel dispersion theories (Lindblom 1986, Liljencrants and Lindblom 1972) predict full use of the vowel space. In this case dispersion would arguably result in some strategy of compensation for this gap, causing a less skewed vowel space by raising the mid back vowel, for instance (Maddieson 1984). Is there compensation for the deviation from a balanced dispersion pattern in Navajo? How marked is this system? The second issue concerns the possible enhancement of the length contrast by a vowel quality change. A quality enhancement of the length contrast has been consistently reported for the high front vowel pair alone. What is the nature of the length contrast for these traditional Navajo speakers? Is it enhanced by a quality difference? How many qualities are there in Navajo?

Previous phonetic studies

There is some variation in the description of the acoustic and phonetic properties of the vowels of Navajo. In Hoijer's (1945) study of Navajo phonology he reports a phonemic four vowel system with both length and nasality contrasts. These four vowels are listed below in (1). Note that there is a gap in the system; there is no high back vowel.

(1) Four Navajo vowels
i
e **o**
a

With contrasts in length and nasality, these four vowels provide sixteen phonemic contrasts, excluding tonal contrasts. The set of these 16 contrasts are listed in (2): long and short for both oral and nasalized vowels.

(2) Navajo vowel contrasts (excluding tone)

ORAL:

SHORT	LONG
i	**iː**
e **o**	**eː** **oː**
a	**aː**

NASAL[1]:

į	**įː**
ę **ǫ**	**ęː** **ǫː**
ą	**ąː**

Hoijer (1945:26-28) also reports four kinds of non-phonemic variation relating to vowel quality: (1) a quality difference between the long and short version of the high front vowel /i/; (2) no such enhancement in the other vowels; (3) variation in the round vowel /o/ that extends into the high back area, the 'missing vowel' area; and (4) considerable variation in the degree of backness of the low vowel. The Hoijer (1945) study is based on Sapir's data from 1927-1939 and his own 1941 fieldwork in the Chaco Canyon area of the reservation. The 1967 Sapir-Hoijer study, again based on Sapir's fieldnotes, reports similar allophonic variation in vowels, but notes a quality difference for only the long and short versions of the high front vowel. In addition, this study notes that the /e/ is a 'lower mid-front position' roughly equivalent to the lax /ɛ/ in English. Young and Morgan (1987) concur with the Sapir-Hoijer (1967) study and, as in the earlier Hoijer study, report the extension of the /o/ into the high back area, 'especially when it is followed by /i/' (1987:xii). Other studies of the Navajo vowels have reported different vowel effects. Maddieson (1984) notes a difference in the long and short versions of both the /i/ and /o/ vowels and records the mid front vowel as in Young and Morgan (1987), in a lower mid position. In a recent acoustic study, McDonough, Ladefoged and George (1993) report a considerable difference between male and female speakers for both the distribution and position of the vowels in the vowel space, and the quality of long versus short vowel pairs. For the female speakers in this study, the long and short vowel distinction in Navajo was enhanced by a quality difference for all but the low vowel. This effect was not repeated in the data from the men. Furthermore the distribution of the vowels for the women and men is quite different as can be seen in Figures 1 and 2 respectively, taken from the McDonough et al. study. For the male speakers, as opposed to the female, the vowels are constrained in three distinct groups and the four front vowels especially are very compacted. In a preliminary speech perception study of Navajo, Johnson, Ladefoged and McDonough (1993) found that listeners chose nearly identical vowel qualities for the long and short vowels, except for /ii/ and /i/. This latter study is in accordance with Young and Morgan (1980, 1987) and others who maintain enhancement for the front vowel pairs alone.

These studies raise a number of important questions concerning language variation and change. We believed a study was needed to provide a point of comparison. Accordingly an acoustic study of the speech of a consistent (monodialectic) and conservative speaker group was undertaken.

---

[1] In deference to the Athabaskan convention of marking tone and nasality, nasal vowels are shown here with a hook beneath the vowel.

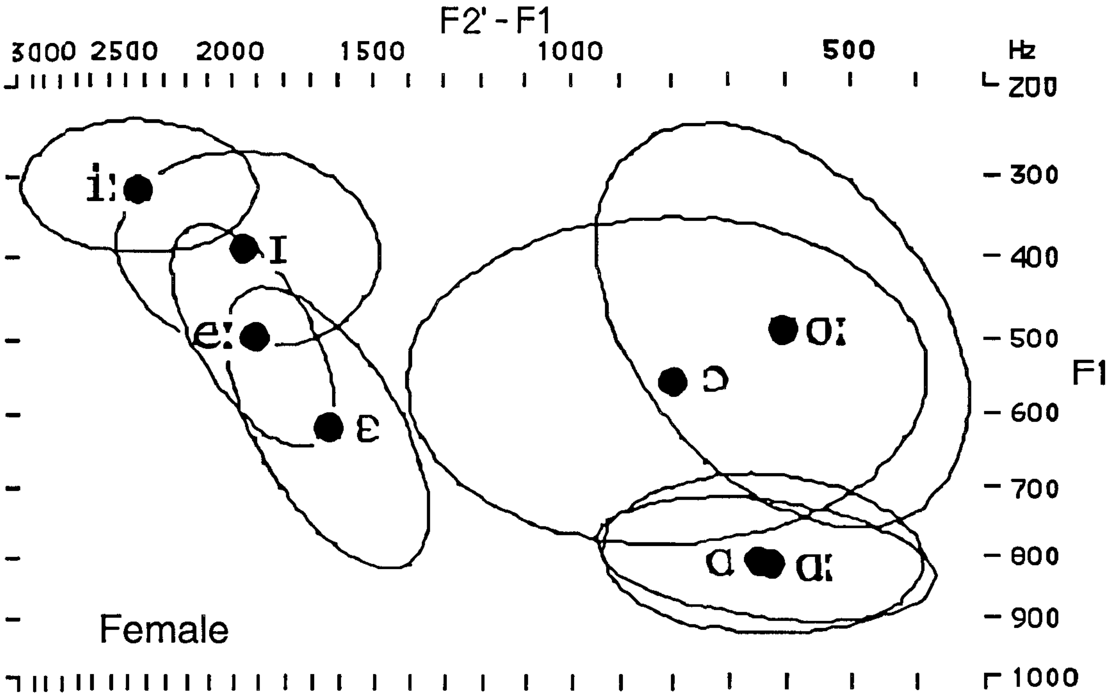


Figure 1. Mean locations of the vowels of the 5 female speakers of Navajo from McDonough, Ladefoged and George (1993). The ellipses enclose areas corresponding to two standard deviations along the first two principal components of variation for each vowel.

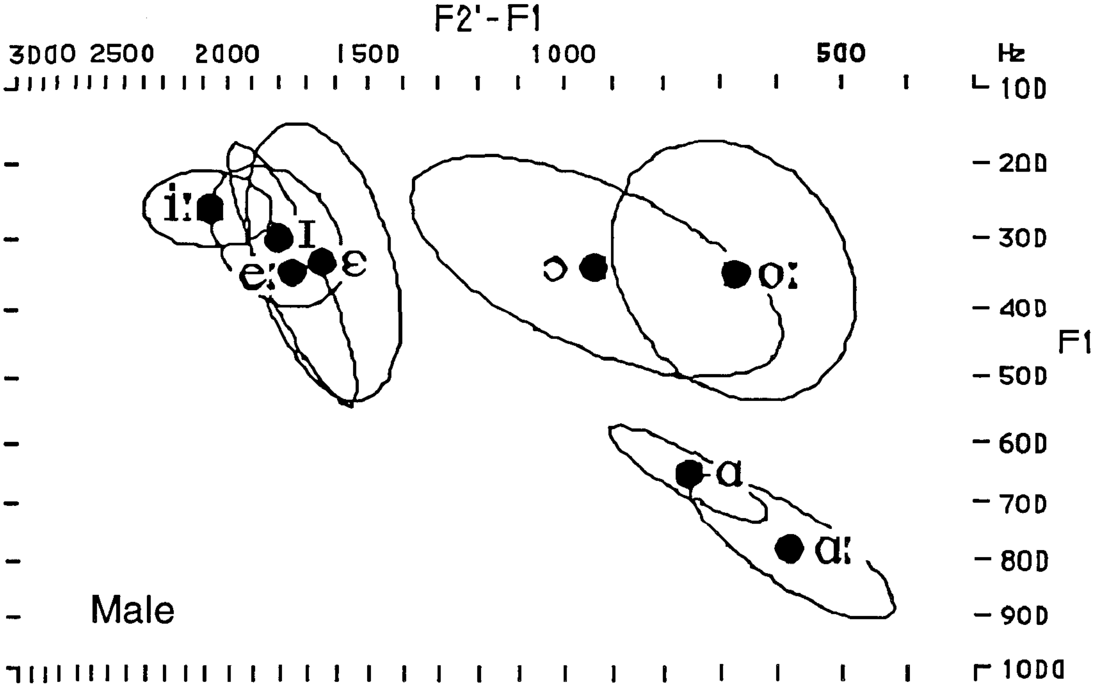


Figure 2. Mean locations of the vowels with ellipses of 2 male speakers from McDonough et al. (1993).

Methods

The study was conducted by recording and analyzing the speech of monolingual or near monolingual speakers of Western Navajo. The first author developed a word list before starting fieldwork, and checked it for accuracy and acceptability with the second author, Martha Austin-Garrison, who is a speaker of this dialect. The speakers were recorded one-on-one, at their homes, repeating the list after Ms. Austin-Garrison. The present report is based on an analysis of 5 speakers, 4 female and 1 male, producing the citation forms of the words. Four of the speakers in the study are monolingual Navajo speakers, and one is a bilingual speaker (Navajo - Southern Paiute) who has been living in the Navajo community for several decades and uses Navajo daily. None of the speakers had received a Western education, none were literate, although some were familiar with Christian hymnals written in Navajo. The speakers were all in their 60's and 70's. Recordings were made using a Marantz 340 cassette tape recorder with a Sony directional microphone.

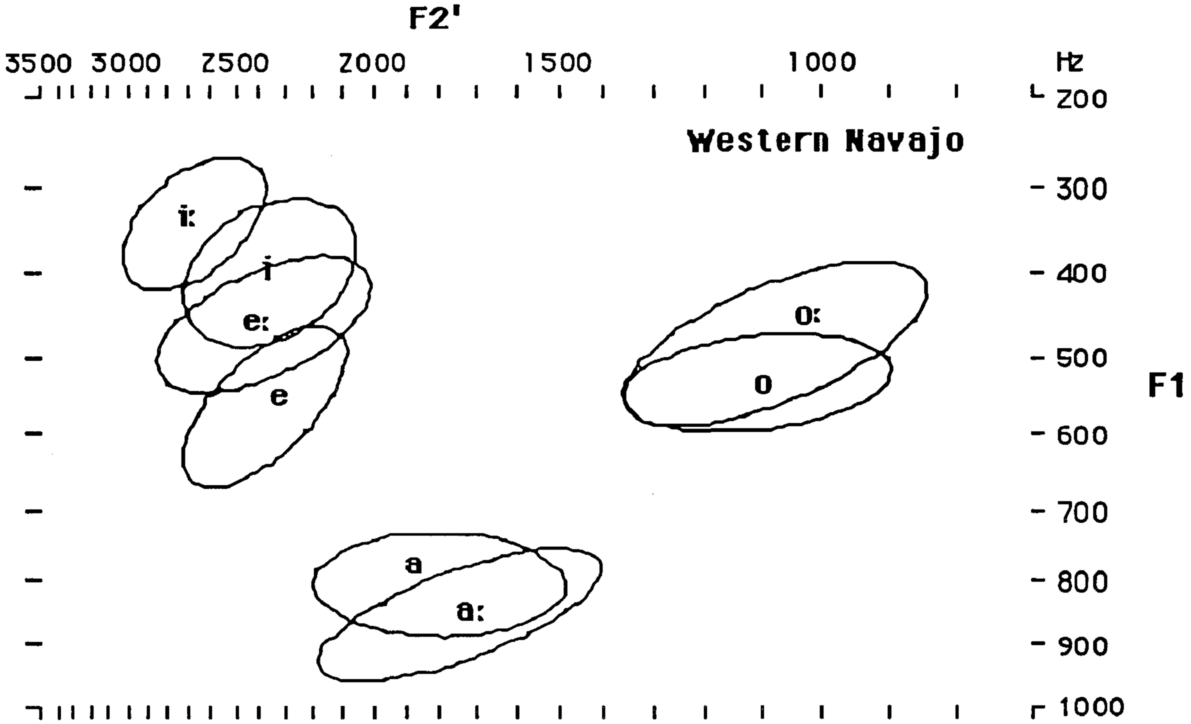


Figure 3. Mean locations and ellipses for long and short vowels of traditional speakers of Western Navajo.

From these recordings we extracted a set of words that exemplified the vowel contrasts. Navajo has a complex verbal morphology and a rich system of inflectional prefixes (Young and Morgan 1987, 1992; Kari 1976, 1990; McDonough 1990). Prefix vowels have been argued to be epenthetic, thus phonemic contrasts are best observed in the stems. The simple nouns, in the word list consist of a prefix, usually /bi/ 'his/hers' and typically a monosyllabic stem. This was the form of around 80% of our words. The remaining words had either disyllabic stems, or were verbal forms. Nevertheless, barring enclitics, the generalization that the stem is the final syllable in a Navajo word is a strong working generalization. In this study we used only the stem vowels for several reasons: the prefix vowels are often severely reduced in speech; these vowels are arguably epenthetic thus their specification is subject to various analyses, including a contextual analysis; and, finally, the complexity of the language's morphology makes it difficult to find minimal pairs illustrating the vowels. All the vowels are in a CV(C) context, but other aspects of the context, including tone, are not strictly controlled.

The data were analyzed at the UCLA Phonetics lab on a Kay Elemetrics Computer Speech Lab (CSL), sampled at 10 kHz. Formants values for F1, F2, F3 were measured on an FFT/LPC display. Duration was measured from a spectrogram. As noted in McDonough, Ladefoged and George (1993), consonants and vowels in Navajo have rather clearly delineated onsets and offsets, making duration measurements straightforward. In the case of vowel final words such as *bito* 'his water' it was possible to measure the duration because such words tend to be closed with either aspiration or a shut down of the glottis resulting in a delineated duration.

We attempted to measure two tokens of every vowel for each of the five speakers, but this was not always possible. The statistics were calculated on a minimum of seven vowel tokens for each vowel. As noted below, one of the concerns of the study is vowel quality differences between the long and short pairs of vowels. To this end, analyses of variance (ANOVA) were performed on the data with long and short vowels as the independent variable and F1 (a measure of height) and F2' – F1 (a measure of backness) as the dependent variables. Formant values for the tokens were plotted on a bark scale with F1 and F2 axes. (F2' is a weighted average of F1, F2 and F3 as in Fant (1973).) The results are discussed below.

Vowel duration

The length contrast is a robust contrast in Navajo. The durations of the long and short vowels for (a) the bilingual and (b) the traditional speakers are listed in the following table. Note that the long vowels are more than twice the length of the short vowels for both groups. It should be kept in mind that these measurements are of isolated disyllabic tokens.

Table 1. Vowel duration means, (standard deviations), in milliseconds.

(a) Bilingual speakers.

	FEMALE	MALE
LONG	264 (90)	264 (80)
SHORT	114 (50)	114 (50)

(b) Monolingual speakers.

	FEMALE/MALE
LONG	288 (81)
SHORT	104 (44.5)

Vowel quality

Figure 3 is a formant chart (F1 against F2' on a bark scale) of the vowels of the traditional speakers of Western Navajo for the four pairs of long and short vowels. The ellipses in the chart enclose all tokens for a given vowel that are within two standard deviations of the mean of that vowel, calculated along the first two principle components of the vowel variation. We note four things concerning this chart: (1) there are three distinct groups of vowels: front, low and back; (2) the front vowels are very close together; (3) the long and short pairs of vowel ellipses overlap considerably; (4) the traditional speakers (Figure 3) and the male bilingual speakers (Figure 2) resemble each other and contrast sharply with the female bilingual speakers (Figure 1) on the three previous points: the distinctness of the vowel groups, the compactness of the front vowels, and the overlapping of the ellipses. We will discuss each of these issues separately below.

Differences in quality between long and short vowels.

The vowel contrasts among the four vowel pairs in Figure 3 were tested by performing an ANOVA with a vowel pair (long and short) as the independent variable; the dependent variables are F1, F2' as measures of height and backness respectively. The means tables for F1 and F2' for the eight vowels are given in Table 2.

Table 2. Mean tables for F1 and F2' of long and short vowels of monolingual speakers.

Means Table for F1
Effect: vowel

	Count	Mean	Std. Dev.	Std. Err.
ii	9	341.889	40.748	13.583
i	9	399.000	46.384	15.461
ee	7	460.429	44.881	16.963
e	8	563.500	54.158	19.148
aa	8	854.250	53.633	18.962
a	8	807.125	41.097	14.530
oo	9	484.556	53.854	17.951
o	8	533.000	32.597	11.525

Means Table for F2'
Effect: vowel

	Count	Mean	Std. Dev.	Std. Err.
ii	9	2684.567	172.466	57.489
i	9	2366.791	183.576	61.192
ee	7	2394.351	232.627	87.925
e	8	2388.818	164.529	58.170
aa	8	1757.713	193.392	68.374
a	8	1810.905	178.864	63.238
oo	9	1083.417	136.711	45.570
o	8	1114.353	127.869	45.208

Table 3 shows the F-values from the eight one way ANOVAs comparing the F1 and F2' values of the four vowel pairs. Note that for the mid front and back vowels /ee,e/ and /oo,o/, any difference between the vowels in the respective pairs is along the F1 axis; neither pair shows any significant centralization of the short vowel of the pair. For the high front vowel pair /ii, i/, a significant difference is found along both the F1 and F2' axis, with a greater difference for F2'. This is the only vowel pair with a significant difference along both axes. Recall that only the high vowel length contrast is reported to be enhanced by a vowel quality difference. For the low vowel pair by comparison, no significant differences are found along either axis. While the low vowel may show variation such as fronting as reported, any variation in the position of the vowel in the vowel space is not due to the enhancement of the length contrast, and therefore in all probability is due to other aspects of the phonetic context. The same may be said for the variation found in the back vowel pair /oo, o/. These statistical results can be visualized by reference to ellipses in Figure 3. Note that in all cases, even those in which the differences in formant frequencies are not statistically significance, the longer vowels are further from a neutral vowel position in the center of the chart. It may be that the short and long vowels have similar targets, but the short vowels do not reach the targets because there is insufficient time to reach these more distant positions.

Table 3. A pair-wise comparison of the four sets of long short vowel pairs. Significant (<.01) differences are double starred and in boldface, marginally significant (<.05) figures are single starred.

	F1	F-VALUES	**F2'**	F-VALUES
ii, i	0.0135*	7.7	**0.0016****	14.3
ee, e	**0.0016****	15.8	0.9580	2.9
aa, a	0.0686	3.9	0.5770	0.33
oo, o	0.0434*	4.9	0.6382	0.23

Figure 4 make these results more visually explicit by means of a bar graph which demonstrates the contrasts between the long and short vowels of each of the four vowel pairs.

In Figure 4, the means of the short vowels are subtracted from the means of the long vowels for each of the four vowel pairs for both F1 and F2'. The significance of the differences between short and long vowels along the F1 and F2' axes are as in Table 3. This graph tells us that there are differences between the long and short vowels of the respective pairs in the direction we expect along the F1 axis; for example, the short low vowel is higher because it is more centralized. Because these bars represent Hz values, some of the bars (notably the long-short bar of the high front vowel pair along F2' axis) are longer than might be expected by reference to Figure 3, in which Bark scales have been used.

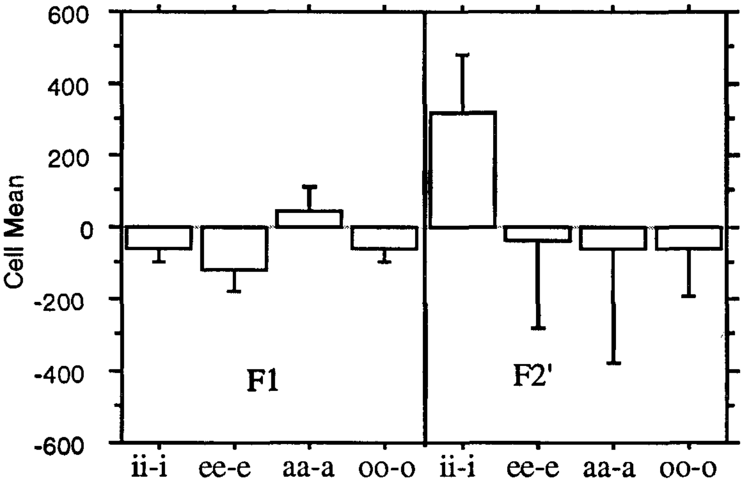


Figure 4. Bar graph of the mean of long minus the mean of short vowels for each of the four vowel pairs for F1 and F2'. This plot demonstrates the enhancement of the long short contrast for the high front vowels /ii-i/ along the F2' axis.

The high front vowel contrast

The two front vowel pairs and the back vowels /oo, o/ show significant differences, but only the high front vowel pair /ii, i/ is a statistically significant difference found along both axes. In the literature, a difference in the quality of the long and short front vowel is consistently reported as [iː, ɪ], a peripheral / central distinction. The mid front and back vowels show statistically significant differences along only a single axis (F1), but they are not perceived as qualitatively enhanced, as both the literature and the perceptual study of Johnson, Ladefoged and McDonough (1993) have shown. Recall the robustness of the length contrast; duration carried the functional load and the variation along F1 is not sufficient to establish a perceptual quality distinction. Differences along the height parameter are not salient and can be understood as timing differences.

The /ii, i/ difference is not mechanical in this way. The vowel targets for the high front vowels cannot be defined along a single articulatory axis. The speakers are aiming at different targets for long and short versions of the high front vowels. These facts provide some evidence that this distinction is perceptually rather than articulatorily based. What is singular about this vowel system is that this enhancement is unique to this single vowel pair.

Dispersion in the vowel space of Western Navajo

Western Navajo presents two problems to dispersion theory. One, there is a high back gap and dispersion theories predict compensation or skewing of the vowels in the vowel space. Two, the four front vowels are contained in a very small area in the high front space (Figure 3) in contradiction to dispersion theory which predicts that the mid vowels would use the available space and be realized lower.

The front vowel space

In the small high front space, there are four vowel contrasts: long and short high vowels /ii,i/, and long and short mid front vowel pairs /ee,e/. These vowels form a single group, so that the pattern we see in Figure 3 is a *triangular* dispersion pattern. The Western Navajo vowel space is characteristic of a three vowel system. The high front area constitutes a Firthian sub domain in the vowel space. The front vowels are maximally dispersed within this area.

The first point to note is that the two long vowels within this space, /ii/ and /ee/ are clearly distinguished from each other, with non-overlapping distributions. The same is virtually true for the two short vowels, /i/ and /e/. It is only if we have to disregard the length differences, which, as we have seen, are considerable, that we need to consider quality differences among these vowels.

A quality enhancement of the length contrast is reported for the high front vowel alone. Nevertheless the mid vowels /ee, e/ also show a significant difference along the F1 axis in Table 3. Given this, we might expect to find at least three quality contrasts in this front space: between the high vowel pair, the mid vowel pair, and between at least one if not both of the vowels of the high and mid vowel sets. To examine the way the vowels disperse within the space, an ANOVA was performed on the front vowel group (the independent variable, vowel, has four levels) with F1 and F2' as the dependent variables. (Note the statistics in Table 2 were performed on each of the vowel pairs separately, because the concern was with a contrast between the vowels of the pairs and not between the vowel pairs themselves, which we have (safely) assumed to be reliably distinct.)

In Table 4 are the results of a Scheffe pair-wise comparisons of these vowels. Note that this analysis is more conservative than the ANOVAs for the sets of long/short vowel pairs, due to the effect of calculating the variation of the four vowels within this space. For instance, the long/short distinction of the high front vowel has lost some of its force; there is no significant difference along the F1 axis; and the significance along the F2' axis is smaller. However, each of these four vowels is significantly different from each of the other in one dimension, except that no significant difference is found between the long mid /ee/ and the short high /i/ vowels. This is exactly what we would expect from the discussion in the previous section. Within the front space there is a hierarchy of descending height, /ii,i,ee,e/, as we see in the bar graph in Figure 5. Among these vowels there is no significant differences found between adjacent vowels except for the /ee,e/ pair. So in comparison to the pair-wise comparisons of Table 3, where both the mid and high vowels showed significant differences in F1, in this analysis only the mid front vowel shows a difference, an arguable result of articulatory undershoot. The contrast between the high vowel pair within this space is found only in F2', as we can see in the graph in Figure 6.

Table 4. Figures for significance for pair-wise comparison of the four vowels in the compact front vowel space.

	F1 (<.0001) (34.6)	F2' (.0034) (5.7)
ii, i	.1039	**.0126****
ii, ee	**.0003****	.0404*
ii, e	**<.0001****	.0277*
ee, i	.1008	.9934
i, e	**<.0001****	.9962
ee, e	**.0024****	>.9999

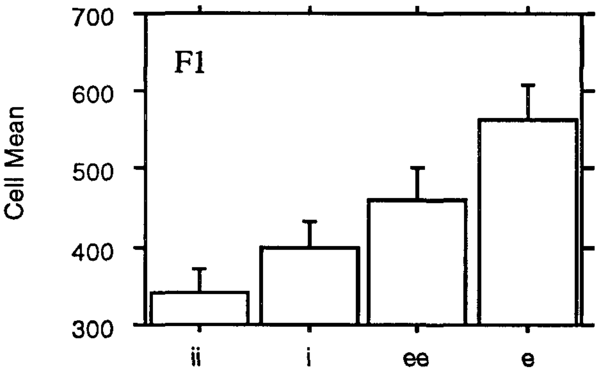


Figure 5. A bar chart of the front vowel group along the F1 axis. Note the descending height among these vowels grouped in the compacted front vowel space. Except for /ee, e/, there are no statistically significant differences between adjacent vowels along the x axis.

These statistics and charts reflect the lack of a significant difference between both the long and short mid vowel pairs along the F2' axis; that is, the mid short vowel is not more centralized than the mid long one (Table 4), and between the high and mid front vowel sets which are only marginally different along the F2' axis. The main part of the contrasts in this compacted space rest with the F1 axis. In Figure 3 these vowels stack up along the F1 axis in a way reflected in Table 5, so we can say that the vowels within this space disperse along articulatory dimensions in a hierarchy of descending height. The F2' bar chart in Figure 6 exemplifies the salience of the quality contrast between the long and short high front vowels.

In summary, within this space four vowels distribute their quality contrasts as follows: the long high vowel /ii/ is statistically higher and slightly less centralized than both the long and short mid vowels. The two short vowels /i/ and /e/ are statistically different along the F1 axis, the short high vowel is higher. There is no significant difference along the F2' axis, that is, the /e/ vowel is not more centralized than the /i/ vowel. For the differences in the two vowel sets /ii, i/ and /ee,e/, in this space the high vowel set uses the F2' axis for contrast, the mid vowel set, the F1 axis. Recall however that only the high front vowel set is reported to exhibit a vowel quality difference between the long and short version. If we were to consider only the articulatory interaction of these four vowels, we would predict quality differences between the vowels of both sets.

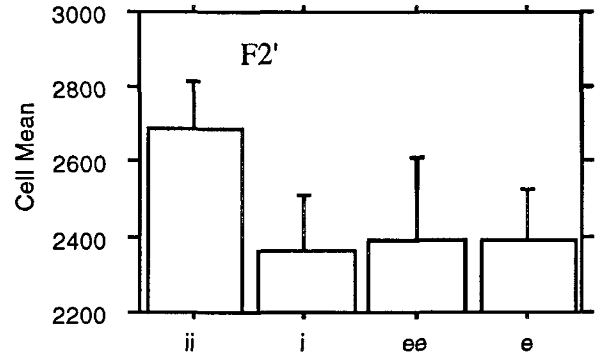


Figure 6. A bar chart of the front vowel group along the F2' axis. The salient contrast between the high front vowel pairs /ii, i/, which is not statistically present along the F1 axis, is emphasized here.

A final point of interest is the existence of a high back gap in the vowel system. Although a high back vowel is present in other Athabaskan languages and in proto-Athabaskan (Krauss and Leer 1976), this vowel has been lost in Navajo. The question arises as to whether or not the variation in the back vowel extends into the 'empty' high back space. To this end a pair-wise comparison of the group of mid front and back vowel pairs was performed. The statistics are reported in Table 5. (We are concerned only with the F1 axis as a measure of height. It is obvious that these vowels will differ along the F2' axis as front and back vowels.)

Table 5. Figures for significance for a pair-wise comparison of front and back mid vowels for F1.

	F1 (.0008) (7.3)
ee, oo	.7975
ee, o	.0524
e, o	.6530
e, oo	**.0192****

We can see that the mid back vowel /oo, o/ are not significantly different from the corresponding mid front vowel /ee, e/. The greatest differences are found, as we might expect, between the long front and short back, and conversely the short front and long back pairs. (Note also that there is no evidence of a quality enhancement of the length contrast for the back vowel set.) In fact we can say that the mid front and back vowel pairs are of basically the same height. This is consistent for all three groups, the female and male bilingual as well as the Western Navajo speakers. With regard to the back vowel's extension into the high back vowel space however, the results are less clear. Recall the compactness of the front vowel space; the front mid vowels are rather high themselves. So we can only say that the back vowel is no higher than the rather high mid front vowel. These results appear in a bar graph in Figure 7. The vowel space is skewed to compensate for the gap because the back vowel is rather high. In this sense the vowel space is balanced and the system represents an adequately, rather than maximally, dispersed vowel space.

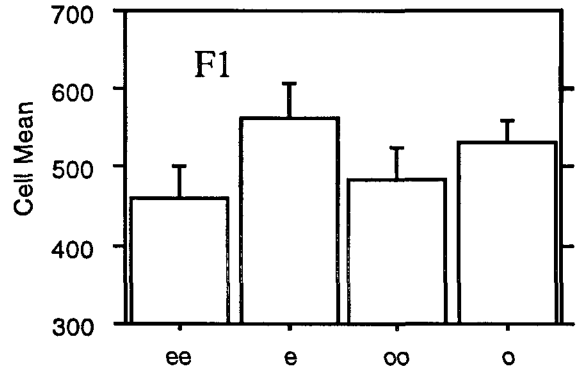


Figure 7. A bar chart of the mid front and back vowels along the F1 axis.

Summary

This data generally concurs with consistent observations concerning enhancement of the length contrast in Navajo. There are four vowels in Navajo, two front, a low, and a back vowel. There is also a vowel quality enhancement of the length distinction for the high front vowel pair /ii,i/. There is a reliable difference between the long and short mid front vowels along the F1 axis. However this distinction does not appear to be salient; it is not reported in the literature. Speakers in a perceptual study (Johnson et al. 1993) chose different vowel qualities only for the high vowel pair. In pair-wise comparison of the vowels (/ii/ against /i/ and /ee/ against /e/ etc.) the high front vowel differed from the other vowel pairs in showing reliable differences along both axes, F1 and F2'. We have further found that the vowels form three rather distinct groups within the vowel space. The Navajo vowel space has a triangular dispersion pattern: front, low and back. The four front vowels crowd into a small rather high area, in effect disregarding the rest of the front space. Within this front group we find at least three reliable differences along the F1 axis; between the high and mid vowels and between the long high and short mid front vowel. The distinction between the two high front vowels *within this space* is realized along the F2' axis. As regards to a third point about the vowel space, the existence of a high back gap, we find the issue difficult to resolve. Because of the front sub domain, the mid front vowel pair is rather high. If the back vowel, a mid vowel, extends into the high back 'empty' vowel space, it does not do so at the expense of differing from the front mid vowel in height; there is no statistical difference along the F1 axis of the front and back mid vowels. Thus the Navajo vowel space is triangular, but the vowels are not in a canonical triangular space.

Also of interest is the distinction this data shows between the mono-dialectical and near monolingual speakers of this study and bilingual speakers of the previous study (McDonough et al. 1993). The previous study of bilingual Navajo speakers showed a large difference between the men and the women of the study in the distribution of vowels in the vowel space. The tendency for the speakers of Western Navajo to conserve the length distinction for the contrast between the vowel pairs for all but the high front vowel is in accordance with the similar tendency among the bilingual men, but it is not in accord with the bilingual women. The bilingual women's vowel space with its enhancement of the length distinction and their pattern of vowel dispersion is arguably related to language interference and language change.

Acknowledgments

This research was supported by grants from the NIH and the American Philosophical Society's Phillips Fund. Our thanks and our appreciation for the interest and cooperation of the people at Navajo Mountain who served as consultants.

References

Fant, G. (1973). *Speech sounds and features*. Cambridge, MA: MIT Press.

Hoijer, H. (1945-49). The Apachean verb I - V. *International Journal of American Linguistics*, 9 (1943)193-203, 11 (1946) 1-13, 51-59, 14 (1948) 247-259, 15 (1949) 12-22.

Hoijer, H. (1945). *Navajo Phonology*. Albuquerque: University of New Mexico.

Kari, J. (1976). *Navajo Verb Prefix Phonology*. Garland.

Kari, J. (1990). Affix positions and zones in the Athabaskan verb complex: Athna and Navajo. *International Journal of American Linguistics*, 55, 424-54.

Krauss, M., & Leer, J. (1976). *Proto-Athbaskan-Eyak *y and the Na-Dene Sonorants* . Alaska Native Language Center, Fairbanks.

Leer, J. (1982). *Navajo and Comparative Athabaskan Stem List.* Alaska Native Language Center, University of Alaska, Fairbanks.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. Ohala & J. Jaeger (Eds.), *Experimental Phonology* (pp. 13-44). Orlando: Academic Press.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language,* 48, 839-862.

Johnson, K., P. Ladefoged and J. McDonough. (1993). "Do women listen to men with men's mouths in mind?"

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

McDonough, J. (1990) *Topics in the Morphology and Phonology of Navajo Verbs.* Ph.D, University of Massachusetts at Amherst.

McDonough, J., P. Ladefoged, H. George. 1993. "Navajo vowels and universal phonetic tendencies" *UCLA Working Papers in Phonetics* 84, 143-150.

Sapir, E. (1914). Notes on Chasta Costa phonology and morphology. *University Museum Anthropological Publications. Philadelphia: University of Pennsylvania,* 11(2).

Sapir, E., & Hoijer, H. (1967). *The Phonology and Morphology of the Navajo Language.*

Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics,* 17, 3-46.

Young, R., & Morgan, W. (1980 and 1987). *The Navajo Language.* Albuquerque: University of New Mexico Press.

A preliminary perceptual study of the vowels of Montana Salish: The method of adjustment as a fieldwork technique

Keith Johnson*
Peter Ladefoged

Montana Salish, a Southern Interior Salish language, has a 5 vowel system /i/, /e/, /ɑ/, /o/, /u/ which is an elaboration of the historical Salishan 4-vowel system /i/, /ɑ/, /u/, /ə/ (Thompson, 1979). In addition, the language, like other Salishan languages, has labialized back consonants. In this paper we present a preliminary study of the vowel system which was aimed at identifying the internal perceptual structure of the vowel categories and in particular the influence of back consonant labialization and vowel stress on the perceptual/phonetic properties of /o/. We begin with a brief description of the method of adjustment perception task and the rationale for its use in field phonetics.

The most common method of instrumental phonetic fieldwork is acoustic analysis. Speakers are recorded in the field and the recorded utterances are later analyzed in the laboratory. Two problems arise with this method. First, speaker-specific characteristics (arising from anatomy and personal style) are confounded with linguistic properties. This is especially problematic when utterances produced by only a few speakers are taken to represent a language. Second, speech is a product of abstract linguistic/phonetic goals and concrete articulatory constraints on vocal movements, and it is difficult to separate the goals from the constraints.

The strategy normally taken to overcome these difficulties with acoustic analysis is to record many utterances produced by a large number of speakers. A critical mass of acoustic data which includes variation in speakers and in speaking styles makes it possible to pull apart speaker and situational variability to some extent, but this sort of critical mass is often difficult to acquire in a field situation. Therefore we have been experimenting with a new approach to the problem of separating the effects of speaker and language, and the effects of linguistic targets and the contingencies of speech production (Johnson, Flemming & Wright, 1993). Our strategy is to have listeners adjust some parameters of a speech synthesizer until a synthetic speech sound matches a particular sound in the language. The results of this method of adjustment perceptual task are speaker-independent in the sense that each listener hears the same synthetic 'speaker'. It is still necessary to use a critical mass of listeners, but all listeners, even those with very different vocal tract shapes, react in a similar way to the synthetic voice, so only a small number of listeners is required. In addition, previous research (Johnson et al., 1993) suggests that when listeners perform the task they tend to 'undo' reduction processes found in normal speech production.

---

* Department of Linguistics, Ohio State University, 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210-1298.

This last claim needs a little elaboration. Johnson et al. (1993) found that American English listeners in a method of adjustment perception task chose vowel sounds which were different from the vowels they produced in normal speech. The vowels chosen in the task had more extreme formant values expanding the perceptual vowel space relative to the measured acoustic vowel space. Johnson et al. also showed that this expanded perceptual vowel space corresponded to vowels produced in hyperarticulated or 'clear' speech. Several experimental manipulations (using groups of listeners with different linguistic training, giving different instructions to the listeners, and varying the number of redundant cues available in the stimuli) had no impact on the results, indicating that the vowel space expansion observed in the task was not an artefact. This finding was interpreted as evidence that the method of adjustment task taps abstract (though physically definable) linguistic goals. Of course, data acquired in the method of adjustment task do not (and cannot) replace the information obtained by acoustic analysis. For instance, the perception task reveals nothing about dynamic aspects of vowels, patterns of vowel reduction in normal speech, or differences between men's and women's speech. Still, this technique does provide a useful control over two confounding variables in traditional acoustic analysis, and thus supplements such an analysis.

One purpose of the present study was to explore the applicability of the method of adjustment in a field situation. Our previous research involved a rather limited group of listeners (university students) so we had no way of knowing how listeners less oriented toward test-taking would perform the task. They might not find the listening task sufficiently interesting to be able to complete it; or there might be other cultural factors which made it difficult for them to understand what to do. Consequently it was not clear whether the method of adjustment would be useful in field phonetics.

Beyond these methodological questions we were interested in the effects of contextual variability on perceptual vowel categories. Miller & Volaitis (1989) found that goodness rating functions, a measure analogous to the method of adjustment, for voice onset time (VOT) are sensitive to rate of speech. This evidence that the internal structure of a perceptual category can be modified by phonetic context challenges the conclusion reached by Johnson et al. (1993) who did not study whether varying the phonetic context of the vowel to be matched had any effect. In this study of Montana Salish, we explored the effect of phonetic context in the method of adjustment task by investigating three variants of /o/; stressed, unstressed and contextually labialized.

Method

The study was conducted in the context of a brief two day visit to the Cultural Center on the Flathead Reservation in Montana, thanks to the assistance of the Director of the Flathead Culture Committee, Clarence Woodcock. The principal purpose of the visit was to record an extensive word list illustrating the phonological contrasts of Montana Salish. The present experiment was conducted during intervals in this process with the help of the three main language consultants: Harriet Whitworth, Felicite ("Jim") McDonald, and Dorothy Felsman, to whom we are very grateful.

Listeners

Three female native speakers of Montana Salish served as listeners in the study. All three were also fluent speakers of English. Their ages ranged from mid 50s to mid 70s.

Materials

Words illustrating the five vowels of Montana Salish in stressed syllables were selected (Table I). Table I lists two additional words; in one the vowel /o/ appears after a labialized consonant and in the other /o/ is in an unstressed syllable. These words were included in the list to investigate the role of stress and contextual labialization on perceptual vowel quality. Table I shows the order in which the words were presented. If lack of familiarity with the task has any impact on the listener's performance we expect this familiarity effect to be largest for the first words in the list.

Table I

Words illustrating the vowels of Montana Salish. The first column lists the vowel which the word was used to illustrate, the second column lists the words in Montana Salish orthography, and the third column lists the English glosses. Words are arranged in the order in which they were presented in the perceptual task.

/é/	č'upélsi	'lonesome'
/í/	píləlš	'a stagger, stumble'
/ʷo/	mxʷól'	'cradle-board'
/ɑ́/	tɑ́m	'it's not, wrong'
/ó/	pólplqən	'thimbleberry'
/ú/	púlsəm	'he killed something'
/o/	olín	'belly'

The computer-produced synthetic vowel sounds used in the experiment have been described in more detail in an earlier paper (Johnson et al., 1993). A set of 298 separate vowel sounds were synthesized using a software formant synthesizer (Klatt & Klatt, 1990). The first vowel formant (F1) varied from 250 Hz to 900 Hz in 15 equal steps, while the second vowel formant (F2) varied from 800 Hz to 2799 Hz in 22 equal steps. The step-sizes were calculated on a psychoacoustic scale (the Bark scale), so they represent equal auditory distances rather than equal acoustic distances. F1 was constrained to be at least 300 Hz below F2; consequently not all possible combinations of F1 and F2 values (e.g. F1 = 900 Hz and F2 = 800 Hz) were used. A regression formula (Nearey, 1989) was used to calculate the value of F3. F4 was fixed at 3500 Hz unless F3 was above 3200 Hz in which case F4 was 300 Hz higher than F3. The bandwidths of F1-F3 were also calculated by regression formulas derived from values used to synthesize vowels in English. The fundamental frequency was 120 Hz with a slight fall at the end of the 250 ms long stimulus.

Procedure

The experiment was run on-line using a Macintosh Powerbook 140. The visual display presented to the listeners was composed of a grid of (15 X 22) 330 square buttons on a Hypercard card and the instruction, for example 'Find the é vowel of č'upélsi', was printed at the bottom of the screen. The words were presented in the order shown in Table I. Clicking on one of the squares caused the corresponding stimulus from the F1/F2 vowel array to be played to headphones or a separate amplifier and speaker. The sampling rate was 22 kHz with 8 bit quantization.

The listener's task was to search through the stimuli to find the computer-synthesized vowel that most closely matched the sound of the specified vowel in that word. Even though the listeners had no previous experience with a computer, all three were quickly able to perform the task. We suggested a strategy to the listeners; namely that they start each trial by sampling different quadrants of the grid and then continue to work within a quadrant to find the preferred stimulus. HW (the first listener) had some difficulty in controlling the cursor with a trackball, so the other two listeners used a mouse. Also, the stimuli were presented over a loud speaker to the first two listeners, but we considered that room noise and the experimenter's reactions to the stimuli might impact performance so listener DF heard the stimuli over headphones. Each listener performed the task once for each of the words listed in Table I, except HW, who had to skip the last word.

Results

Before discussing the results it is important to note that this is a preliminary study. As the method section indicated, we collected a very scant amount of data; only one repetition of the perceptual task for the seven example words, with only three speakers of the language. Therefore, this section is merely a record of how we went about trying to use a new perceptual technique, what lessons we learned, and what types of linguistic phonetic questions we think can be answered using this method. The results suggest that a more extensive study would indeed shed valuable light on the phonetic structure of Montana Salish.

The results are shown in Figure 1. The ellipses show 65% bivariate confidence intervals for the mean formant values. Between-listener variance was greatest for the vowels /i/ and /ɑ/, and was also rather large for /o/ if the different allophones are pooled together. On the other hand, the listeners chose very similar values for the vowels /u/ and /e/. Note that, with the exception of /i/, the range of formant values observed in this experiment was comparable to or smaller than the measured male formant values reported by Flemming, Ladefoged, & Thomason (1994).

If the words illustrating /i/ and /ɑ/ had been presented first in the list we could have attributed the between-listener variation to lack of familiarity with the task. However, the first word in the list illustrated /e/ (see Table I), and, as the figure shows, the listeners chose very similar formant values for this vowel. It is therefore not obvious that the large variability seen for /i/ and /ɑ/ is the result of uncertainty on the first few trials in the experiment. It may also be important that the listeners did not generally disagree with each other; their choices for the vowels /e/ and /u/ were similar even though listening conditions (loudspeaker vs. headphones) and response methods (trackball vs. mouse) differed. This suggests that the method of adjustment task may be used in field situations to provide a concise speaker-independent view of a language's vowel space. However, before we can confidently rely on the method of adjustment it is necessary to perform the same experiment with several other speakers and to have each speaker repeat the task several times for each word in order to get some idea of the range of between- and within-listener variation one is likely to encounter.

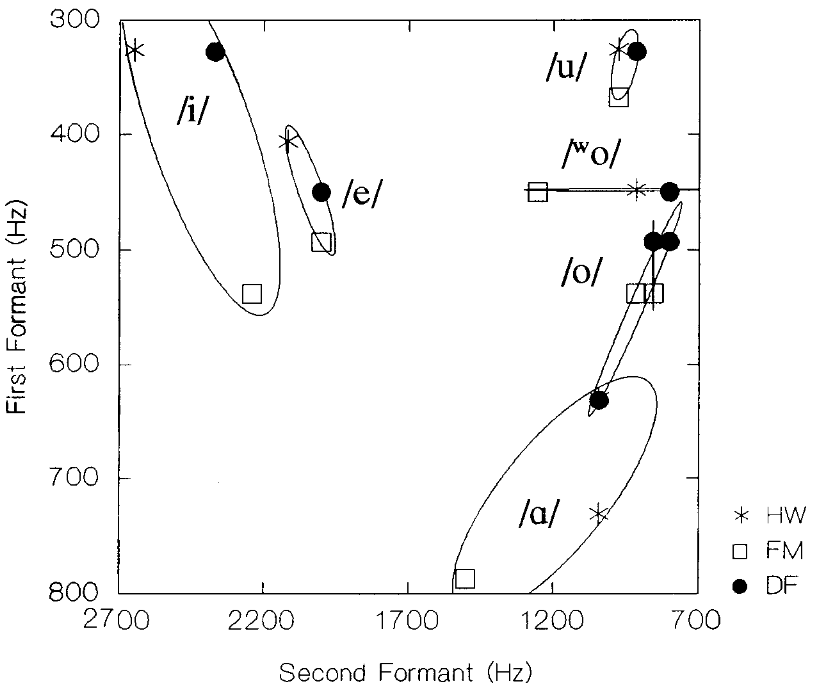


Figure 1. Results of method of adjustment study of Montana Salish vowels. Formant values of the synthetic vowels chosen for the words listed in Table I are plotted by listener.

We were also interested in the three variants of /o/. The stressed and unstressed variants are not separately labeled in Figure 1 because they overlap with each other. The stressed /o/'s are connected by the diagonal ellipse and the unstressed /o/s are connected by the vertical ellipse/line. This overlap of stressed and unstressed versions of /o/ indicates that even though vowel quality is (to our ears) much reduced in unstressed syllables in normal productions, listeners tended to equate the stressed and unstressed versions of /o/ in the perceptual task. Interestingly however, HW and FM seem also to have maintained a distinction between two other variants of /o/; one with a preceding labialized consonant and one without. Although DF chose a lower F1 for /ʷo/, her choices for the three versions of /o/ had very similar formant values. So, HW and FM seem to have kept /ʷo/

and /o/ separate while DF did not. This difference between the speakers is correlated with a difference in how they listened to the synthetic vowel sounds. HW and FM listened to the sounds over a loudspeaker, in a room that was not always quiet, while DF listened to the sounds over headphones and in a less busy room. Obviously, further research is needed before any conclusion can be drawn.

Discussion

From this very brief study we learned several methodological lessons. (1) The method of adjustment is a feasible tool for phonetic fieldwork. The listeners seemed to find the task at least moderately interesting (as compared with the complete and total boredom induced by more common speech perception tasks), and the results were provocative and interpretable. (2) To get usable data it will be necessary to collect several repetitions of each vowel from each speaker. Some estimate of within-speaker variation is needed, and we need to see if variability decreases as experience with the perceptual task increases. (3) A hearing screening is needed. Our listeners were above 50 years of age and may have had some hearing loss. (4) Headphones are preferable to loudspeakers.

The results of the present experiment on the different allophones of /o/ are provocative. Miller & Volaitis (1989) reported an effect of speaking rate on the internal structure of phonetic categories. Johnson & Armstrong (in preparation) also found an effect of F0 on the vowel space obtained in the method of adjustment. The effect of contextual labialization found in this study is, however, of a different sort. Speaking rate and F0 (or speaker identity) persist over long stretches of speech, while coarticulatory effects have much smaller domains (see Johnson, 1991 for an example of the differential perceptual effects of these two types of phonetic 'context'). Therefore, it is perhaps not surprising that the internal structure of a perceptual category might be affected by global nonlinguistic properties such as rate and speaker. However, if the method of adjustment taps a somewhat abstract level of representation, as suggested by Johnson et al. (1993), we have to consider why contextual labialization in the visually presented word could have an effect on the method of adjustment result (see also Volaitis & Miller, 1992).

The overlap of stressed and unstressed /o/ in the listeners' responses suggests that these two versions of the phoneme are considered to be the "same" and may have the same phonetic target (Johnson, et al., 1993). The pattern shown by listeners HW and FM, who kept /o/ distinct /ʷo/ leads to the conclusion that either the variant of /o/ in labialized context has become phonemicized in Montana Salish or that this experimental technique taps knowledge at a subphonemic level. The pattern shown by listener DF suggests that more data is needed before any clear answer will emerge. It is possible that some listeners match sounds at a somewhat abstract level of representation, whereas others pay more attention to phonetic details.

References

Flemming, E., Ladefoged, P. & Thomason, S. (1994) Phonetic structures of Montana Salish.

Johnson, K. (1991) Differential effects of speaker and vowel variability on fricative perception. *Language and Speech* **34**, 265-279.

Johnson, K. & Armstrong, A.A. (in preparation) Memory for speaker-specific vowel information in speech.

Johnson, K., Wright, R. & Flemming, E. (1993) The hyperspace effect: phonetic targets are hyperarticulated. *Language* **69**: 505-528.

Klatt, D. & Klatt, L. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J. Acoust. Soc. Am.* **87**, 820-857.

Miller, J.L. & Voliatis, L.E. (1989) Effects of speaking rate on the perceived internal structure of phonetic categories. *Percept. Psychophys.* **46**, 505-512.

Nearey, T.M. (1989) Static, dynamic & relational properties in vowel perception. *J. Acoust. Soc. Am.* **85**, 2088-2113.

Thompson, L.C. (1979) Salishan and the Northwest. In *The Languages of Native America: Historical and Comparative Assessments*. (L. Campbell & M. Mithun, Eds.) Austin: Univ. Texas Press.

Volaitis, L.E. & Miller, J.L. (1992) Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *J. Acoust. Soc. Am.* **92**, 723-735.

Phonetic Structures in Jalapa Mazatec

Daniel Silverman, Barbara Blankenship, Paul Kirk, and Peter Ladefoged

0. Introduction

Mazatecan belongs to the Popolocan branch of the Otomanguean language family (Gudschinsky 1958, Grimes 1988). It is composed of twenty three speech communities which are spoken in Mexico by approximately 125,000 people in the northeastern section of the state of Oaxaca, as well as in southern Puebla and western Veracruz.

This paper describes the phonetic inventory of the dialect spoken in San Felipe Jalapa de Díaz (henceforth Jalapa Mazatec). Jalapa Mazatec is spoken in the vicinity of San Felipe Jalapa de Díaz, district of Tuxtepec, by 6000-8000 people, many of whom are also fluent in Spanish. A practical orthography has been developed for Jalapa Mazatec, and some Christian texts and native folk tales have been published. Both Mazatec and Spanish reading are taught in local schools.

Data for this study were collected in March and April, 1993, in Jalapa de Díaz. Six male and six female native speaking adults were recorded. The primary corpus was a list of 335 words spoken in isolation. Two of the male speakers recorded the complete set on two occasions, giving us fourteen sets of recordings in all. Additionally, palatographic data were collected from one speaker. Airflow data for selected phonemes were also recorded.

This paper is organized as follows. Section 1 provides a brief overview of Jalapa Mazatec phonology and morphology. Section 2 presents the vowel system. Section 3 presents the consonant system.

1. Brief Overview of Phonology and Morphology

The Jalapa syllable is maximally CCGV. Consonant clusters consist of either fricative-stop sequences, or homorganic nasal-voiceless stop sequences. Despite the simplicity of the syllable structure, the availability of laryngeal augmentation greatly expands possible syllable types. Obstruents may be contrastively voiced, voiceless stops may be contrastively aspirated, and sonorants may be contrastively aspirated or glottalized. Additionally, vowels obligatorily possess tone, and may be augmented by breathiness or creakiness.

There are both diachronic and synchronic reasons to believe that breathiness and creakiness are affiliated with the vowel, and not with the onset consonant. Breathy vowels cannot be posited as a feature of proto Mazatec (Kirk 1966:38-44). Thus, Jalapa Mazatec is unique in that it is the only one of the twenty-three distinct Mazatec speech communities that developed breathy vowels. The breathy vowels in Jalapa Mazatec developed from Proto-Mazatec (PMaz) disyllabics of the form *-V.hV in which the laryngeal /h/ consonant margin of the second syllable was retracted through the vowel of the first syllable and the vowel of the second syllable coalesced with this first vowel. The development of these breathy vowels was conditioned by two further factors: (1) the syllable margin of the first syllable of these disyllabics had to be voiced, and (2) the vowels in both syllables contiguous to the /h/ had to be identical. If the tones in these PMaz disyllabics were the same in both syllables, they were reduced in length to those of a single tone; if the tones in these PMaz disyllabics were different, then they were coalesced into a tone glide. In present day Jalapa Mazatec whistle speech, these PMaz disyllabics are whistled with a single whistle pulse. Examples of the development of Jalapa Mazatec breathy vowels can be seen from the following PMaz cognate sets (cognate set numbers are from Kirk 1966 with tones indicated by the format of this paper) PMaz 406 *ntja^{1}hu^1 ‘stone’ > early Jalapa Mazatec *ndjo^{1}ho^1 > present day Jalapa Mazatec [ndjo̤^{1}]; PMaz 301 *ntʃe*^{2}he^4 ‘thief’ > early Jalapa Mazatec *ndʒæ^{1}hæ^2> present day Jalapa Mazatec [ndʒæ̤^{12}]; PMaz 400 *ntu^{1}hwi^{13} ‘your soap’ > early Jalapa Mazatec *ndu^{1}hwi^{12} > present day Jalapa Mazatec [ndi̤i̤^{12}]; PMaz 329 *nĩ^{1}hĩ^1 ‘dry ear of corn’ > present day Jalapa Mazatec [nĩ̤^{1}]. For a number of other cognate sets exhibiting the development of Jalapa Mazatec breathy vowels see Kirk 1966:45-46.

In a similar way present day Jalapa Mazatec contrasts disyllabics VʔV with monosyllabic creaky vowels. Compare [tʃu^{1}ʔu^2] 'bedbug' with [tsṵ^{3}] 'blouse'. However, a large number of PMaz disyllabics of the form *-VʔV have been reduced to monosyllabics with creaky vowel in Jalapa Mazatec. The [ʔ] retracted through the first vowel of the proto disyllabic, producing a creaky vowel with coalescence of the vowel in the second syllable. The development of these creaky vowels is thus parallel to the manner in which breathy vowels developed in Jalapa Mazatec with the retraction of [h]. The development of present day Jalapa Mazatec creaky monosyllabic vowels from PMaz disyllabics is evidenced by the following cognate sets: PMaz 45 *tsu^{4}ʔu^4 'blouse' > [tsṵ^{3}], PMaz 40 *tsĩ^{4}ʔã^3 'penis' > [tsḭ̃ã̰^{2}], PMaz 65 *tʃa^{2}ʔu^2 'bad' > [tʃo̰^{2}], PMaz 174 *kã^{4}ʔã^3 'alone' > [kã̰^{2}]. For further cognate sets exhibiting this development pattern see Kirk 1966:51.

Jalapa Mazatec roots are predominantly monosyllabic, while the rich inflectional system is by and large subsyllabic. Given this dichotomy between segmental poverty and morphological richness, each segment of the Mazatec word potentially accommodates a great number of linguistically significant components.

Further information about Mazatec phonology (historical and synchronic) and morphology can be found in Pike and Pike 1947, Gudschinsky 1958, Kirk 1966, Kirk 1970, Schram and Pike 1978, Schane 1985, Steriade 1993, and Silverman *in prep.*

2.0 Vowels

Jalapa Mazatec contains a basic five vowel system, as shown in (1).

(1)

i	**u**
	o
æ	**a**

Minimal pairs exemplifying these contrasts are presented in (2).

(2)

si^2	'dirty'
sæ^2	'he sings'
sa^3	'moon'
so^2	'you (pl) sing'
su^2	'lukewarm'

Tonal, phonatory, nasal, and length contrasts greatly expand the vowel inventory. Tone obligatorily accompanies every vowel. There are three tones, low (1), mid (2), and high (3). Some tonal contours (12, 32 23, 21, 31, 131) have also been recorded. These are found primarily in morphologically complex environments. In (3) are examples of each level lexical tone pattern.

(3)		
	ʃa^{3}	'work'
	ʃa^{2}	'mountain lion'
	ʃa^{1}	'mould'

Breathiness or creakiness may accompany Mazatec vowels. In either case, non-modal phonation is most prominent in the first portion of the vowel. Examples of breathy and creaky vowels follow.

(4)	BREATHY VOWELS		CREAKY VOWELS	
	ŋgi̤^{2}	'he went'	**sḭ**^{3}	'holiday'
	jæ̤^{1}	'boil' (noun)	**t**^{h}**æ̰**^{2}	'sorcery'
	ki^{2}**ŋga̤**^{23}	‘he fastened’	**tʃa̰**^{3}	'load, burden'
	ki^{2}**ŋgo̤**^{2}	‘you (pl) will fasten’	**tʃṵ**^{3}	‘blouse’

Breathy vowels do not co-occur with voiceless stop onsets. In such environments however, a contrastive glottal abduction may occur at stop release, resulting in an aspirated stop followed by a modally phonated vowel.

All five vowels may be lexically oral or nasal. Vowels following nasal consonants are obligatorily nasal. Examples of both contrastively and redundantly nasalized vowels are shown in (5).

(5)	CONTRASTIVELY NASALIZED VOWELS		REDUNDANTLY NASALIZED VOWELS	
	sĩ^{2}	'tasty'		
	sæ̃^{1}	'ghost'	**nĩ**^{1}**m̥æ̃**^{3}	'corn'
	sã^{2}	'acid, sour'	**mã**^{2}	'is able'
	sõ^{21}	'song'		
	sũ^{2}	'level, on'		

Phonatory and nasal contrasts cross-classify. That is, a given vowel may be both breathy and nasal, or both creaky and nasal. Some examples are given in (6).

(6)	BREATHY NASALIZED VOWELS	
	nĩ̤^{1}	'ear of corn'
	^{n}**dʒæ̤̃**^{2}	'visibility'
	nã̤^{23}	'my tongue'
	nõ̤^{2}	‘your (pl) tongue’
	^{n}**dʒṳ̃**^{2}**nĩ**^{1}	'tomorrow'

CREAKY NASALIZED VOWELS	
ʃḭ̃^{1}	'man'
k^{w}a^{2}sæ̰̃^{2}	'he entered'
kã̰^{2}	'single, widowed'
tʃõ̰^{1}	'lightning'
nṵ^{3}	'vine'

Finally there is the possibility that there is a ballistic/controlled contrast in Jalapa Mazatec syllables. Previously thought limited to related Chinantec and Amuzgo, ballistic syllables have been variously described as possessing a fortis release of syllable-initial consonants, with a surge of energy, culminating in a weakened, breathy release (see Merrifield 1963, Bauernschmidt 1965, Mugele 1982, Silverman, 1994, *in preparation*). Controlled, or plain syllables do not possess these characteristics. Schram and Schram (personal communication) have pointed out to us a number of contrasts in Jalapa that might be worth investigating to see whether they manifested the ballistic/controlled contrast.

(7)	BALLISTIC SYLLABLES		CONTROLLED SYLLABLES	
	***su^{2}**	'warm'	**su^{2}**	'blue'
	nĩ^{2}*ntu^{2}	'slippery'	**ni^{2}ntu^{2}**	'needle'
	***tsæ^{2}**	'guava'	**tsæ^{2}**	'full'
	***hũ^{2}**	'you (pl)'	**hũ^{2}**	'six'

2.1 Modal Vowels

Two tokens of each vowel were recorded for each male speaker and each female speaker. Words from which formant structures were culled are listed in (8).

(8)				
	si^{2}	'dirty'	**hi^{2}**	'you (sg)'
	sæ^{2}	'he sings'	**hæ^{2}**	'finished'
	sa^{3}	'moon'	**ha^{3}**	'men'
	so^{2}	'you (pl) sing'	**ho^{2}**	'two'
	su^{2}	'lukewarm'	**tʃu^{1}tu^{2}hu^{2}**	'dove'

Figures 1 and 2 show vowel formant plots for two tokens each by three male and five female speakers, respectively. Note the wide distribution of F2 values for *u* in female speakers. This suggests that either lip rounding, tongue backness, or both, are varying in the production of this vowel.

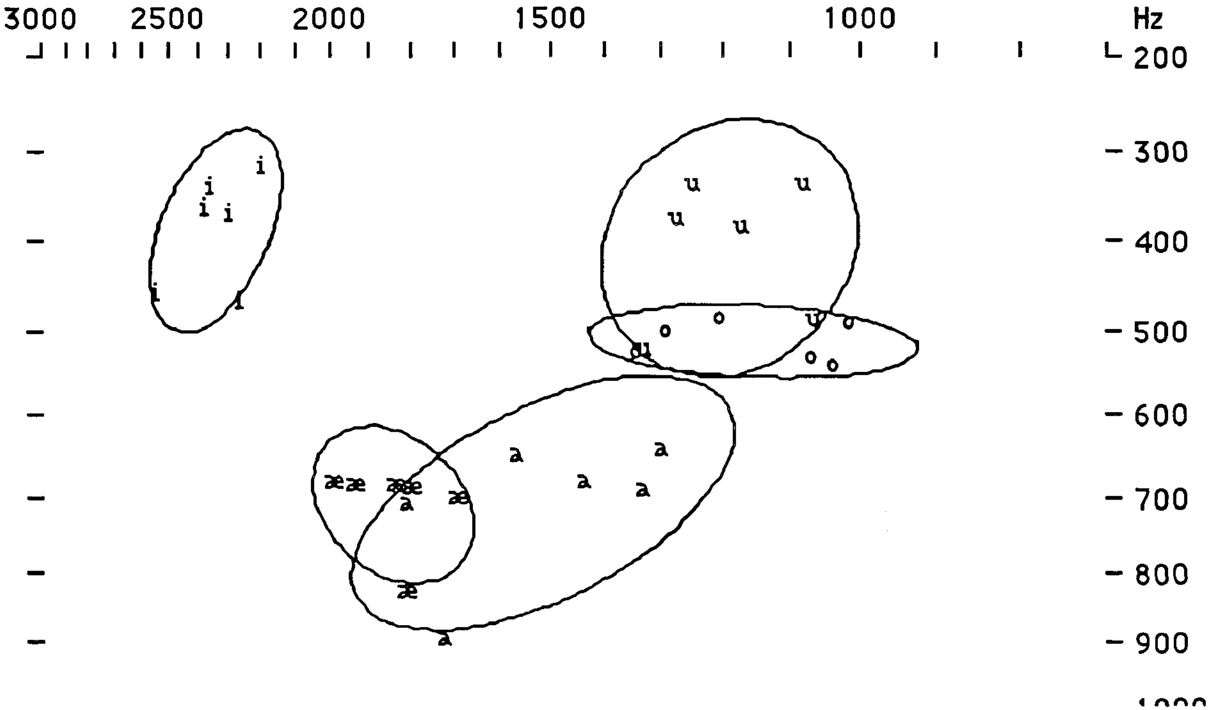


Figure 1. F1 versus F2 for two tokens each by three male speakers.

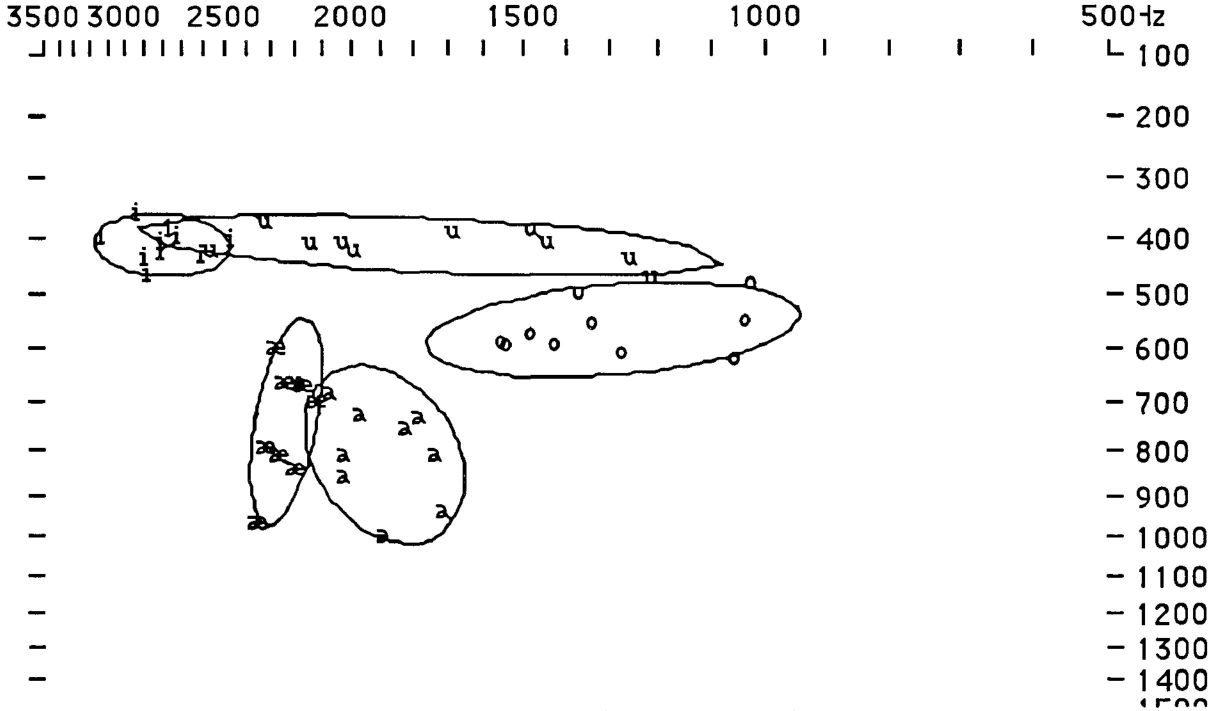


Figure 2. F1 versus F2 for two tokens each by five female speakers.

2.2 Nasalized Vowels

Vowel nasalization has proven difficult to characterize acoustically. In addition to a general weakening of F1, nasalized vowels have been reported to possess both a resonance (or nasal pole) and an anti-resonance (or nasal zero). However, the frequency of nasal poles and zeros seemingly display cross-linguistic, cross-speaker, and cross-vowel quality variation (see Smith 1951, House and Stevens 1956, Hattori, Yamamoto, and Fujimura 1956, Fant 1960, and Fujimura 1962). As we have noted, all five plain and breathy vowels have nasal counterparts, but we will leave the determination of how these contrasts are made to a later analysis.

2.3 Breathy Vowels

Jalapa Mazatec breathy vowels manifest their breathiness primarily during the first portion of the vowel. Following this period, breathiness is substantially reduced. This description holds true for the majority of recorded tokens, although there are exceptions. In some instances strong breathiness continues for the duration of the vowel. In others, voiceless aspiration is followed by modal phonation.

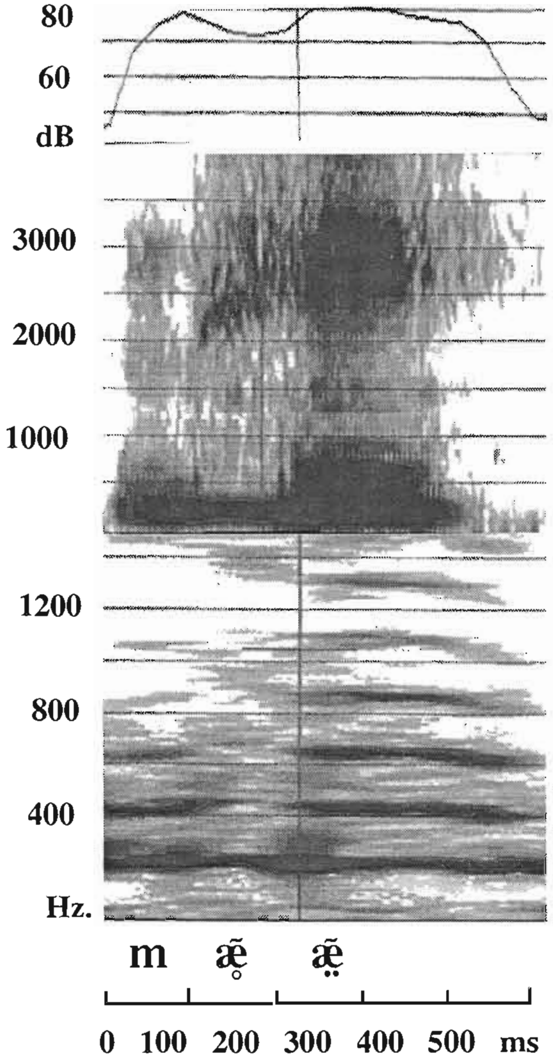


Figure 3. Energy contour, wideband, and narrowband spectrograms for mæ̤̃^{21},‘he wants’

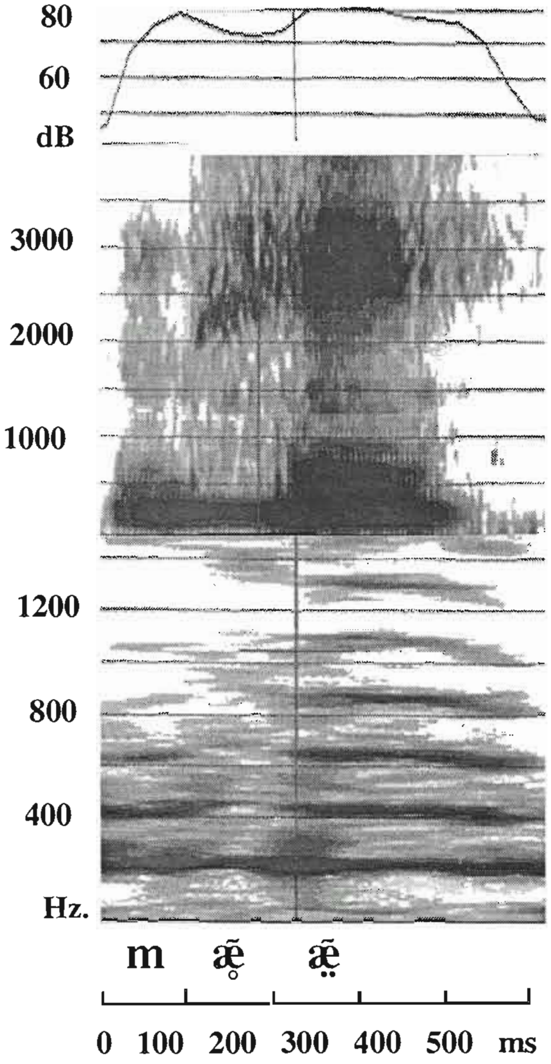


Figure 4. Energy contour, wideband, and narrowband spectrograms for nda̤^{23}, ‘hard’

Figures 3 and 4 provide examples of both wideband and narrowband spectrograms of canonical breathy vowels, along with energy contours. The narrowband spectrogram shows that strong breathy phonation often weakens the harmonic structure, making the harmonics less prominent than in the latter portion of the vowel. The point of transition from strong breathy phonation to weak breathy phonation was determined by analyzing narrowband spectrograms in conjunction with energy contours, under the assumption that overall energy is greater during weak breathy phonation than during strong breathy phonation. This assumption is confirmed by observing that the onset of a salient harmonic structure coincides with an overall energy increase. Note finally that weak breathy phonation is often accompanied by a moderate increase in fundamental frequency. The subglottal pressure falls during the glottal abduction that results in breathiness. This reduction in subglottal pressure may account for both the decreased intensity and the moderate decrease in pitch which accompany breathy phonation, as subglottal pressure correlates positively with both intensity and pitch.

Table 1 presents data from eleven speakers (six male, five female) producing eight forms with breathy vowels, resulting in 88 tokens. The first column lists the mean duration (in ms.) of the strong breathy portion of these vowels for each of the eight forms. The second column lists the mean duration of the weak breathy portion for each of the eight forms. Finally, the third column shows the percentage of vowel duration which possesses strong breathy phonation. The bottom row lists means across both speakers and forms. These figures show that strong breathiness typically persists for a little less than half the duration of the vowel as a whole.

Table 1. The extent of breathy phonation in phonologically breathy vowels.

WORD	BREATHY DURATION (ms)	TOTAL DURATION (ms)	% BREATHY
ᵑgi̤^{2}	136.4	306.2	37.6
ki^{2}ŋga̤^{23}	096.2	225.3	42.7
ki^{2}ŋgo̤^{2}	096.3	247.2	39.0
ⁿdja̤^{1}	111.7	262.4	42.6
ⁿda̤^{2}a^{3}	108.4	290.0	37.3
no̤^{2}	107.7	263.4	40.8
βo̤^{2}	119.5	228.8	52.2
ⁿdṳ^{2}	099.3	223.5	44.3
MEAN	1094	255.6	42.8

2.4 Creaky Vowels

Creaky vowels, like breathy vowels, manifest their non-modal phonation primarily during the first portion of the vowel. Often, auditory impression suggests a full glottal closure marking the end of this creaky period, followed by modal phonation. However, wideband and narrow band spectrograms do not always reveal these characteristics.

Ladefoged, Maddieson, and Jackson (1988) showed that creaky vowels have a characteristic spectral tilt: compared to modally phonated vowels, the amplitude of H1 in creaky vowels is reduced relative to that of H2. The amplitude of H1 and H2 were measured, both during the initial strong creaky portion of the vowel, as well as during its latter portion. Data were collected from ten speakers (four male and six female), each uttering six distinct tokens, for a total of 60 forms measured. The difference H2 - H1 was subsequently calculated for each token, both during the initial portion and the latter portion. Table 10 presents these differences for each form, averaged across speakers. The bottom row lists averages across both speakers and forms.

Table 2. The relative intensities (in dB) of the first and second harmonics in creaky vowels.

word	H2 - H1 INITIALLY	H2 - H1 FINALLY	DIFFERENCE
sḭ^{3}	10.3	.03	10.27
tʰæ̰^{2}	6.27	-1.25	7.52
tʃa̰^{3}	6.99	-3.66	10.65
tʃõ̰^{2}kũ^2	9.30	5.48	3.82
tsṵ^3	10.03	3.6	6.43
ndæ̰^{1}	4.58	-3.28	7.86
mean	7.86	.15	7.71

During the initial portion of the vowel, the difference in amplitude between H2 and H1 across forms varies between 4.58 and 10.3 dB. During modal phonation, this difference varies between 5.48 and -3.66 dB. But within forms, H2 - H1 initially is always greater than H2 - H1 finally; the difference between these differences varies between 3.82 and 10.27 dB, for an average of 7.71 dB. These results support the findings of Ladefoged, Maddieson, and Jackson (1988), in that creakiness possesses a characteristic spectral tilt involving a weakening of H1 relative to H2. Our impression is that the creaky portion of the vowel is, like the breathy portion, a little under half the total duration, but we do not have a reliable way to determine the duration of creakiness in comparison to the duration of the vowel as a whole.

2.5 Ballistic Syllables

As mentioned earlier, there is a possibility that the ballistic syllable phenomenon, which has been reported in the related Otomanguean languages of Chinantec and Amuzgo, might also occur in Jalapa Mazatec. Descriptively, ballisticity in these languages has been described as a prosodic phenomenon affecting the entire syllable. Its primary features include:

1) fortis release of syllable-initial consonants
2) a gradual surge and rapid decay in intensity
3) post-vocalic aspiration

Mugele (1982) proposes that ballisticity is phonologically characterized by an increase in subglottal pressure. In contrast, Silverman (1994, in prep.) offers phonetic and phonological evidence that a laryngeal abduction is the phonologically relevant articulatory gesture.

Previously, ballisticity has not been reported in Mazatec. However, Schram and Schram (p.c) suggested to us that the forms presented in (9) were worth investigating from this point of view. For ease of discussion we will refer to one set as having ballistic syllables and the other set as having controlled syllables. As there are no IPA symbols for this distinction, we will use an asterisk before a ballistic syllable. Note that this usage is unrelated to the convention of denoting a reconstructed form by an asterisk, as was done in section 1 above.

(9)

BALLISTIC SYLLABLES		CONTROLLED SYLLABLES	
***su^2**	'warm'	**su^2**	'blue'
ni^2*ntu^2	'slippery'	**ni^{2}ntu^2**	'needle'
***tsæ^2**	'guava'	**tsæ^2**	'full'
***hũ^2**	'you (pl)'	**hũ^2**	'six'

We now consider in turn each of the three primary descriptive features of ballisticity with respect to the forms in (9).

1) Fortis release of syllable-initial onsets. While the supposed fortition which accompanies ballisticity remains an impressionistic description with no reliable articulatory or acoustic correlates, there are two likely interpretations: a) syllable onset duration b) VOT.

a) Onset consonant duration. Table 3 shows the onset consonant duration for both ballistic and controlled syllables, averaged across three speakers.

Table 3. Durations of syllable onsets (ms).

	BALLISTIC	CONTROLLED
su	207.2	152.3
nĩntu	25.6	45.9
tsæ	69.2	45.9
hũ	196.3	170.0

As can be seen, there is no systematic difference between onset duration in ballistic versus controlled syllables. However, ballistic syllable onsets are marginally longer in the three monosyllabic words.

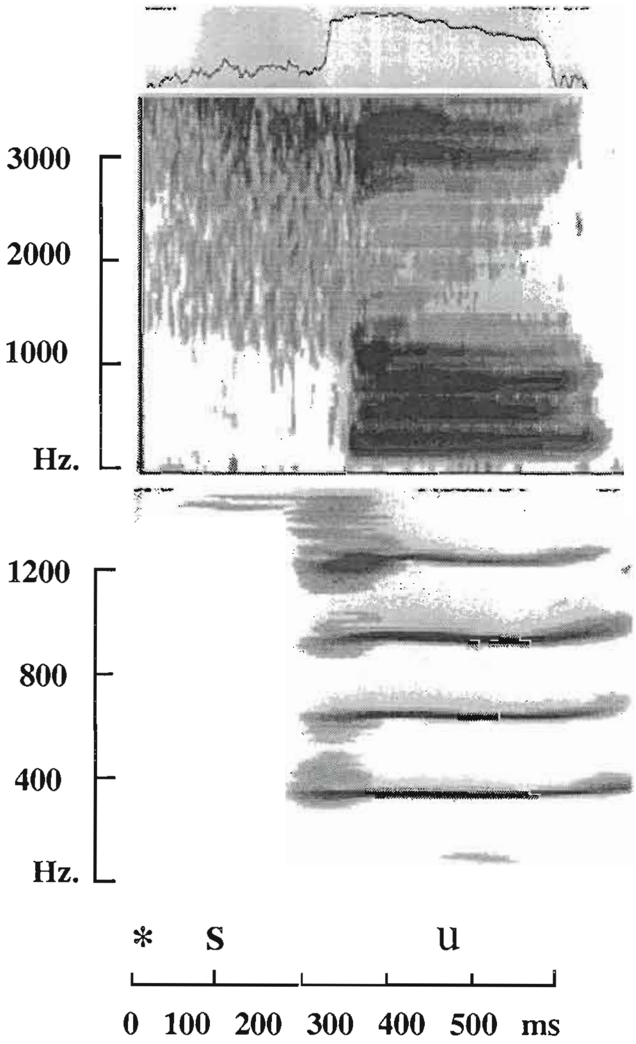


Figure 5. Energy contour, wideband and narrowband spectrograms for $*su^2$, 'blue'.

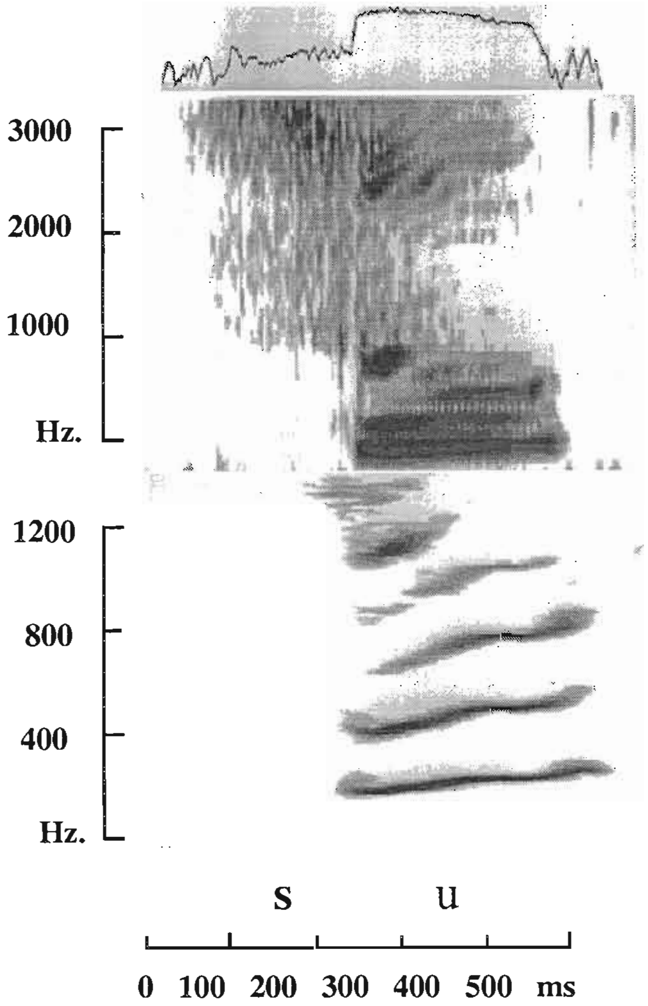


Figure 6. Energy contour, and wide and narrowband spectrograms for su^2, 'warm'.

b) VOT. Wideband spectrograms in Figures 5, and 6, and Figures 7 and 8 — pairs which near-minimally contrast in so-called ballisticity — suggest no difference in VOT between ballistic and controlled syllables.

2) Intensity. A comparison of energy contours for both ballistic and controlled syllables yielded no systematic difference. Furthermore, energy levels measured at the vowel midpoint revealed no systematic difference between ballistic syllables and controlled syllables, although ballistic syllables often possess a marginally greater intensity. Table 4 reports these values for four ballistic/controlled pairs, averaged across three speakers.

Table 4. Intensity (dB) at the vowel midpoint in ballistic and controlled syllables.

	BALLISTIC	CONTROLLED
su	83.35 dB	79.07 dB
ńintu	82.97	82.21
tsæ	82.32	82.34
hũ	82.53	82.33

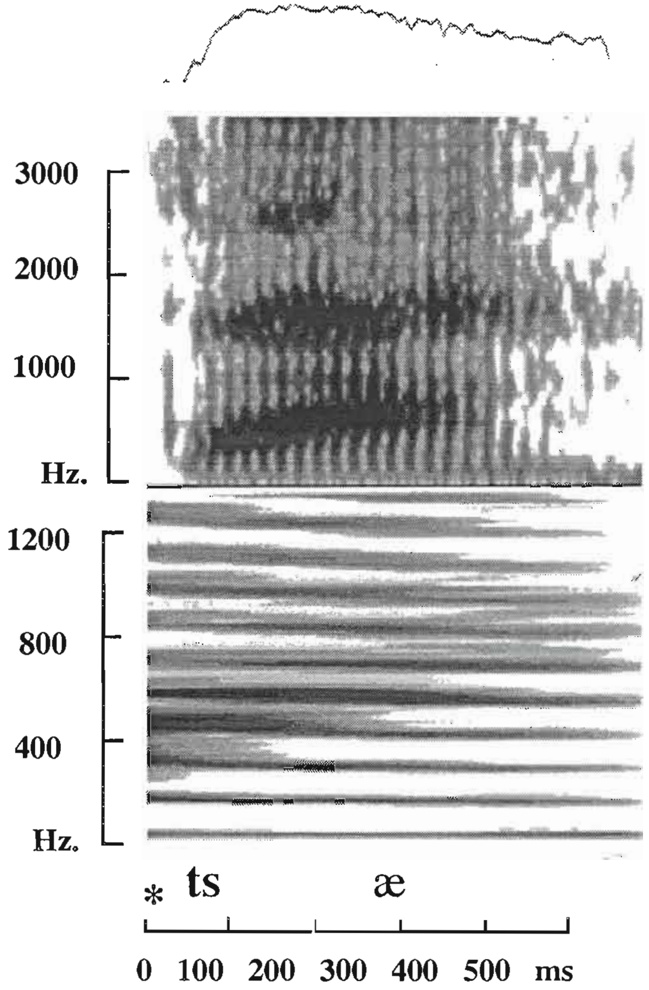


Figure 7. Energy contour, and wide and narrowband spectrograms for *tsæ^2, 'guayaba'

3) Post-vocalic aspiration. Cessation of voicing towards the end of isolated tokens is a common feature within the Jalapa Mazatec corpus. However, there is no tendency toward a greater degree of aspiration in ballistic syllables.

Thus, while native speakers distinguish the forms in (9), the distinction does not conform to the traditional ballistic/controlled contrast. Instead, these contrasts appear to involve a combination of tone and length differences. Table 5 shows the frequency in the middle of the vowel of and duration measures for the four contrastive pairs indicated, for two male speakers.

Table 5. Frequency (Hz) and duration (ms) of the vowel in four pairs of Mazatec syllables.

	BALLISTIC				CONTROLLED			
	***su**	**nĩ*ntu**	***tsæ**	***hũ**	**su**	**nĩntu**	**tsæ**	**hũ**
male 1								
F0	177	197	166	165	162	153	165	153
DURATION	16.1	10.6	14.1	17.3	23.9	17.8	21.8	24.4
male 2								
F0	183	186	179	188	167	159	165	164
DURATION	21.3	16.7	16.4	21.7	26.7	19.6	28.1	23.8

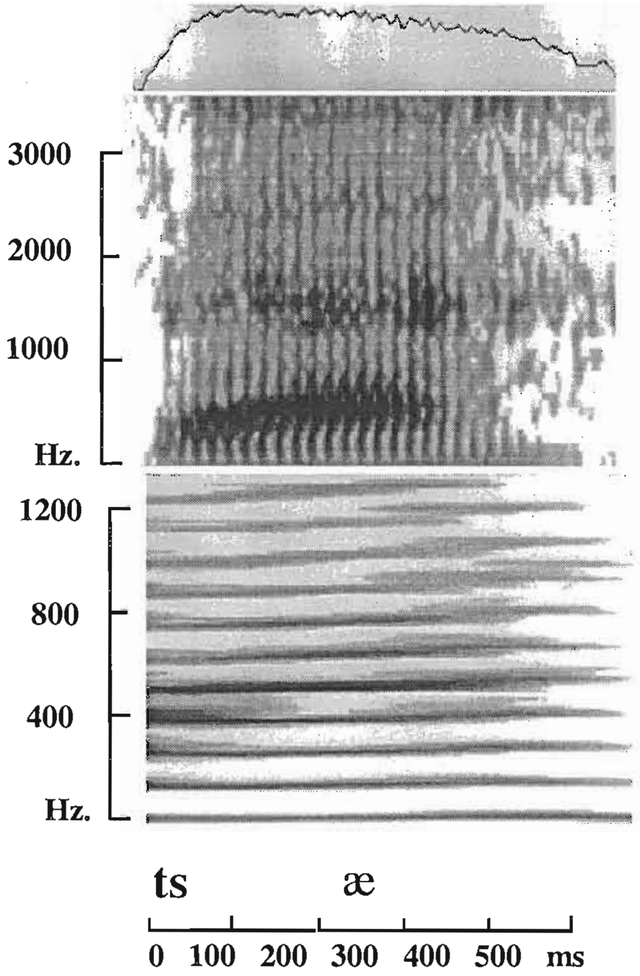


Figure 8. Energy contour, and wide and narrowband spectrograms for tsæ^2, 'full'

As can be seen, the so-called ballistic/controlled contrast apparently involves systematic differences in frequency and vowel duration. In each case the fundamental frequency is higher and the duration less in the ballistic syllable than in the corresponding controlled syllable. Schram and Pike (1978) in fact characterize the contrast in question as one involving length.

It is instructive to compare the F0 values in Table 5 to those in Table 6, which shows F0 averaged values for the three contrastive tonal patterns, High, Mid, and Low, for four female speakers and one male speaker.

Table 6. Mean F0 (Hz) for each of the three tones.

word		High	Mid	Low
tjo	FEMALES	272	232	206
tjo	MALE	217	167	132
ha	FEMALES	255	232	199
ha	MALE	220	177	128

A comparison of Tables 5 and 6 shows that tonal contrasts involve a greater difference in pitch than is in evidence in so-called ballistic syllables. Thus ballisticity seemingly does not involve a lexical contrast in pitch.

Let us then consider in more detail the possibility that ballistic syllables involve a length contrast. While vowel length contrasts are clearly evident in morphologically complex environments, the data set provides no instances of morphologically simple length contrasts. This contrasts with many other Otomanguean languages, in which vowel length is phonemic (see, e.g., Rensch 1976). If so-called ballistic syllables exemplify a length contrast, this otherwise unexplained gap is accounted for: ballistic syllables may in fact be morphologically simplex short syllables. Given that the canonical Mazatec root is monosyllabic, we might expect a prevalence of bimoraic words, as increased duration is preferred within the open class of the lexicon, so that lexical contrasts may be rendered salient (see Silverman, *in preparation*n). Observed pitch increases in ballistic syllables may enhance the salience of short roots, as pitch increases render more prominent their acompanying segmental material.

For now, the data remain too limited to draw confident conclusions regarding the so-called ballistic/controlled contrast. However, length distinctions combined with enhancing pitch increases would seem a promising approach.

3. Consonants

The consonant phonemes are shown in Table 7. The plosives occur in five places of articulation: bilabial stop (limited to loanwords), dental stop, dental affricate, palatal affricate, and velar stop. There is also a glottal stop. Each of the plosives has a three-way contrast in voice onset time (VOT) with the voiced phoneme redundantly prenasalized. Moreover, the velar series may possess a contrastive labial offglide. Similarly, the dental series may possess a contrastive palatal offglide.

Table 7. Jalapa Mazatec consonants. Segments found only in loan-words are in parentheses.

	LABIAL	DENTAL	DENTAL AFFRICATE	POST-ALVEOLAR	VELAR	GLOTTAL
PLOSIVES	(pʰ)	tʰ	tsʰ	tʃʰ	kʰ	
	(p)	t	ts	tʃ	k	ʔ
NASALS	m̥	n̥		ɲ̊		
	m	n		ɲ		
	m̰	n̰		ɲ̰		
FRICATIVES		s		ʃ		h
APPROXIMANTS	w̥			j̥		
	w	(l)		j		
	w̰			j̰		

Nasal consonants occur in three places of articulation, bilabial, dental, and palatal. Each of these can be voiceless, voiced, or glottalized. Similarly, the approximants **w** and **j** exhibit a three-way contrast between voiceless, voiced and voiced glottalized articulation. Aspirated and glottalized sonorants may only co-occur with plain vowels.

Liquids are rare. [l] occurs primarily in loan words. A tap **ɾ** is present only in the clitic **ɾa** 'probably', and is thus marginal.

There is an alternation between **ɸ** before front vowels and **w̥** beore back vowels, both in the full consonant (e.g., [ɸæ^2] ‘it is finished’, [w̥a^1] ‘Juan’) and in the labial element of the aspirated labial velar (e.g., [k^{ɸ}æ^1] ‘file’, [k$^{\text{w̥}}$a^1] ‘will happen’). The same alternation occurs between w/β and w̰/β̰ as full consonants.

A number of consonant clusters also occur. Both the aspirated and voiceless unaspirated plosives series can be preceded by homorganic nasals, giving **(mp^{h}), nt^{h}, nts^{h}, ntʃ^{h}, nk^{h}, nk$^{\text{w̥}h}$**, and **(mp), nt, nts, ntʃ, nk, nk^{w}**, There is no similar contrast between pre-nasalized and non-pre-nasalized voiced plosives. This series is voiced in initial position, and redundantly prenasalized when intervocalic. In addition to single segments, There are also voiceless consonant clusters with **s** followed by a dental, velar, or labial-velar, and **ʃ** followed by velar or labial-velar stops.

Surface clusters involving laryngeals are treated as single segments. Note that the aspirated plosives are the only laryngeally augmented onset consonants that may co-occur with creaky vowels: **tʃõ̰^{2}** (fifteen). Contrariwise, surface post-aspirated fricatives never co-occur with creaky vowels; they only co-occur with plain vowels. Given this distributional asymmetry, we regard all fricatives as plain.

3.1 Stops

The duration differences between the voiceless and the aspirated stops are summarized in Table 8. The measurements are means of 6 speakers, each saying a single word, except in the case of **k^{h}**, for which there were two words. Both the closure durations and the VOTs are comparable with those in other languages that have a three way voicing contrast. Also as is usual, the VOTs for the velar stops are longer than those for the alveolars.

Table 8. Durations of alveolar and velar medial stop elements (ms).

	Closure	VOT	Total
t^{h}	97	63	160
k^{h}	95	80	175
t	111	11	122
k	93	23	116

As we have noted, the voiced stops are prenasalized in initial position. The durations of the different portions are as shown in Table 9. Again, the measurements are means of 6 speakers, each saying a single word in the case of the bilabial stop, and two words in the case of the alveolar and the velar stops. The voiced stop portion was not always voiced throughout the oral closure; the voiced nasal portion was sometimes followed by a voiceless stop which was released with a very brief VOT (less than 10 ms).

Table 9. Durations of the oral and nasal closures of initial voiced stops (ms).

	Nasal	Oral	Total
^{m}b	66	42	108
^{n}d	68	60	128
^{ŋ}g^{w}	50	21	61

3.2 Nasals

Spectrograms of the contrasting nasals are shown in Figure 9. The voiceless bilabial nasal at the top of the figure is similar to that found in Burmese and many languages of South East Asia (Bhaskararao and Ladefoged, 1993). The first part of the nasal is voiceless, but before the articulatory closure is released, voicing commences. The first upward pointing arrow in the illustration marks this point; the second arrow indicates the release of the articulation (the opening of the lips in this case) and the start of the vowel. In our sample the voiced portion exceeded half of the total closure duration in extreme cases, but averaged around one fourth of the total. The nasal with modal voicing in the middle of the figure is similar to that found in most languages of the world.

The glottalized nasal in the lower part of the figure has a number of creaky voice pulses preceding modal voicing for the portion immediatley before the vowel. Glottalized nasals (and, indeed, all the glottalized sonorants) were very variable in their articulation. Sometimes, as in the example illustrated here, there were creaky voice pulses before the nasal. Sometimes there was a single glottal pulse and then a long glottal closure before a modally voiced nasal. On some occasions there was almost no modal voicing, and much of the nasal was pronounced with creaky voice; on other occasions the glottalization spread even further so that the neighboring segments had creaky voice. It was usually impossible to measure the glottal portion of glottalized nasals in word-initial position, but in six tokens of initial glottalized nasals (two words spoken by each of three male speakers), there was some glottal activity noticeable from 241 to 359 (mean 296) ms before the short modally voiced nasal. This long interval with glottal activity may be due to these words' being produced as citation forms.

Table 10 summarizes duration measurements for the nasals. The figures for the glottalized nasals are unreliable, because (as noted above) they may be realized in several ways. The measurements are means for six speakers each saying a single word, except in the case of **m** for which two words were available.

Table 10. Durations of nasal elements (ms).

	Voiceless Nasal	Voiced Nasal	Total	Position in Word
m̥	101	37	138	medial
n̥	73	35	108	medial
ɲ̊	81	18	99	medial
m		75	**75**	initial
n		93	**93**	initial
ɲ		97	**97**	initial
	Glottal Closure			
m̰	70	43	113	medial
n̰	75	38	113	medial
ɲ̰	107	42	149	medial

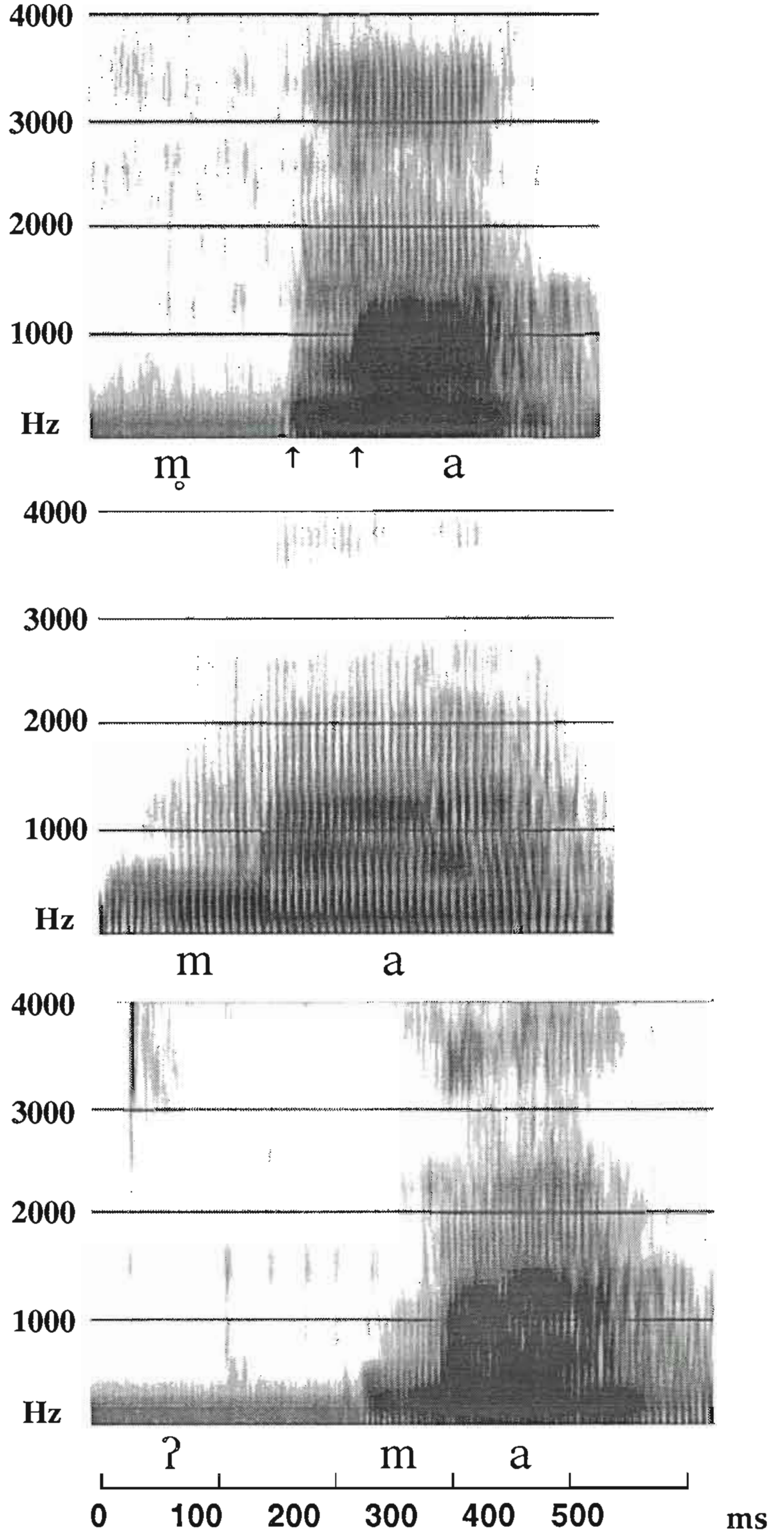


Figure 9. Spectrograms of **m̥a**^{2}, ‘black, **ma**^{2}, ‘is able, **m̰a**^{2}, ‘hidden.

4. Conclusion

The sound system of Jalapa Mazatec is unusual in possessing a great number of linguistically significant laryngeal contrasts in both its consonant and vowel inventories. Moreover, the distribution of these laryngeal contrasts within the syllable abides by unusual co-occurrence restrictions.

The present approach to Jalapa Mazatec segmentation is surely not the only viable candidate (see, for example, Pike and Pike 1947, and Steriade 1993 for alternative accounts). We nonetheless hope that our instrumental analyses provide a foundation on which future investigations—both phonetic and phonological—may be based.

References

Bauernschmidt, A. (1965) "Amuzgo Syllable Dynamics," *Language* 41-3:471-483).

Bhaskararao, P. and P. Ladefoged (1991) "Two Types of Voiceless Nasals," *Journal of the International Phonetic Association* 21.2:80-88.

Fant, G. (1960) *Acoustic Theory of Speech Production*.. Mouton, The Hague, The Netherlands.

Fujimura, O. (1962) "Analysis of Nasal Consonants," *Journal of the Acoustic Society of America* 34:1865-1875.

Gudschinsky, S.C. (1959) *Proto-Popotecan: A Comparative Study of Popolocan and Mixtecan*.. Indiana University Publications in Anthropology and Linguistics. *International Journal of American Linguistics* Memoir 15. Waverly Press, Baltimore.

Grimes, B. (1988) *Ethnologue*. Summer Institute of Linguistics, Dallas.

Hattori, S., K. Yamamotto, and O. Fujimura (1958) "Nasalization of Vowels in Relation to Nasals," *Journal of the Acoustic Society of America* 30:267-274.

House A.S. and K. Stevens (1956) "Analog Studies of the Nasalization of Vowels," *Journal of Speech and Hearing Disorders* 21:218-232.

Kingston, J. (1985) "The Phonetics and Phonology of the Timing of Oral and Glottal Events," Ph.D. dissertation, University of California at Berkeley.

Kingston, J. (1990) "Articulatory Binding," in J. Kingston and M.E. Beckman, eds., *Papers in Laboratory Phonology II--Between the Grammar and Physics of Speech*.. Cambridge University Press, 406-434.

Kirk, P.L. (1966) *Proto-Mazatec Phonology*. Ph.D. dissertation, University of Washington.

Kirk, P.L. (1970) "Dialect Intelligibility Testing: The Mazatec Study," *International Journal of American Linguistics* 36:205-211.

Ladefoged, P., I. Maddieson, and M. Jackson (1988) "Investigating Phonation Types in Different Languages," in O. Fujimora, ed., *Voice Production, Mechanisms, and Functions*. Raven Press, Ltd., New York, 297-317.

Merrifield, W.R. (1963) "Palantla Chinantec Syllable Types," *Anthropological Linguistics* 5.5:1-16.

Mugele, R.L. (1982) "Tone and Ballistic Syllable in Lalana Chinantec," Ph.D. dissertation, University of Texas at Austin.

Pike, K.L. and E.V. Pike (1947) "Immediate Constituents of Mazateco Syllables," *International Journal of American Linguistics* 13:78-91.

Rensch, C.R. (1976) *Comparative Otomanguean Phonology*. Indiana University, Bloomington.

Schane, S. (1985) "Vowel Changes of Mazatec," *International Journal of American Linguistics* 51:576-578.

Schram, J.L, and E.V. Pike (1978) "Vowel Fusion in Mazatec of Jalapa de Díaz,"*International Journal of American Linguistics* 44:257-261.

Silverman, D. (1994) "A Case Study in Acoustic Transparency: [spread glottis] and Tone in Chinantec," *Proceedings of NELS 24*, University of Massachusetts, Amherst.
Silverman, D. (in prep.) *Acoustic Opacity and Acoustic Transparency in Chinantec.* Ph.D. dissertation, University of California, Los Angeles.
Steriade, D. (1992) "Complex Onsets as Single Segments: The Mazateco Pattern," ms., UCLA.

Acoustic Characteristics of Tiwi Coronal Stops

Victoria Balboa Anderson and Ian Maddieson

1. INTRODUCTION

Tiwi is the language of the idigenous people of Bathhurst and Melville Islands, located to the north of Darwin in the Northern Territory of Australia (see Figure 1). There are today about 1500 speakers, making this one of the largest communities speaking an Australian language. Like most Australian languages, Tiwi is characterized by a rich set of contrasts among coronal consonants, including four phonetically distinct types of coronal stops. The principal objectives of the present study are to determine what acoustic characteristics distinguish among these coronal stops, and, further, to determine whether there are acoustic characteristics which can consistently be attributed to the distinctive features [anterior] and [apical], which have been used to classify them into two pairs. Secondary goals are to shed light on the articulatory gestures used in Tiwi coronals on the basis of inferences from the acoustic patterns found and through comparison with the work of investigators who have described comparable articulations in other languages on the basis of instrumental data, and to examine certain aspects of the relationship between the acoustics and the articulation of such sounds.

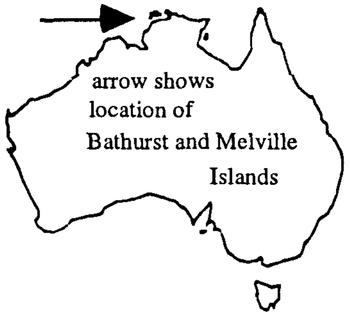


Figure 1. Map of Australia showing location of Bathurst and Melville Islands.

The paper is organized as follows. After brief background sections on Tiwi and coronal-internal oppositions in the world's languages and in Australian languages, probable articulatory properties for each of the Tiwi stop categories will be discussed in the light of work in progress by A. Butcher. These articulatory descriptions will then be related to acoustic characteristics which might accompany them, based on previous work on articulatory-acoustic relations, and a range of specific resulting hypotheses will be described. Next, measurements of durations, burst amplitudes and spectra of Tiwi stop categories will be presented in turn, and discussed, both as they relate to the hypotheses developed, and to the possible articulations of the stops.

1.2. Background

1.2.1 Tiwi: General background and relevant phonology

Tiwi is one of the 10% or so of surviving Australian languages which is still being learned by children. However, it has been in contact with English since about 1911, and the language has been undergoing an accelerating process of change in all aspects of its structure, as sentence structures, morphological simplicity and even phonemes are borrowed from English (Lee 1987). The language that will be described and analyzed here is "Traditional Tiwi" (Lee

1987), a stable variant used by those speakers who were over 30 years of age by the mid-1980's.

In the phonological typology for Australian languages used by Dixon (1980), Tiwi is a "single laminal language"; that is, it has three phonemic coronal stops, two of them (alveolar, post-alveolar) regarded as being apical in articulation, and the third (dental) regarded as laminal. Apical stops are made with the tip of the tongue, or its underside; laminal stops are made with the blade of the tongue. Tiwi also has a notably distinct allophone of the laminal stop, conditioned by a following high front vowel. This allophone, described as palatal, has a longer period of frication after the release. There are thus four clearly distinct phonetic types of coronal stops in the language. We will use the symbols /t̪/, /t/, /ṭ/, [ṯ] .for these four types of stops. As we shall note below, all four classes of stops are often contrastive in other Australian languages.

The two published phonologies of Tiwi disagree somewhat on how the four coronal stops are made. Lee (1987) refers to /t̪/ as "laminal dental", /t/ as "alveolar", /ṭ/ as "post-alveolar", and [ṯ] as "laminal palatal". Osborne (1974) describes /t̪/ as "apico-dental", /t/ as "apico-alveolar", /ṭ/ as "apical, slightly above the alveolar ridge" (post-alveolar), and [ṯ] as "alveopalatal". Osborne's description thus suggests a highly marked situation in which three apical stops contrast in the same language; however, it is possible that by "apico-dental" he is referring to the use of tip and blade together (i.e. a dentialveolar articulation, see Figure 2 below), or he may simply not be distinguishing between tongue tip and blade. Corresponding nasals exist for the coronal stops. Laterals and rhotics show only two-way distinctions; /l/ vs /ḷ/, and /r/ vs /ṛ/. The overall Tiwi consonant inventory, as interpreted by Maddieson (1984), is shown in Table 1. Note that there is no contrast of voicing among stops nor a fricative series, both also common properties of Australian languages. Lee (1987) largely corroborates Maddieson's analysis. However, Osborne (1974) and Oates (1967) had considered /ṭ/ a combination of /t/ with a preceding postvocalic /ɹ/, reflecting its distributional restrictions to medial position.

Table 1: Tiwi consonant phonemes, after Maddieson (1984).

	Labials	Dentals	Alveolars	Postalveolars	Velars
Stops	**p**	**t̪**	**t**	**ṭ**	**k**
Nasals	**m**	**n̪**	**n**	**ṇ**	**ŋ**
Laterals			**l**	**ḷ**	
Rhotics			**r**	**ṛ**	
Approximants	**w**			**j**	**ɣ**

The three phonemic series of coronals contrast in word-medial position, but in word-initial position apical post-alveolars are absent. As noted above, dentals and palatals are in complementary distribution, with palatals occuring before /i/, and dentals before the other vowels of the language, namely /a, o u/. Words illustrating the Tiwi coronal stops are shown in Table 2.

Table 2: Tiwi words illustrating coronal stops, including allophonic distribution.

	Vowel Environment	Dental/Palatal	Alveolar	Postalveolar
Initial	High front vowel	**ṯiráka** *wallaby*	**tíwi** *people*	----
	Other vowel (a,o,u)	**t̪ampináḷa** *cliff*	**taŋini** *stick*	----
Medial	High front vowel	**pikáṯi** *swordfish*	**aḷitiwíyi** *kangaroo(f)*	**milikúṭi** *big toe*
	Other vowel (a,o,u)	**pót̪a** *bone*	**kurumúta** *upper arm*	**wóṭa** *bush*

No stops occur in the coda position of the syllable, including word-final position. There are a limited number of intervocalic consonant clusters consisting of liquids preceding stops. There are also phonetic homorganic nasal + stop sequences which can be viewed either as prenasalized stops, or as an additional series of clusters. We do not know of any evidence in Tiwi that would bear persuasively on this choice. Primary stress is penultimate in Tiwi. In the collected word list there is no marked secondary stress to consider.

1.2.2. Coronal stops in Australian and other languages

Coronal stops are nearly ubiquitous in the world's languages. Of the 452 languages whose phonemic inventories are currently represented in UPSID (Maddieson 1984, Maddieson & Precoda 1992) only Hawaiian lacks a coronal stop. Most languages (81.4% of the sample) have just one distinctive place of articulation for coronal stops. Dart (1991) shows that in languages with one coronal place, actual place of articulation may vary among individual speakers within a given range on the palate as well as on the active articulator. For example, in both English and French, coronal stops are articulated from immediately behind the teeth to just behind the alveolar ridge despite being traditionally described as alveolar in English and dental in French. Similarly, the part of the tongue used varies. In Dart's examination of 20 American English speakers and 21 French speakers, about 78% of the English speakers used an apical articulation, and about 22% used a laminal articulation. The French data was split in the reverse direction (27%, 73%.) Thus, while it is possible to study characteristics common to coronals as a group using languages like English and French, it is not possible to study properties of the apical/laminal or anterior/non-anterior distinctions without direct articulatory verification.

The UPSID corpus includes 67 languages (14.8%) with coronal stops at two distinctive places of articulation, usually apical alveolar and laminal dental (cf. Ladefoged & Maddieson 1986, 1994). Only about 3.5% of languages make use of three or four distinctive coronal stops, and this rare type occurs most commonly in the Australian language family. Of the 3-coronal languages, 10 of 12 are Australian, and of the 4-coronal languages, all are Australian. The usual four-way distinction, regarded as due to the cross-classification of values of the two features, [anterior] and [apical], is as shown in Table 3. Note that if Osborne's (1974) description of Tiwi as having three apicals were correct, it would necessitate the use of an additional feature such as [alveolar], to represent the contrast between two [+apical, +anterior] segments. Table 3 also provides some transcription equivalents in different traditions.

Table 3. Classificatory use of the features [apical] and [anterior].

Feature classification	[-apical] [+anterior]	[+apical] [+anterior]	[+apical] [-anterior]	[-apical] [-anterior]
Possible articulation: Australian convention	"laminal dental"	"apical alveolar"	"apical post-alveolar"	"laminal palatal"
IPA symbol	**t̪**	**t**	**ʈ**	**t̻**
Australianist symbol	**th**	**t**	**rt**	**ty**
symbol in this paper	**t̪**	**t**	**ṭ**	**t̠**

Despite the general agreement on the phonological classification, previous phonetic descriptions of coronal articulations in Australian languages have shown some variation. It is unclear if this is because the languages differ, or because the descriptions are inconsistent. Jernudd (1974) published palatograms for three male speakers of Gunwinjgu, but no linguographic data. This study showed different ranges of inter-speaker variability for the coronal sounds. More recently, Butcher (in progress) has gathered both static palatographic and linguographic data and electro-palatographic (EPG) data for one speaker of each of six Australian languages. We summarize the preliminary results below from this work which Professor Andrew Butcher, of Flinders University, Adelaide, has generously made available to UCLA. Although Tiwi is not one of the languages examined by Butcher, the general picture of the articulations found may provide some insight into the likely articulations in Tiwi.

Before we discuss Butcher's results, a brief definitional note is in order. Following Butcher, "tongue tip" will be defined as the point along the rim of the tongue at the midsagittal line, plus 5mm on the dorsal and ventral (sublaminal) sides of the tongue. Variation exists between British and American phonetic tradition in usage of the term "tongue blade". While British tradition (e.g. Catford 1977) defines the blade to include an area 10 to 15 mm back from the rim of the tongue, American (and Australian) tradition defines the blade as a larger area, approximately the front third of the upper surface of the tongue (up to approximately 20 mm back from the "tip".) Keating (1991) argues for the more extensive view of the blade, on the basis that unambiguous (anterior) coronals use this area of the tongue in their articulations. Usage in this paper will follow American and Australian tradition on this point.

1.2.3. Apicals.

Sounds called "apical alveolars" in the Australian languages examined by Butcher are marked by consistency of articulation across both languages and trials. Linguograms and palatograms together show these articulations to be uncontroversially apical and alveolar; the tongue apex makes a narrow band of contact (from 2-7 mm at the midline between the central incisors, measured from photographs of the palate) just at or in front of the alveolar ridge. Little allophonic variation is observed among the articulations. EPG shows evidence of a very rapid closing movement (10-20 milliseconds) and a stable, static hold. (The closing movement is defined as the interval from the time of initial movement of the tongue away from the vowel toward the tectum, until the maximum number of electrode contacts has been reached. But note that the onset of movement away from vowel position is necessarily imprecise because of the nature of EPG records. Only contact patterns are measured, not actual movement.)

Articulations characterized as "apical post-alveolar" show greater variability across the speakers examined, and allophonically within them. Typically, a 5-12 mm band of contact is made in the post-alveolar or prepalatal area, most often sublaminally. (Butcher also suggests that apical articulations may be characteristic of rapid speech and sublaminal articulations of careful speech.) Postalveolars have slower closing movements than alveolars (30-90 ms) and the hold, once closure has been made, is dynamic; the tongue moves forward so that release is often from an alveolar position. This movement prompts the question of whether postalveolars should be described as flaps. By the criterion used in the TIMIT database (Zue & Seneff 1988) for defining flaps, i.e. lack of release burst, they are not. Moreover, they are much longer in total mean duration (100ms) than the flaps in TIMIT (mean=29ms, Byrd 1993.)

According to Dixon (1980), most Australian languages, like Tiwi, do not display the alveolar/post-alveolar distinction word-initially. Butcher's data shows that the articulation for neutralized initial apicals has characteristics of both the alveolars and post-alveolars in medial position, though they are significantly longer than either in duration. Neutralized apicals are like post-alveolars in that they form a band of constriction in the 5-10 mm range, and move forward between closure and release. On the other hand, they are not usually sublaminal even in careful speech, and are intermediate between the alveolar and post-alveolar categories in their point of constriction on the tectum.

1.2.4. Laminals.

Butcher finds that "laminal dentals" involve the dental and alveolar region of the tectum, but may be formed in one of three ways: 1) interdental and tongue-tip up (i.e. the tip is visible between the teeth), 2) interdental and tongue-tip down, (i.e. the tip is behind the lower teeth so that the blade is visible between the teeth) or 3) dentialveolar and tongue-tip up (i.e. both tip and blade make contact with the surface behind the upper teeth.) The tip-down articulation appears to be least common for the languages examined. Even among the languages with two contrasting laminals, there is a range of variation. Moreover, within a given language all three types of articulation can occur. Variation may in part be attributed to variation in speech style (with interdentals being careful variants, and dentialveolars rapid variants.) The interdental articulation is easily determined by visual inspection. (Note that the productions of the recorded Tiwi dentals were not interdental, according to the direct observation of the second author.) Whether tip-up or tip-down, "laminal dentals" create a 13-20 mm band at closure, from the teeth to the alveolar zone, so that they are about twice as distributed in the midsagittal plane as apicals.

Articulations described as "laminal palatal" appear to be articulated with the tongue tip down in the majority of cases observed by Butcher, and are similar whether they have phonemic or merely allophonic status. They form a 9-13 mm band of constriction in the alveolar and post-alveolar, or post-alveolar and prepalatal areas. Note that this contact is narrower than for "laminal dentals." However, additional contact behind the occlusion may be broader at the sides of the tongue, showing evidence of a raised tongue body. An additional factor which probably affects stop durations of "laminal palatals" is indicated by Butcher's EPG data: at closure, the tongue is initially braced at the teeth, and contact is extended from front to back. At release, contact is "peeled" away, from back to front. (For an account of how bracing facilitates tongue movements, see Stone 1991.) This type of articulation is likely to involve slower movements than those in which the tongue tip meets and pushes off the tectum.

Butcher makes clear that, apart from apical alveolars, there are significant variations in articulation in each of the stop categories, due to idiosyncratic or language-based differences. These aside, Figure 2 broadly summarizes the articulations that Butcher finds representative of the four categories of Australian stop, as drawn by the authors of the present paper. As a point of departure, we will suppose that Tiwi articulations for the coronals stops are as Butcher describes, since we have no reason a priori to expect otherwise.

Figure 2. Midsagittal sketches of the articulations of different types of Australian coronal stops, based on descriptions in Butcher (in progress).

2. MATERIALS AND HYPOTHESES

2.1. Materials

The experimental materials for this study were collected by the second author during fieldwork at Nguiu, Bathurst Island, in 1988, and consist of medium-quality audiotapes of two female and three male native speakers of Traditional Tiwi. The speakers were in their 50's and 60's at the time that the recordings were made, and represent the most conservative type of speech used in the community. Younger speakers were not used, because, as Lee reports, under influence from English the phonetic realization of the palatal allophone [ṯ] for younger speakers is the affricate [tʃ]. The five speakers will be referred to as B, G, (both female), D, R, and E (male). They were recorded in different groups during three recording sessions in which they repeated selected words in isolation, in response to a prompt in English. Two sessions were recorded indoors, and suffer from some reverberation in the room; the third session was recorded outdoors and is interrupted by bird calls and engine noise. Because of the short time available for constructing a word-list when these recordings were made, and because few minimal pairs exist (Tiwi words are often three to five syllables in length) materials are imperfectly balanced for vowel environment, length and other factors. Moreover, for a variety of reasons the speakers provided between six and zero tokens of each word (normally about two). These factors limit the extent to which some desirable comparisons can be made in a controlled way.

In the discussion of results below, relative duration, relative amplitude, and shape of the burst spectrum will each be considered in a separate section. The allophonic palatal will be treated as a separate category so as to show its characteristics in these domains.

2.2. Hypotheses

Two kinds of hypotheses will be entertained, focussing either on the characteristics of individual stop places of articulation or on acoustic characteristics attributable to their component features. We do not expect that duration components, relative amplitudes and spectral properties, each taken alone, will distinguish the four categories. But we have every reason to believe that some combination of duration, relative amplitude and spectral properties taken together, will reliably distinguish all four stops. We will call this general idea Hypothesis 1.

Our other hypotheses focus on acoustic properties that might be associated with the feature opposition [+anterior] (at or in front of the middle of the alveolar ridge) and [-anterior] (behind the middle of the alveolar ridge), or the feature opposition [+apical] and [-apical]. We hypothesize that acoustic bases for the groupings of [+apical] consonants, [-apical] consonants, [+anterior] consonants and [- anterior] consonants can be found. In testing for such groupings, the following specific hypotheses will be considered. Hypothesis 2: A stop's value for the feature [anterior] will affect durations of stop closure. [-Anterior] articulations will show shorter closures than [+anterior] articulations. Hypothesis 3: A stop's value for the feature [apical] will affect durations of stop release. Laminals ([-apical]) will have longer release durations than apicals. Hypothesis 4: Laminal stop bursts will have lower amplitudes than apical stop bursts. Hypothesis 5: The [-anterior] stop categories will be characterized by burst spectra which have high amplitude energy concentrated in a narrow frequency range, whereas [+anterior] stops will show a wider distribution of high amplitude energy.

2.2.1. Acoustic background to hypotheses: Duration

It has often been said that the further back the place of articulation is for a stop, the shorter the closure duration, and the longer the VOT. Thus, investigators have usually found that labials have longer closures than velars (Fischer-Jørgensen 1964, Zue 1976, Keating 1984a) and that VOT is shorter for labials than velars (Fischer-Jørgensen 1954, Keating, Westbury & Stevens 1980, Crystal & House 1988.) Often, provided the effects of language-specific processes such as the American English "flapping" of coronals are set aside, coronal stops show intermediate durations between labials and velars (Stathopoulos & Weismer 1983, Ren 1985). If "backness in the mouth" influences durations, we might expect anterior coronals to have longer closure closure and shorter release than non-anterior coronals. However, on further analysis such a conclusion is not so clear. Explanations of closure differences have focused on the fact that intraoral pressure reaches a peak value more quickly when the supraglottal cavity behind the constriction is small than when it is large. VOT differences are thought to depend on the speed of articulatory release. The tongue tip moves fast, but the slow-moving tongue body means that it takes longer to reinstate a pressure differential and begin voicing in velars. Since the lips are more elastic than the tectum, more compression is likely during labial holds than tongue/tectum holds (also yielding longer durations), but the elastic rebound from this compression contributes to a fast release for labials. For these several reasons, a front-to-back ordering often obtain for closure and release of labials, coronals and velars, but we cannot necessarily expect this to hold for coronal-internal comparisons. Place of articulation *per se* is not important; what do seem to be important are wall area behind the constriction (which in coronals is influenced by the tongue shape as a whole, not just point of constriction) and mass of the active articulator.

If articulator mass influences stop component durations, wemight expect the tongue blade to move more slowly than the smaller, lighter tongue tip in reaching the tectum, in forming a wide seal, and in breaking that wide contact. (This seems especially likely for palatals, in light of Butcher's observations indicating that the tongue first anchors itself on the tectum and then rolls back and forward again while closure is maintained.) Thus, in relation to the surface area of the tongue making tectal contact, [-apical] seems to imply a greater duration of closure. However, another factor is involved in the [apical] distinction. Svarny and Zvelebil's (1955) x-ray data shows that Tamil laminals involve higher jaw positions for their articulations than apicals. On this basis, holding [anterior] constant for paired comparisons, we would expect dentals to have briefer closures than alveolars, and palatals to have briefer closures than post-alveolars, because their higher jaw positions would entail smaller supraglottal cavity volumes. These potentially opposing effects of articulator mass and jaw height make it difficult to predict how apicality will affect closure duration overall. Thus a hypothesis based on [apical] (Hypothesis 3 above) could predict longer durations of release for [-apicals], but not relative durations of closure.

2.2.2. *Acoustic background to hypotheses: Voicing*

Weismer (1980) found a constant duration of the voiceless portion (closure plus release, or VOT) for the English voiceless stops /**p**, **t**, **k**/ regardless of place of articulation. He interpreted this as due to the presence of a consistent glottal opening (devoicing) gesture, whose time course is uninfluenced by place. While total voiceless duration remained constant, the ratios of voiceless closure to voiceless release identified place of articulation. If Tiwi also incorporates a glottal opening gesture during stops, such ratios would be an important aspect of duration to consider. Tiwi stops are described by both Osborne (1974) and Lee (1987) as generally voiceless and unaspirated. However, since voicing is not distinctive in Tiwi, it may be premature to assume that a glottal opening gesture is involved in producing these stops.

2 2.3. *Acoustic background to hypotheses: Spectral shapes of stop bursts*

Stop places have also been identified on the basis of the spectral pattern of the stop burst. Adapting acoustic featural descriptions proposed by Jakobson, Fant and Halle (1952), Blumstein and Stevens (1979) found 85% identifiability of the spectra of labials as diffuse-falling, alveolars as diffuse-rising and velars as compact. "Falling" (or "grave") denotes a downward-tilting slope on a line drawn from peak to peak and "rising" (or "acute") denotes an upward slope on a spectrum calculated using a 25.6 ms half-window beginning at stop release (thus including frication), and a frequency range of 0 to 5 kHz.

Lahiri, Gewirth & Blumstein (1984) later used this metric to attempt to group the contrasting dentals and alveolars of Malayalam together, and find unifying characteristics of coronals as opposed to labials. Only 71% of alveolars fit into the diffuse-rising template, and only 57% of dentals did. We infer from this, and from the illustrations they provide, that the spectra of Malayalam alveolars and dentals were fairly different from each other. (Lahiri, et al (1984) and Keating and Lahiri (1993) also examined the relation of burst spectra to the spectra of following vowels, but hypotheses based on these studies will not be addressed here.) Dart (1991), examining contrasting apicals and laminals in Malayalam and 'O'odham, found trends toward flat spectra for dentals, rising spectra for palatals, a low-frequency energy peak for retroflexes (both sublaminal and apical) and more highly peaked spectra for alveolars than dentals. (She focused a very short FFT window over the burst, so as to exclude frication from the spectrum.) These results, too, indicate that contrastive coronals can show different spectral signatures. Fant (1960) ascribes spectral differences among stops to differences in the cavity in front of the occlusion; burst spectra can therefore give us an acoustic "picture" of the space in front of a stop constriction. There is reason to believe that the four Tiwi stop categories involve front cavities of sufficiently different shapes that their spectra will differ appreciably. This is encouraging in terms of Hypothesis 1, but can we relate these differences to features?

Stevens and Keyser (1989) note that [-anterior] articulations cause a major peak in the midfrequency range, because of the high ratio of the length of the front cavity to the length of the back cavity. For these articulations, either F2 and F3, or F3 and F4 are closely approximated, combining to form one peak. However, they suggest that [-anterior] will only be strongly salient in [-sonorant] [-coronal] segments, since the high frequency energy present in [+coronal] spectra will mask the midfrequency peak caused by the front cavity. In light of Lahiri et al. and Dart's findings that [+coronal] spectra can instead be diffuse-flat, we hypothesize that the Tiwi [-anterior] stops will contain an isolated (therefore potentially salient) midfrequency peak (Hypothesis 5).

2.1.3. Acoustic background to hypotheses: Burst amplitude.

Keating, Westbury & Stevens (1980) show that English voiceless alveolar stops have an early amplitude peak after release of closure, followed by a sharp drop-off of noise, while velars show much later peaks of energy following release, and gradual dissipation of noise. Release energy for distributed stops like velars is diffused over a relatively long period following the break of closure, while the energy of a less distributed stop is dissipated in a much shorter length of time. Extrapolating to coronals, we expect the burst energy of apicals to be of greater amplitude and to be contained in a shorter period than in laminals. Jongman, Blumstein & Lahiri (1985) found this result in Malayalam. At the burst release, Malayalam apical alveolar stops were reliably higher in amplitude than laminal dentals (articulations are confirmed by Dart, 1991). We project similar differences for Tiwi apicals and laminals (Hypothesis 4.) Note that quieter bursts and longer VOTs for laminals are attributed to the same cause: slow movement of the articulator prolongs obstruction of the supraglottal cavity, and prevents a rapid escape of air.

3. DURATION MEASUREMENTS

3.1. Data and methods

The Kay Elemetrics Computer Speech Laboratory (CSL) was used to digitize and measure portions of words containing coronal stops in various vowel and stress environments. Only stops in intervocalic or initial environments were considered; coronals in consonant clusters were excluded. For the examination of duration, all words of the specified form were measured so as to balance out effects of different factors over the largest available sample. The 38 words measured are listed in the Appendix. For some analyses, subsets of the data were used in order to isolate certain effects from potentially confounding factors. Selected examples isolating vowel environment, stress environment and position in word are shown in Table 4 (a-c).

Table 4. Subsets of words used to isolate particular factors in the duration analysis. The syllable in question is shown in **boldface**.
(a). Vowel environment.

	Following Vowel ---->			
Preceding Vowel ⇓	**i**	**a**	**o**	**u**
i	aḷi**ti**wíyi *kangaroo (f)*	ki**ta**ṭawíni *bread*	----	----
a	yá**ti** *one*	tiŋá**ta** *beach*	----	pat̪ipa**ṭú**wu *turtle egg*
o	----	pó**t̪a** *bone*	----	----
u	kumú**ti** *beard*	kurumú**ta** *upper arm*	----	tuṭu**ṭú**ṭi *move around*

(b). Stress. (Primary stress is penultimate).

Stress Environment	**Tiwi exemplar (vowel environment constant)**	**English Gloss**
pre-stress	aḷi**ti**wíyi	*kangaroo (f)*
stress	muri**tí**wa	*neck*
post-stress	poalí**ti**	*beard*

(c). Position in word.

Position in Word	**Tiwi exemplar (vowel environment and stress constant)**	**English Gloss**
initial	**t̪a**mpináḷa	*cliff*
medial	kita**ṭa**wíni	*bread*
final	pirá**t̪a**	*rice*

2.1.2. *Duration measurements*

Because there is some evidence that salient frequency information for obstruents can be contained in frequencies as high as 12 or 16 kHz (Sands 1991 for clicks), material was initially sampled at 40 kHz, to make use of this range. However, a consistent band of noise on the recordings between 8 and 11 kHz meant that analysis had to be limited to information below 8 kHz, and waveforms were subsequently downsampled to 20 kHz to maximize resolution of information in the 0 to 8 kHz range.

Duration measurements were made using simultaneously displayed, time-aligned waveform and amplitude windows, and two spectrogram windows. Time-cursors can be linked across all four windows so that events and times can be reliably matched. An example of the four time-aligned windows is shown in Figure 3. A very wideband (586 Hz) spectrogram was displayed (lower left) in order to obtain fine resolution of the time axis. This made possible a reliable determination of the points at which stop bursts and voicing occurred. A 293 Hz bandwidth spectrogram (lower right) resolved formants. In some cases the energy display facilitated determination of onset or offset of voicing, as well as the position of the burst. Duration measurements for VTV sequences (where T = any coronal stop)were divided into six increments, as shown in Figure 4.

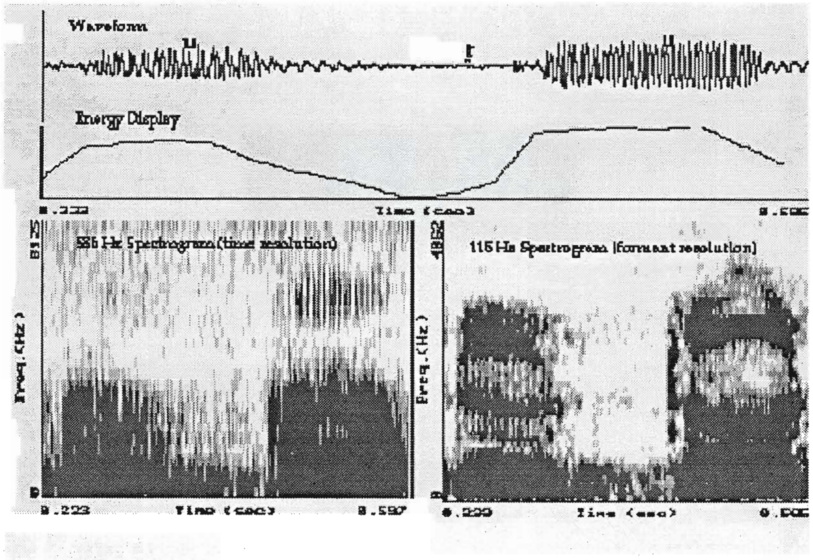


Figure 3. Example of display used for measuring duration on CSL.

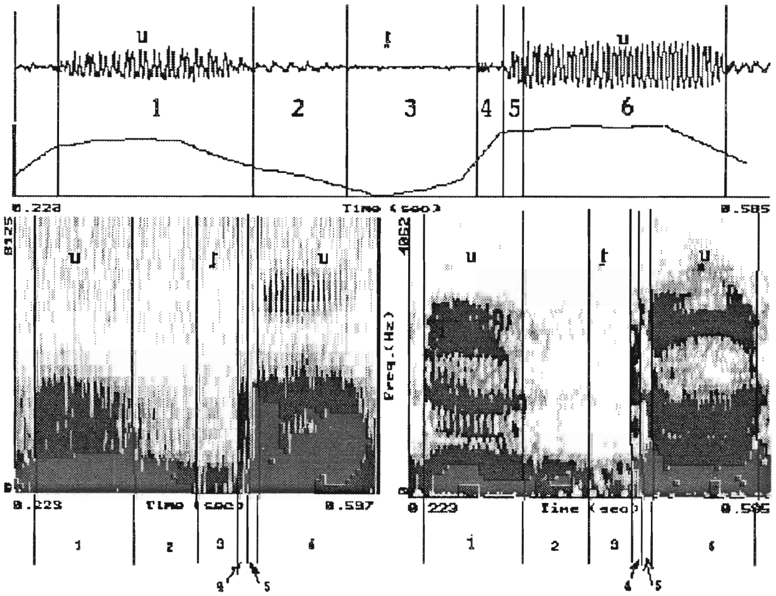


Figure 4. Increments used in measuring durations of Tiwi VTV intervals

Increment 1 is the duration of the vowel preceding the stop in question, from onset to offset of visible formants. Increment 2 includes just the voiced portion of the stop closure, from the point at which the higher formants of the preceding vowel disappear, to the end of visible voicing. The third measure contains the voiceless, silent portion of the closure. Voiced and voiceless portions of stop duration were measured separately in order to determine whether degree of voicing plays any role in differentiating these stop categories. We also entertained the idea that obstruent voicing in a language with no phonemic voicing distinction may be in completely free variation, or may be a matter of the speaker's preference. (Note that these first three measures are not applicable to initial stops.) The fourth section measures the time from the beginning of the transient indicating release of tongue contact with the roof of the mouth, to the onset of vibrations of the vocal folds. This is voice onset time (VOT), as traditionally understood (Lisker & Abramson 1964.) The fifth component is the voiced section of the release. This measurement consists of the period during which the vocal folds have begun vibrating, but frication noise is still apparent on the spectrogram. Increment 6 is the duration of the vowel following the stop in question, once again from onset to offset of formants. In some cases one or more of these increments had values of zero.

This data was not normalized for speech rate variation within speakers. First, there was no carrier phrase with which to compare speech rate over multiple utterances. Normalization of consonants with respect to their syllables was considered, but was rejected because of the unbalanced nature of the data set. As mentioned earlier, several factors which potentially affect duration varied across tokens. Syllables of interest appeared in various stress environments, various vowel environments, and various positions in the word. In addition, the number of syllables in a given word, which also affects syllable durations (Lehiste 1970), varied across tokens. A 3-factor analysis of variance (ANOVA) for stop category, speaker and position in word showed a very significant statistical effect of position in word on duration of syllables. Penultimate syllables were greatest in length because they were always stressed, followed by word-final syllables (which are subject to prepausal lengthening, Klatt 1975.) Syllables in antepenultimate or earlier positions were shortest in duration. To balance for each of these factors would have yielded excessively small subsets of words. Moreover, if durations of neighboring vowels affect consonant identification, we would want to catalog this variation, rather than removing it from consideration by using it as a normalizing tool.

However, to screen the data for gross variations of speech rate, the median was calculated for each speaker for all VTV intervals measured (for TV in the cases of consonants in initial position, and GVTV sequences). Total VTV (or TV) interval was considered the best index of speech rate among the measures taken. Assuming that duration differences arising from gross variation in speech rate would be of a much greater order than those contributed by factors like stress, intrinsic vowel duration and word position, tokens whose measurements fell more than two standard deviations away from speaker medians were considered outliers for speech rate and excluded.

3.2. Results

Unless otherwise noted, each of the ANOVAs mentioned below is a 2-factor analysis, with stop category and speaker identity being the two factors. Speaker identity was included so as to account for a significant probable source of variance; this factor in every case turned out to be significant. Interaction terms between the factors will not be mentioned unless statistically significant. For post-hoc comparisons between means, Fisher's Protected Least Significant Difference (PLSD), corrected in this statistics package for unbalanced data sets, is used.

3.2.1. Duration of word-initial vs intervocalic consonants

As mentioned above, Butcher found initial apicals to be longer in duration than intervocalic ones. To determine whether initial and intervocalic stops showed different durations for Tiwi, a 3-factor ANOVA was performed with stop category, speaker identity, and initial or intervocalic position as main effects. Only release durations can be compared in our data, as the onset of closure cannot be determined in the initial stops. The result is shown in Figure 5; in this and following figures showing data in histogram form the error bars show one standard deviation (95% confidence level). Initial stops are significantly longer in total release duration than intervocalic tokens for the three categories compared (post-alveolars were excluded from consideration, as they do not occur in initial position), but the order of categories remains the same. Since a significant effect was found for initial vs intervocalic position, each of the tests below which included initial consonants was performed separately for combined initial and intervocalic data and for just intervocalic (VTV) data. No significant differences in the patterns of results were found, so the more inclusive result is reported.

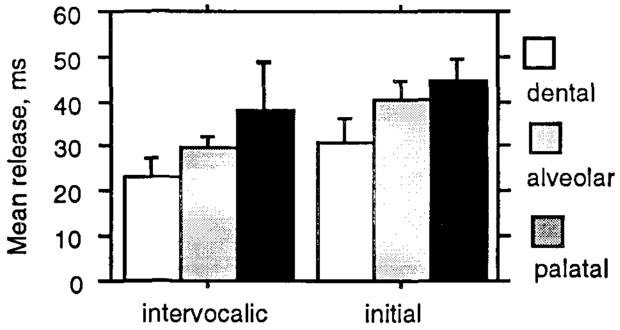


Figure 5. Mean release durations of inter-vocalic vs initial stops, by coronal stop category.

3.2.2. *Duration of closure*

Mean values for the total closure duration of four stop categories are shown in Figure 6. Total closure is the sum of both voiced and voiceless sections of stop closure, from disappearance of formants of a preceding vowel to the moment of the stop burst. Although the main effect of category on closure duration was highly significant (F[3, 313]=8.180, p<.0001), only the post-alveolar category is significantly different from each of the others in post-hoc comparisons of means. Other categories could not be statistically distinguished. Thus the main result is that post-alveolars are shorter in closure duration than the other categories. Dentals and palatals also tend to have slightly larger standard deviations than alveolars and post-alveolars (a trend which will continue in most of the tests below), reflecting greater articulatory variation.

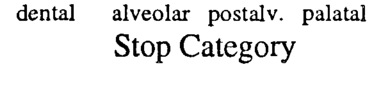


Figure 6. Mean total closure durations for Tiwi consonant types.

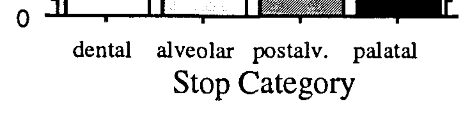


Figure 7. Mean release durations for Tiwi consonant types.

3.2.3. *Duration of release*

Total release refers to the duration of the stop from its burst to the onset of formants for the following vowel, and so includes VOT as well as any short voiced portion which reflected the presence of a consonantal constriction. Mean durations of releases for the four coronal stops are shown in Figure 7. Post-alveolars had the shortest mean durations of release, followed by dentals, alveolars and palatals. ANOVA showed a very significant effect of stop category on release duration (F[3, 448]=40.89, p<.0001), and in this case each pairwise comparison between categories was statistically differentiable from the others.

3.2.4. Durations of voiced and voiceless components of stops

Mean durations of the voiced portion at the onset of intervocalic coronal stop closures are shown in Figure 8. Duration of voiced closure does not distinguish stop categories (p=.9607). Values are very similar for all four stops types; voicing persists during 45-50 ms of the closure. These numbers are in the range of Keating's (1894b, 1984c) results modeling passive devoicing of voiced stops, such as English /d/, and suggest that Tiwi stops are not produced with an accompanying glottal opening gesture. Rather, voicing may cease due to passive factors such as the equalization of supraglottal and subglottal pressure. The lack of a glottal opening gesture means that the ratio of voiceless closure to voiceless release is not an appropriate metric for distinguishing Tiwi stops.

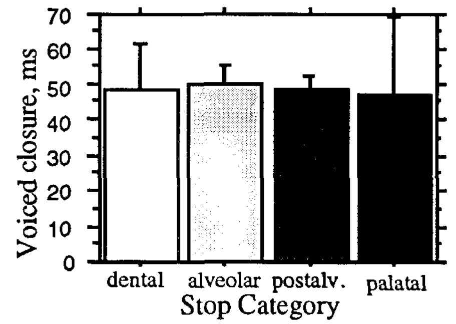


Figure 8. Mean durations of voiced portions of Tiwi stop closures, by category.

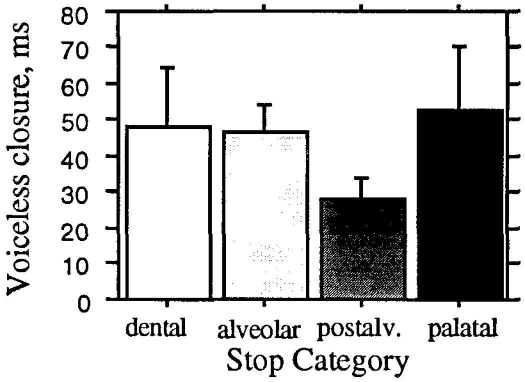


Figure 9. Mean durations of voiceless portions of Tiwi stop closures, by category.

Mean durations of the voiceless portion of intervocalic coronal stop closures are shown in Figure 9. Post-alveolars are significantly different in pairwise comparisons from every other group, but none of the other categories was significantly distinguishable. Note the similar pattern in Figure 9 and Figure 6, which shows total closure duration. Clearly the voiceless portion of closure contributes most of the differences in closure among the stop groups.

Mean durations of the voiceless portion (VOT) of the stop releases are shown in Figure 10. All of the stop categories were significantly differentiable from each other, except for the dental and alveolar categories. Dentals and palatals showed higher standard deviations than alveolars and post-alveolars. Among individual speakers, VOTs for post-alveolars and palatals always stood in the same relation — palatals showing the longest VOTs and post-alveolars the shortest. But the relationship between dental and alveolar categories depended on the speaker. As shown in Figure 11, Speaker G has considerably longer VOT for alveolars than the other speakers, and a substantial difference between alveolar and dental values. The other four speakers have more equal VOT's for alveolars and dentals. Comparison of Figure 10 with Figure 7 above shows that VOT forms the major contribution to total release duration, as compared with voiced release, for which results are shown in Figure 12.

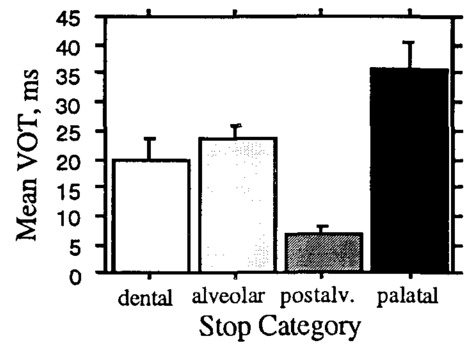


Figure 10. Mean VOT durations for Tiwi stop categories.

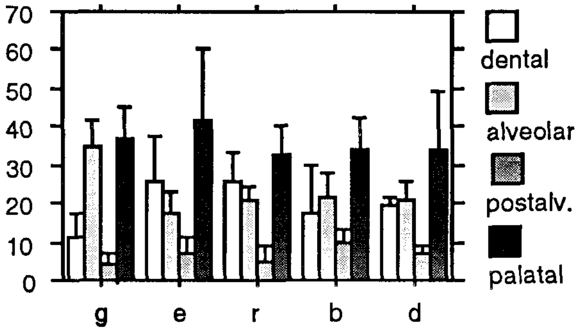


Figure 11. Individual speaker means of VOT for each consonant type.

Mean durations of the voiced release component of stops are shown in Figure 12. (Recall that this is the portion at the onset of the post-consonantal vowel where voicing has started but where signs of the effect of the consonantal constriction, such as some residual frication, are still apparent. Not all stops displayed such a portion and demarcation of this interval was sometimes problematical.) The post-alveolar category was significantly different from the dental and palatal categories. None of the other pairs of comparisons yielded a statistically reliable difference.. Some trade-off between components of release duration is thus apparent, as the post-alveolar VOT is significantly shorter than those of dentals and palatals. Note that while neither VOT nor voiced release duration alone separates dentals from alveolars, the total release duration for alveolars is significantly longer than that for dentals (Figure 7.)

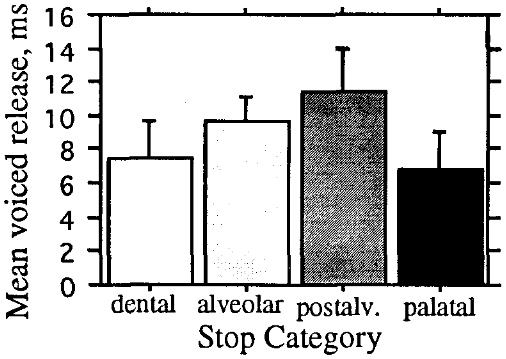


Figure 12. Mean voiced release durations for Tiwi stop categories.

3.2.5. Durations of surrounding vowels

Fischer-Jørgensen (1964) found that, for Danish, place of articulation distinctions can be reflected in the relative durations of a preceding vowel. To test for the possibility that differences in vowel duration signal a stop's identity to Tiwi speakers, the durations of both preceding and following vowels were analyzed.

An ANOVA testing the effect of stop category (and speaker identity) on the duration of the vowel preceding the coronal in question showed no significant differences. However, the effect on the following vowel (V2) was highly significant. In this case the ANOVA for vowel durations included the position of the syllable in the word as a factor, as well as preceding stop category and speaker identity. Speaker identity was marginally significant (p=.0423), reflecting some difference in characteristic speaking rates, while both stop category and position in word were highly significant (p<.0001). (For position in word, stressed syllables are longest, followed by word-final syllables. Syllables falling before the penultimate syllable are shortest.)

The mean durations of V2 after each coronal stop category are shown in Figure 13, together with the total release duration, yielding the total syllable duration (TV). The longest duration for V2 is after post-alveolars, the shortest after palatals, and the dental category shows a longer mean V2 than alveolars. The V2 duration of each of the pairs was distinguishable at a p<.0001 level, excepting dental vs alveolar and dental vs post-alveolar, which were not significantly distinct. More important, however, is the trend of inverse relationships we see between release duration of the stop, and duration of V2. Recall that for total release, post-alveolars are shortest, and palatals longest, with alveolars showing longer mean release durations than dentals (Figure 7.) Because of this trade-off between release duration and following vowel duration, TV duration across categories is very nearly constant. The apparent exception to this constant syllable duration are syllables with palatal onsets, which remain shorter than the others.

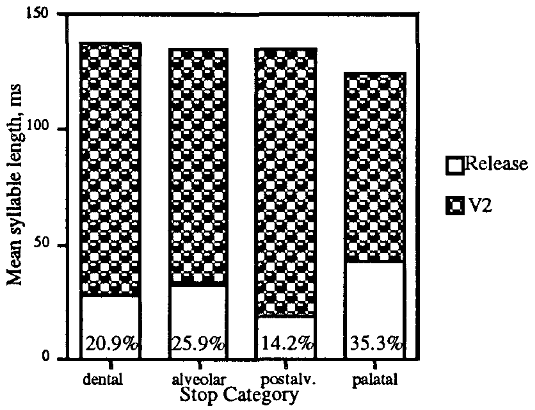


Figure 13. Mean TV syllable lengths, by proportions of release and V2

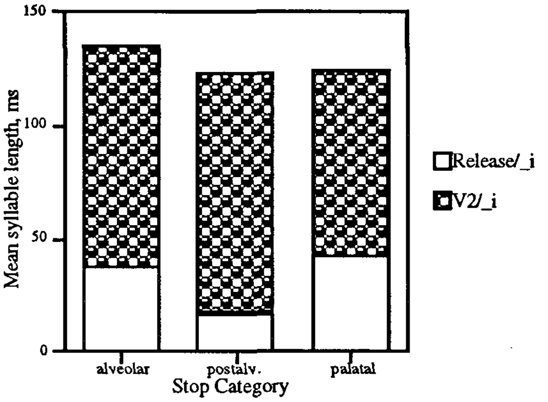


Figure 14. Mean lengths of /Ti/ syllables

However, recall that palatals only precede /i/; this short syllable duration was suspected to be a result of the short intrinsic duration of /i/. Figure 14 shows relative contributions to syllable durations of stop release and following vowel for just the /i/ environment (which excludes dentals.) Here, although alveolars and post-alveolars show an unexpected significant difference, palatals fall between them, and are indistinguishable from either group. Thus, when intrinsic vowel length is controlled for, syllables containing palatals are not significantly different from other categories of stop.

3.3. Summary and discussion of durations

The general pattern seen in the duration results is that while the apical post-alveolars and the laminal palatals can often be clearly distinguished in pairwise comparisons, the post-alveolars are not typically grouped together with the apical alveolars, nor are the palatals typically grouped with the dentals. Where there is no significant difference between the two apicals or the two laminals, it is often the case that three or four of the coronal stop categories do not differ. There is even less evidence for a grouping of the two [+anterior] place categories, or the two [-anterior] categories. Clearly, there is no direct relationship between the classificatory features [anterior] and [apical] and any particular durational component. Rather, the evidence points to an intermediate level of phonetic parameters (Keating 1988). The features [anterior] and [apical] interact to determine a general oral tract shape and size. They may work “cooperatively” or “antagonistically” in creating an overall oral tract configuration. For example, [-apical] and [-anterior] values both contribute to a small oral cavity, but whereas [+anterior] suggests a larger oral cavity behind the closure, an articulation simultaneously specified for [-apical] will probably involve a raised tongue and jaw, which work to decrease the size of the oral cavity. It is this overall oral tract configuration, as well as an articulator’s speed of movement and specification for movement during the articulation (as for the post-alveolar), that predict more closely the resulting acoustics. Neither [anterior] nor [apical] can predict the acoustics of duration without referring to effects of the other feature. The kind of complex relationship assumed can be summarized by the schema in Figure 15.

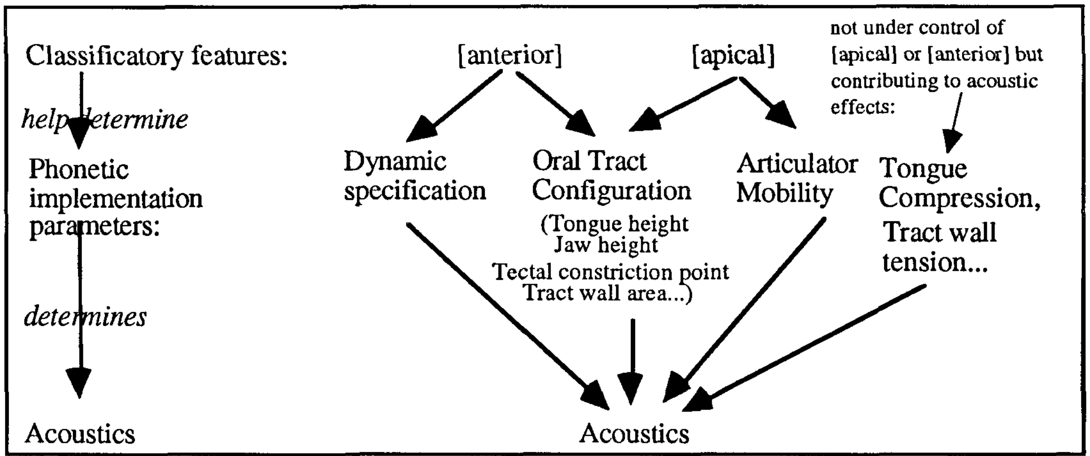


Figure 15. Schematic representation of the suggested relation between [anterior] and [apical], the phonetic implementation component, and acoustic output.

Let us review the duration results in light of the specific hypotheses concerning duration outlined above. Hypotheses 2 and 3 respectively predicted longer closure durations for anterior coronals than for non-anterior coronals, and longer release durations for laminal coronals than for apicals. Since mean total closure was very similar for three of the stop categories (Figure 6), Hypothesis 2 fails. However, it fails only for palatals, correctly predicting long closures for dentals and alveolars, and short closure for post-alveolars. The unexpected length of the palatal closure may be related to the time required to make a rolling movement of the tongue similar to that indicated by Butcher’s data. This particular type of dynamic specification may over-ride the relationship between anteriority and closure duration. A possible effect of anteriority on relative offset of voicing after closure was also anticipated, but was not observed (Figure 8).

Hypothesis 3 correctly predicts a short release for post-alveolars and a long one for palatals. It fails, however, in differentiating dentals from alveolars, which are similar in release duration and intermediate between the post-alveolar and palatal categories. In this case, it seems that more specific characteristics of the stops than their classification in terms of [apical] may be responsible for the apparent partial success of the hypothesis. While the low mass of the tongue tip probably facilitates a quick release for apical post-alveolars, early onset of voicing may be most effectively enhanced by a forward movement of the tongue during the closure of these stops, expanding the supraglottal cavity so that the pressure differential required for voicing is reached quickly. Recall that the ‘voiced release’ portion was longest for post-alveolars, indicating that voicing starts at an early point in the release movement while the aperture is still somewhat constricted (Figure 12). For the palatals there is a long (‘distributed’) constriction of (tip and) blade behind the alveolar ridge, and probably a higher jaw and higher tongue body position behind the occlusion. We assume the supraglottal cavity is the smallest of all the coronal types, and the rolling pattern of closure formation from front-to-back actually reduces the cavity size during the first part of the closure. Breaking this contact and re-establishing conditions for voicing takes longer under these conditions.

Alveolars have longer VOTs than post-alveolars, probably because they are static during closure; they do not benefit from the decrease in pressure afforded by a forward movement of the articulation during closure. Alveolars are shorter in VOT than palatals due to the light mass and greater mobility of their articulator, and the probable lack of narrowing of the oral tract behind the constriction. The reason for the similar release durations of dentals and alveolars is not clear. As Tiwi has no aspiration or voicing contrasts, we posit that the patterns of release and V2 durations in coronal stops are due to biomechanically and aerodynamically determined effects. Thus, the constant TV syllable length reflects the fact that the same articulatory gestures are used for the vowel regardless of the preceding consonant. The proportion of the syllable counted acoustically as part of the release is dependent on where non-fricated, high-amplitude voicing can begin, as determined by considerations of supraglottal volume, rate of aperture increase, etc. These considerations must balance out in the case of the two anterior coronals.

The duration results for Tiwi generally parallel results from other languages in which multiple coronals articulations have been examined, and in which the articulatory gestures used have been studied by palatographic techniques. The Tiwi post-alveolars stand out as shorter in duration than each of the other categories in most tests. Short durations have also been found for Malayalam sublaminal and apical post-alveolars (Dart 1991) and Toda sublaminal post-alveolars (Shalev, Ladefoged & Bhaskararao 1993). Tiwi palatals are distinguished from other categories by long VOT. Malayalam laminal alveolars and 'O'odham laminal or apicolaminal alveolars are also distinguished by long VOT (Dart 1991).

The difficulties in differentiating Tiwi dentals and alveolars by duration are not isolated. Dart (1991) found no differences in closure duration between dentals and alveolars in both Malayalam or 'O'odham, and longer VOT (for the laminal <u>inter</u>dental) in the case of only one of three speakers of Malayalam. The speakers who used laminal dentialveolars did not have significantly different VOT's. In Toda, VOT does not distinguish laminal dentialveolars from apical alveolars (Shalev et al. 1993.) In Dahalo, it is the (confirmed) apical alveolars that have very significantly longer VOTs than laminal dentals! (Maddieson, Spajić, Sands & Ladefoged 1993). One possible explanation for the Dahalo case, noted by the authors, is that placing frication on the alveolar is an active gesture designed to distinguish it sufficiently from the dental. In Iaai (Maddieson & Anderson, forthcoming), it is post-alveolars which have longer periods of frication than dentals, which adds weight to the idea of frication as a potential <u>active</u> enhancing characteristic with which to separate coronal categories. Tiwi does not choose this option, however.

4. RELATIVE AMPLITUDE

4.1. Procedure

4.1.1. Data

For examinations of relative amplitude and spectral shape, subsets of words were chosen for the purpose of comparing stop categories in nearly balanced environments. Recall that the data set as a whole was unbalanced in several respects. Not every speaker produced every token in the word list, numbers of repetitions of tokens varied among speakers, and vowel and stress environments, and position in words, were not equally matched across stop categories. The subset of words shown in Table 6 was chosen for the examination of relative amplitude. Where more than one word occurs in a cell this is because insufficient tokens were available of a single item but a well-matched substitute word was available. Wherever possible, at least two tokens of each stop in each environment were obtained for each speaker. Because a larger data set was available for coronal stops in unstressed syllables, stops in stressed (penultimate) syllables were avoided. Analyses also considered whether the stop fell in an initial, medial or final syllable.

Table 6: Data subset used for relative amplitude study. Numbers of tokens measured are shown after each word. For clarity, glosses are not included here, but can be found in the Appendix.

	Vowel Environment	alveolar	post-alveolar	dental	palatal
Initial stops	#_a	**ta**ríṉa (7) **ta**ŋíni (5)	--	**t̪a**mpináḷa (10)	
	#_i	**ti**ŋáta(9) **ti**nimúli(2)	--		**ṯi**táka (4) **ṯi**ráka (7)
	#_o	**to**kampíni(4) **to**kampúŋa (3) **to**ŋuláka(3)	--	--	
	#_u		--	**t̪u**kwantáŋa (5)	
Intervocalic stops	a_i	yá**ti** (12)	pulá**ṭi** (11)		pika**ṯi** (10) paṯipat̪úwu (2)
	a_a	tiŋá**ta**(10)	kita**ṭa**wíni (9)	pirá**t̪a** (10)	
	u_a	kurumú**ta** (9)			
	o_a		wó**ṭa** (7) values for 4/5 speakers	pó**t̪a** (8)	

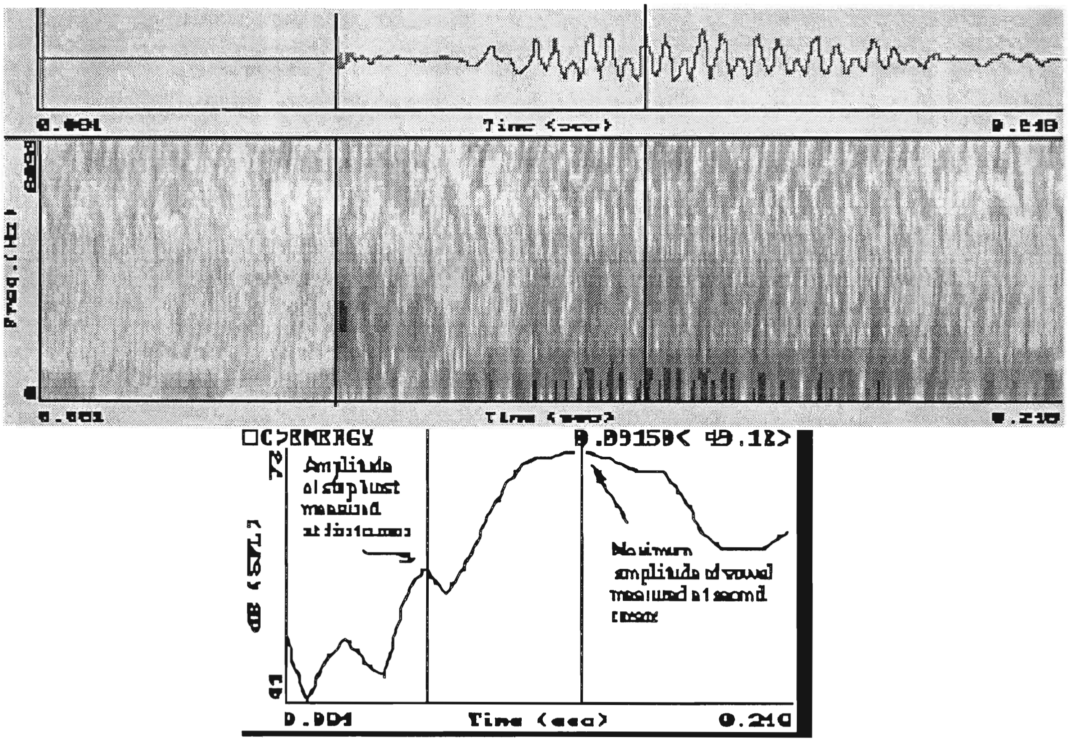


Figure 16. Time-aligned CSL windows used for measuring relative amplitude of stop bursts from waveforms

For each token, energy measurements were taken on CSL of the peak amplitude of the burst transient for the consonant in question, and of the maximum amplitude of the tautosyllabic vowel. As in the duration portion of the study, simultaneously displayed and time-linked waveform, spectrogram and energy windows facilitated finding locations of burst transients and vowel maxima. Figure 16 illustrates the display used. Amplitude information for the energy display was averaged over short unsmoothed windows of 10 ms with time advances through the displayed data at 5 ms intervals, to capture small increments or decrements in energy. In most cases the burst peak was prominent enough to be unambiguous in each of the displays, and the local maximum amplitude which most nearly corresponded to the location of the burst peak on the waveform was recorded. In ambiguous cases in which a ramping effect from burst to vowel was observed, the value at the moment of the burst (on the waveform) was recorded. For vowels, the maximum energy displayed by the amplitude window was recorded.

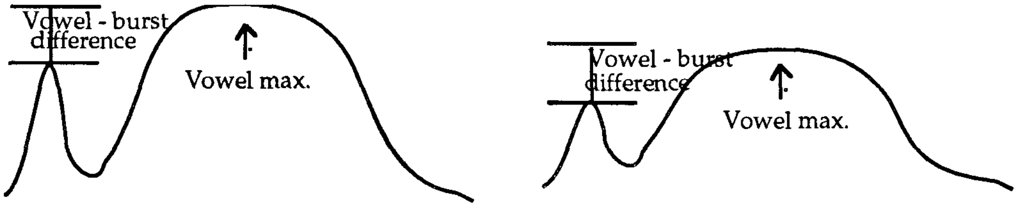
4.1.2. Analysis and interpretation of data

In the discussion below, the decibel (dB) values reported are not the raw measures taken from bursts, as in the diagram above, but rather differences between the peak amplitudes at tautosyllabic vowel maxima, and peak amplitudes at burst transients. This derived measure was employed as a way to normalize for a speaker's amplitude; it better reflects the energy of the burst compared with the overall amplitude for a given utterance. This idea, based on Jongman, et al. (1985) departs from their procedure in using differences in dB between vowel and burst amplitude, rather than vowel amplitudes divided by burst amplitudes. We feel this procedure reduces the effect of outliers.

Note that because amplitude differences between vowel and stop burst are used, rather than raw values for bursts alone, a large value for "vowel minus burst" (hereafter "vowel-burst") for a given token means either a low-amplitude burst, a high-amplitude vowel, or a combination of these two conditions. Alternate possibilities are illustrated in Figure 17.

Figure 17. Two scenarios yielding a large difference in "vowel-burst" amplitudes.

On the other hand, a small difference in vowel-burst amplitude means either a high-amplitude burst with respect to the following vowel, or a low-amplitude vowel (or again a combination of these factors). This is illustrated in Figure 18. Since both amplitudes being measured may vary independently, it is important to make comparisons between stop categories in well-matched subsets of data, e.g. keeping vowel quality fixed.



A: High amplitude burst condition B: Low amplitude vowel condition

Figure 18. Two scenarios yielding a small difference in "vowel-burst" amplitudes.

As in the duration portion of the study, statistical analyses for each of the effects being tested below will reflect the two factors (independent variables) stop category and speaker identity, unless otherwise specified. In all cases, a highly significant main effect for speaker identity (at the p<.0001 level, usually) was found, unless otherwise reported. Speaker identity is probably a more important effect here than it was for duration, because of the less than optimal signal-to-noise ratio for these recordings. Because recordings were made on three separate occasions and in two different locations the factor "speaker identity" for this data comprises not only a difference in speaker, but the particular noise conditions that obtained when that speaker was recorded. Different recording conditions potentially contribute not only to both speaker identity as a significant category but also to significant interactions with stop category. (A high noise level could mask differences and yield apparent groupings across categories for a given speaker.) Thus, while Speaker D showed the greatest range of relative amplitude values, probably because a good signal-to-noise ratio allowed for small burst measurements, Speaker B showed the smallest range of values, consistent with the poorest signal-to-noise ratio.

3.2. Results

3.2.1. Relative amplitudes of word-initial vs intervocalic stops

Because significant durational differences were found for word-initial vs intervocalic stops (suggesting, following Butcher, that initial apicals may differ in their articulations from medial, contrastive apicals) the amplitude data was also tested along this dimension. Again, post-alveolar stops were necessarily excluded from the test because they do not occur initially. No significant difference was found between initials and intervocalics for vowel-burst differences, F(1,114)=1.086, p=.2995. Initial stops are not significantly higher or lower in amplitude than intervocalic stops. This result is plotted in Figure 19.

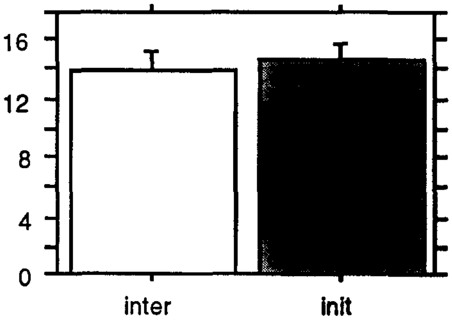


Figure 19. Mean vowel-burst amplitude differences for intervocalic vs initial tokens.

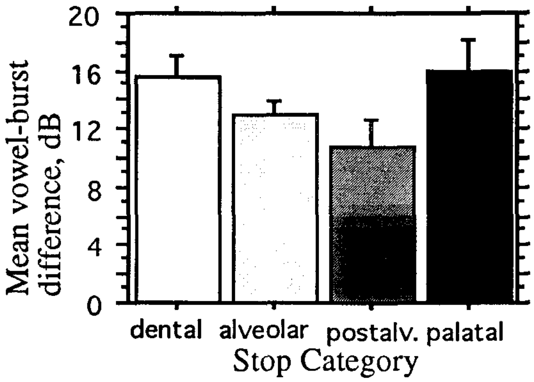


Figure 20. Vowel-burst differences in amplitude for Tiwi stop categories.

4.2.2. Relative amplitudes of burst categories

A two-factor ANOVA over all the data showed that vowel-burst differences distinguished all pairs of stop categories except dentals and palatals, which act as a group here. Postalveolars had the smallest vowel-burst differences, palatals and dentals the largest differences, and differences for alveolars fell in between, as can be seen in Figure 20. The main effect of stop category on vowel-burst amplitude was highly significant, F(3,127)=12.096, p<.0001. In paired post-hoc comparisons of means, over all vowel and word-position environments, each of the four categories was distinguishable from each of the others at the 95% significance level, with the exception of the palatal and dental categories.

For this test there was also a significant interaction effect between speaker and place (F[12,127]=2.904, p=.0014), taken to be an artifact of recording environment, as explained above. Despite the significant interaction term, the post-alveolar consistently showed small normalized differences across speakers; i.e. potentially the largest burst amplitudes. Palatals and dentals showed the largest differences across speakers (smaller potential burst amplitudes). The alveolar was variable across speakers.

Intervocalic environment has been shown not to be conducive to louder bursts (section 4.2.1.) Thus, the low vowel-burst difference for post-alveolars is not attributable to its restriction to medial environments. In order to determine with certainty which of the two scenarios in Figure 18 accounted for the small vowel-burst differences for post-alveolars, two additional tests were performed. These are described below.

The intervocalic alveolar, dental and palatal stops in this data subset appear in the word-final syllable. However, the post-alveolars are in antepenultimate position in 9 of 27 tokens (33%). (These are the post-alveolars in the word /kitaṭ**a**wíni/). A 3-factor ANOVA including speaker, stop category and position in the word as factors showed that statistically important differences do exist among amplitude maxima for vowels occupying different word positions; word-final syllable nuclei are lower in amplitude than antepenultimate nuclei. Thus, to verify that post-alveolars showed small vowel-burst differences by virtue of high-amplitude bursts rather than low-amplitude vowels, a second analysis was performed restricted just to word-final syllables. The effect remained statistically robust for the post-alveolars.

Effects of intrinsic vowel amplitudes were considered (/**i**/ is expected to be lower in energy than /**a**/, other factors being equal). Narrowing the data set even further, to control for vowel environment, the intervocalic environments **aTi** and **aTa** were separately examined. Results for **aTi** environments are shown in Figure 21 (recall that this environment excludes dentals). A two-factor ANOVA showed that each speaker had the smallest relative differences for the post-alveolar category, while speaker variability in the behavior of the alveolar and palatal categories made it impossible to conclude anything further. Two speakers showed greater differences for alveolars than palatals, two showed greater differences for palatals than alveolars, and one showed roughly equal mean differences, as shown in Figure 22 below.

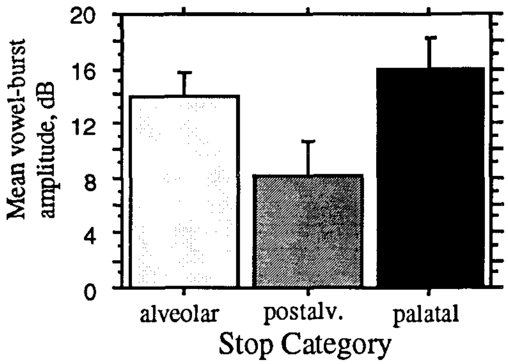


Figure 21. Mean vowel-burst amplitude differences for stop categories in **aTi** environments.

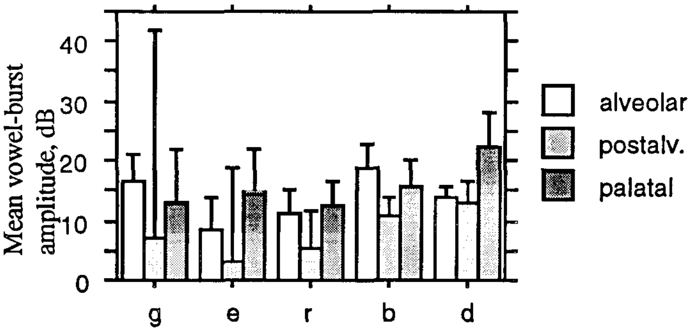


Figure 22. Mean vowel-burst amplitude differences for stop categories in **aTi** environments, by speaker.

For the comparison of relative amplitudes in **aTa** environments, a two-factor ANOVA showed that dentals have significantly higher vowel-burst amplitude differences than either alveolars or post-alveolars. Alveolars and post-alveolars did not differ significantly in this environment. As Figure 24 shows, four of the five speakers showed this trend of low amplitude dentals, while for Speaker B (for whom the poorest recording conditions obtained) the mean vowel-burst difference for post-alveolars just exceeded that for dental, and showed a large standard deviation.

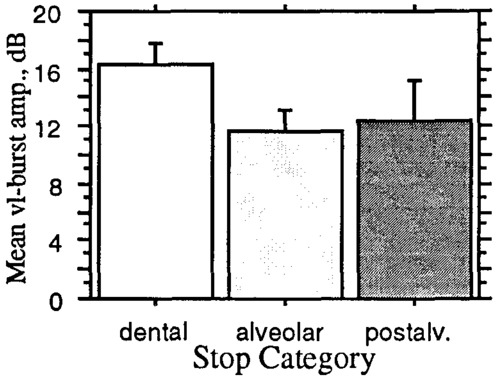


Figure 23. Mean vowel-burst amplitude differences for stop categories in **aTa** environments.

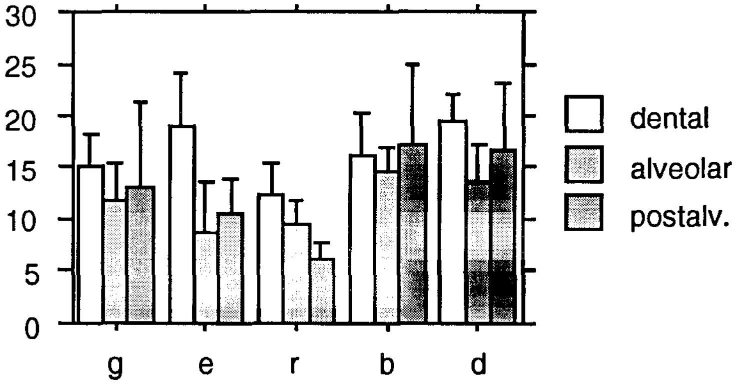


Figure 24. Mean vowel-burst amplitude differences for stop categories in **aTa** environments, by speaker.

Since vowel environment is controlled for in the latter two tests, it is safe to assume that significant differences arise from amplitude differences of bursts, rather than difference in vowel amplitudes. Thus, from Figure 21 we conclude that post-alveolars have the highest amplitude bursts among the stops, in /i/ environments. From Figure 23 we conclude that dentals have the lowest amplitude bursts, in /a/ environments. What is surprising is that alveolars are as quiet in amplitude as palatals in /i/ environments (Figure 21), but are up in the range of post-alveolars, in /a/ environments (Figure 23); that is, they maintain a constant relative amplitude (-12 dB) with respect to the following vowel. Compare the case for the post-alveolar, which shows a smaller difference (-8 dB) in the quieter /i/ environment, and a larger difference (-12 dB) in the environment of the higher amplitude vowel.

In sum, relative amplitude strongly distinguishes post-alveolar bursts as the loudest in /i/ environments. In /a/ environments, dental bursts are the lowest in amplitude. However, the variable behavior of alveolar bursts in the two environments means that they group with the quieter palatals on one hand (before /i/) and the louder post-alveolars on the other (before /a/). As in the case of duration, post-alveolars differentiate themselves far more clearly from the laminals than do the alveolars.

4.3. Discussion of amplitude results

Hypothesis 4 suggested that laminals would be lower in amplitude than apicals at the moment of the burst, because of the comparatively gradual release of the articulator. When looked at overall, this appeared to be the case. Figure 19 shows laminals grouping together, and while apicals are statistically significantly different from each other, they are both lower in vowel-burst differences than the laminal group. However, this result is misleading. When we look at vowel environments separately (Figures 21, 23) post-alveolars do stand out as loudest, and dentals as quietest, but alveolars covary with the amplitude of the following vowel. In /i/ environments they are indistinguishable from palatals; in /a/ environments they are indistinguishable from post-alveolars. These results seem to indicate that the post-alveolar is more resistant to coarticulation with the following vowel than is the alveolar. This is surprising, in light of Butcher's finding that among the four stops it is the apical alveolars that show the greatest stability in their articulations. It may be that for the alveolar in /i/ environments, the tongue blade and body behind the occlusion, while not making tectal contact, are nevertheless raised, making the general oral tract configuration more similar to that of palatals. But the post-alveolar, characterized by a highly specified articulatory program involving forward movement, may permit less coarticulation of the tongue blade and body with /i/. Here again, classificatory features seem to have less to predict about actual acoustics than tongue shape in general.

Our results show partial agreement with those of Shalev et al. (1993) for Toda, and Jongman et al. (1985) for Malayalam; dental bursts are quieter than alveolars. This at least suggests that the articulations in Tiwi are apical alveolar and laminal dental, similar to the Toda and Malayalam articulations.

5. SPECTRAL SHAPES

5.1. Procedure

5.1.1. Data

The subset of data used to examine burst spectra, shown in Table 7, consisted of coronals preceding the vowels /a/ and /i/ in either initial or intervocalic position. Spectra for the three male speakers were examined since there are more males in the set than females and the recordings for the male speakers had, in general, better signal-to-noise ratios than those for the female speakers. By examining data from only one sex, the problem of making comparisons across speaker-dependent frequency range differences was reduced.

Although the vowel environment **Ta** was chosen for the examination of dentals, alveolars and post-alveolars, as the most neutral of the available possibilities, the environment **Ti** was by necessity also used, in order to include a comparison of the palatal with alveolars and post-alveolars. Spectral information for bursts was obtained from CSL by centering a 12.8 ms Hamming window around the burst transient, (yielding a 78 Hz frequency resolution) and displaying a Fast Fourier Transform (FFT) spectrum of the frequency range 0-8000 Hz. Pre-emphasis (.90) was used to boost high frequencies. The procedure is illustrated in Figure 25. As in Dart (1991), a very brief time window was used in order to facilitate capture of just the burst transient itself, rather than a significant portion of frication. It was considered that spectral shape at the moment of the burst would yield a more accurate representation of front cavity shape at the moment of release than would a spectrum including frication (in which the cavity shape is changing.) Spectra were output as data files, each consisting of the amplitudes of 100 frequency values at 80 Hz intervals, which allowed the spectra to be superimposed and averaged.

Table 7: Data subset used for spectral shape of bursts

	Vowel Environment	alveolar	post-alveolar	dental	palatal
Initial stops	#_a	**ta**rin̠a **ta**ŋíni	--	**t̪a**mpinaḷa	
	#_i	**ti**ŋáta **ti**nimuli	--		**t̠i**táka **t̠i**ráka
Intervocalic stops	a_i	yá**ti**	pulá**ṭi**		piká**t̠i** pa**t̠i**pat̠úwu
	a_a	**ti**ŋáta	kita**ṭa**wíni	pirá**t̪a**	
	u_a	kurumú**ta**			
	o_a		wó**ṭa**	pó**t̪a**	

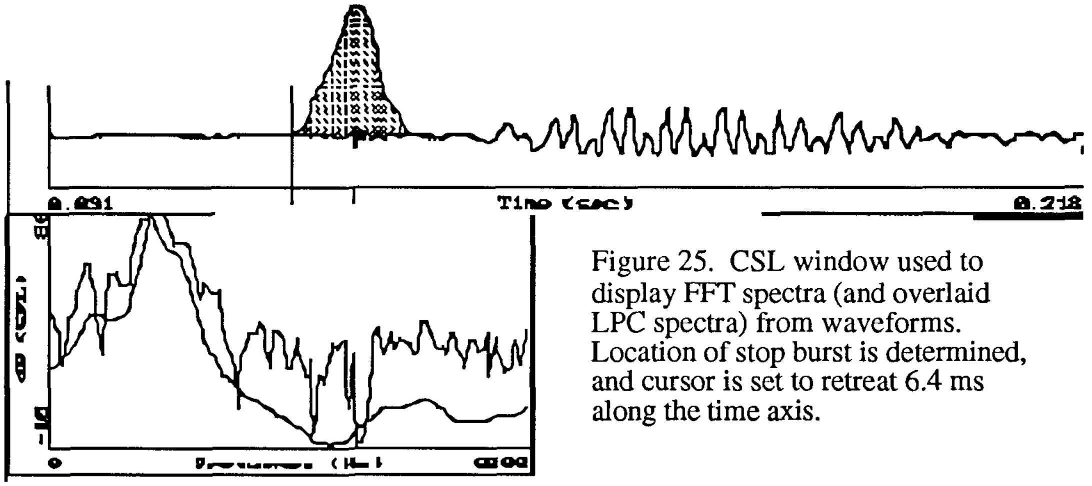


Figure 25. CSL window used to display FFT spectra (and overlaid LPC spectra) from waveforms. Location of stop burst is determined, and cursor is set to retreat 6.4 ms along the time axis.

5.1.2. Analysis

While spectral information was not subjected to statistical analysis in this study, individual tokens for a given category of burst and speaker were overlaid and examined together so as to assure that the spectra shown in the resulting composite averages accurately represented the general shape of spectra for each stop category for each speaker.

Initially it was assumed that preceding vowel environments would not affect burst spectra, because most stops involved a period of silent closure before the burst. However, a first examination of the data showed that a preceding rounded environment can indeed affect the burst spectrum, by lowering amplitudes, especially of frequencies under 1200 Hz as compared with their unrounded counterparts. Presumably there is perseveration of rounding from the preceding vowel through the consonant burst. Figures 26a and 26b show cases of this effect. In each of the graphs below, abscissas are shown on a logarithmic scale to better approximate the perceptual weighting of these frequencies. Graphs were smoothed using a five-point moving average of values, for the hundred increments between 0 and 8000, which is why frequency values do not extend below 200 Hz. Ordinates show amplitudes of each averaged point in dB.

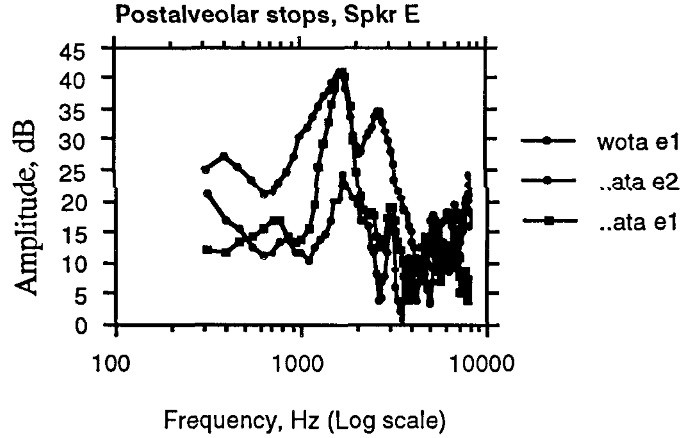


Figure 26a. Postalveolar stop spectra, Speaker E, rounded and unrounded vowel environments. Note low peak at 1200 Hz for stop burst in /wo̦ta/.

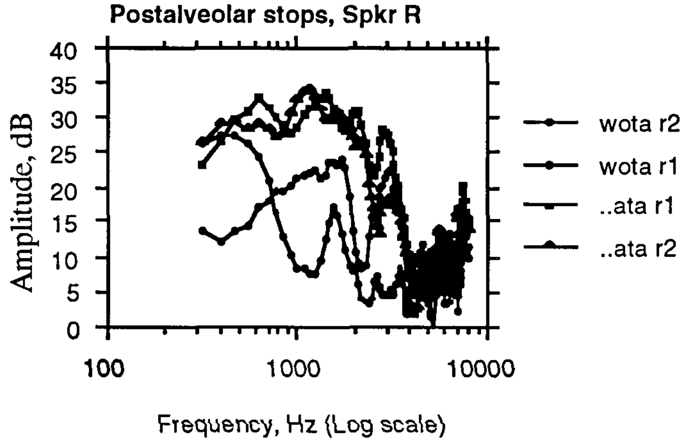


Figure 26b. Postalveolar stop spectra, Speaker R, rounded and unrounded vowel environments. The two lower cuves are for the stop burst in /wo̦ta/.

Cases which departed from the norm in amplitude but retained the general shape of other exemplars, (as in the lowest amplitude curve in Figure 26a) were retained, but cases which clearly departed in both amplitude and general shape (as in the two low amplitude curves in Figure 26b) were excluded from averaged data.

A second strong effect of vowel environment was found for each of the stop categories. Spectra for bursts followed by /i/ were very characteristic of the spectrum for /i/ itself. Hence, it was not possible to characterize palatal spectra separate from other stops.

5.2. Results

Spectral averages for the Ta environment are shown in Figure 27 for the three male speakers. Many of the spectra in Figure 27 show a low frequency peak of high energy below 1 kHz. While this is not broad-band noise, this peak did appear to represent either an external noise source, or an extra resonance unrelated to the coronal articulations, characteristic to different extents of different speakers. Note that it is most characteristic of Speaker D, to a lesser extent of Speaker R, and least characteristic of Speaker E. (Initially, a plausible alternative appeared to be that this low frequency energy peak represented F1 of the following /a/, but it also occurs in the spectra of Speaker D for Ti environments, where there is also a clear F1 peak at about 300 Hz.) Since it appears to related to characteristics of the speaker or the recording environment, this peak was disregarded as a potential differentiating characteristic between coronals. Moreover, above about 7 kHz all of the spectra behave similarly, rising in response to the band of external noise between 8 and 11 kHz mentioned earlier. Thus, usable differentiating information was located between 1 and 7 kHz. This range is, however, the region to which the human auditory system is most attuned.

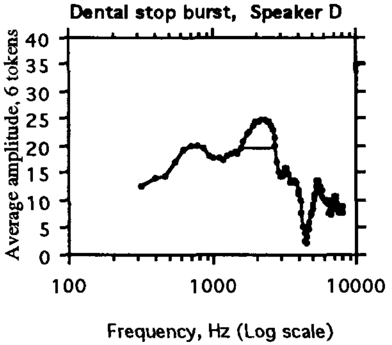
Each of the spectral averages has its major zero (low amplitude area) between 3 and 4 kHz. This seems to be a consistent factor which identifies the group "coronal" as a whole in Tiwi. The consistency of this result is surprising, considering that Dart (1991) found zeroes ranging from 4 to 7 kHz for Malayalam coronals. Although the spectra have yet to be compared with non-coronals, this zero seems to be a robust contender for a characteristic of Tiwi coronals in general.

These composite spectra, with their midfrequency zero and few peaks, are not easily classified as either "diffuse" or "rising" in Jakobson, Fant and Halle's terms. A more useful measure in differentiating among dental, alveolar and post-alveolar categories seems to be the bandwidth of the first major peak above 1 kHz (centered in each case approximately between 1.7 and 2 kHz, which we take to be the approximate F2 locus). To obtain comparable measures of bandwidth for each category of stop, the frequency bandwidth (measured by hand from composite spectra) is shown at 6 dB down from the maximum amplitude of the peak, as a horizontal line in each of the panels in Figure 27.

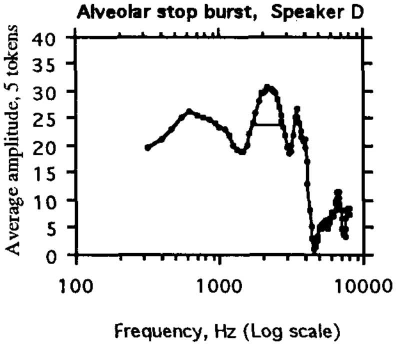
For Speakers D and E in the top two rows, the alveolar has a broader first peak bandwidth than the post-alveolar (approximately 700 Hz as opposed to 600 Hz for D and 1900 vs 1800 Hz for E), and a second high amplitude excursion immediately following the first. For these speakers, the post-alveolar has an isolated, narrow-bandwidth major peak, without a second high-amplitude peak. This may indicate that F2 and F3 are joined in the post-alveolar spectrum, forming a single peak. The contrast between Speaker R's alveolars and post-alveolars in the bottom row does not pattern with the two other speakers' as there is no narrow-bandwidth peak in the post-alveolar. This may be because there is masking noise from other sources.

Dental bursts for each speaker are marked by the smaller prominence of any peak; i.e. amplitude excursions are not so extreme. The first peak, as well as the second "hitch" or plateau following it are broader in bandwidth than either of the apicals (2000 and 1800 Hz respectively for D and E), implying a greater diffusion of energy.

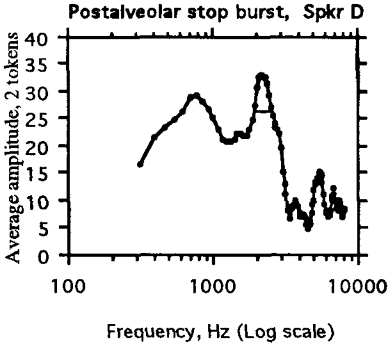
Dental Alveolar Post-alveolar



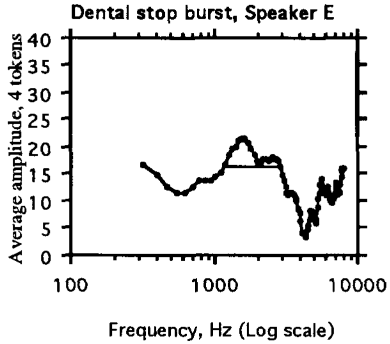
Bandwidth measured at: 19 dB
Bandwidth: 900 Hz
Frequency of peak amplitude: 2000 Hz



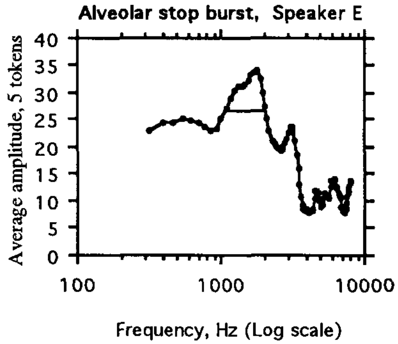
Bandwidth measured at: 24 dB
Bandwidth: 700 Hz
Frequency of peak amplitude: 2100 Hz



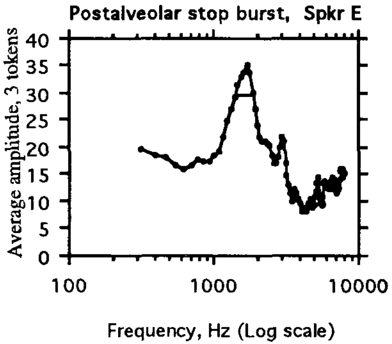
Bandwidth measured at: 26 dB
Bandwidth: 600
Frequency of peak amplitude: 2000



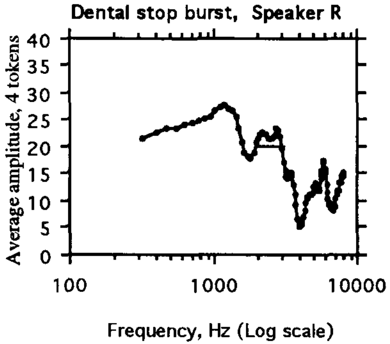
Bandwidth measured at: 16 dB
Bandwidth: 2000 Hz
Frequency of peak amplitude: 1800 Hz



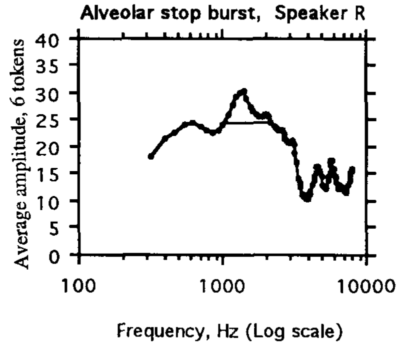
Bandwidth measured at: 26 dB
Bandwidth: 1000 Hz
Frequency of peak amplitude: 1900 Hz



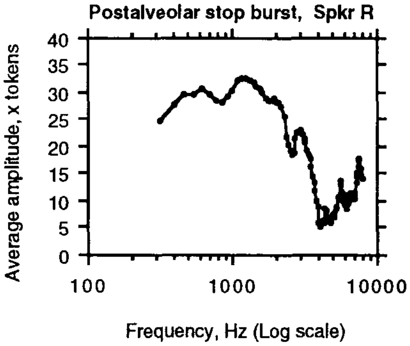
Bandwidth measured at: 30 dB
Bandwidth: 500 Hz
Frequency of peak amplitude: 1800 Hz



Bandwidth measured at: 20 dB
Bandwidth: 1100
Frequency of peak amplitude: 2000



Bandwidth measured at: 24 dB
Bandwidth: 1100 Hz
Frequency of peak amplitude: 2000 Hz



Bandwidth measured at: not measured

Frequency of peak amplitude: 1700 Hz

Figure 27. Mean spectra of coronal stop bursts before /a/ for three male speakers.

5.3. Discussion of spectral measurements

Hypothesis 5 suggested that [-anterior] stops would show a narrower concentration of energy than [+anterior] stops. As it happens this data did not allow us to characterize palatals apart from the vowel /i/. The energy for the ([-anterior]) post-alveolars is indeed less distributed than in either of the [+anterior] segments. However, there seems to be less a discrete split between [-anterior] and [+anterior] than a continuum from dental to post-alveolar. There is a range of differences in energy concentration from dentals (least concentrated) to post-alveolars (most concentrated.) Again, this seems less related to the classificatory feature [anterior] than it does to a more scalar measurement of anteriority which reflects the size of the cavity in front of the constriction.

Recall that Stevens and Keyser (1989) suggest that [-anterior] can only be strongly salient in [-coronal] segments. However, Lahiri et al. (1984) found that coronal spectra can be diffuse-flat. As such, a high amplitude mid-frequency peak can, and does, seem to characterize the Tiwi alveolars and post-alveolars. Dart (1991) found relatively less peaked spectra for apicolaminal dentialveolars, increasingly peaked spectra for apical alveolars and a lower frequency peak for (apical or sublaminal) post-alveolars, in both Malayalam and 'O'odham. These results are broadly consistent with the Tiwi data.

Perkell, Boyce and Stevens (1979) observe that a sublaminal cavity lowers acoustic resonances. For the Tiwi post-alveolars, we surmise that the resonance most affected is F3; spectrograms show clear evidence of a common locus for F2 and F3 (these formant transitions look very much like velars.) Halle and Stevens (1989) in examining Polish fricatives, show an amplitude zero at 4 kHz for retroflex /ʂ/, but they attribute the peak to a merging of F3 and F4. For Tiwi post-alveolars the meeting point seems to be uncontroversially F2 and F3. It is this common locus which creates the isolated, dominant peak we observe in post-alveolar spectra.

6. CONCLUSION

The contrast between coronal stops of different categories, as in Tiwi, is relatively unusual among the world's languages. Languages with such features allow an opportunity to examine acoustic concomitants of regular but subtle differences in place of articulation, shape of the active articulator, and configuration of the vocal tract as a whole.

Let us summarize the acoustic characteristics of the Tiwi coronals and relate them to probable articulations. The dental allophone of the laminal phoneme ([-apical] [+anterior]) has intermediate duration, a low burst amplitude and a relatively diffuse-energy spectrum. This is expected for an articulation with a wide constriction in the front of the mouth. The post-alveolar phoneme ([+apical] [-anterior]) is well differentiated in nearly every domain; it is short in duration, high in burst amplitude, and has a narrow distribution of spectral energy. This is consistent with employment of a small, forward-moving articulator without narrowing of the oral tract behind it. Given the lowering of F3, there is probably a significantly large front cavity involved, suggesting that the articulation in Tiwi is probably sublaminal. The palatal allophone of the laminal phoneme ([-apical] [-anterior]) is differentiated by long VOT and associated frication, showing evidence of a long constriction. The alveolar phoneme ([+apical] [+anterior]) appears as the least distinctive, as it often grouped with one or more of the other categories. However, it differs from the post-alveolar in voiceless closure duration, VOT, burst amplitude and bandwidth of its major spectral peak. These differences are consistent with it being a more forward articulation than the post-alveolar, probably not involving movement during closure. It differs from the palatal in VOT, suggesting a lighter articulator than the palatal. Its greater relative amplitude than the dental points to a difference in constriction size between them.

What do Tiwi speakers use to hear differences between these relatively rare oppositions? Three of these stops contrast phonologically. Let us consider each contrast in turn. In vowel environments other than /i/, post-alveolar /ṭ/ is distinguished from both dental /t̪/ and alveolar /t/ in having a higher burst amplitude, being shorter in voiceless closure and VOT, and having an isolated narrow-bandwidth spectral peak. Dental /t̪/ is distinguished from alveolar /t/ by a lower amplitude burst and more distributed spectrum. When the vowel /i/ follows, dental /t̪/ is realized as palatal [ṯ]. Postalveolar /ṭ/ is shorter in voiceless closure and VOT and higher in burst amplitude than both alveolar /t/ and palatal [ṯ]. It has a narrower spectral energy distribution than alveolar /t/. The alveolar and palatal are distinguished by VOT. Thus, Hypothesis 1 is validated: although the durations, relative amplitudes and spectral shapes of bursts cannot by themselves differentiate all the stop categories, taken together they do make the required distinctions.

Earlier, Dart (1991) looked for, but did not find any "one property (that) consistently differentiates the apicals from the laminals." Despite increased focus on components of duration and average spectra, as well as the inclusion of a study of relative amplitudes, the result is the same in this study. Similarly, the present study does not find any one property that separates anterior from non-anterior coronals.

Thus, the more specific Hypotheses 2 through 5 failed in their predictions. There was no straightforward mapping between the features [apical] and [anterior] and acoustic correlates in any of the tests presented above. Instead, these classificatory phonological features seem to be mediated by a phonetic implementation layer which arises (partly) from feature specifications. The phonetic parameters, or determinants, in this layer include oral tract configuration, articulator mass and movement during closure location. It is this intermediate layer which relates in a direct way to acoustics.

In the larger picture, these results indicate support for a theory in which classificatory features relate directly to production, but not to perception; there could be no "feature detectors" which decode simple or invariant relationships between features and acoustics, because, in this case, no such relationships exist. The perceptual apparatus is likely to be significantly more sophisticated, retrieving classificatory feature values from a relatively opaque signal.

For stops, duration information contributes to an acoustic "picture" of the cavity behind the constriction. Spectral information yields a picture of the cavity in front of the constriction. Burst amplitude and VOT give a picture of the active articulator. Formant transitions, not studied here, would further distinguish the stops.

The absence of simple correlations between features and acoustic properties is likely to be highlighted by extending the present study to nasals, which have the same articulatory contrasts as the stops in Tiwi. Nasals, however, do not have VOTs, nor burst spectra, nor burst amplitudes comparable to stops. Therefore, these properties cannot relate to the articulatory features of nasals. Instead, different properties, such as nasal formants and zeroes, must be examined.

We have assumed that [apical] and [anterior] are indeed phonologically relevant in Tiwi, as determined by rule governed behavior, particularly distributional patterns. Unless this assumption is wrong, the results found here lend weight to the idea of the primary importance of articulation rather than acoustics in defining distinctive features.

Acknowledgments

The authors gratefully acknowledge the assistance of Marie Godfrey, the Tiwi Land Council, and the Catholic Church at Nguiu in making the collection of the data on which this paper is based possible. We also thank Phil Rose of Australian National University and John Henderson of the Institute for Aboriginal Development for their help in establishing contacts. Our greatest debt is, of course, to the Tiwi speakers who generously shared their knowledge of their language with us. An earlier version of this paper was presented as a master's thesis by the first author.

References

Blumstein, S. and Stevens, K. (1978) Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64, 1358-1368.

Butcher, A. (in progress) *Phonetics of Australian Languages,* ms.

Byrd, D. (1993) 54,000 American Stops. *UCLA Working Papers in Phonetics*, 83: 97-116.

Catford, I. (1977) *Fundamental Problems in Phonetics*, Bloomington: Indiana University Press

Crystal, T. and A.S. House (1988) Segmental durations in connected speech signals: current results. *Journal of the Acoustical Society of America* 83(4), 1553-1573.

Dart, S. (1991) Articulatory and Acoustic Properties of Apical and Laminal Articulations. *UCLA Working Papers in Phonetics*, 79.

Dixon, R.M.W. and B.J. Blake, eds. (1991) *The Handbook of Australian Languages,* vol.4, Melbourne: Oxford University Press Australia

Dixon, R. M.W. (1980) *The languages of Australia.* Cambridge: Cambridge University Press.

Dixon, R. M. W. (1970) Proto-Australian laminals. *Oceanic Linguistics,* 9, 79-103.

Fant, G. (1960) *Acoustic Theory of Speech Production.* Mouton, The Hague.

Fischer-Jørgensen, E. (1954) Acoustic analysis of stop consonants. *Miscellanea Phonetica* 2, 42-59.

Fischer-Jørgensen, E. (1964) Sound duration and place of articulation in Danish *Zeitschrift für Sprachwissenschaft und Kommunikationsforschung* 17, 175-207

Halle, M. and Stevens, K.N. 1989. (Polish fricatives)

Jakobson, R., Fant, G. and Halle, M. (1952) *Preliminaries to speech analysis.* Cambridge: MIT Press

Jernudd, B. (1974) Articulating Gunwingju laminals. *Papers in Australian Aboriginal Languages, Linguistic Communication #14,* Melbourne: Monash University (ed. B.J. Blake).

Jongman, A., Blumstein, S. and Lahiri, A. (1985) Acoustic properties for dental and alveolar stop consonants: a cross-language study. *Journal of Phonetics* 13,235-251.

Keating, P. (1984a) Phonetic and phonological representations of consonant voicing *Language* 60, 286-319.

Keating, P. (1984b) Aerodynamic modeling at UCLA. *UCLA Working Papers in Phonetics,* 59.

Keating, P. (1984c) Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics,* 59, 29-34.

Keating, P. (1988) The phonology-phonetics interface. In *Linguistics: The Cambridge Survey* (1), 281-302. Cambridge: Cambridge University Press.

Keating, P. (1991). Coronal Places of Articulation. in *The Special Status of Coronals , Phonetics and Phonology* 2 (C. Paradis and J.-F. Prunet, eds.) San Diego: Academic Press

Keating, P. and Lahiri, A. (1992) Fronted velars, palatalized velars, and palatals. *UCLA Working Papers in Phonetics* 81, 76-105.

Keating, P., Westbury, J. and Stevens, K. (1980) Mechanisms of stop consonant release for different places of articulation. *Journal of the Acoustical Society of America* 67, Supp. 1:S93

Kewley-Port, D. (1983) Time-varying features as correlates of place of articulation in the articulation of stop consonants. *Journal of the Acoustical Society of America*, 73, 322-335.

Klatt, D. (1975) Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3, 129-140.

Ladefoged, P. and Maddieson, I. (1986) Some of the Sounds of the World's Languages, *UCLA Working Papers in Phonetics,* 64.

Ladefoged, P. and Maddieson, I. (1994, ms). *The Sounds of the World's Languages.* To be published by Blackwells, Oxford.

Lahiri, A. L. Gewirth and S. Blumstein, (1984) A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study *Journal of the Acoustical Society of America* 76, 391-404.

Lee, J. (1987) *Tiwi Today: a study of language change in a contact situation.* Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University

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

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

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

Maddieson, I. and Precoda, K. (1992) *UCLA Phonological Segment Inventory Database,* UCLA

Maddieson, I., Spajic, S., Sands, B. and Ladefoged, P. (1993) Phonetic Structures of Dahalo. *UCLA Working Papers in Phonetics*, 84.

Oates, W. J. (1967) Syllable patterning and phonetically complex consonants in some Australian languages. *Papers in Australian Linguistics* no. 1:29-52. (Pacific Linguistics, series A, no. 10)

Osborne, C. R. (1974) *The Tiwi Language: Grammar, myths and dictionary of the Tiwi Language spoken on Melville and Bathurst Islands, northern Australia* Canberra: Australian Institute of Aboriginal Studies.

Perkell, J., S. Boyce & K. Stevens (1979) Articulatory and acoustic correlates of the [s-ʃ] distinction, in J.J. Wolf and D.K. Klatt, eds, *Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America*, Acoustical Society of America, New York

Ren, H-M. (1985) Linguistically-conditioned duration rules in a timing model for Chinese. *UCLA Working Papers in Phonetics* 62: 34-49.

Shalev, M., Ladefoged, P. and Bhaskararao, P. (1993) Phonetics of Toda. *UCLA Working Papers in Phonetics,* 84.

Stathopoulos, E. and Weismer, G. (1983) Closure duration of stops consonants. *Journal of Phonetics*, 11: 395-400.

Stevens, K. (1989) On the quantal nature of speech. *Journal of Phonetics,* 17, 3-45.

Stevens, K., Keyser, S.J., and Kawasaki, H. (1986) Toward a phonetic and phonological theory of redundant features. in *Invariance and Variability in Speech Processes,* J.S. Perkell and D.H. Klatt, eds., pp. 426-463. New Jersey: Lawrence Erlbaum Associates.

Stevens, K. and J. Keyser (1989) Primary features and their enhancement in consonants *Language,* 65(1),81-106

Stone, M. (1991) Toward a model of three-dimensional tongue movement. *Journal of Phonetics,* 19 (3/4), 309-320.

Svarny, O. and Zvelebil, K. (1955). Some remarks on the articulation of the "cerebral" consonants in Indian languages, especially in Tamil. *Archiv Orientalni,* 23, 374-407.

Weismer, G. (1980) Control of the voicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations. *Journal of Phonetics,* 8, 427-438.

Zue, V. (1976) Acoustic Characteristics of Stop Consonants: A Controlled Study. Ph D. dissertation MIT, distributed by Indiana University Linguistics Club, Bloomington, IN.

Zue, V. & Seneff, S. (1988) Transcription and alignment of the TIMIT database. Proceedings of the second meeting on advanced man-machine interface through spoken language.

Appendix: List of words used in duration study

Tiwi word	*English Gloss*
aḷitiwíyi	*female kangaroo*
erepuṭára	*lip*
kitaṭawíni	*bread*
kumúti	*beard*
kurumúta	*upper arm*
milikúṭi	*big toe*
miṇapúṭi	*cold (A)*
muritíwa	*neck*
ŋawatíwi	*people (pl., excl.)*
pat̲ipat̪úwu	*turtle egg*
pat̪iwáni	*dead*
pikát̲i	*swordfish*
pirát̪a	*rice*
pit̪ára	*eyes*
poalíti	*beard*
pót̪a	*bone*
puláṭi	*milk*
taŋaríma	*camp*
taŋíni	*stick*
tarín̲a	*back of head*
tinimúli	*eyebrow*
tiŋáta	*beach*
tíwi	*people*
tokampíni	*bird (m)*
tokampúŋa	*bird (f)*
toŋuláka	*canoe*
tuṭuṭúṭi	*move around*
tuulúa	*head*
t̪ampináḷa	*cliff*
t̪apára	*moon*
t̪apinári	*morning*
t̪ukwantáŋa	*back*
t̲ikapúṭi	*cold (N)*
t̲iráka	*wallaby*
t̲itáka	*swordfish*
t̲íṭi	*bad, ugly*
wóṭa	*bush*
wuratíŋa	*forehead*
yáti	*one*

Phonetic Structures of Iaai

Ian Maddieson and Victoria Anderson

1. Introduction

Iaai [jaːi] is a one of the twenty-five or so indigenous languages of New Caledonia, a French "overseas territory" in the South Pacific. This territory consists principally of the large island called Grande Terre and an archipelago of three smaller islands to the east known as the Loyalty Islands. Iaai is spoken by about two thousand people on Ouvéa, the northernmost of the Loyalty Islands. Its closest linguistic relatives are the other Loyalty Islands languages, Drehu (Lifou) and Nengone (Mare). All the New Caledonian languages are generally assigned to a separate New Caledonia group of the Oceanic branch of Austronesian languages.

The most authoritative description of Iaai is Ozanne-Rivierre (1976), which outlines the phonology and morphology of the language in some detail and provides a syntactic sketch. The earlier outline grammar by Tryon (1968) is superceded by this work. A bilingual Iaai-French dictionary was published by Ozanne-Rivierre in 1984, based in large part on texts collected by Jean Guiart in 1948 but supplemented by more recent materials. There are no prior studies which have focussed on the phonetics of the language, and in particular no published instrumental phonetic studies of Iaai.

A second Austronesian language, belonging to the Polynesian group, is also spoken on Ouvéa. This language is called Faga Ouvéa (also West Uvean) and is closely related to the languages of Futuna and the Polynesian enclaves in southern Vanuatu. It is this language which has given the island its present administrative name. In the second half of the eighteenth century Polynesian-speaking settlers arrived and occupied the northern and southern extremities of the island, where their descendants remain dominant. At both the northern and the southern language boundaries there are areas of bilingualism, but Iaai remains the dominant language of the central part of Ouvéa, including the adminstrative capital of Fayaoué [fajawe].

Iaai has been influenced by a variety of other languages in addition to Faga Ouvéa. It contains a considerable number of lexical borrowings from English as English-speaking missionaries from the London Missionary Society established themselves on Ouvéa in 1858 and opened schools as early as 1859. The orthography established by these Protestant missionaries and continued by their successors from the Paris Mission remains in use, reinforced by its use in the Bible translation of 1901, although attempts have been made to introduce a more consistent and better-suited orthography (Wahéo and Wahéo 1987). The missions employed catechists from Samoa, providing another source of Polynesian influence. As in the rest of New Caledonia, the current medium of instruction in schools on Ouvéa is French, and most speakers of Iaai are reported to be fairly fluent in French. A substantial number of Iaai speakers live in Nouméa and other parts of Grande Terre, where command of French is close to essential for economic success.

Data

For the present study of the phonetic structures of Iaai, selected words were chosen from Ozanne-Rivierre's grammar and dictionary to illustrate all of the contrastive segments of the language. These words were checked for currency with Tai Wahéo and then recorded by two male and three female native speakers in two separate sessions. The resulting corpus of data consists of high quality analog audio recordings of 140 words spoken in isolation by these five speakers. To minimize background noise a Shure SM10A close-recording headset microphone was used for all audio recordings. In addition, simultaneous video-audio recordings were made for all five speakers, focusing on labial articulations. To examine coronal consonant articulations in more detail, palatographic data was collected for 1 male speaker and the 3 female speakers. The data were collected by the first author in February 1993 in Nouméa, New Caledonia.

A number of aspects of this language are of particular phonetic interest. For an Austronesian language, Iaai has a relatively large vowel inventory consisting not only of ten different vowel qualities, but also a phonemic length distinction. In distinguishing these vowels, large differences in lip rounding and spreading are used and these are independent of the front-back distinction. Moreover, there are interesting limitations on the distribution of certain vowels according to the consonant context. The consonant inventory is also quite extensive. The language has three coronal places of articulation, dental, retroflex and pre-palatal, for stops and nasals. In the stops, these three places appear to be acoustically differentiated along lines which differ from most other languages of the world which make use of such distinctions. The Iaai consonant inventory also contains voiced and voiceless sonorants which have phonemic status. What follows will present a general phonetic survey of the language, with emphasis laid on these various aspects of particular interest.

2. Vowels

2.2 Vowel contrasts and distribution

Iaai has a maximal system of ten distinctive vowel qualities. Phonologically distinct vowel qualities are shown on a conventional vowel chart in Figure 1:

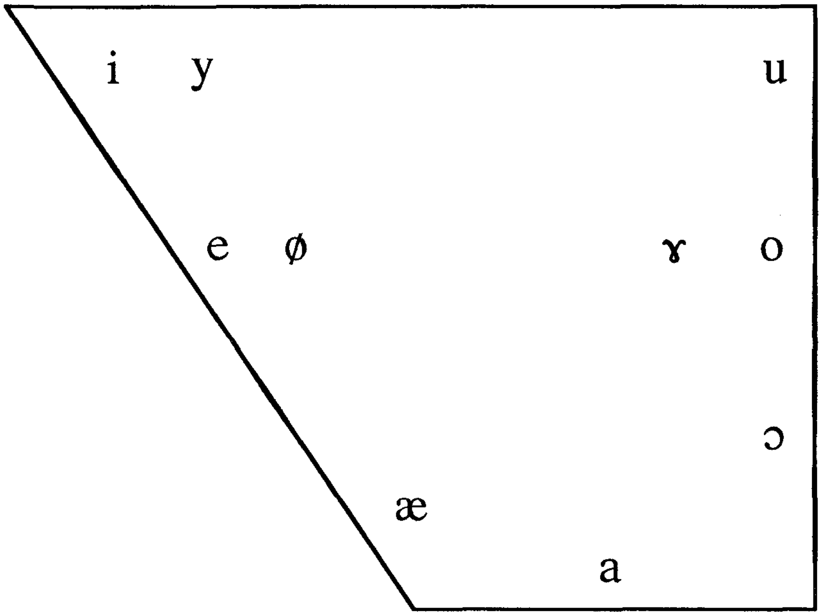


Figure 1: Distinctive vowel qualities of Iaai.

The placement of symbols on this chart follows closely the descriptions given by Ozanne-Rivierre (1976), with the exception that what is transcribed /ɤ/ here is represented as /ə/ in her analysis and described by her as a central rather than a back mid unrounded vowel. The formant pattern of this vowel seems to indicate that it is a back vowel similar in height to /o/, /e/ and /ø/.

Lip position pictures certainly show it to be quite unrounded. These properties of the vowel will be discussed in more detail below. We also prefer to use the symbol [æ] for the lowest of the front vowels on the chart, rather than [ɛ]. This symbol indicates its phonetic quality more precisely, as Ozanne-Rivierre noted.

Each of the vowels occurs distinctively long and short. There are no auditorily salient differences in the qualities of the vowels depending on their length. Words illustrating the twenty quality and length distinctions are shown in Table 1 below:

Table 1: Word illustrating vowel quality and length distinctions in Iaai.

Vowel (IPA)	Exemplar (IPA)	Exemplar (Orthography)	Gloss
i	**ito**	*ito*	‘round hut’
ii	**ṭii**	*trii*	‘tea’
y	**yɤɤ**	*ûöö*	‘tree’
yy	**yy**	*ûû*	‘quarrel’ (n)
e	**ṭeṇ**	*trenr*	‘rich’
ee	**eet**	*eet*	‘fishing line’
ø	**møk**	*mok*	‘ill, dead’
øø	**møøk**	*mook*	‘close eyes’ (v)
æ	**ɥæk**	*vëk*	‘four’
ææ	**m̥ææk**	*hmëëk*	‘heavy’
a	**at**	*at*	‘person’
aa	**aat**	*aat*	‘wounded’
ɔ	**θɔm**	*thâm*	‘peck’ (v)
ɔɔ	**θɔɔn**	*thâân*	‘cook in stone oven’
o	**oṭ**	*otr*	‘lobster’
oo	**tooṭ**	*tootr*	‘marry’ (v)
ɤ	**ɤṭ**	*ötr*	‘pot’
ɤɤ	**lɤɤ**	*löö*	‘banana leaf’
u	**u**	*u*	‘yam’
uu	**uu**	*uu*	‘fall’ (v)

There are some notable limitations, largely associated with labial consonants, on the distributions of certain of these vowels. The low vowels /**a**/ and /**æ**/ are largely in complementary distribution. The low front vowel /**æ**/ is restricted to occurrence after the labial consonants /**b**, **m̥**, **m**, **p**, **f**, **ɥ**,**ɥ̥** and the vowel /**y**/. Ozanne-Rivierre notes that /**æ**/ occurs to the exclusion of the low central vowel /**a**/ after /**ɥ**/, /**ɥ̥** and /**y**/, that is, after consonants and vowels with rounded lips and a high front tongue position. On the basis of a search of the dictionary, it seems also to be the case that /**a**/ occurs to the exclusion of /**æ**/ after /**bw**, **mw**, **m̥w**, **w**/, a fact that Ozanne-Rivierre does not comment on. However, after /**b**, **m̥**, **m**, **p**, **f**/ the vowels /**æ**/ and /**a**/ contrast in comparable environments, although no exact minimal pairs are available. Among illustrative pairs are **m̥an** ‘oily’ and **mæn** ‘strength’, **maaṭ** ‘tern’ and **mæṭ** ‘ripe; weak’, **fat** ‘gift to chief’, **fææṭ** ‘pull alongside’. The subset of consonants /**b**, **m̥**, **m**, **p**, **f**, **ɥ**,**ɥ̥** and the vowel /**y**/ also condition a fronted [œ]-like variant of the lower mid back rounded vowel /ɔ/.

The mid front rounded vowel /**ø**/ occurs in only a few words — little over half a dozen are listed by Ozanne-Rivierre. In each case this vowel is followed by a velar consonant and in all except one of these words it is preceded by a labial consonant, usually **/m/**. The words, apart from those listed in table 1 are **ɲamøk** 'mourning cries', **eekemøk** 'burial ground', and **møxon** 'shortness of breath', all three of which might be etymologically related to **møk** 'ill, dead' and **omøk** (a species of edible fish), **bebøŋ** 'rotten' and **ɲøøk** 'dedicate'. The last is the only word which does not have a labial consonant before the vowel in addition to the following velar.

We will return to the question of what might account for some aspects of these distributional restrictions in a later section after the consonant inventory has been described.

2.2 Acoustic analysis

In order to characterize the vowels of Iaai acoustically two set of measurements were made. For the first, an average of three words containing each vowel quality and of each length were selected from the audio recordings of each of the five speakers. The selected words largely fall into two series; those with a labial preceding the vowel, and those with a coronal consonant following the vowel. Vowels in the environment of labial consonants were included since this is where the biggest distinction among vowels exists. By including labial environments for each vowel wherever possible we hoped to control for differential effects of surrounding consonants on the acoustic pattern of the vowels. However, the data are not fully balanced for consonant environment, and some caution should be exercised in interpretation of the results.

Waveforms were digitized and measurements taken using the Kay Elemetrics Computer Speech Laboratory. Measurements for the first three vowel formants were taken during the steady-state portion of the vowel, which often but not always corresponded with the area of greatest vowel amplitude. An FFT averaged over 51.2 msec, and an overlaid LPC averaged over 30 msec, both centered at the middle of the steady-state, were used to determine vowel formants, except in cases where the duration of any steady-state portion was less than 51.2 msec. In these few cases, a 25.6 msec FFT and 20 msec LPC window were used, to avoid including formant transition effects contributed by consonants. For both men and women an LPC filter order of 12 coefficients was usually optimal for identifying formants. Some problematic cases were examined with different settings.

Measurements were also made to examine the intrinsic pitch of vowels using a further time-aligned window containing a narrowband spectrogram. The frequency of the tenth harmonic, or the next lowest reliably identifiable harmonic, was recorded at the vowel steady-state, clear of the raising or lowering effects of adjoining consonants. This frequency was then divided by the harmonic number to obtain an accurate estimate of F0. Cases in which the auditory impression suggested that a marked intonation pattern had been used, such as a rising question contour, were excluded.

Measurements for vowel duration were taken using a simultaneously displayed, time-aligned waveform, energy curve and wideband spectrogram. The most reliable duration measurements are those of vowels between stops or clearly demarcated nasals or laterals but the data contained too few such instances to make a sufficient number of measurements. Consequently the duration measurements include a variety of contexts, including both open and closed syllables and vowels with and without a preceding consonant. However, a number of glide-vowel sequences were excluded because of the difficulty of segmenting them.

The second set of measurements were made from the audio track of the video-recordings. In this case frequencies of the first three formants were determined in a similar fashion to that outlined above using the Kay CSL. The purpose in this case was to correlate the formant measurements with lip position measures made from the video images. These results will be described following the discussion of the consonant articulations.

2.2.1 Vowel formants

The acoustic pattern of the vowels in an F1 vs F2 space for the female and male speakers respectively are shown in Figures 2 and 3. These plots show a vowel symbol for each individual token measured, with an ellipse drawn centered on the mean and having radii of 2 standard deviations along the first two principal components of the distribution. While the axes are marked in Hz they are scaled on the Bark scale. Only nine vowels are plotted since only a few tokens of the restricted vowel /ø/ in a single context are available. Long and short vowels are not distinguished in the figures since there are no consistent effects of length on the formant values. That is, long vowels do not consistently have a higher or lower F1 or F2, nor are they more peripheral or more central in the acoustic vowel space than their short counterparts. Analyses of variance for female and male speakers separately showed that /o/ has a significantly higher F1 (i.e. is perceptually lower) than /oː/ for the three female speakers but not for the men, /y/ had a significantly lower F2 than /yː/ for the men but not for the women, and /æ/ had a significantly lower F2 than /æː/ for the women, which was matched by a trend in the same direction among the men. Both groups of speakers also showed a trend for a higher F1 in /ɤ/ than in /ɤː/.

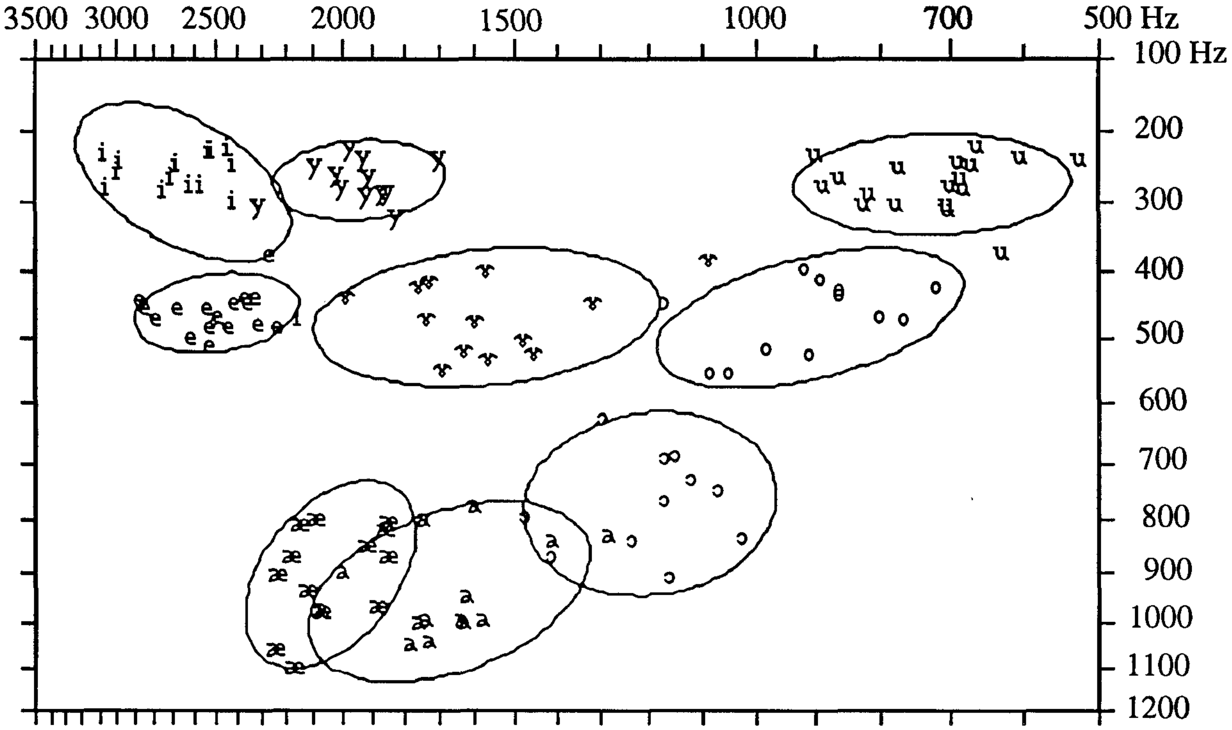


Figure 2: F1 vs F2 for female speakers, long and short vowels together.

Figure 2 shows that these nine vowels are well separated in the F1/F2 space for the women speakers. There is a very distinct demarcation between high, mid and low vowels in the F1 dimension, but the F2 values are distributed in a more gradient fashion. The ellipses for the three non-low front vowels /i, y, e/ are quite compact on the Bark scaling used. (Note that the ellipse plotted for /i/ would be considerably smaller and have a more horizontal orientation but for a single token which falls in the range of F1/F2 values for /e/. Auditorily this vowel token, in the word orthographically represented as *hwii*, sounds like a mid rather than a high vowel and it might well be that this word for this speaker indeed has an instance of the vowel /e/ rather than /i/.) In the F2 dimension, the mid back unrounded vowel /ɤ/ lies almost equidistant from the back rounded /o/ and the front unrounded /e/ vowels, whereas the front rounded /y/ is much closer to front unrounded /i/ than to back rounded /u/. The low vowels /a/ and /æ/ show considerable overlap, but recall that the contrast between these vowels is somewhat marginal.

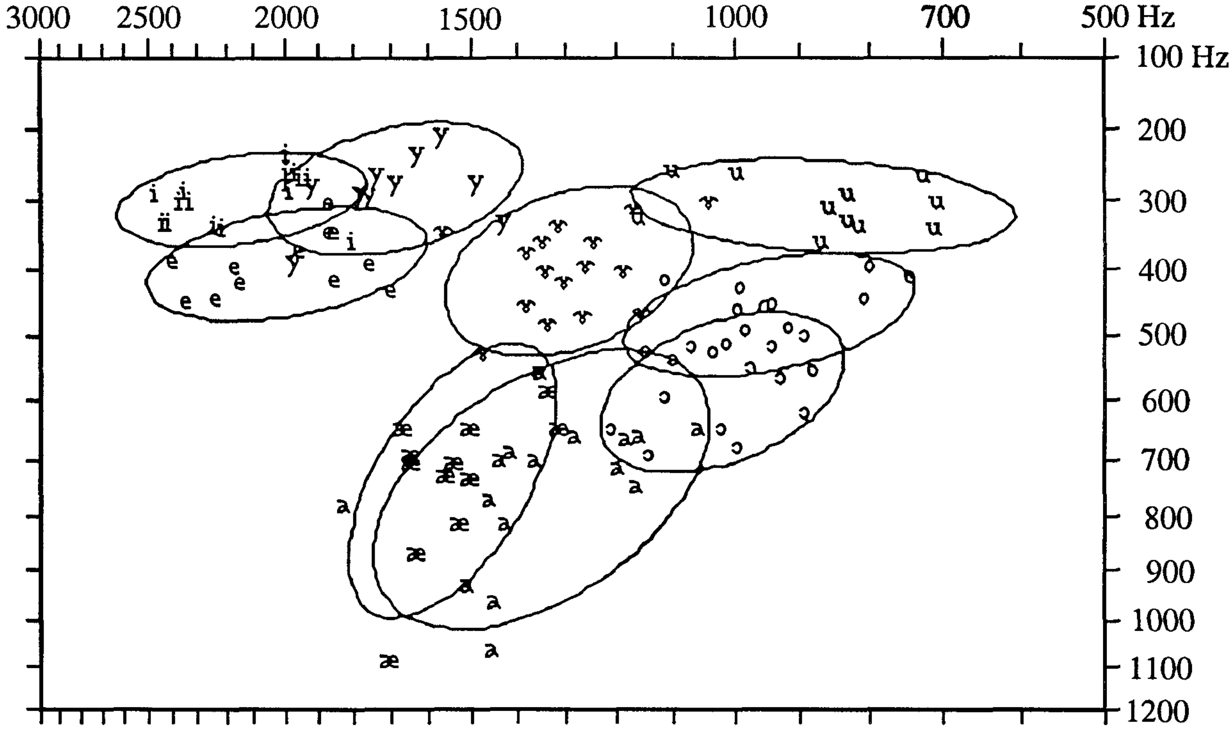


Figure 3. F1 vs F2 for male speakers, long and short vowels together.

The measurements made on the vowels of the two male speakers, shown in Figure 3, display a greater degree of overlap in their distributions. There are several factors which contribute to this effect. The two male speakers show some individual difference in their typical F2 for non-low front vowels /i, y, e/, most visible in the appearance of two clusters of tokens for /i/. One speaker has a generally higher F2 for vowels in this area than the other, but no comparable difference is observed between the speakers for vowels in other regions of the F1/F2 plane. For a different reason, /u/ shows a greater dispersion of values than for the women; this is because the short variants have more coarticulation with coronal consonants for these two speakers, resulting in perceptually fronted variants. Even allowing for these factors, it is still apparent that there is less separation, particularly in the F1 dimension, between the vowels of these male speakers. In particular /ɔ/ is less clearly a low vowel, and the distinction between /o/ and is /ɔ/ markedly reduced. However, the relative relationships between the three high vowels and the three mid vowels with respect to F2 are comparable to those seen for the female speakers.

2.2.2. *Intrinsic pitch of vowels*

Intrinsic pitch was examined for the vowels, to test whether Iaai conforms to the general tendency observed in the world's languages for vowel height to show a positive correlation with fundamental frequency, for example for high vowels to have higher F0s than low vowels when other things are equal. Vowels were grouped into three categories, high, mid and low, including long and short variants together but omitting mid front rounded /ø/ as in Figures 2 and 3 above. Mean values of F0 for different vowel heights are plotted in Figure 4 and listed in Table 2.

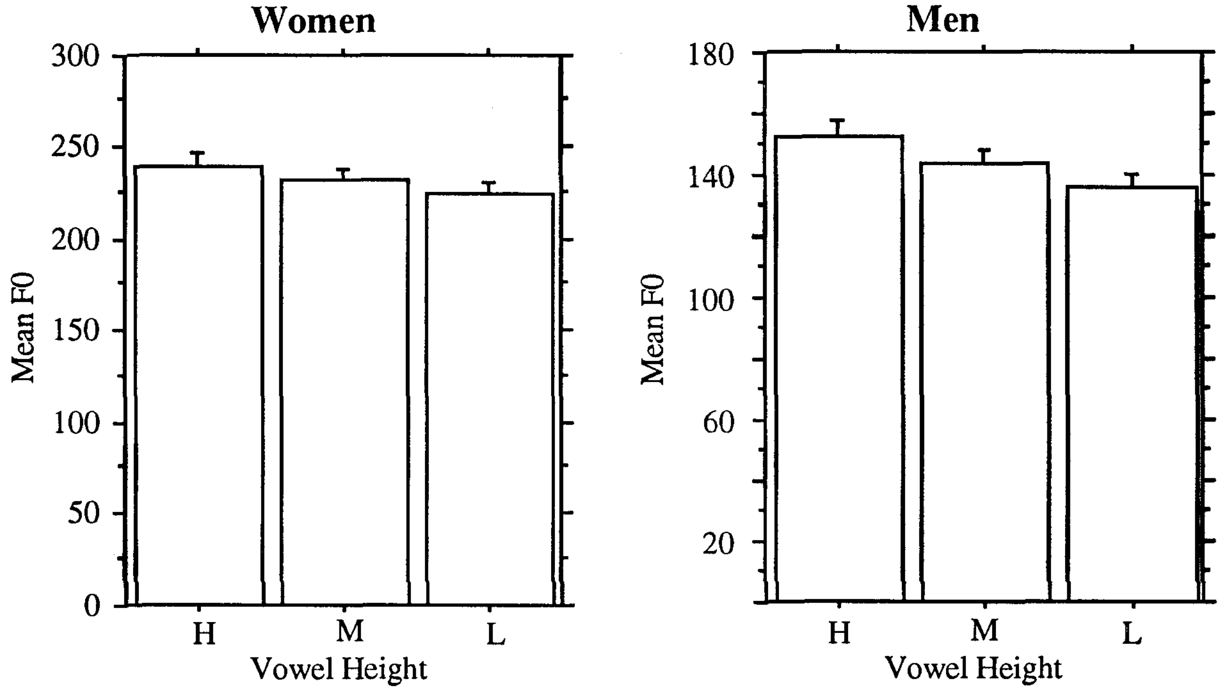


Fig 4. Intrinsic pitch of Iaai vowels according to vowel height categories for 3 female and 2 male speakers. Error bars show 95% confidence interval

Table 2. Vowel Height and F0 (values rounded to nearest integer)

	Women			Men		
	n	*F0*	*s.d.*	*n*	*F0*	*s.d.*
High	**43**	240	22	37	152	15
Mid	**41**	233	15	40	144	12
Low	**37**	224	19	37	136	13

Analysis of variance shows that there is a significant effect of vowel height category on the fundamental frequency. This effect is more apparent in the male speakers than in the females, but is significant for both; for the women F (2, 118) = 6.304, $p = .0025$, and for the men F (2, 111) = 13.84, $p < .0001$. Each pairwise comparison of adjacent heights is significant for the male speakers at better than the .05 level but only the comparison between high and low classes reaches significance for the female speakers. The magnitudes of these differences are comparable to those found in studies of relatively careful speech in other languages.

2.2.3 Vowel duration

An indication of the durational differences between long and short vowels can be obtained from the mean durations plotted in Figures 5 and 6. Since the environments in which the durations were measured include a variety of segmental and syllabic contexts which are not fully matched across the vowel categories, these data cannot provide reliable estimates of inherent durational differences. Nonetheless, it is striking that the long versions of /y/ and /ɤ/ are longer than most others; these are the two vowels in this set with values of lip-rounding opposite to those normally found cross-linguistically. The very large variance in the duration of /u/ for male speakers suggests that some words expected to have a lexical long high back vowel may be pronounced with a short vowel.

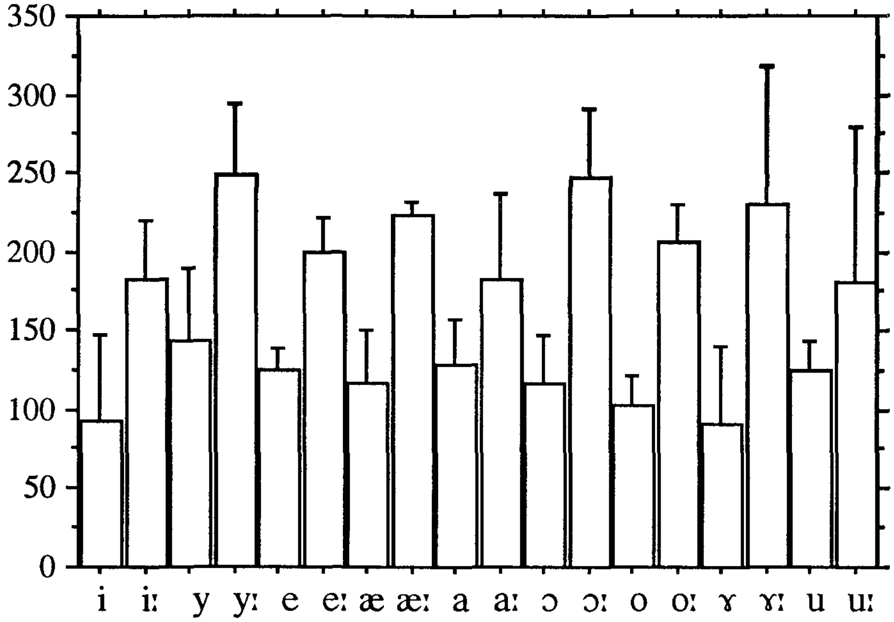


Figure 5. Mean durations in ms of nine long and short vowels for the three female speakers. Error bars show 95% confidence level.

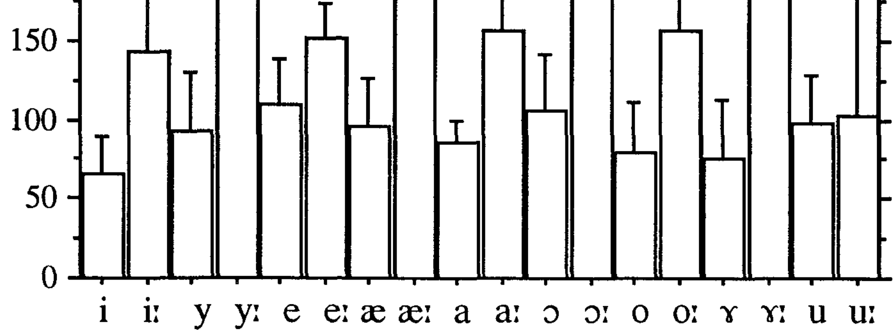


Figure 6. Mean durations in ms of nine long and short vowels for the two male speakers. Error bars show 95% confidence level.

Aggregating the durations of the long and short vowel classes does provide a good estimate of the overall difference in their length. Long vowels are about twice the duration of short. The results are shown separately for men and women in Table 3. Because of the unreliable estimate of the duration of /u/ for the males, only eight vowels are included in this comparison for the male speakers.

Table 3. Mean durations in ms of long and short vowels for female and male speakers.

	Short	Long	Difference	Ratio
Women (9 vowels)	116.1	211.1	95.0	1.85
Men (8 vowels)	89.4	175.1	85.6	2.0

Lip position

A second set of vowel formant measurements was made on words designed to illustrate the range of lip positions for the 10 vowels, which were videotaped. Data for three of the speakers have been analyzed so far. The words used and the formant values measured for these three speakers are shown in Table 4.

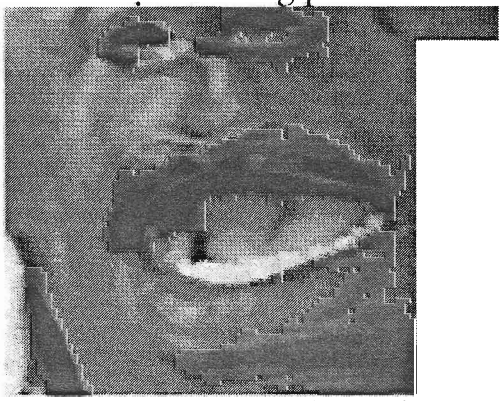
		F1			***F2***			***F3***		
Word	*Vowel*	*Spkr H*	*Spkr B*	*Spkr J*	*Spkr H*	*Spkr B*	*Spkr J*	*Spkr H*	*Spkr B*	*Spkr J*
trii	i	253	289	256	2866	2492	2563	3328	3294	3102
ûû	y	265	261	268	1799	1895	1752	2946	2395	2607
eet	e	462	444	494	2018	2466	2450	3483	3126	3001
møøk	ø	386	492	511	1382	1588	1525	2264	2361	2739
vëët	æ	741	845	743	1479	1905	2247	2358	2939	2947
aat	a	708	996	1056	1140	1556	1554	2281	2519	2883
thâân	ɔ	640	893	691	1135	1423	1106	2329	2635	2955
otr	o	458	491	489	1079	1491	958	2230	2648	2504
ötr	ɤ	468	486	504	1425	1468	1615	2176	2670	3041
trutru	u	328	281	346	829	1399	812	1347	2423	3180

The lip positions are illustrated in the panels of Figures 7 and 8 for two of the speakers studied. These figures were produced by viewing the time-coded videotape frame-by-frame and determining the frame in which the lips reached the culminating position of the gesture for the particular vowel. This frame was then digitized and a number of distances between lip points were measured in the transverse plane. These distances were a) Lip Height — distance between the lower surface of the upper lip and the upper surface of the lower lip at the center; b) Lip Width — the vertical distance between the point of contact of the lips at the left and the point of contact on the right; c) Lip Corner distance — the horizontal distance between the corners of the lip, i.e. the lateral margins of the vermilion border; d) Side Contact distance — the horizontal distance over which the lips are in contact between the corners and the aperture. This measure was simply calculated by subtracting b from c for each measurment. Also e) Lip Area — the area of the visible opening enclosed by the lips was measured. Since these data are measured from a frontal view, there is no quantitative data on lip protrusion. However, some idea of the degree of protrusion associated with the various rounded vowels can be obtained from inspecting the video frames. Some useful information on the position of the tongue can also be obtained for those vowels with a more open jaw position.

/i/ ṭii “tea” /y/ yy “dispute” /uu/ kaluu “fall down”

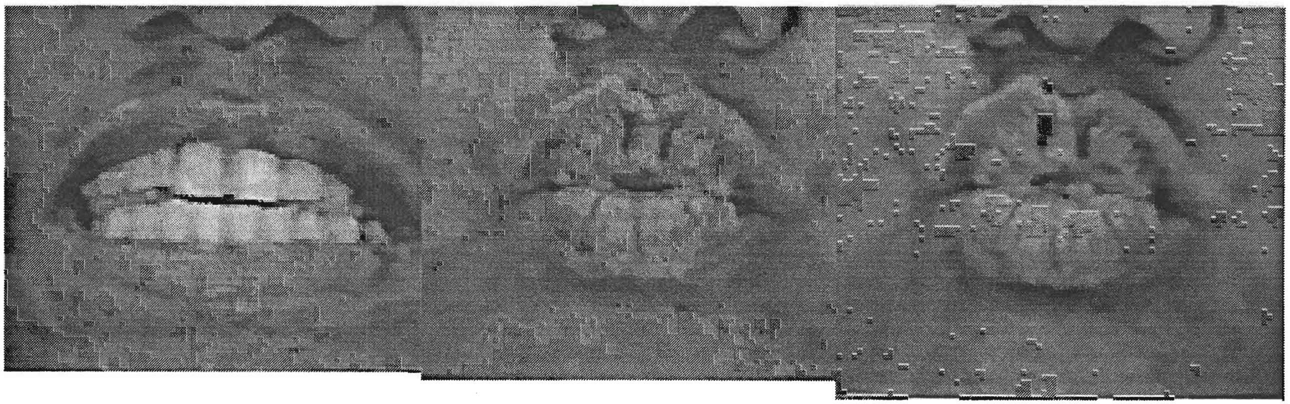
/e/ eet̪ “fishing net” /ø/ møøk “close eyes” /o/ oṭ “lobster”

/æ/ mææk “heavy” /ɤ/ ɤṭ “cooking pot” /ɔ/ t̪ɔɔn̪ “oven”

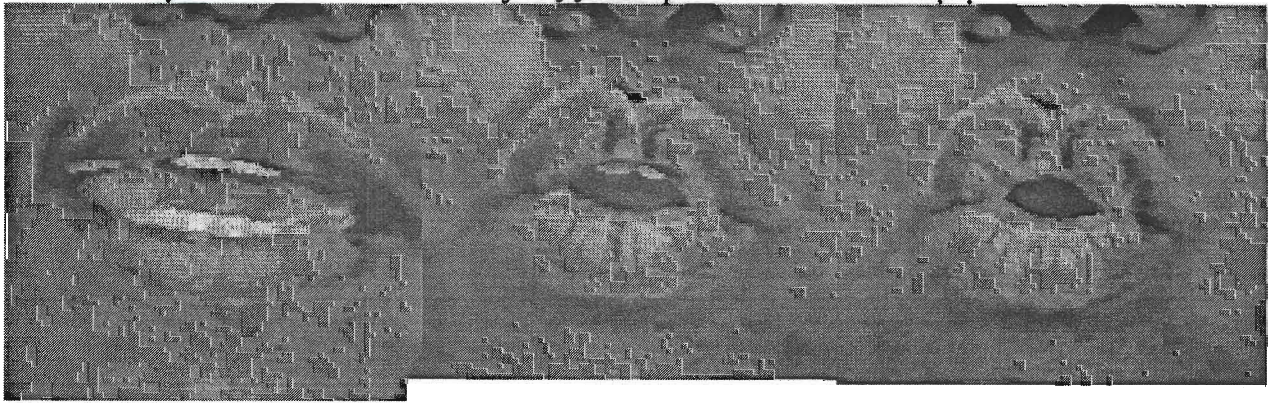


/a/ aat̪
“wounded person”

Figure 7. Lip positions for ten Iaai vowels: male speaker H.

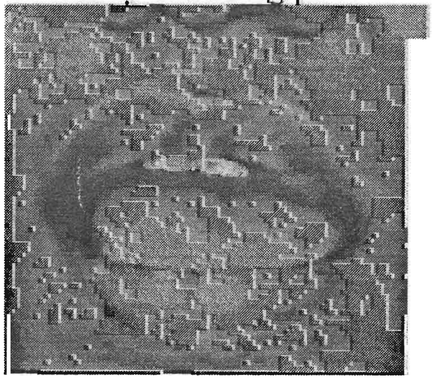


/i/ ṭii “tea” /y/ yy “dispute” /u/ ṭuṭu “fall down”



/e/ eet̪ “fishing net” /ø/ møøk “close eyes” /o/ oṭ “lobster”

/æ/ ɥææt̪ “war” /ɤ/ ɤṭ “cooking pot” /ɔ/ t̪ɔɔn̪ “oven”



/a/ aat̪
“wounded person”

Figure 8 Lip positions for ten Iaai vowels: female speaker J.

Lip measurement results for three speakers, one male and two female, are shown in Table 5. In order to normalize across the different speakers, all measurments were converted to standardized scores (with a within-speaker mean of zero and standard deviation of 1) before means were calculated and statistical tests performed. This transformation of the data, to a good approximation, sets the value of a neutral lip position to zero. The values in Table 5 are the means of the within-speaker standardized scores. From this one can see, for example, that the area of the aperture between the lips in pronouncing /ɤ/ is very close to the mean lip area, /y/ and /u/ have the smallest area, indicated by the large negative number, and /a/ has the largest area.

Table 5. Mean normalized lip measurement values for the ten Iaai vowels, from 3 speakers.

Word	*Vowel*	*Lip Height*	*Lip Width*	*Lip Corners*	*Lip Sides*	*Lip Area*
trii	i	0.53	0.70	0.89	0.31	0.69
û	y	-1.41	-1.15	-0.63	0.70	-1.14
eet	e	0.19	1.26	1.27	-0.08	0.51
møøk	ø	-0.45	-0.75	-1.10	0.12	-0.78
vëët	æ	1.22	0.87	0.79	-0.62	1.12
aat	a	1.38	1.03	0.45	-1.18	1.51
thâân	ɔ	0.61	-0.05	-0.61	-0.71	0.14
otr	o	-0.75	-1.01	-0.84	0.55	-0.92
ötr	ɤ	-0.15	0.40	0.70	0.12	-0.01
trutru or kaluu	u	-1.17	-1.29	-0.92	0.79	-1.12

Another view of the lip measurement data is shown in Table 6, which ranks the vowels in terms of each of the measures, with the vowel having the lowest value appearing at the top of the column and the vowel with the lowest measure at the bottom.

Table 6. Rank order of the vowels on each of the normalized lip measurement. Vowels are ordered from smallest to greatest value.

Rank	*Lip Height*	*Lip Width*	*Lip Corners*	*Lip Sides*	*Lip Area*
1	y	u	ø	a	y
2	u	y	u	ɔ	u
3	o	o	o	æ	o
4	ø	ø	y	e	ø
5	ɤ	ɔ	ɔ	ø	ɤ
6	e	ɤ	a	ɤ	ɔ
7	i	i	ɤ	i	e
8	ɔ	æ	æ	o	i
9	æ	a	i	y	æ
10	a	e	e	u	a

Examining Tables 5 and 6 we see the following generalizations. The vertical distance between the lips (Lip Height) is least for rounded non-low vowels, and greatest for unrounded low vowels. Although a three-way classification of vowels by height (high, mid, low) predicts a significant amount of the variance in the lip height, the four mid vowels show quite substantial differences with rounded mid vowels having higher lip position than unrounded ones. Most strikingly, the unrounded high vowel /i/ is more open than /e/ at the lips. The difference between /i/ and /e/ is made by raising the jaw independently of the lips, as can be seen from the position of the teeth in the video frames. The overall pattern is broadly similar for the horizontal measure of the lip opening (Lip Width), with the vowels receiving similar rank orders, except that /e/ is wider than /i/, and indeed any other vowel, and /ɔ/ has substantially less width than the other low vowels. As Table 7 shows, the correlation coefficient between Lip Height and Width is .82, but whereas Lip Height does not divide rounded from unrounded vowels Lip Width does. The distance between the outside corners of the lips (Lip Corners), however, is the best of our measures at effecting such a separation. All five rounded vowels have negative values of at least -0.61, and all the unrounded vowels have positive values of 0.45 or greater. Moreover, the Lip Corner distance relates well only to the classification of vowels by rounding and not to classification by height as well. It is likely that this measure is the best index of lip protrusion when measurements are only taken in a flat plane, transverse to the body. Drawing the corners of the lips closer together is a consequence of protruding them.

The measurement of contact at the sides of the lips (Lip Sides) was added following the suggestion of Goldstein (1991:98) that "rounded vowels must be produced with contact along the sides". Although rounded high vowels have the greatest amount of side contact this measure does not serve to separate rounded and unrounded vowels in general. This measure does not separate the Iaai vowels into rounded and unrounded classes, and little of the variance in the Lip Sides measure can be predicted from the classification of vowels by rounding, F (1, 29) = 2.87, p = .1011) (nor by height or backness either).

Naturally enough, both Lip Height and Width measurements are very highly correlated with Lip Area (.95 and .92 respectively). Since /y/ and /u/ have the smallest height and width, they have the smallest area of lip opening, having almost identical mean normalized values. In this respect Iaai differs from a number of other languages with a similar pair of vowels, such as French, Swedish, Cantonese and Finnish, where the lip area for /y/ is considerably larger than that for /u/, and is comparable to that for /i/ (Linker 1982). Iaai also has a larger than expected area for /i/. The lip area measure broadly separates rounded from unrounded vowels with vowel height ranking vowels within those groups. The salient exception occurs with /ɔ/ and /ɤ/. These two vowels have similar area, but only because they have different degrees of Lip Height and Width that balance each other out. Not unexpectedly, none of the lip measures predict the backness of vowels.

Table 7. Correlation matrix between the normalized lip measures and also the within-speaker normalized formant values of the first three formants.

	L Area	*L Height*	*L Width*	*L Corners*	*L Sides*	*F1*	*F2*
L Height	.95						
L Width	.92	.82					
L Corners	.68	.51	.84				
L Sides	-.5	-.58	-.44	-.1			
F1	.64	.74	.5	.19	-.59		
F2	.39	.26	.54	.62	.03	-.27	
F3	.43	.39	.53	.59	-.15	-.11	.64

The correlation analysis shown in Table 7 indicates that the normalized F1 correlates most highly with Lip Height. Standard acoustic theory predicts a relationship between vowel height and F1; the more open a vowel is the higher the F1 frequency. F2 and F3 both correlate most highly with the Lip Corner measure. This measure is associated with rounding and is hypothesized to be related to lip protrusion; low values indicate protruded lips. Since increasing the effective vocal tract by protruding the lips lowers the frequency of these higher formants, the correlation is presumably attributable to this component of their variation. Since these formants are also very sensitive to the location of constrictions inside the oral cavity, the strength of the correlations with this lip measure are quite striking.

A few other observations can be made on the vowel articulations visible in Figures 7 and 8 concerning tongue and jaw positions. A very front tongue position, with the tongue tip appearing pressed against the lower teeth, is visible in both /e/ and in /æ/. The tongue is also quite visible in /a/, indicating that this vowel is central rather than back. It is not possible to resolve whether the tongue is central or back in /ɤ/ from these pictures, but the probability seems greater that it is back. The position of the teeth in /ɤ/ show that the jaw tends to be higher in this vowel than in the other mid vowels.

3. Consonants

Table 8 shows the consonants that we provisionally accept for Iaai. The set of consonants is the same as recognized by Ozanne-Rivierre, although we categorize a few of them in slightly diffent terms. In particular, the segments that we transcribe as /ɥ ɥ/ are here categorized as approximants rather than as fricatives. As Ozanne-Riviere noted, these segments are indeed somewhat intermediate between the these two traditional categories and might be described as either. Some frication is sometimes audible but it is weak in amplitude (see Figures 10 and 12 below). We however also choose to emphasize that there seem to be two articulatory strictures involved in the production of this pair of consonants. In addition to the labial constriction - the location at which some frication can often be observed, there is also a raising of the front of the tongue. Hence they are transcribed with symbols that represent labial-palatal approximants; another way to represent this aspect of their production in the transcription would be to use the complex symbols /ɸʲ, βʲ/. This palatal articulation is surely connected to the distributional constraints affecting these labial consonants which were discussed in the presentation of the vowel system. It is possible that all the "plain" labials in front or fronted vowel environments should be considered to be palatalized, and opposed to labials with audible rounding or in back vowel environments by a contrast of secondary articulations. We do not feel that we have enough evidence at the present time to address either the phonetic or phonological questions raised by the consonant/vowel interdependencies in Iaai, but will provide a brief comment below.

Table 8. Symbols for consonants in Iaai.

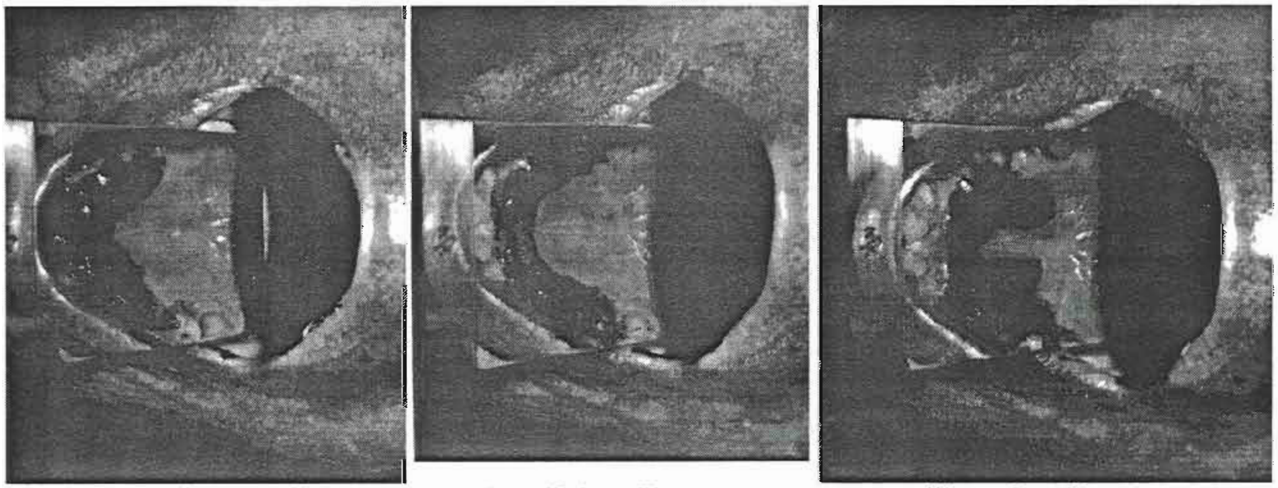
	Labial	Labialized Labial	Dental	Alveolar	Retroflex	Pre-palatal	Velar
Plosive	p b	bʷ	t̪ d̪		ṭ ḍ	c ɟ	k g
Nasal	m̥ m	m̥ʷ mʷ	n̪̥ n̪		ṇ̥ ṇ	ɲ̥ ɲ	ŋ̥ ŋ
Fricative	f		θ ð	s		ʃ	x
Central Approx.	ɥ̥ ɥ	w̥ w					h
Lateral Approx.				l̥ l			
Flap					ɾ̣		

Note: /ɾ̣/ appears only intervocalically

Unlike the majority of New Caledonian languages, the voiced stops of Iaai are not accompanied by any prenasalization. Two other marked characteristics of the consonant system are the contrast of several places of articulation within the coronal region, and the occurrence of a large variety of voiceless sonorants. We will focus our discussion of the Iaai consonants on these areas.

3.1 Coronal stops

Palatograms of the three coronal series of stops were made of four speakers, using the technique described in Ladefoged (1994). Linguagrams were also obtained from one male speaker. Illustrative palatograms of the difference in coronal places of articulation for this speaker are shown in Figure 9. The figure shows views in a mirror placed in the mouth to show the upper surface; blackened areas indicate where the tongue made contact during the articulation. The dental in /at̪ / has a relatively large contact area entirely covering the upper front teeth and the alveolar ridge. The linguagram confirms that the contact is laminal, or more precisely apico-laminal (Dart 1991), and includes a considerable extension of the contact laterally back toward the molar teeth on both sides. The 'retroflex' stop in /ḍa/ involves a much narrower band of contact which falls entirely behind the teeth toward the back part of the alveolar region. The linguagram shows this contact is strictly apical, with only the narrow anterior-facing surface of the tongue tip and a small area on the upper surface of the tip involved. There is markedly less lateral contact behind the front closure than for the dental, indicating that the center of the tongue is lower in this articulation. For the pre-palatal stop in /ca/ there is a broad contact area from the back of the alveolar ridge to a point about at the location of the second molars. The linguagram shows that the contact is strictly laminal, with no contact on the frontmost part of the tongue (about the first 1 cm).



at̪ *at* "person" ḍa *dra* "blood" ca *ca* "leg, foot"

Figure 9. Palatograms illustrating the three coronal stop places of Iaai.

The palatograms obtained from the other speakers, who are of a slightly younger generation, tended to show a less clear articulatory distinction in the location of the contacts on the palate and in the tongue contact area between these coronal places of articulation. However, all speakers observed maintain a three-way acoustic distinction. The dental place is characterized by a lack of frication of the release, and if voiceless and prevocalic, by a very short voice onset time (VOT). Both the palatal and retroflex places are characterized by a noisy and sustained release. The noise of the release seems to be concentrated in the area of the third formant resonance for the palatal and the fourth formant for the retroflex. These characteristics are illustrated in the three waveforms and spectrograms in Figure 10, from the same male speaker who provided the palatograms in Figure 9..

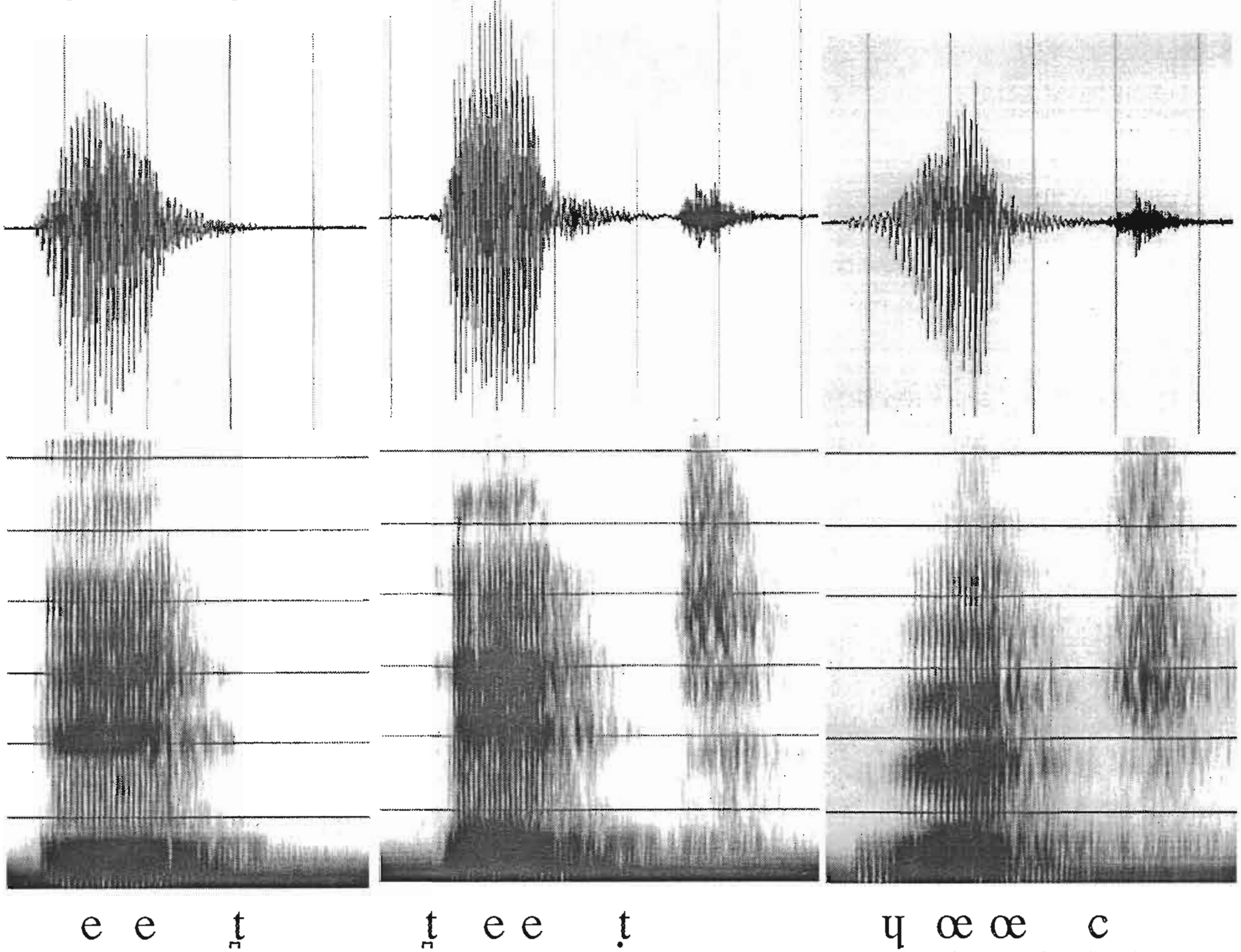


Figure 10. Waveforms and spectrograms of final voiceless dental, retroflex and palatal stops in /eet̪/ “fishing net”, /t̪eeṭ/ “excrement” and /ɥœœc/ “light (a fire)”, spoken by a male speaker. Calibration lines are at 100 ms intervals on the waveforms and at 1000 Hz intervals in the spectrograms.

The very weak release of the dental can be seen both in the final consonant of /eet̪/ “fishing net”and in the initial of /t̪eeṭ/ “excrement”. It contrasts very obviously with the energetic noisy release of the other two voiceless coronal stops in final position. The duration of this noise portion tends to be longer in the palatal than in the retroflex. In two words with matched vowel environments, the five speakers had a mean noise duration of 157 ms in /l̥it̪/ “(be) black” versus 204 ms in /ic/ “drink”. In onset position a mean VOT of only 13 ms was measured in /taŋ/ “basket” versus 75 ms in /caan/ “foot (3sg possessed)”. Additional examples of final /t̪/ and /ṭ/ are shown in Figure 12. In a number of languages, laminal dental stops are accompanied by a long noisy release, whereas apical alveolar or retroflex stops have short release. This appears to be typical for Australian languages (see Anderson and Maddieson 1994, this volume) and for Dravidian languages (e.g. Dart 1991), so that it has been proposed as a universal pattern (Stevens, Keyser and Kawasaki 1986). However, an opposite distribution has been observed in several African languages (e.g. Maddieson, Spajić, Sands and Ladefoged 1993) in addition to Iaai.

3.2 Voiced and voiceless sonorants

Nasals and lateral and central approximants in Iaai demonstrate a voicing contrast. A minimal pair illustrating the contrast between voiced and voiceless lateral approximants in the words /lɤŋ/ "night" and /l̥ɤŋ/ "(be) black" is shown in Figure 11. Note that the major resonance above 3000 Hz in the voiced lateral is matched by a corresponding resonance in the voiceless excitation at the onset of /l̥ɤŋ/. This voiceless portion is followed by about 80 ms of voiced lateral before the vowel. Although quantitative studies are incomplete on this point, the strong impression is that a relatively long transitional voiced portion is typical for the voiceless approximants in Iaai. This may be a residue of the origin of these devoiced segments from the coalescence of a sequence of /h/ and a sonorant. Although the voiceless continuants appear in many non-derived forms, a significant number of pairs of words with voiced and voiceless sonorant initials are derivationally related to each other. As Ozanne-Rivierre points out, there is also a parallel alternation between words with initial vowels and initial /h/. Some examples of these two types are listed in Table 9.

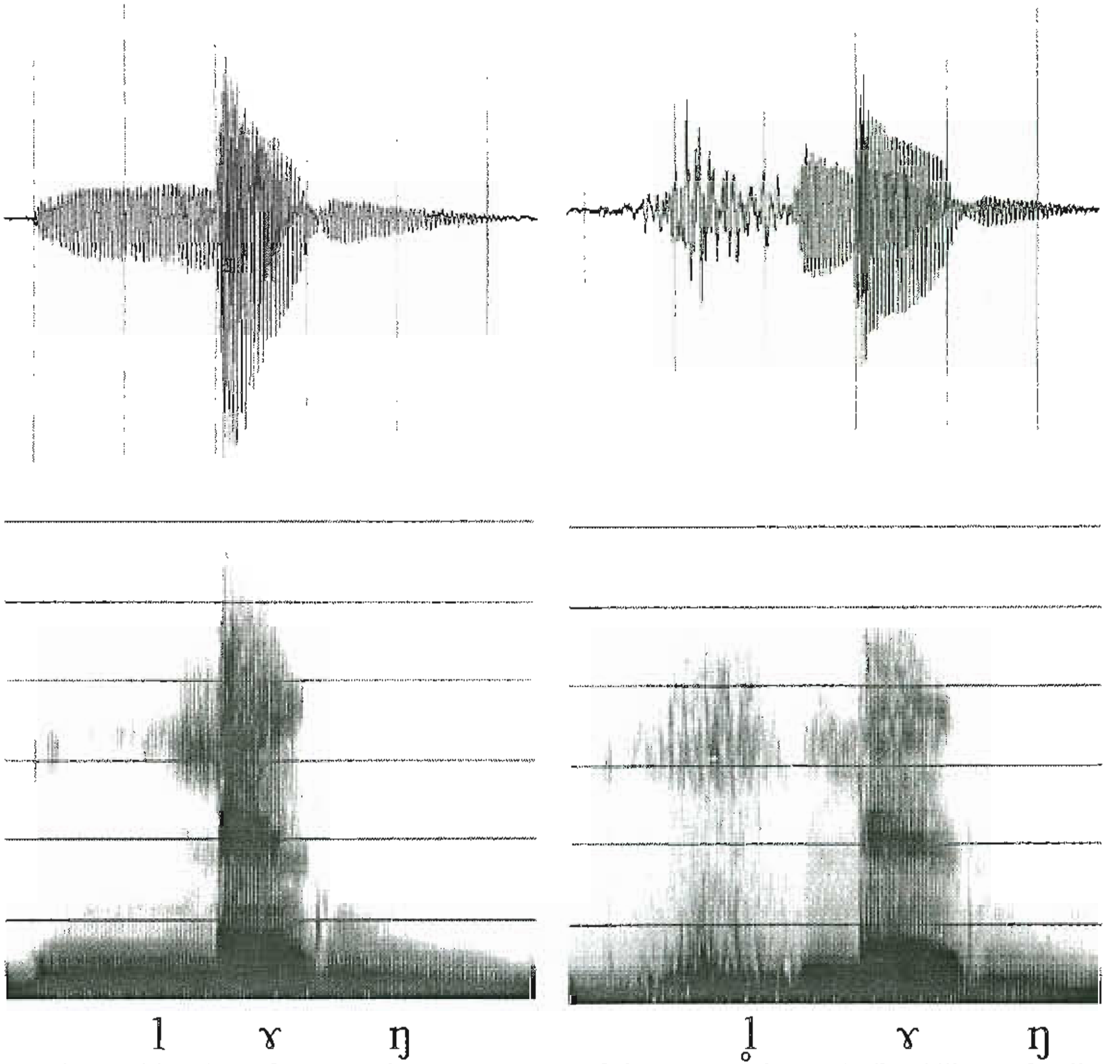


Figure 11. Waveforms and spectrograms of the contrasting pair /lɤŋ/ "hear, feel" and /l̥ɤŋ/ "close, stop up" spoken by a female speaker.

Table 9. Some examples of alternation of voiced and voiceless sonorants and initial vowel and /h/ from Ozanne-Rivierre (1976, 1984).

weɲe	"name (vb, determ.)"	w̥eɲii	"name (vb, indet.)
wiɤɤ	"turn, change" (+ obj)	w̥iɤɤ	"turn, change" (incorp obj)
naŋ	"brandish"	n̥ɤŋ	"brandish" (incorp obj)
litr	"night"	l̥iṭ	"(be) black"
yca	"choose, select"	hyce	"choose, select" (incorp obj)
an	"eat"	han	"eat" (incorp obj)

3.3 Labial consonants

Samples of some of the labial consonants are shown in Figure 12 (see also /ɥœœc/ "light (a fire)" in Figure 10. Note that there are quite extensive and slow second formant transitions during the vowels in the first two words, falling in /møk/ and rising in /mwɤɤṭ/. Since in a labial nasal the tongue is free to anticipate the articulatory position of the following vowel and the releasing movement of the lips is rapid, these transitions are unlikely simply to be due to the effect of the labial articulation. In the case of /mwɤɤṭ/ a secondary articulation of labialization (actually labio-velarization) is recognized in the transcription, involving an /u/-like accompaniment to the nasal and producing a quasi-diphthongal vowel portion in the word. In /møk/ the onset F2 is considerably higher than in /mwɤɤṭ/ and continues a resonance that is visible in the nasal at about 1500 Hz. The comparable resonance in /mwɤɤṭ/ is only at about 1000 Hz. /ɥææt/ shows a similar onset F2 frequency. The fall in F2 over the whole of the vowel in /møk/ cannot be satisfactorily attributed to the effect of the final velar. Although velars in rounded vowel environments often have a low F2 close to that of the vowel, there is no reason for the F2 to have been so much higher at the onset of the vowel unless the labial nasal has some property which is responsible for it. Both the timing and the frequency aspects of this pattern can be explained by the assumption that there is a secondary palatal articulation accompanying the onset nasal. Iaai might thus have a system similat to that found in some Micronesian languages, where there are no truly "plain" labials, but every labial consonant has a secondary articulation attached to it. The illusion of occurrence of plain labials arises when the secondary articulation matches closely with the articulatory and acoustic attributes of the adjoining vowel.

4. Concluding remarks

This paper has provided some observations on the phonetics of the vowels and consonants of the Iaai language which make clear that this language has some properties of more than ordinary interest. We also feel that this examination has provided some ideas for further analysis of some of the phonological characteristics of the Iaai system. In both of these areas we are very aware that there is much more to be done.

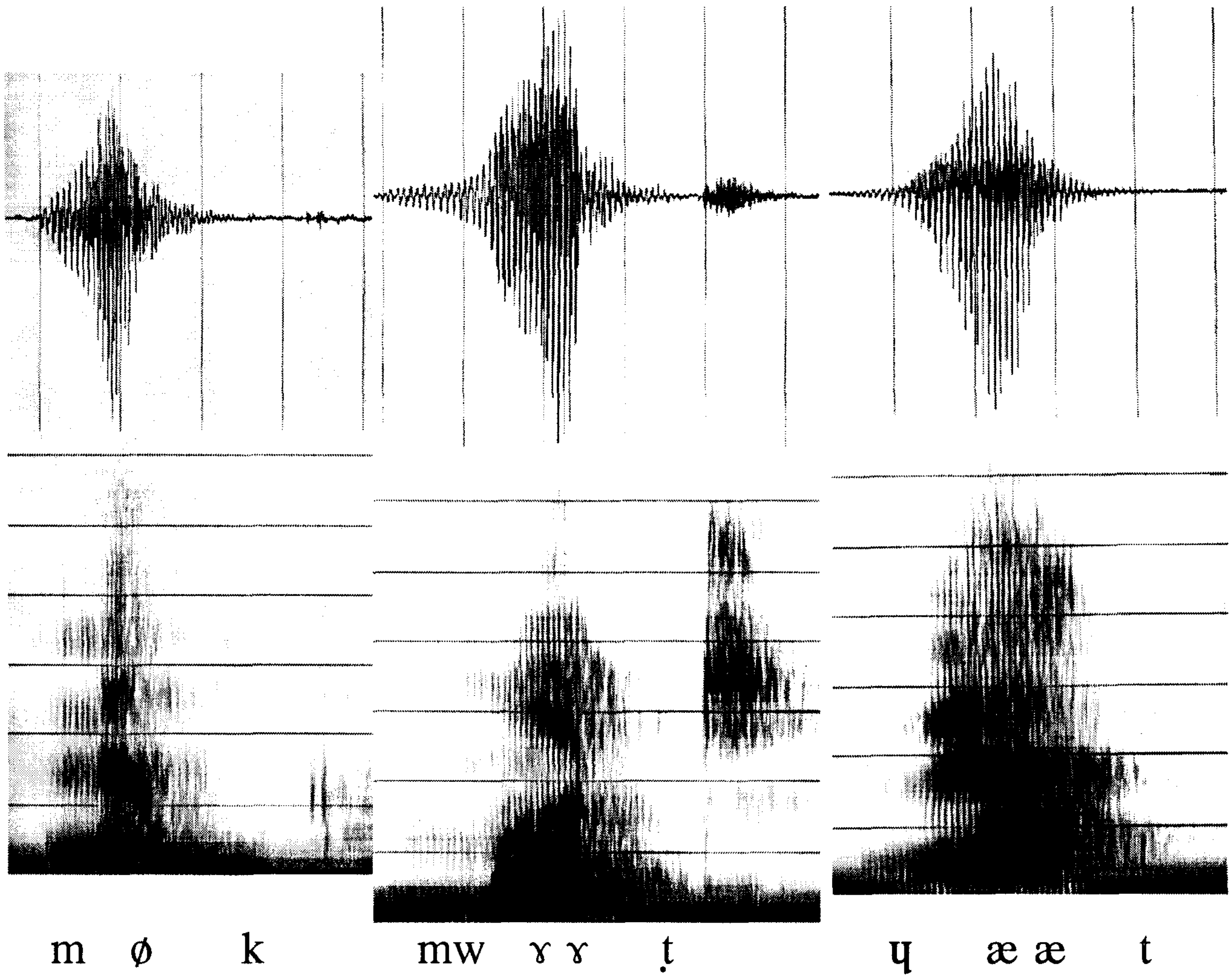


Figure 12. Waveforms and sp[ectrograms illustrating labial consonants in /møk/ "dead, ill", /mwɤɤṭ/ "live" and /ɥææt̪/ "light (a fire)", spoken by a male speaker.

References

Anderson, V. B. and Maddieson, I. 1994. Acoustic characteristics of Tiwi coronal stops. *UCLA Working Papers in Phonetics* 87: 131-162.

Dart, Sarah. 1991. *Articulatory and Acoustic Properties of Apical and Laminal Articulations* (= UCLA Working Papers in Phonetics 79).

Goldstein, Louis. 1991. Lip rounding as side contact. *Proceedings of the 12th International Congress of Phonetic Sciences, Aix-en-Provence, Vol 1*: 97-101.

Linker, Wendy. 1982. *Articulatory and Acoustic Correlates of Labial Activity inVowels: A Cross-Linguistic Study.* (= UCLA Working Papers in Phonetics 56.)

Maddieson, I., Spajic, S., Sands, B. and Ladefoged, P. 1993. Phonetic Structures of Dahalo. *Afrikanistische Arbeitspapiere* 36: 3-53.

Ozanne-Rivierre, Françoise. 1976. *Le Iaai: Langue mélanésienne d'Ouvéa: Phonologie, Morphokgie, Esquisse Syntaxique.* SELAF, Paris.

Ozanne-Rivierre, Françoise. 1984. *Dictionnaire Iaai-Français.* SELAF, Paris.

Stevens, K., Keyser, S.J., and Kawasaki, H. 1986. Toward a phonetic and phonological theory of redundant features. In *Invariance and Variability in Speech Processes,* J.S. Perkell and D.H. Klatt, eds., pp. 426-463. New Jersey: Lawrence Erlbaum Associates.

Wahéo, T. and Wahéo, J. 1987. *Langue Iaai Ouvéa – Propositions d'Écriture (Hua Setr Hwen Iaai).* Bureau des Langues Vernaculaires, Centre Territorial de Recherche et de Documentation Pédagogiques de Nouvelle Calédonie, Nouméa.

Acknowledments

The prime debt is owed to the five speakers of Iaai who provided the data on which this study is based, most especially to Tai Wahéo who was an enthusiastic participant in all stages. Stephan Schooling went out of his way to provide introductions and to make transportation and all other practical matters as smooth as possible, as well as being a stimulating and knowledgeable partner for linguistic discussions. The staff of the ADCK and CTRDP in Nouméa were also very helpful. Zareh Movsessian and Mona Patel assisted with some of the analysis of the videotaped sessions. The first author would also like to thank Françoise Ozanne - Rivierre and Jean-Claude Rivierre for encouraging his first tentative steps in the study of New Caledonian languages.