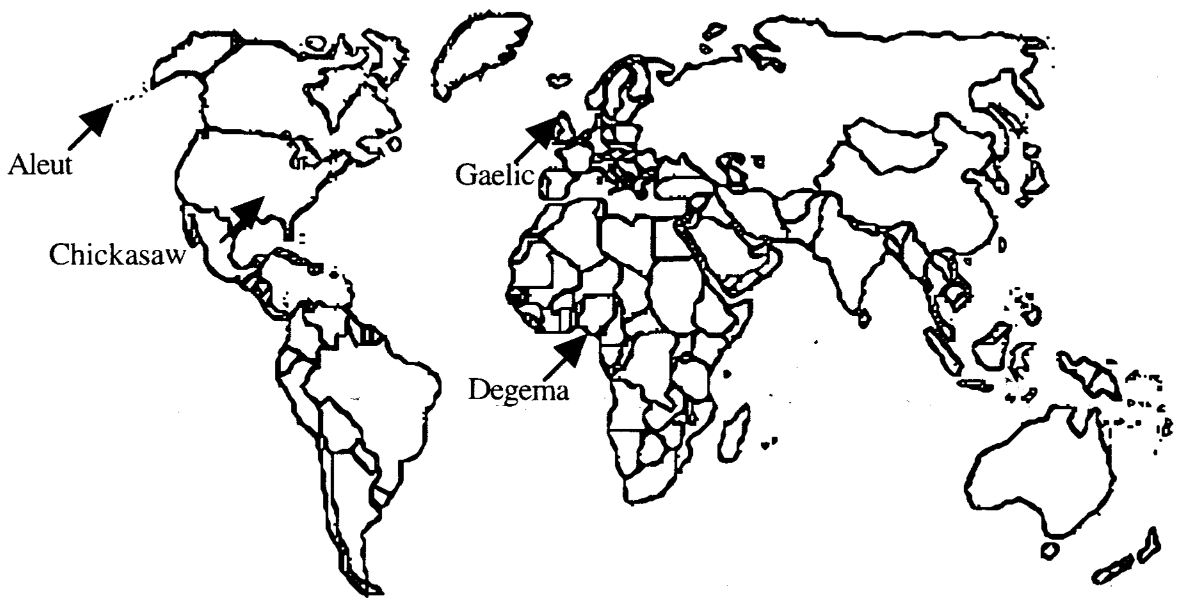
Title
WPP, No. 95: Fieldwork Studies of Targeted Languages V

Publication Date
1997-12-01

Fieldwork Studies of Targeted Languages V



UCLA WPP Volume 95

December 1997

The UCLA Phonetics Laboratory Group

Soo-Woong Ahn
Adam Albright
Victoria Anderson
Marco Baroni
Barbara Blankenship
Taehong Cho
Sandy Disner
Gorka Elordieta
Christina Foreman
Susan Meyers Fosnot
Cécile Fougeron
Sean Fulop
Matt Gordon
Bruce Hayes
Chai-Shune Hsu
Sun-Ah Jun
Pat Keating
Robert Kirchner
Jenny Ladefoged
Peter Ladefoged
Hyuck-Koon Lee
Ian Maddieson
Elizabeth Menard
Aaron Shryock
Donca Steriade
Sumiko Takayanagi
Henry Tehrani
Kimberly Thomas
Moto Ueyama
Jie Zhang

As on previous occasions, the material which is presented in this volume is simply a record for our own use, a report as required by the funding agencies which support the Phonetics Laboratory, and a preliminary account of research in progress for our colleagues in the field.

Funds for the UCLA Phonetics Laboratory are provided through:
USPHS grant 5 T32 DC00029
NSF grant SBR 9319705
NSF grant SBR 9511118
and the University of California

Correspondence concerning UCLA Working Papers in Phonetics should be addressed to:

Phonetics Laboratory
Department of Linguistics
UCLA
Los Angeles, CA 90095-1543

The General Editor of UCLA Working Papers in Phonetics is Ian Maddieson.

UCLA Working Papers in Phonetics

Number 95

December 1997

Fieldwork Studies of Targeted Languages V

Table of Contents

Further analysis of Degema vowel harmony — 1
Sean Fulop, Ethelbert Kari and Peter Ladefoged

Variations and universals in VOT:
evidence from 17 endangered languages — 18
Taehong Cho and Peter Ladefoged

The phonetic structures of Chickasaw — 41
Matthew Gordon, Pamela Munro
and Peter Ladefoged

Some phonetic structures of Aleut — 68
Taehong Cho, Alice Taff, Moses Dirks
and Peter Ladefoged

The effect of stress on vowel length in Aleut — 91
Lorna Rozelle

Intonation of declaratives and questions
in Unangan (Eastern Aleut) — 102
Alice Taff and Jacob Wegelin

Phonetic structures of Scottish Gaelic — 114
Peter Ladefoged, Jenny Ladefoged,
Alice Turk, Kevin Hind and St. John Skilton

Further analysis of Degema vowel harmony

Sean A. Fulop Ethelbert Kari Peter Ladefoged

Abstract

This paper is an elaboration of the results section of (Fulop 1996). The background material in that paper is omitted here. Degema is an Edoid language of Nigeria whose ten vowels are organized phonologically into two sets of five. The two sets are thought to be differentiated by the degree of tongue-root advancement. This paper examines the acoustic nature of these vowels as represented in field recordings of six speakers. The most consistent acoustic correlate of the tongue root contrast was found to be the first formant frequency. Additionally, a comparison of corresponding advanced and retracted vowels using a normalized measure of relative formant intensity demonstrated that this correlate could also distinguish them in 43% of the cases. Finally, a perceptual study was conducted which demonstrates that Degema speakers do not classify their vowels very well using formant frequencies as the sole acoustic variable.

1 Introduction

Nearly all the languages of the world use degrees of both tongue height and front-back position of the tongue body as contrasting articulatory features of their vowel sounds. These, together with lip rounding, are sometimes referred to as the *major* vowel features (e.g. Ladefoged and Maddieson 1996). Many languages possess vowel inventories which make essential use of features other than these (a common example being nasalization). One of the most perplexing of the *minor* vowel features is Advanced Tongue Root (ATR). This feature has generated considerable interest since its use in a great many African languages was first elucidated in the 1960s. Its many enigmatic properties include the fact that it is largely restricted to the Niger-Congo and Nilo-Saharan language families of Africa, and that it is invariably employed in a phonological process of vowel harmony.

[0]Thanks are due to Bruce Hayes, Russ Schuh and Ian Maddieson for important comments during the development of this work. Special thanks to Kay Williamson for all her help in Nigeria. Thanks also to Pat Keating for important input, Matt Gordon, Aaron Shryock, Richard Wright, Dan Silverman, and all those in the Phonetics Lab who were willing to chat and give advice. This work was supported in part by NSF grant SBR 9319705 to P. Ladefoged and I. Maddieson.

This paper will discuss the feature ATR, presenting the results of an acoustic investigation of the Niger-Congo language Degema in support of several points. The results provide evidence for three acoustic correlates of the ATR contrast, namely the frequencies of the first and second formants, and the intensities of the first and second formants relative to each other. We will discuss the correspondence of the third correlate to the theoretically predicted damping of F_1 due to viscous air flow in a narrow pharyngeal constriction.

2 The feature ATR

[This section of the paper is omitted in this version. See (Fulop 1996).]

3 Acoustics of ATR in Degema

Degema is a Niger-Congo language, more specifically classified as Edoid. The Degema are one of three groups of Edoid-speaking people in Nigeria's Rivers State; the three languages spoken in this small region comprise the Delta-Edoid subgroup. The two dialects of Degema are Usokun and Atala. These are not substantially disparate, showing only small differences in lexical tones and in the phonetic realization of one consonant (Elugbe 1989). Degema exhibits a 10 vowel harmony system involving the two mutually exclusive sets discussed above in more general terms:

[+ATR]	i, e, ə, o, u
[−ATR]	ɪ, ɛ, a, ɔ, ʊ

In a given word, all the vowels will be of one set or the other.

3.1 Data and methods

The speech of six male Degema speakers was recorded in the field on a portable DAT recorder. The utterances are one and two-word sentences. Two of the speakers provided eight tokens of each vowel, while the remaining speakers provided four tokens of each vowel. The vowels occurred in minimal sets (verb paradigms) between [m] and an alveolar stop (there are the same number with following [t] as with following [d]).

To analyze the data, FFT spectrograms were computed using a 150 Hz bandwidth filter and a Hamming window. LPC formant histories were superposed on these, computed with a filter order of 12 over a 10 ms window. This information was used to estimate the first three formant frequencies.

Additionally, the amplitudes of the formants were estimated using a narrow-band short-time FFT. The formant amplitudes in dB were measured as the energy in the most prominent harmonic clearly within each formant band. This technique is by nature approximate; the

Table 1: F_1 difference: Significance and effect sizes

	Vowel quality	a	e	i	o	u
Speaker 1	FP significance	**0**	**0**	**0.0012**	**0**	**0**
	Effect size	**0.89**	**1.0**	**0.85**	**1.0**	**1.0**
Speaker 2	FP significance	0.49	**0**	0.90	**0**	**0**
	Effect size	0.66	**1.0**	0.47	**1.0**	**1.0**
Speaker 3	FP significance	0.46	**0**	**0.0013**	**0**	**0**
	Effect size	0.62	**1.0**	**0.84**	**1.0**	**1.0**
Speaker 4	FP significance	0.23	0.23	0.37	0	0.50
	Effect size	0.75	0.78	0.72	0	0.66
Speaker 5	FP significance	0.30	**0**	**0.012**	**0**	**0**
	Effect size	0.28	**1.0**	**0.88**	**0.97**	**1.0**
Speaker 6	FP significance	**0**	**0**	**0**	**0**	**0**
	Effect size	**1.0**	**1.0**	**0.94**	**1.0**	**1.0**

"true" amplitude can only be discovered this way if a harmonic occurs at the center frequency of the resonance. Otherwise, the harmonic amplitudes will be slightly less.

3.2 Results and analysis

3.2.1 Formant frequencies

Figures 1 and 2 show the vowels for each speaker plotted in an acoustic vowel space. The ellipses are computed to enclose two standard deviations of the token dispersion along the major and minor axes. Dark ellipses are used for [+ATR] vowels, and light grey for [−ATR] vowels. A few vowel groups were found by boxplot analysis to have probable outliers; the offending tokens have been removed from the plots to give tighter ellipses that are more accurately located.

As expected, the difference between the [+ATR] and [−ATR] vowels is often evident in the value of F_1; the retracted tongue root vowels frequently show a distinctly higher F_1 than their advanced tongue root counterparts.

Table 1 gives the significance of the difference between the first formant frequency of the [+ATR] as against the [−ATR] vowels; Table 2 is parallel but deals with the second formant frequencies. The significance (commonly called p) values are computed by two-sided Fligner-Policello tests on each pair of vowels. We chose to perform the Fligner-Policello test because statistical research has shown there to be differences in performance which, under certain conditions, can save the experimenter from costly errors of both type I and type II (Wilcox 1996). A further discussion of these statistical issues is included as an appendix.

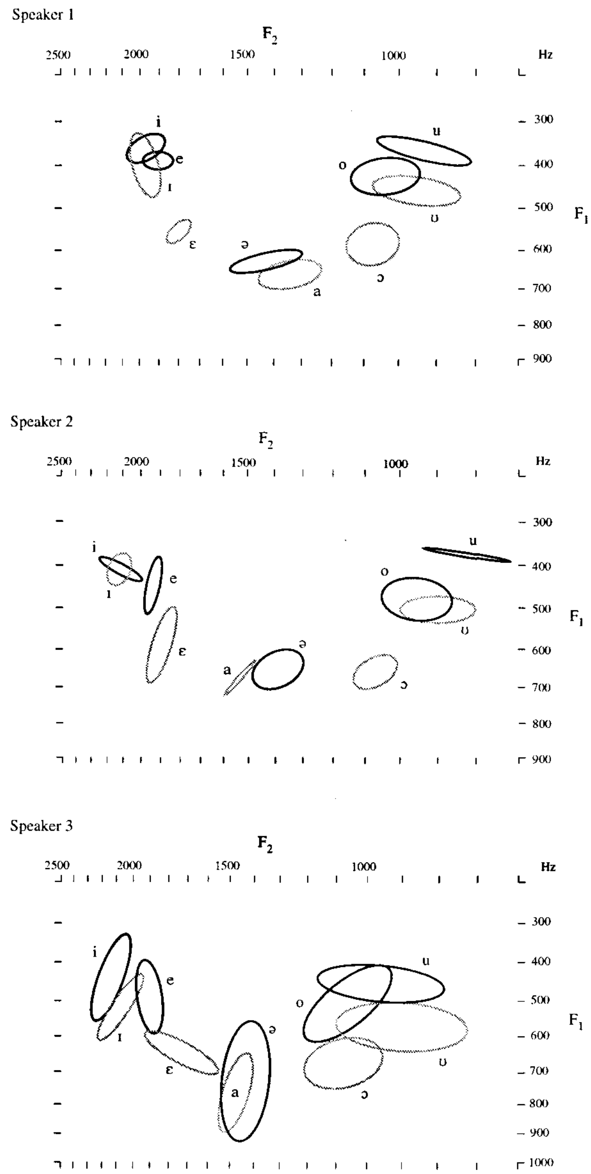


Figure 1: Conventional F_1–F_2 vowel formant plots for speakers 1–3

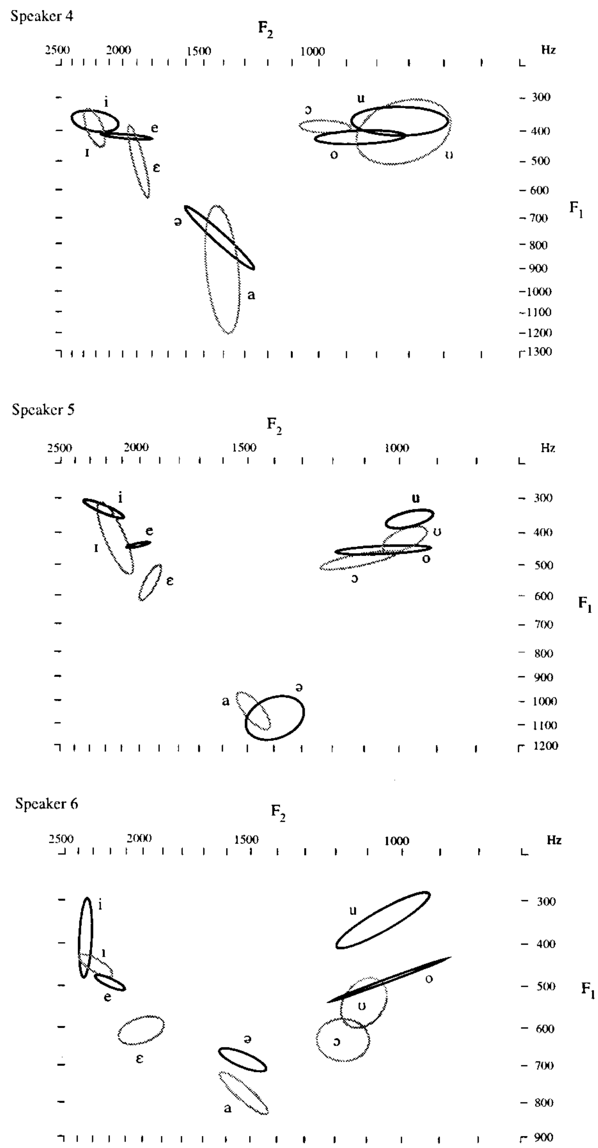


Figure 2: Conventional F_1–F_2 vowel formant plots for speakers 4–6

Table 2: F_2 difference: Significance and effect sizes

	Vowel quality	a	e	i	o	u
Speaker 1	FP significance	**0.0028**	**0**	0.72	0.18	0.93
	Effect size	**0.84**	**1.0**	0.56	0.30	0.48
Speaker 2	FP significance	**0**	0.12	1.0	**0**	0.076
	Effect size	**1.0**	0.81	0.50	**1.0**	0.19
Speaker 3	FP significance	0.096	**0**	0.31	0.85	0.17
	Effect size	0.27	**1.0**	0.66	0.47	0.70
Speaker 4	FP significance	0.81	0.45	1.0	**0.00077**	0.81
	Effect size	0.56	0.69	0.50	**0.94**	0.56
Speaker 5	FP significance	**0.012**	**0.00077**	0.30	0.50	0.49
	Effect size	**0.13**	**0.91**	0.72	0.34	0.34
Speaker 6	FP significance	0.90	**0**	**0.016**	**0**	0.36
	Effect size	0.47	**1.0**	**0.88**	**0**	0.28

The first number in each cell in Table 1, the FP-test significance, gives the probability that it is mistaken to say "the two distributions differ, and in particular the median F_1 of the [+ATR] vowels is less than that of the [−ATR] vowels." The second number, the effect size, gives us an indication of the degree to which the two vowels differ in their F_1 values, measured as the probability that a randomly selected [+ATR] token will have a lower F_1 than a randomly selected [−ATR] token. For example, when the [+ATR] and [−ATR] vowels of each pair are perfectly distinguishable in this way and [+ATR] shows a lower F_1 than [−ATR] this probability has the value 1.0. If the two types of vowels are not distinguishable in this way (as when their ellipses in Figures 1 and 2 overlap), this probability has a value closer to 0.5. When the vowels are perfectly distinguishable, but with the direction of the difference opposite to that expected, the effect size is zero.

Three of the vowel pairs, /e, ɛ/, /o, ɔ/, and /u, ʊ/, are almost perfectly distinguished by F_1 alone for five of the six speakers. A fourth pair, the high front vowels /i, ɪ/, are distinguished by F_1 by four of the six speakers. Speakers 1 and 6 are the only ones who distinguish the final pair, /ə/ from /a/, by F_1 frequency. Speaker 4's results deserve further comment in that he does not distinguish any of these vowels in the usual way. As the statistics show, he makes no reliable difference in F_1 except for /o, ɔ/. In this case he behaves in a very odd way, making the vowels distinguishable, but in the opposite direction to other speakers, with F_1 for /o/ having a higher value than for /ɔ/.

The FP-test significance in Table 2 gives the probability that it is mistaken to say "the two distributions differ, and in particular the median F_2 of the [−ATR] vowels is less than that of the [+ATR] vowels." The direction of the effect is less important here, since we are only interested in whether or not any vowel pairs are distinguishable using F_2 alone. Eleven of the 30 vowel pairs have significantly distinct F_2 values. The direction of the difference is not entirely consistent, but 9 of the 11 significantly different [$-$ATR] vowels have a smaller F_2 than the advanced vowels. A lower F_2 indicates greater tongue retraction as hypothesized.

3.2.2 Normalized relative formant amplitude

When formant amplitude and/or bandwidth are used as spectral shape measures, it is important to account for the fact (shown by Fant 1960) that these values are correlated with (and are at least grossly predictable from) the frequency of a formant, owing to the properties of the human vocal tract mechanism. Hess (1992) avoided the resulting difficulties by comparing the bandwidths of vowels with similar formant frequencies. If an investigator desires to compare spectral attributes among vowels of different formant frequencies (as would be the case in a comparison of phonologically corresponding [$\pm$ATR] vowels), s/he must take steps to eliminate the possibility that evident spectral shape differences can be accounted for by their correlation with formant frequency differences.

A normalizing procedure Our auditory impressions are that Degema retracted vowels have a somewhat brighter sound than their advanced counterparts. This impression could be caused by a difference in spectral shape between the two varieties, but it would have to be in addition to that which would anyway occur when there is a change in F_1. In order to investigate this possibility, a measure of the normalized relative intensity of the first two formants was computed from the data at hand.

The amplitudes of the formants in a vowel depend on a number of factors. The first is the bandwidth of the resonance of the vocal tract that produces the formant. The relation between the bandwidth and the formant amplitude is given in equation (1), in which $dB(f)$ is the amplitude of each frequency, f, through the frequency range of interest, which is taken to be from 100 to 4,000 Hz. For each formant, F is a resonant frequency of the vocal tract, and b is the bandwidth.

$$dB(f) = 20\log_{10}\frac{F^2 + (b/2)^2}{\sqrt{(f-F)^2 + (b/2)^2} \times \sqrt{(f+F)^2 + (b/2)^2}} \tag{1}$$

Curves for three formants as specified by this equation are shown in Figure 3. The spectrum of a vowel is the sum of these three curves, plus two other curves which we will consider below. As may be seen in Figure 3, when we add just the curves representing the first three formants together, the second formant peak is lowered by the negative contributions of the first formant, and the third formant peak is lowered even more by the negative contributions of both the first and the second formants. There is also a converse effect, in that the peak of each formant is raised by the contribution of each of the formants above it. The first formant peak is raised only slightly, but the second formant peak is noticeably raised by the

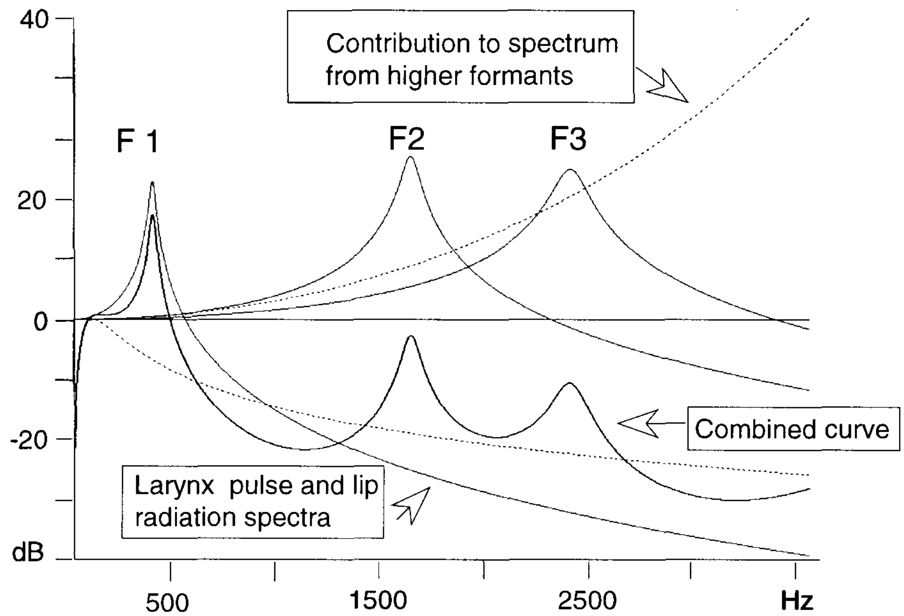


Figure 3: Curves representing three formant frequencies, the contribution of higher formants, and the larynx pulse and lip radiation. The heavier curve, representing the spectrum of a vowel, is the sum of these other curves.

positive contributions of the third formant at this frequency. To get a true summation of the first three formants we must include the raising contributions of all the higher formants. When three formants are specified, the equation for the additional contributions is as in (2), assuming a total vocal tract length of 17.5 cm and thus a first resonance of 492 Hz for a neutral vowel (Fant 1960:50). These raising contributions are represented by the upper dashed curve in Figure 3.

$$dB(f) = 0.72(f/492)^2 + 0.0033(f/492)^4 \tag{2}$$

The lower dashed curve in Figure 3 corresponds to the spectrum of the pulse produced by the vocal folds as modified by the lip radiation, which also affects the relative amplitudes of the formants. The shape of the larynx pulse depends on the phonation type. It is often taken as having a spectrum that falls at a rate of -12 dB per octave. It can be combined with a fixed curve representing the acoustic effects that occur when the vibrations of the air in the vocal tract are converted into a sound wave propagated from the lips. These effects contribute a positive slope of $+6$ dB per octave, so that the combined curve typically has a slope of -6 dB per octave. The equation we will use is shown in (3), where g is a variable representing differences in phonation type. This is the equation for the lower dashed curve in Figure 3, in which $g = 1.0$.

$$dB(f) = g\left(-20\log_{10}\left(2\frac{f/100}{1+\sqrt{f/100}}\right)\right) \tag{3}$$

When we consider the relative amplitudes of the formants in pairs of Degema vowels, we need to take all these factors into account. The relative amplitudes, $A_1 - A_2$, may differ in a pair of vowels because the bandwidths are different, or because the frequencies in one pair are not the same as in the other, or because there are difference in the glottal slope. (We will take it that the contributions of higher formants remain constant.) The bandwidths and glottal slope are two aspects of what we have called spectral timbre; since we already know how the vowels differ as regards formant frequency and are now interested in these other factors, the desire is to in some sense "factor out" the effects of formant frequency differences across different vowels. In order to normalize the relative formant amplitudes in this way between [+ATR] and [−ATR] vowels in Degema to facilitate the comparison of the other aspects of their spectral timbres, a program was written to calculate what the relative amplitudes would be if the bandwidths of the formants, the spectral slope and the contributions of higher formants had remained constant, and only the formant frequencies varied. If there are non-systematic differences between the observed and calculated values of $A_1 - A_2$ for each vowel, then we may hypothesize that some other factor(s) are not being held constant by the speakers from one vowel to another, viz. either the bandwidths or the glottal slopes. The default bandwidths were taken to be $B_1 = 30$, $B_2 = 80$ and $B_3 = 150$. The spectral slope was as defined by (3), with $g = 1.0$, and the contributions of the higher formants were as defined by (2).

For each vowel token, the modeled $A_1 - A_2$ value was subtracted from the observed value. This provides a measure of the discrepancy between measurement and model for the relative intensity of the first two formants. It can be regarded as a measure of the relative formant amplitude, corrected so as to take into account differences due to variations in the formant frequencies. Our next task is to see whether these normalized relative amplitudes differ between advanced and retracted vowels.

Results Table 3 gives the results of 30 FP-tests, one for each [±ATR] pair of vowels. The alternative hypothesis in each case is that the normalized relative formant amplitude of the retracted vowel is less than that of the advanced vowel. When this is so, it can be concluded that the first formant contributes relatively less, and the second formant relatively more, to the spectral shape of the retracted vowel.

The results are mixed between the null hypothesis (that advanced and retracted counterparts do not have different spectral shapes) and the alternative hypothesis. Speaker 1 shows the expected difference only for /a/ and /o/ vowel pairs; Speaker 2 for /e/, /o/, and

Table 3: Relative amplitude difference: Significance and effect sizes

	Vowel quality	a	e	i	o	u
Speaker 1	FP significance	**0.046**	0.36	0.71	**0**	0.96
	Effect size	**0.76**	0.64	0.56	**0.92**	0.49
Speaker 2	FP significance	0.80	**0**	1.0	**0**	**0**
	Effect size	0.56	**0.94**	0.50	**1.0**	**1.0**
Speaker 3	FP significance	0.85	**0**	0.45	0.48	0.16
	Effect size	0.53	**0.95**	0.63	0.61	0.70
Speaker 4	FP significance	0.44	**0**	1.0	**0.012**	**0**
	Effect size	0.67	**1.0**	0.50	**0.88**	**0.94**
Speaker 5	FP significance	0.23	0.61	**0.012**	**0**	0.23
	Effect size	0.25	0.63	**0.88**	**0.94**	0.75
Speaker 6	FP significance	**0**	0.44	0.61	**0**	0.81
	Effect size	**1.0**	0.31	0.63	**1.0**	0.44

/u/; Speaker 3 only for /e/; Speaker 4 for /e/, /o/, and /u/; Speaker 5 only for /i/ and /o/; and Speaker 6 only for /a/ and /o/. It is interesting to note that Speaker 4, who was the one speaker who did not use F_1 to distinguish the /e/, /o/, and /u/ pairs, is one of the two speakers who distinguishes these pairs by spectral shape. In all, 13 of the 30 vowel pairs reveal a difference in spectral shape between advanced and retracted through the relative formant amplitudes, with the /o/ pair showing this effect for five of the six speakers. It is also important that the reverse of the alternative hypothesis is never indicated. In other words, it is certainly not the case that there is a spectral shape difference that is opposite to the theoretical expectation.

Discussion It is established above that F_2 is at times a relatively greater contributor to the spectra of retracted tongue root vowels in Degema. Switching perspective, it may be said that F_1 is at times a lesser contributor to the spectral intensity of these same [−ATR] vowels. Stated in this way, the results can be seen to partially agree with those of Hess (1992), who found that Akan [−ATR] vowels showed a larger F_1 bandwidth and thus a decreased F_1 intensity.

What could be the cause of the spectral shape differences in Table 3? Earlier discussion has addressed the fact that formant bandwidth is directly related to damping in the acoustic system; the greater the bandwidth, the more the damping. By modeling the vocal tract configuration of the vowel /i/ as a series of connected acoustic cylinders, and performing calculations on an electrical analog developed by Dunn (1950), van den Berg (1955) was able to distinguish three different contributions to the resistance (damping) of the supraglottal vocal tract. These are:

1. Radiation resistance, due to the coupling of the mouth opening with the air outside, which affects primarily F_2 and higher formants.

2. Friction damping, due to the viscosity of the air within any approximation involving the tongue or pharynx. This is present to an appreciable degree only in vowels with small openings, and affects primarily F_1.

3. Cavity wall coupling; the vibration of the vocal tract walls in sympathy with that of the air consumes energy from the acoustic signal. This also mainly affects F_1.

There are also important damping effects associated with the degree of opening of the glottis and the mode of vibration of the vocal folds. We surmise that these are not the source of spectral shape differences between [+ATR] and [−ATR] vowels, since such differences have not been found to distinguish more than three of the vowel pairs for any speaker.

In the Degema tongue root contrast, a difference in lip opening between advanced and retracted vowels would cause a change in the radiation resistance. Additionally, the articulation of some retracted vowels may be sufficiently constricted that the back cavity resonance could suffer increased damping due to the increased air friction within the approximation. The reader is referred to Mason (1948:118) for a mathematical treatment of acoustical damping in a narrow constriction. Some combination of these effects must underlie the consistent spectral shape difference between [o] and [ɔ]. The spectral shape differences in the other vowel pairs is much more sporadic across speakers, and we cannot say exactly what the articulatory cause is in any particular case.

In earlier discussion it was suggested, following Hardcastle (1976), that isometric tension of the pharyngeal constrictors is likely to accompany pharyngeal constriction. By increasing the tension of the throat walls in this manner, the vocal tract damping resulting from coupled vibration of the walls would be reduced. We also mentioned van den Berg's (1955) indication that this would primarily affect resonances below 1000 Hz. Given that F_1 is in this frequency range, the effects of pharyngeal tension in [−ATR] vowels should be opposite to our observation of *increased* damping of F_1 relative to F_2.

No such opposite spectral shape effect was observed; this is probably because the damping characteristics of the vocal tract walls cannot be changed by the speaker over as wide a range as the constriction damping or the radiation resistance. So, the pharyngeal wall coupling factor will remain much steadier than the constriction damping factor during pharyngeal constriction in retracted tongue root vowels. Additionally, the damping contributed by the vocal tract walls is also larger for smaller constrictions, whether the walls are stiff or slack. This amounts to saying that any *decrease* in F_1 damping caused by increased pharyngeal tension will be effectively swamped by the *increase* in damping owing to a marked narrowing of the pharynx.

4 Degema vowel perception

4.1 Procedure

We were not able to test the way in which differences in spectral slopes and bandwidths affect the perception of ATR differences, but we were able to test the importance of the formant frequencies in the perception of Degema vowels. We constructed a computer model that would synthesize vowels. The subject and the experimenter (E. K., a native speaker of Degema) viewed a computer display of the kind shown in Figure 4, which included a written instruction to match a particular vowel. All the subjects were literate in English, and had no difficulty in understanding the task, The program synthesized a vowel in accordance with a first and second formant frequency as specified by a mouse click in one of the large squares. The third formant frequency could be set by clicking in a separate scale, but we have found that inexperienced subjects usually find the addition of this variable confusing. Accordingly the frequency of the third formant was determined automatically from the first two in accordance with an algorithm specified by Broad and Wakita (1977). The fourth formant was fixed at 3500 Hz. The program produced a vowel with a natural intonation contour; the mean pitch could be varied, but was kept fixed throughout this experiment. Once the subject had determined the approximate area for an appropriate vowel quality, clicking in the Detail button produced a set of smaller squares in that neighborhood (as shown in the figure), and the subject was asked to make a more precise judgment. On clicking in the Next button, the formant frequencies of the vowel chosen were recorded in a log file, and the program displayed a written form of the next vowel to be matched.

In each session subjects were asked to match all ten vowels, which were presented to them in a random order. Each subject took part in two sessions. The subjects were not very experienced in the use of computers (most of them had never operated one), nor in performing perceptual experiments (none of them had ever attempted a similar task), so they required considerable assistance from the experimenter in performing the task. Nevertheless, the vowel matches were the subject's own choice, as the experimenter remained strictly neutral, carefully avoiding biasing them in any way.

4.2 Results

Ellipses enclosing the vowels selected by the five listeners are shown in Figure 5. There are clear cut results for only two of the five vowel pairs. For all five listeners, the [+ATR] /e/ and /o/ are higher (i.e. they have a lower F_1) than their [$-$ATR] counterparts. The five listeners do not behave alike when trying to synthesize any of the other vowels. This may be because there are no consistent differences in formant frequency in these other pairs of vowels. The perception results indicate that, with the exception of the mid vowels, formant frequency alone is a poor indicator of the vowel category for native speakers.

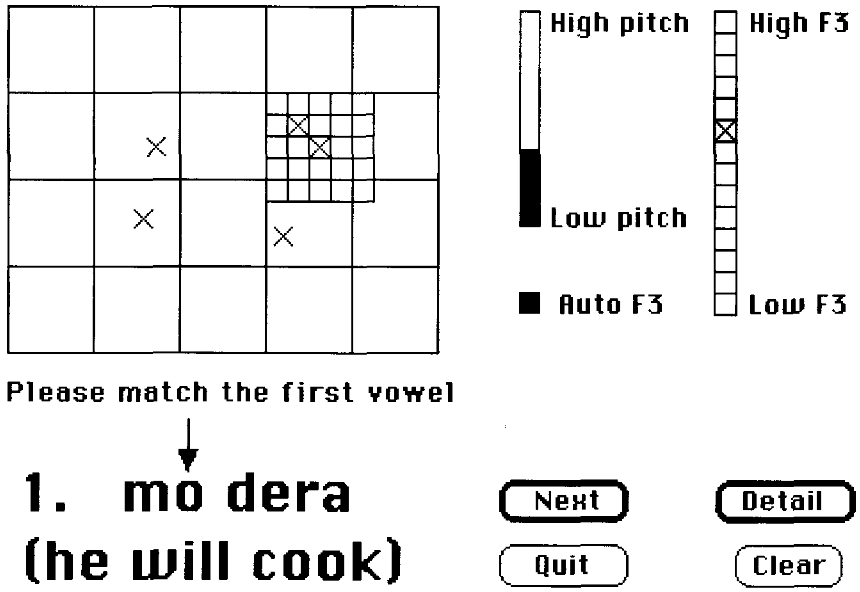


Figure 4: Display used by subjects and the experimenter when trying to find a match for the first vowel in *módera* 'he will cook.'

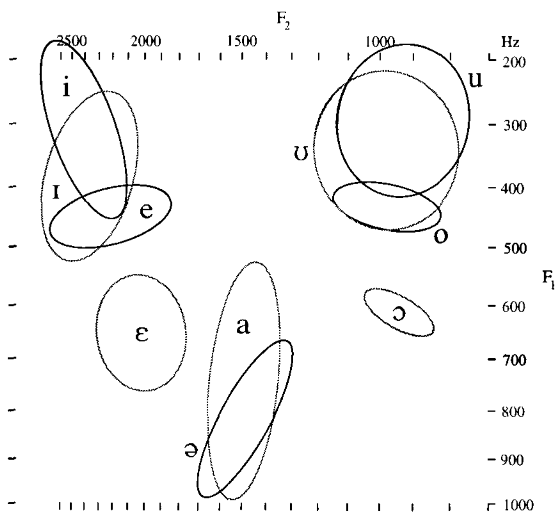


Figure 5: Formant space of Degema vowels synthesized by 5 listeners.

5 Summary

This paper began by tracing some of the history of the linguistic investigation of ATR contrasts in African languages, and discussing Stewart's original work at some length. The general nature of the ATR distinction has been shown to be unlike a tense/lax distinction; Lindau's characterization of tongue-root advancement as pharyngeal expansion seems to be a better generalization.

The expected acoustic correlates of pharyngeal expansion/contraction have been discussed, through consideration of the acoustic properties of the vocal tract and the relevant actions within. The expected effects, based on both mathematical models and past empirical results, are the lowering of F_1 for [+ATR] vowels and a possible difference in vowel spectral shape, by which we here mean the relative amplitudes (bandwidths) of the formants, between the two tongue root postures.

In an investigation into the acoustic properties of the ATR contrast in Degema, the expected difference in F_1 was largely confirmed in that all vowel pairs except /ə, a/ are distinguished in this way by most speakers. F_2 is a much poorer indicator of this kind of contrast, but the three pairs /ə, a/, /e, ɛ/, /o, ɔ/ are distinguished by at least three speakers each.

A procedure for obtaining a normalized measure of relative formant intensity was described. A comparison of corresponding advanced and retracted vowels using the measure demonstrated that relative formant intensity could distinguish them in 43% of the cases, but only the vowel pair /o, ɔ/ is distinguished in this way by a majority of speakers. The articulatory mechanisms which could underly differences in normalized relative formant intensity were discussed; these are thought to be a change in the size of the lip opening, and a narrowing of a constriction within the vocal tract as a result of tongue root retraction.

A perceptual study was conducted which demonstrates that Degema speakers do not classify their vowels very well using formant frequencies as the sole acoustic variable. It is thus suggested that the changes in spectral shape contribute to the overall perception of the [−ATR] vowels, though the effects are shown to be sporadic. A difference in spectral shape may be detectable by the Degema speakers, who could listen for a change in relative formant intensities as was measured here. They may also listen for the formant bandwidths directly; indications are that both factors should be perceptually salient. The presence of several acoustic correlates is thought to contribute to the Degema's ability to distinguish the ten vowels shown in Figures 1 and 2. The sporadic nature of the correlates may be compensated for somewhat by the additional constraints on vowel distribution provided by the ATR harmony system.

6 Appendix: Statistical analysis

The data in the present study provide something of a challenge for statistical analysis; speakers 1 and 3 provided eight tokens of each vowel and the remainder provided only four tokens. While the best-known method for comparing two groups is Student's t-test (which is the same thing as an Analysis of Variance for two groups), applying it here would not be the best choice for a number of reasons. Student's t-test assumes a normal distribution of the data and compares the means of the two groups. This procedure is not reliable when there are departures from normality (which cannot be detected with small sample sizes) or with skewed distributions since the mean is not a resistant measure of location (i.e. it can be thrown off by a single datum). Two difficulties can plague t-test results in a case like this. The first is loss of control over Type I errors; in other words, false significance of small differences. The second is loss of *power* (control over Type II errors). When power is lost, we can obtain limited significance and be led to accept the null hypothesis, when in fact our alternative hypothesis is the correct one.

As an alternative to the t-test it is often prudent to consider using a measure of location other than the mean when two groups are being compared. The median, for example, is much more resistant; it is unaffected by a change in as much as 50% of the data. A problem still arises with small samples, however; it turns out that tests which explicitly compare measures of location require a sample size of at least 15 in order to confidently control Type I errors and power at 5% significance levels (Wilcox 1996). Of particular interest as an alternative here are methods of comparison based on ranks. The Mann-Whitney-Wilcoxon test (Mann and Whitney 1947), for example, allows us to test the hypothesis that two groups have identical distributions without relying on a direct estimate of some measure of location. The test is theoretically sensitive to differences between the medians, but no estimate of the median is used (Wilcox 1996). It also allows the computation of an equally important property revealed by the data, the *effect size.*

Let P be the probability that an observation randomly sampled from the first group is less than a randomly sampled observation from the second group. If two groups have identical distributions, $P = 1/2$; the Mann-Whitney-Wilcoxon procedure tests an alternative against the null hypothesis that $P = 1/2$. One difficulty with the procedure is that it assumes the two groups have equal shapes and therefore equal variances, a condition that is not met by our data. As a result, we employ a related but improved procedure that eliminates this requirement, the Fligner-Policello test (Fligner and Policello II 1981). Again, the null hypothesis is $H_0 : P = 1/2$. As P approaches 0 or 1, the two-tailed procedure is more likely to show a significant difference in distributions. According to Wilcox (1996), the Mann-Whitney-Wilcoxon test (and by extension the Fligner-Policello test) has much higher power than methods for comparing means like the t-test. The tests were carried out using Minitab statistics software, and macro procedures written by Wilcox.

We have determined values that indicate the significance of the effect (the familiar p

values), but these are meaningless without some indication of the importance of the effect (i.e. its size) and a corresponding interpretation of the significance. The value that is often used to report significance tells us the chance of being mistaken in accepting our alternative hypothesis. The Fligner-Policello test enables us to report significance results and draw conclusions from them as such, and we are also provided with an excellent measure of the size of the effect, the value of the probability P that is used as a test statistic. In fact, Cliff (1993) argues for this statistic as an excellent measure of effect size in general.

References

van den Berg, J. 1955. Calculations on a model of the vocal tract for the vowel /i/ (meat) and on the larynx. *Journal of the Acoustical Society of America* 27:332–337.

Broad, D. J., and H. Wakita. 1977. Piecewise-planar representation of vowel formant frequencies. *Journal of the Acoustical Society of America* 62:1467–1473.

Cliff, N. 1993. Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin* 114:494–509.

Dunn, H. K. 1950. The calculation of vowel resonance and an electrical vocal tract. *Journal of the Acoustical Society of America* 22:740–753.

Elugbe, Ben O. 1989. *Comparative Edoid: Phonology and Lexicon.* University of Port Harcourt Press.

Fant, Gunnar. 1960. *Acoustic Theory of Speech Production.* The Hague: Mouton. Reissued 1970.

Fligner, M. A., and G. E. Policello II. 1981. Robust rank procedures for the Behrens-Fisher problem. *Journal of the American Statistical Association* 76:162–168.

Fulop, Sean A. 1996. An acoustic investigation of the tongue root contrast in degema vowels. *UCLA Working Papers in Phonetics* 93:13–46.

Hardcastle, W. J. 1976. *Physiology of Speech Production.* London: Academic Press.

Hess, Susan. 1992. Assimilatory effects in a vowel harmony system: an acoustic analysis of advanced tongue root in Akan. *Journal of Phonetics* 20(4):475–492.

Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages.* Oxford: Blackwell.

Mann, H. B., and D. R. Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics* 18:50–60.

Mason, W. P. 1948. *Electromechanical Transducers and Wave Filters.* New York: D. van Nostrand.

Wilcox, Rand R. 1996. *Statistics for the Social Sciences.* San Diego: Academic Press.

Variations and universals in VOT: evidence from 17 endangered languages

Taehong Cho and Peter Ladefoged

Voice Onset Time (VOT) is known to vary with the place of articulation. For any given place of articulation there are also differences from one language to another. Using data from multiple speakers of 17 languages, all of which were recorded and analyzed in the same way, we show that most, but not all, of the within language place of articulation variation can be described by universally applicable phonetic rules (although the physiological bases for these rules are not entirely clear). The between language variation is also largely (but not entirely) predictable by assuming that languages choose one of three possibilities for the degree of aspiration of voiceless stops. Some languages, however, have VOTs that are markedly different from the generally observed values. The phonetic output of a grammar has to contain language specific components to account for these results.

1. Introduction

When a pattern recurs in hundreds of languages it may seem inevitable. For example, many phoneticians have noticed that vowels are usually longer before voiced than before voiceless stops (Halle and Stevens 1967, Chen 1970, Lisker 1974, Maddieson and Gandour 1977, Maddieson 1997, inter alia). It is also a common observation that high vowels in stressed monosyllables are shorter than low vowels in comparable syllables (Lindblom 1967, Lehiste 1970, Lisker 1974, Westbury and Keating 1980, Maddieson 1997). But neither of these patterns is inevitable. A language that at one time had a contrast between long and short vowels, could lose this possibility and keep just the long high vowels and the short low vowels. A language of this kind might be slightly more difficult to learn, but it would not be impossible.

There are, however, other kinds of phonetic events that have inevitable consequences. Whenever the tongue goes from a raised position in the front of the mouth to a low position in the back, the frequency of the first formant will go up and that of the second formant will go down. Similarly, if there is no compensatory adjustment, stretching the vibrating vocal folds will always raise the pitch of a voiced sound. Again, other things being equal, whenever a contraction of the internal intercostal muscles occurs to produce a stressed syllable, then the syllable will have a higher pitch and an increase in loudness.

In discussing phonetic universals we should keep these two kinds of phonetic events distinct. It is physically impossible to move the tongue from a high front to a low back position without raising F1 and lowering F2. It is perfectly possible to reverse the usual vowel length differences between high and low vowels, although the resulting gestures may be more difficult to make. In this paper we will discuss variations in VOT (Voice Onset Timing), and try to show which if any of them are inevitable consequences of some physiological adjustment, and which are simply the most favored (perhaps the easiest) articulatory gestures.

It is well known that VOT varies with place of articulation. The principal findings are that: (1) the further back the closure, the longer the VOT (Fischer-Jørgensen 1954, Peterson and Lehiste 1960); (2) the more extended the contact area, the longer the VOT (Stevens, Keyser and Kawasaki 1986); (3) the faster the movement of the articulator, the shorter the VOT (Hardcastle 1973). These patterns can be observed in Lisker and Abramson's (1964) classic cross-linguistic study of VOT — although they themselves did not go into detail concerning variations of VOT conditioned by place of articulation. Tables 1 and 2 show that, in their data, velar stops always have a longer VOT. Furthermore, in both aspirated and unaspirated stops, VOT is shortest before bilabial stops and intermediate before alveolar stops, with the exception of the unaspirated stops in Tamil and the aspirated stops in Eastern Armenian.

Table 1. Summary of VOT (ms) in unaspirated stops reported by Lisker and Abramson (1964)

	Dutch	Puerto Rican Spanish	Hungarian	Cantonese	Eastern Armenian	Korean	Tamil
/p/	10	4	2	9	3	18	12
/t/	15	9	16	14	15	25	8
/k/	25	29	29	34	30	47	24

Table 2. Summary of VOT (ms) in aspirated stops reported by Lisker and Abramson (1964)

	Cantonese	English	Eastern Armenian	Korean
/ph/	77	58	78	91
/th/	75	70	59	94
/kh/	87	80	98	126

In early forms of generative phonology, such patterns were considered to be attributable to low level (automatic) phonetic implementation rules, constrained by physiological (biomechanical) factors, and thus not a necessary part of the grammar of any one language. This is the view expressed by Chomsky and Halle (1968) in the *Sound Pattern of English* (SPE). In SPE, for any given language, once binary features have been converted into scalar featural values, the physical output is completely determined by universal phonetic implementation rules.

It has been known for many years that the SPE view is not completely correct, and that there are language specific phonetic rules which must be part of the grammar of each language (Pierrehumbert 1980, 1990; Keating 1984, 1985, 1990; Cohn 1993 among others). In particular, Keating (1985) convincingly shows that three assumed phonetic universals, intrinsic vowel duration, extrinsic vowel duration, and voicing timing, are not automatic results of speech physiology. They are not universal attributes of sounds, but are at least in part determined by language specific rules. Keating did not have data available that would have enabled her to asses the extent to which variations in VOT were due to language specific factors as opposed to universal physiological or aerodynamic principles. We will survey the VOT variations in 17 languages, and try to assess these factors.

In order to see if place dependent VOT patterns found in any given language fall out from the physiological-aerodynamic factors, we will first discuss the commonly assumed explanations for these variations. Next, in section 2, we will survey VOT data from 17 languages, all of which were recorded and analyzed in the same way. Finally, in section 3, we will consider the significance of the data summarized in this survey. We will discuss the place of timing rules for VOT patterns within the grammar of a given language, and we will also address the issue of whether there is a point in the continuum of VOT distinguishing voiceless unaspirated from voiceless aspirated stops.

1.2. Explanations of VOT patterns

There have been several explanations in the literature for the voice onset differences found in the studies reported above. These explanations depend on a number of factors, including general laws of aerodynamics, articulatory movement velocity and differences in the mass of the articulators. In addition, there is an alternative analysis that suggests there is a temporal adjustment between the closure duration and the VOT (Weismer 1980, Maddieson 1997).

General laws of aerodynamics

Many phoneticians (cf. Hardcastle 1973, Maddieson 1997) have suggested that one of the factors which contribute to the voice onset differences is the relative size of the supraglottal cavity behind the point of constriction. There are two ways of considering this. Firstly, the cavity behind the velar stop has a smaller volume than that behind the alveolar or bilabial stops. Secondly, the cavity in front of the velar stop has a larger volume than that in front of the alveolar or bilabial stops.

In order to produce voicing there needs to be a difference in air pressure across the vocal folds (van den Berg 1958). If the air in the oral cavity is at a pressure similar to that below the vocal folds, there cannot be airflow between them and they will not vibrate. From the first point of view, the notion that the cavity behind the velar stop has a smaller volume than that behind the alveolar or bilabial stops, it may follow that the velar stop has a greater pressure behind it at the beginning of the release phase. The air in the lungs and the vocal tract has to be considered as a single volume, which is smaller in a velar stop than it is in an alveolar or bilabial stop. During an utterance, the air is compressed by the action of the respiratory muscles. If the volume being compressed is small, a given reduction in size will produce a greater increase in pressure. As a result the air pressure in the vocal tract may be higher for a velar stop. (We should note that this is an unverified hypothesis. Although there have been many measurements of the pressure of the air in the vocal tract, we do not know of any that report systematically higher pressures for velars.) If the oral air pressure is higher for a velar stop, it will take longer for it to fall and allow an adequate transglottal pressure for the initiation of the vocal fold vibration.

The second point of view considers the fact that there is a larger body of air in front of the velar stop. This body of contained air will act like a mass that has to be moved before the compressed air behind the velar closure can be released into the open air. Irrespective of whether there is or is not a higher air pressure behind velar closures, the drop in the pressure of the air in the vocal tract will be slower for velars. As a result it will take relatively more time to attain the crucial transglottal pressure difference required to initiate voicing.

Movement of articulators

In addition to noting the effect of differences in the relative size of the supraglottal cavities, Hardcastle (1973) also postulates that the voice onset difference can be due in part to the fact that the tip of the tongue and the lips move faster than the back of the tongue. This notion is supported by a cineradiographic study of VC and CV articulatory velocities by Kuehn and Moll (1976), who report that the articulatory movement is fastest for the tongue tip, intermediate for the lower lip, and slowest for the tongue body. This may be partly due to the relative masses of the articulators involved; the tongue tip is smaller and lighter than the lips or the body of the tongue. It is also due to the fact that jaw movements affect lip and tongue movements in different ways; tongue dorsum movement is least affected by jaw movement, while lower lip movement is accelerated by jaw movement.

In line with this latter point, Maddieson (1997) suggests that one of the reasons for the difference in VOT between English stops /p/ and /k/ is the distance from the pivot point of the jaw rotation. A schematized representation of the effect of jaw rotation is shown in Figure 1. As illustrated in the figure, because the pivot of jaw rotation is further from the lip than the tongue body, the movement of the lower lip will be greater than that of the tongue body for a given angular motion of the jaw (see also Vatikotis-Bateson and Ostry 1995). When the articulator is the lower lip, the compressed air behind the constriction escapes at a faster rate, resulting in a shorter time before building up an appropriate transglottal pressure for the initiation of voicing.

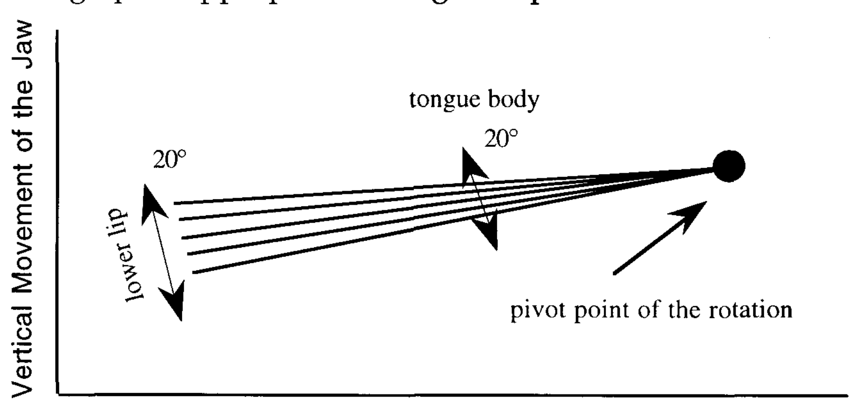


Figure 1. Schematic representation of the effect of jaw rotation. A 20 degree shift in jaw angle separates the lips apart more than the tongue back and velum.

As Maddieson (1997) notes, this explanation does not account fully for the place-related difference in VOTs between bilabial and alveolar stops. Recall that Kuehn and Moll (1976) report that the tongue tip moves faster than the lower lip. If the articulatory velocity is the primary physiological factor for the voice onset difference, we would expect that the VOT would be shorter for apical alveolar stops than for either bilabials or velars, which is not the general finding. This suggests there may be some other factors accounting for the place-related voice onset difference.

The extent of articulatory contact

The VOT variations can also be partly accounted for in terms of the extent of the contact area between the articulators. As velar stops are produced with a constriction between the rounded upper body of the tongue (the dorsum) and the similarly rounded soft palate, the contact area is more extended than that in bilabial and alveolar stops. There is a similar difference in contact length between laminal and apical stops which almost always accompanies dental vs. alveolar stop contrasts (Ladefoged and Maddieson, 1996). In general, stops with a more extended articulatory contact have a longer VOT.

Stevens (in press) provides an aerodynamic explanation for these differences. His main point is that the rate of change in intraoral pressure following the release depends on the rate of increase in cross-sectional area at the constriction. This is significantly different for different places of articulation, primarily due to the differences in the extent of articulatory contact. When there is a long narrow constriction the Bernoullli effect causes the articulators forming the constriction to be sucked together. Because the velar stop has extensive contact between the tongue body and the palate, there is a larger Bernoulli force so that the change in cross-sectional area is relatively slow compared with that for the bilabial or alveolar stops. Consequently, the decrease in intraoral pressure after the closure is gradual for the velar and rapid for the bilabial. Stevens' aerodynamic data show that the volume velocity of airflow at both the constriction and the glottis increases roughly in proportion to the rate of the decrease in intraoral pressure for the first 50 ms immediately following the release of the closure. Schematized curves of airflow and intraoral pressure at the release of voiceless stops appear in Figure 2.

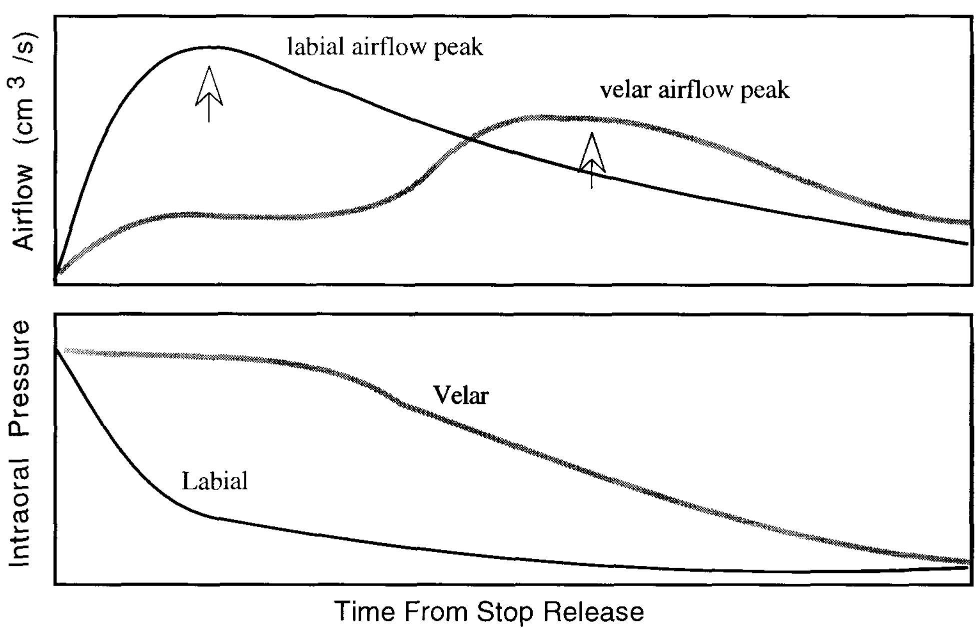


Figure 2. Schematized curves of airflow and intraoral pressure at the release of voiceless stops, based on data in Stevens (in press).

The timing of the vocal folds vibration is determined by the two inter-related aerodynamic factors shown in the figure: (1) the rate of decrease in intraoral pressure and (2) the rate of increase in volume velocity of the airflow. In his discussion of these relationships Stevens was not concerned with differences in laminality. He noted that alveolar stops can be considered to be produced at an intermediate rate of change of both the intraoral pressure and the volume velocity of airflow, assuming that the alveolar contact area is longer than the bilabial, but shorter than the velar.

Change of glottal opening area

In addition to the factors described above, Stevens (in press) ascribes differences in VOT among voiceless aspirated stops to the different degrees of glottal opening area that accompany the different places of articulation. For the aspirated stops, the glottis is already open well before the release to allow for aspiration. After the release this glottal opening must be reduced to reach approximately 0.12 cm^2 in order to initiate vocal fold vibration. Stevens suggests that the glottal opening area after the release will decrease less rapidly for the velar than for the alveolar or for the labial stop because the intraoral pressure for the velar stop drops more slowly. A build-up of intraoral pressure induces an outward displacement (i.e. abduction forces) on the glottal folds (as well as on the walls of the vocal tract). In addition, during the closure interval the stiffness of the walls of both the glottis and the vocal tract is increased, presumably to counteract the increased intraoral pressure. The decrease of intraoral pressure immediately following the release causes an inward force of the walls of the glottal folds, coupled with a corresponding relaxation of the stiffness. However, the stiffness is still preserved to some degree, thus inhibiting the vibration of vocal folds immediately following the release. On the basis of these assumptions, Stevens posits that the glottal area decreases somewhat more rapidly following the release of bilabial or alveolar stops than for the velar stops, since the decrease in intraoral pressure following the release of the bilabial or the alveolar stop is more rapid, and there is a more rapid formation of the adduction forces along with a more rapid relaxation of the stiffness. Thus, the voice onset occurs somewhat earlier for a labial or alveolar than for a velar voiceless aspirated stop.

Temporal Adjustment between stop closure duration and VOT.

The stop closure duration for bilabial stops is, in general, longer than that of either alveolar or velar stops, which may be due to different degrees of the air pressure in the cavity behind the constriction (Maddieson 1997). We already noted that a smaller cavity behind the constriction will cause a more rapid build-up of the intraoral air pressure, reaching equity with supraglottal air pressure in a relatively shorter time. Based upon this aerodynamic principle and results of an experiment by Ohala and Riordan (1979), Maddieson (1997:630) posits that "if the consonant gesture is timed in some way that directly relates to the time of the pressure peak, then broadly speaking, the further back in the oral cavity a stop closure is formed, the shorter its acoustic closure duration will be." This provides an inverse relationship between the closure duration and the observed VOT variation. Weismer (1980) reports that for word initial English /p/ and /k/, the interval from the onset of the stop closure to the voice onset is the same. Based upon this result and other evidence cited by Weismer, Maddieson (1997) suggests another possible alternative account of the place-dependent VOT:

"There is an abduction-adduction cycle of the vocal cords for voiceless stops which is longer in duration than the closure and has a constant time course, anchored to the onset of closure (p. 621)." In other words, the duration of the vocal fold opening is considered to be fixed, and when the closure duration is relatively longer, the following VOT becomes relatively shorter (and vice versa). Figure 3 is a schematic representation from Maddieson (1997:622) showing this relationship. Umeda (1977) and Lisker and Abramson (1964) also discuss the same type of durational relationship between closure and aspiration.

	Duration of vocal fold opening	
Bilabial :	Closure	Aspiration
Velar :	Closure	Aspiration

Figure 3. Schematic representation of place differences in aspirated stops from constant vocal fold abduction plus different closure duration. (From Maddieson, 1997:622).

Summary of reported causes of VOT variations due to place of articulation

In summary, the literature indicates that the following physiological/aerodynamic characteristics account, to some extent, for the variations of VOT associated with a difference in place of articulation.

(1) The volume of the cavity behind the point of constriction.
The relatively smaller volume of the supralaryngeal cavity in velar stops causes a greater pressure, which will take longer to fall and allow an adequate transglottal pressure for the initiation of the vocal folds vibration.

(2) The volume of the cavity in front of the point of constriction.
The relatively greater mass of the contained air in front of velar stops causes a greater obstruction to the release of the pressure behind the velar stop, so that this pressure will take longer to fall, resulting in a greater delay in producing an adequate transglottal pressure.

(3) Movement of Articulators.
A faster articulatory velocity (e.g. the movement of the lower lip as compared to the tongue dorsum) allows a more rapid decrease in the pressure behind the closure and thus a shorter time before building up an appropriate transglottal pressure.

(4) Extent of articulatory contact area.
The more extended contact area in laminal dental and velar stops results in a slower release because of the Bernoulli effect pulling the articulators together. Because the articulators come apart more slowly there is a longer time before an appropriate transglottal pressure is produced.

(5) Change of glottal opening area (for voiceless aspirated stops).
The glottal opening area after the release will decrease less rapidly for the velar than for the alveolar or labial because the intraoral pressure drops more slowly for the velar.

(6) Temporal adjustment between closure duration and VOT.
There is a trade-off between the closure duration and the VOT so that there is a fixed duration of vocal fold opening.

Characteristics (1) – (4) hold better for unaspirated or slightly aspirated stops. They are based on a general principle of aerodynamics: objects such as the vocal folds will vibrate only when there is a sufficient pressure difference across them, and sufficient flow between them. This principle holds, however, only if the vocal folds are adducted so that they are in a suitable position to vibrate. In the case of aspirated stops this does not occur for a considerable period after the release. Place effects on the transglottal pressure occur in the first few milliseconds after release. Even for velar stops the tongue body is expected to have lowered 4 to 5 mm by 50 ms after the release (Maddieson 1997). It is therefore unlikely that in any aspirated stop the supraglottal pressure will be high enough to affect the voicing initiation more than 50 ms after the release when the vocal folds are sufficiently adducted.

On the other hand, (5) will hold for aspirated stops, and (6) will hold for both unaspirated and aspirated stops. The characteristics in (5) explain, though indirectly, why the vibration of the vocal folds are suppressed even after the adequate transglottal pressure is attained. Recall that the stiffness in walls of both vocal folds and vocal tract are maintained to some degree following the release, which presumably inhibits the vocal fold vibration (Stevens, in press). The explanation in (6) also seems to account better for the variation of the aspirated stops. It depends on notions of speech timing rather than any aspect of the aerodynamic mechanism varying with different places of articulation.

2. VOT variations in 17 endangered languages

To discuss the factors underlying variations in VOT we need a body of data from a number of widely different languages, all of which have been collected and analyzed in the same way. Without such controls it is possible that any observed differences between languages may be due to the procedures used to obtain the data. The UCLA endangered languages project provided a suitable body of data. In this project all the recordings of endangered languages were made in the field in a standardized way by one or other of the two Principal Investigators, Peter Ladefoged and Ian Maddieson, with the exception of the Tiwi and the Hupa data, which were recorded by Victoria Anderson and Mathew Gordon, respectively, graduate student members of the UCLA Phonetics Lab. The recorded material always included lists of words illustrating the segmental contrasts in each language in various contexts. In this paper we are concerned only with the recordings of voiceless unaspirated and voiceless aspirated stops which were always recorded in initial position in contrasting words before a low vowel of the **a** type. (Contrasts before other vowels such as **i** were also recorded, but will not be considered here as the vowels differed from language to language. For similar reasons, contrasts that were elicited in sentences rather than as citation forms will nor be considered.) In all cases several speakers of each language were recorded (the exact number for each language is given in the tables below), all of them being adult native speakers who used the language in their daily life. The recordings were made on high quality equipment as described by Ladefoged (1997). They were analyzed in the UCLA Phonetics lab by Graduate Research Assistants working under the direct supervision of the Principal Investigators. In all cases the analysis procedure provided for observation of the waveform and spectrogram of each utterance, usually by means of the Kay CSL system. VOT was measured as the interval between cursors placed at the onset of release (the final release, if there was more than one) and the first complete vibration of the vocal folds as indicated on the waveform.

Greater control could have been taken, in that no special instructions were given to subjects concerning rate of speech, which was usually that of a typical fieldwork elicitation session in which citation forms are being repeated. The lack of such controls is unlikely to have biased our results. Although we cannot establish it by a rigorous procedure, we believe that the overall differences between languages that emerged are not artifacts of the different circumstances in which the recordings were made. It seems to us, for example, that the longer VOTs recorded in Navajo are a characteristic of that language and not due to a procedural artifact. We should also note that most of the languages investigated are moribund (Navajo is not; it is still spoken by children as a first language), but they are no different from any other languages in respect to their general phonetic structures. They are in danger of dying because of external circumstances forced upon the speakers, not for any internal linguistic reason.

In this paper we will consider only the mean VOT as reported in the papers published on each language. For further details on the speakers and more detailed statistical analyses the individual papers should be consulted.

2.1 Languages investigated

1. *Banawá* (Ladefoged, Ladefoged, and Everett 1997)

Banawá is an Arawan language, spoken by about 75 speakers in the rain forest of northern Brazil (Ladefoged et al., 1997). There are fully voiced stops at three places of articulation, bilabial, dental and palatal, but voiceless unaspirated stops at only two places of articulation, dental and velar.

2, *Wari'* (MacEachern, Kern, and Ladefoged 1997)

Wari' is spoken along the Pacaas Novas river in Western Rondonia, Brazil by about 1300 speakers (Everett and Kern in press). It is categorized as a Chapacuran, Madeira language and is described as an isolate (Grimes 1988). In addition to bilabial, dental and velar stops, there are also labialized velar stops.

3. *Tsou* (Wright and Ladefoged, 1994)

Tsou is an Austronesian language, spoken on Alishan in Central Taiwan by about 3000 people. Unaspirated stops occur at four places of articulation, bilabial, alveolar, velar, and glottal (Tung 1964, Wright and Ladefoged 1994). We do not have any data on the VOT after glottal stops.

4. *Defaka* (Shryock, Ladefoged and Williamson 1995)

Defaka is an Ijoid language belonging to the Atlantic-Congo group of the Niger-Congo family. Voiceless unaspirated stops occur at three places of articulation: bilabial, alveolar, and velar.

5, *(Western) Apache* (Potter, Dawson, de Reuse, and Ladefoged to appear)

Western Apache is an Athabaskan language, spoken by about 15,000 speakers who live on the San Carlos and White Mountain Reservations in Arizona (Potter et al,. to appear). Aspirated and unaspirated pulmonic stops occur at three places of articulations: bilabial, alveolar and velar. As the aspirated bilabial stops can be found in only a few words, many of which are clearly loan words, they are not included in this analysis.

6. *Navajo* (McDonough and Ladefoged 1993)

Navajo is an Athabaskan language, closely related to Apache. It is spoken by a large number of people on the Navajo reservation and in surrounding areas. It is not an endangered language, but it is included in this survey as it was investigated using the same protocol as was used for the endangered languages. Navajo stops can be divided into two groups: affricates and non-affricates. There are three places of articulation, at each of which aspirated and unaspirated stops occur along with glottalized (ejective) stops (which will be discussed separately).

7. *Khonoma Angami* (Blankenship, Ladefoged, Bhaskararao and Chase 1993)

Khonoma Angami is one of the smaller dialects of Angami, a Tibeto-Burman languages spoken in the Naga Hills in the northeastern parts of India (Blankenship et al., 1993). In the Khonoma dialect, there are unaspirated and aspirated bilabial, dental and velar stops.

8. *Toda* (Shalev, Ladefoged and Bhaskararao 1993)

Toda is a Dravidian language spoken by about 1,000 people in the Nilgiri Hills in Southern India (Shalev, Ladefoged and Bhaskararao 1993). In addition to bilabial and velar stops, there are three coronal stops: a dental stop made with the blade of the tongue contacting the teeth and the front part of the alveolar ridge, an alveolar stop made with the tip of the tongue contacting the center of the alveolar ridge, and a retroflex stop in which the underside of the tongue curls up and back to make post-alveolar contact. (Ladefoged and Maddieson 1996).

9. *Tiwi* (Anderson and Maddieson 1994)

Tiwi is a language spoken by the indigenous people of Bathhurst and Melville Islands in the Northern Territory of Australia (Anderson and Maddieson 1994). The data for this language are restricted to the four stops in the coronal region. The first two of these are similar to the laminal dental and apical alveolar stops in Toda described above. The Tiwi retroflex stop differs from that in Toda in that it is the tip rather than the underside of the tongue that is used in the articulation. The final stop in this region is a laminal palatal. We do not have data on the velar and labial stops.

10. *Dahalo* (Maddieson, Spajić, Sands and Ladefoged 1993)

Dahalo is a Cushitic language, one of the branches of the large Afro-Asiatic language family, spoken by a small population in the northern coastal area of Kenya. There are bilabial, dental, alveolar, and velar stops (as well as epiglottal and glottal stops which we will not consider here). The palatographic study reported by Maddieson et al. (1993) shows that the dental stop is laminal while the alveolar stop is apical.

11. *Jalapa Mazatec* (Silverman, Blankenship, Kirk and Ladefoged 1994)

Jalapa Mazatec is a language of the Popolocan branch of the Otomanguean language family (Gudschinsky 1958, Grimes 1988). It is spoken in Mexico by approximately 125,000 people in northeastern Oaxaca, in southern Puebla and in western Veracruz (Silverman, et al., 1994). The language is endangered in the sense that it is changing rapidly due to the influence of Spanish. Many distinctions are no longer made by younger speakers. Stops occur at three places of articulation: bilabial, dental and velar; there are also dental and palatal affricates not considered here. Bilabial stops occur only in loan-words.

12. Montana Salish (Flemming, Ladefoged, Thomason 1994)

Montana Salish is an Interior Salishan language, spoken on the Flathead reservation in Northwest Montana by approximately 70 speakers (Flemming et al. (1994). Excluding the affricates, Montana Salish stops occur at four places of articulation, bilabial, alveolar, velar and uvular. At each of these places there is a voiceless unaspirated stop (and an ejective, which we will discuss separately).

13. Eastern Aleut

14, Western Aleut (Cho, Ladefoged, Dirks and Taff, forthcoming)

Eastern Aleut is the language of the peoples of the Pribilof Islands, in the Bering Sea; Western Aleut is spoken in the nearby Aleutian Islands. The two languages might be considered simply as different dialects; they do, however, have substantial phonological and other differences. There are alveolar, velar and uvular stops in both forms of Aleut. Bilabial stops occur only in loan words.

15, Hupa (Gordon 1996)

Hupa is an Athabaskan language spoken by less than 100 speakers who live on the Hoopa Valley Indian reservation, near Eureka, northwest California. There are bilabial, denti-alveolar, velar and uvular unaspirated stops, but aspirated stops occur only in the denti-alveolar and the velar regions. There are also ejectives at the same places of articulation as the unaspirated stops.

16. Tlingit (Maddieson, Bessell, and Smith 1996)

Tlingit is the indigenous language spoken by the people in most of the islands of the Alexander archipelago and in contiguous parts of the North American mainland. The best estimate of the population is about 2000 at most (Krauss 1979, cited by Maddieson et al. 1996). Tlingit has alveolar, velar and uvular stops, but no labial stops. There are also glottalized (ejective) stops in the same places of articulation .

17. Yapese (Maddieson 1997)

Yapese is a language spoken by about about 8000 people in the island cluster of Yap in the Western Pacific. It is an Austronesian language, but its precise place within the family is quite controversial (e.g. Ross 1995, cited by Maddieson 1997). Yapese is specially interesting for the phoneticians becuase it has a large number of glottalized consonants in its inventory of sounds, including stops, fricatives, nasals, and contral and lateral approximants.

2.2 VOT data

The mean VOT (ms) of the stops in the 17 languages is shown in Table 3. When a language contrasts unaspirated and aspirated stops, the latter are shown in a second line. As noted in the description of the individual languages, Wari' has plain and labialized velar stops, the latter being shown as the second entry in the velar column. The data for Tiwi bilabial and velar stops are not available in our sources; the mean VOT of Tiwi laminal palatal stops is shown in the velar column.

The first point to note is that, with the exception of language 10, Dahalo, velar stops have the longest VOTs in all of the 12 languages that do not have contrasts between velar and uvular stops; and in the remaining five languages either velars or uvulars have the longest VOT. Even Dahalo follows this trend if we disregard the alveolar stops. At first glance, disregarding the unusual pattern of Dahalo VOT values may seem problematic. In post-hoc analyses of a one-factor ANOVA of Dahalo the alveolar stops were significantly different from all the others ($p < .01$). Mean VOT for the alveolar stops is 275 % and 157 % of that for the dental and the velar stops respectively. These differences cannot be accounted for by the aerodynamic characteristics that we have described so far. The volume of the cavity behind the constriction must be less for the alveolar than for the velar stops. Furthermore, the fact that a palatographic study showed that the dental stop is laminal while the alveolar stop is apical would lead to the expectation that the laminal dental would be produced with a longer VOT than that of the apical alveolar, as was the case in the articulatory and acoustic study of English, French, Malayalam, and 'O'odham by Dart (1991). Maddieson et al. (1993) note that Dahalo alveolar stops are distinguished from the other stops by their relatively long friction and aspiration duration, which serve as significant cues. It seems that Dahalo speakers have simply chosen to use a slower articulatory velocity for the alveolar stop than for the velars and as a result have a longer VOT for their alveolar stops. If we disregard the Dahalo alveolar stops because of their affrication, then it is true that velar or uvular stops always have the longest VOTs.

Table 3. Mean VOT (ms) of the stops in 17 languages studied in the UCLA endangered languages project .

	Language (speakers)	bilabial	dental	alveolar	retroflex	velar	uvular
1.	Banawa (5)		22			44	
2.	Wari' (10)	19	26			50 - 58	
3.	Tsou (13)	11		17		28	
4.	Defaka (6)	18		20		30	
5.	Western Apache (8)	13		15		31	
	Western Apache (aspirated)			58		80	
6.	Navajo (7)	12		6		45	
	Navajo (aspirated)			130		154	
7.	Khonoma Angami (6)	10	9			20	
	Khonoma Angami (aspirated)	83	55			91	
8.	Toda (13)	17	22	25	12	28	
9.	Tiwi (5)	--	20	25	7	36	
10.	Dahalo (3)	20	15	42		27	
11.	Jalapa Mazatec (6)			11		23	
	Jalapa Mazatec (aspirated)			63		80	
12.	Montana Salish (5)	22		24		48	55
13.	Eastern Aleut (13)			59		75	78
14.	Western Aleut (11)			76		95	92
15.	Hupa (3)	11	16			44	27
	Hupa (aspirated)		82			84	
16.	Tlingit (4)			18		28	30
	Tlingit (aspirated)		120			128	128
17.	Yapese	20	22			56	

Figure 4 presents the data in Table 3 in another way, grouping all the aspirated consonants together, and neglecting the Dahalo alveolar stops. Some of the differences are small and, in the cases in which the language names are parenthesized, are probably not significant (our sources do not always provide statistical analyses); but all the differences are in the same direction.

An interesting point that emerges from the data in Table 3 and Figure 4 is the similarity of VOT differences due to place of articulation in aspirated and unaspirated stops. The difference between unaspirated velar or uvular stops and unaspirated bilabial or coronal stops is very similar to the difference between aspirated velar or uvular stops and aspirated bilabial or coronal stops. The mean for the difference between the unaspirated stops is 20.9 ms and that for the aspirated stops is 18.2 ms. In four out of the 6 languages that have an aspirated — unaspirated contrast the difference is greater for the unaspirated stops, but two languages make greater place differences among the aspirated stops. Even Navajo, which has aspirated stops with an exceptionally long VOT, has longer VOTs for velar aspirated stops than for alveolar aspirated stops. From a physiological or aerodynamic point of view, there must be two different explanations for this coincidence. Any appeal to the aerodynamic conditions shortly after the release can apply to unaspirated stops, but not to aspirated stops (certainly not the Navajo stops); and any explanation that considers the special characteristics of aspirated stops cannot apply to unaspirated stops.

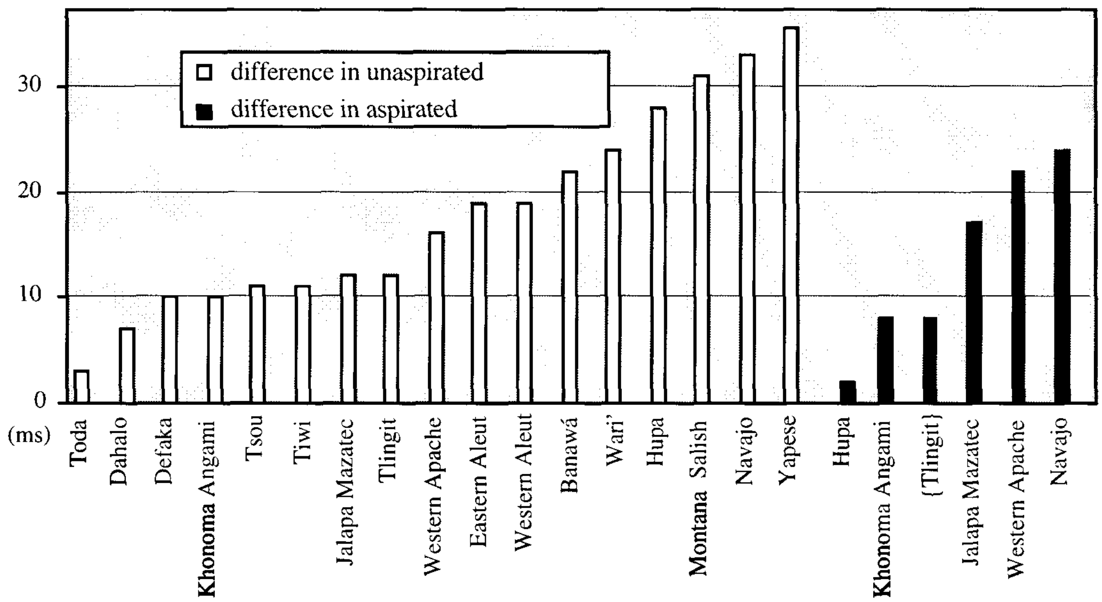


Figure 4. The difference (ms) between the VOT for the velar or uvular stop (whichever has the longest VOT) and the bilabial or coronal stop (whichever is the longer) in the 17 languages investigated.

Figures 5 and 6 show that the range of VOTs associated with dental stops overlaps with that of alveolar stops. The volume of air behind the closure is much the same in the laminal dentals and the apical alveolars, Accordingly, the laminal dentals (which have a more extended contact area) might have been expected to have a slower release, and hence a significantly longer VOT. But it seems that the length of the contact is not an important source of differences in VOT for the coronal stops in these languages.

In both of the languages that contrast dental, alveolar and retroflex stops, the VOT is significantly shorter for the retroflex stops ($p < .01$). It is not just the place of articulation, but also the movements of the articulators that are important in considering the aerodynamic conditions. Shalev, et al. (1993) pointed out that in the Toda retroflex stops the forward movement of the tongue tip and blade during the release draws air upward from the pharynx, which effectively lowers the supralaryngeal pressure. The same explanation applies in Tiwi.

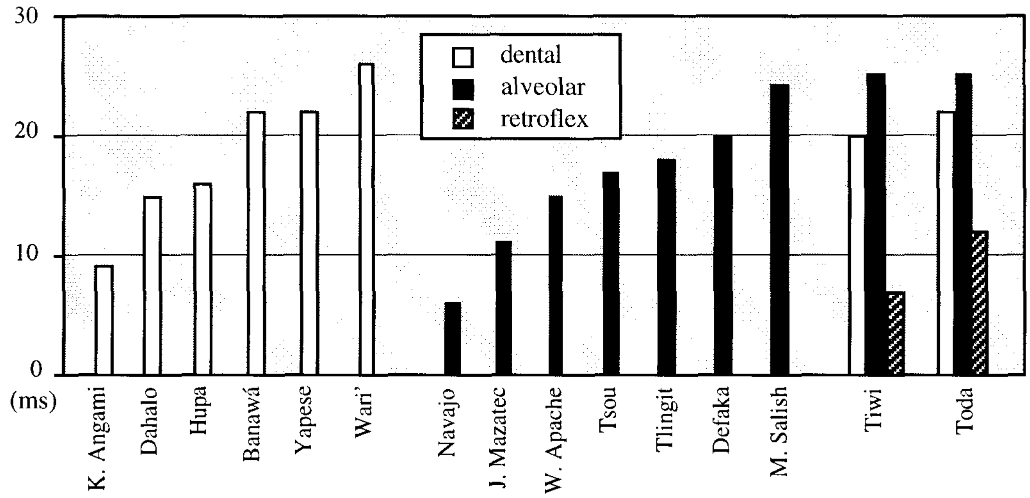


Figure 5. VOTs (ms) for the unaspirated coronal stops. (The Dahalo alveolar stops, which have anomalous VOTs, have been omitted.)

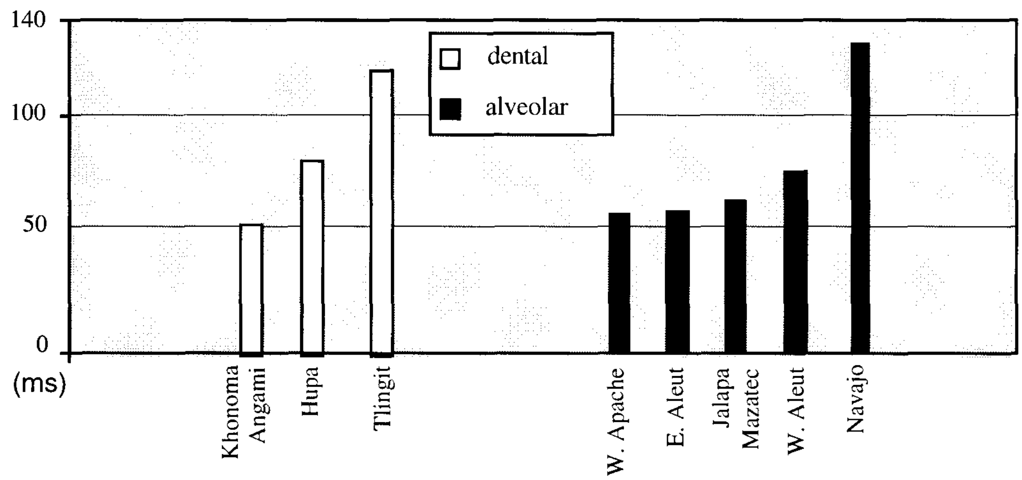


Figure 6. VOTs (ms) for the aspirated coronal stops.

We will now consider the relation between bilabial stops and coronal stops. Many of the languages investigated do not have bilabial stops, and accordingly we are left with only 10 languages to compare as shown in Figure 7. We have omitted the retroflex stops in this comparison because of the special aerodynamic conditions which apply to them. In three cases, Khonoma Angami, Navajo and Dahalo, the coronal stops have a shorter VOT, but in the other 8 languages there is some evidence for the prediction that the further back the articulation the longer the VOT. We do not have data for a full statistical comparison of the significance of these differences, but the mixed nature of these results shows that the situation is complicated and that more than the position of the constriction in the vocal tract has to be considered. Presumably the way in which the articulators are moved differs among these languages, and these movements affect the aerodynamic conditions in much the same way as in the case of the retroflex stops, but to a lesser extent. It is also possible that different languages have different contact lengths for alveolar stops that have been categorized as apical alveolar. Alveolar stops in languages such as Khonoma Angami or Navajo may be produced with a relatively short length of apical contact, which presumably induces a relatively rapid change in intraoral pressure immediately following the release. This assumption remains to be corroborated by palatograms and linguograms of languages in which VOT is consistently shorter for alveolar stops than for bilabial stops. It could be that there is an articulatory/aerodynamic explanation of the VOT variations in these languages; but it does not depend simply on the place of articulation. The movements of the articulator must also be specified. We should also note that it has been reported that in many languages the difference in VOT between bilabial and alveolar stops is not significant, or shows substantial overlapping (e.g., Cho 1996, Silva 1992, Abramson and Lisker 1971, Lisker and Abramson 1964).

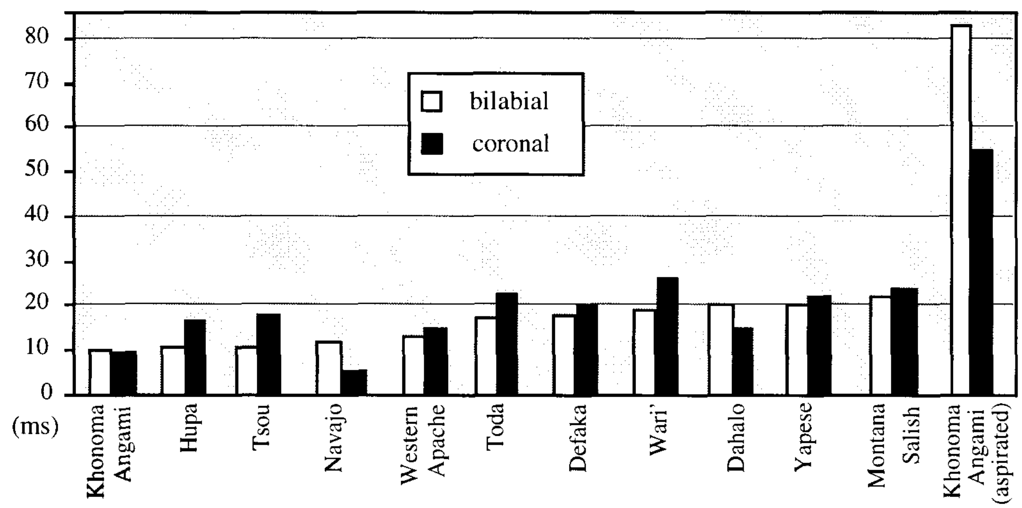


Figure 7. VOT (ms) for bilabial stops contrasted with dental or alveolar stops. In languages with both dental and alveolar stops, the one wth the shorter VOT is shown.

The differences in VOT between velars and uvulars in languages that contrast these two types of sounds are shown in Figure 8. There is little consistency in these data. The only statistically significant difference is between the velar and uvular unaspirated stops in Hupa, where VOT is shorter for uvular than for velar stops. In accounting for the inconsistent variations between velar and uvular stops, we suggest that although the volume of the cavity behind the constriction is smaller for uvulars than for velars, the uvular stop might be produced by a constriction with relatively shorter contact. The first of these two factors might result in a shorter VOT for velars, and the second in a shorter VOT for uvulars. The trade off between these two factors apparently varies from language to language.

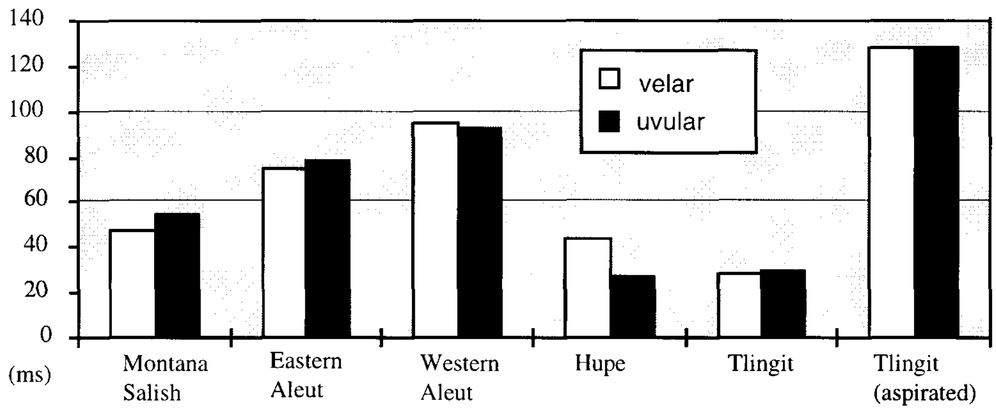


Figure 8. VOTs (ms) for velar and uvular stops.

2.5 Ejectives

The data for the VOTs in languages with ejectives are interesting, in that the aerodynamic conditions associated with these sounds are different from the conditions that occur in plosives. Ejectives are produced with a glottic airstream mechanism in which the air in pharyngeal cavity is compressed by the upward movement of the closed glottis (Ladefoged 1993:138). The same articulators are involved in the production of ejectives as in plosives, but at the time of the articulatory release the vocal folds are pressed tightly together and above their usual position, rather than being potentially in a position such that voicing might occur given an appropriate transglottal pressure drop. Table 4 shows a summary of the mean VOTs for the six languages in our data set that have eectives. Among these languages, only Yapese has regularly contrasting bilabial ejectives.

Table 4. Voice onset time (ms) for ejectives in six languages.

	# of speakers	bilabial	alveolar	velar	uvular
Montana Salish	5		65	86	81
Navajo	7		108	94	
Apache	8		46	60	
Yapese	3	60	64	78	
Tlingit	4		95	84	117
Hupa	3		93	80	89

The mean VOTs for velar ejectives are longer than those for the alveolar ejectives in three of these languages, Salish, Apache, and Yapese. In Navajo, Tlingit and Hupa, however, the mean VOTs for velars are shorter than those for alveolars; these three languages also did not have a significant difference between velar and coronal aspirated stops, although for these apirated stops the velars were slightly longer than the dentals.

VOTs in uvulars show irregular patterns: in Tlingit, VOT is longest for uvular ejectives, but in Salish and Hupa uvulars are intermediate between the other two stops. At the moment we do not know if there is a physiological mechanism that accounts for the effect of the place on VOT in the production of ejectives. The irregularities in VOTs found in these six languages may be due partly to a differences among the languages in the degree of the upward movement of the closed glottis. Another possibility is that there may be a different timing between the oral release and the glottal release in different languages. There are also major differences in the data for Apache and Navajo in Table 4. Stops in these languages have very different VOTs, although they are closely related languages belonging to the same language group, southern Athabaskan. Mean VOTs for alveolar and velar in Navajo are about 234% and 156% of those in Apache.

Unaspirated vs. aspirated stops

Languages differ in the values of VOT that they choose as the basic value for an unaspirated or an aspirated stop. Let us consider for the moment just the velar stops in these 17 languages. Figure 9 shows the complete set of values for both aspirated and unaspirated velar stops, a total of 22 mean values. It would be possible to draw an arbitrary line at, say, 50 ms, and suggest that this separates aspirated from unaspirated stops. But it is not at all clear that there are just two phonetic categories from which languages can choose. The data do not lend themselves to a statistical clumping procedure, but it would certainly be plausible to say that there are four phonetic categories, one around 30 ms representing unaspirated stops, another around 50 ms for slightly aspirated stops, a third for aspirated stops at around 90 ms, and a fourth for the highly aspirated stops of Tlingit and Navajo.

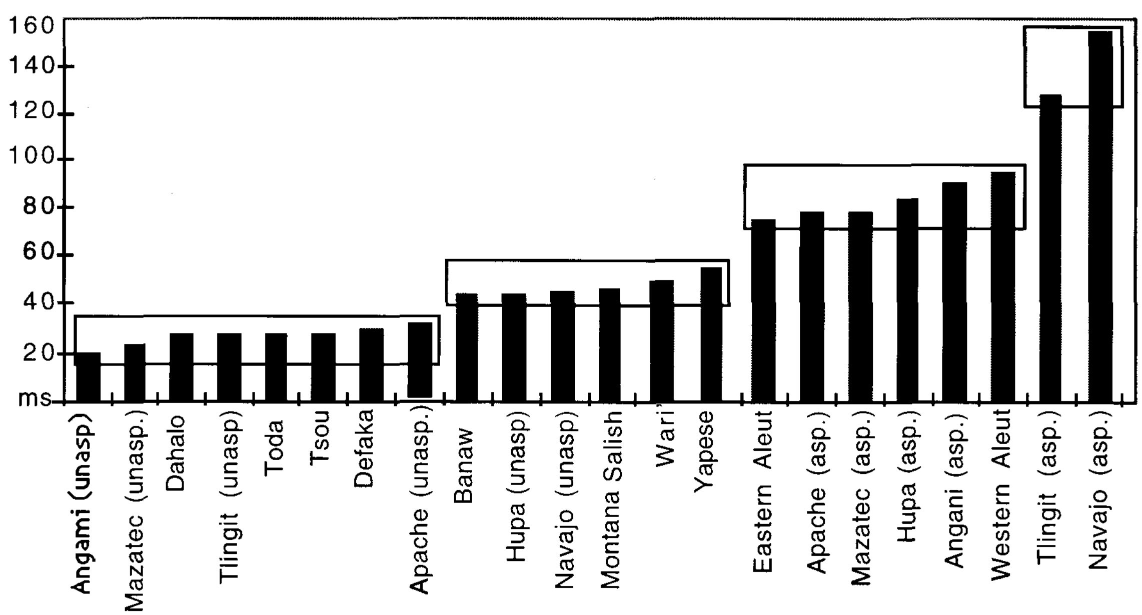


Figure 9. Mean VOTs (ms) for velar stops across languages. The rectangles enclose four regions, representing what might be called unaspirated stops, slightly aspirated stops, aspirated stops and highly aspirated stops

There does not seem to be any phonological reason why there might be four groups as suggested. They do not reflect differences dependent on the number of contrasts in voicing that each language has. Banawa, for example, has only a single velar stop, with no contrast in voicing; the mean VOT for this stop is 44 ms, placing it in the second group. But Western and Eastern Aleut also have only one velar stop; their mean values are 78 and 95 ms, making them fully aspirated stops. Similarly it does not matter whether a language contrasts voiceless unaspirated stops with aspirated stops. Both Angami and Hupa make these contrasts. But the Angami voiceless unaspirated stops have much shorter VOTs than their Hupa counterparts, so that they appear in different groups in Figure 9. The Angami aspirated stops are in the same group as their Hupa counterparts, but have slightly longer VOTs.

Another interesting point is that the highly aspirated stops of Tlingit and Navajo, which are in the fourth group in the figure, have glottalized as well as unaspirated counterparts. Table 5 shows mean values for languages with more than two types of stops. VOT values for glottalized velar stops for these two languages fall in between the VOT values of unaspirated and aspirated stops. Our conjecture is that the languages with more than two types of stops tends to enhance the contrastiveness among stops by dispersing VOT values along the VOT continuum. VOT patterns for stops in Apache is also in the same direction, although the values are relatively closer together. A similar pattern can also be found in Korean which also has three types of stops: unaspirated, aspirated, and tense (or fortis). Mean values for unaspirated and aspirated velar stops in Korean reported by Lisker and Abramson (1964) are 47 ms and 126 ms, respectively (see Tables 1 and 2 above). Tense stops in general have been reported to have around 25 ms (Cho 1994). However, not all languages follow this tendency; Hupa does not show this pattern. Thus, languages may employ different strategies for contrasting three-way distinction of stops.

Table 5. VOT (ms) of stops in languages with more than two stop categories. Values for Korean are from Lisker and Abramson (1964) and Cho (1994)

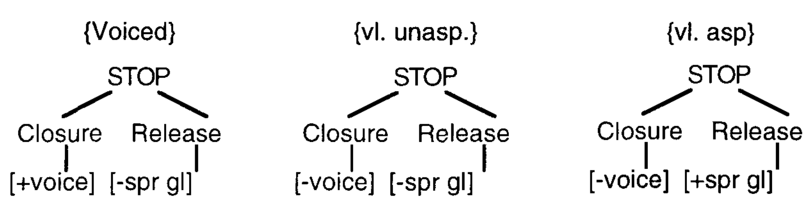
	tense **k***	unaspirated **k**	glottalized **k'**	aspirated **k^h^**
Navajo	-	45	94	154
Tlingit	-	28	84	128
Apache	-	31	60	80
Hupa		44	80	84
Korean	25	47	-	126

Discussion

From the data we have presented it appears that we might be able to account for nearly all the VOT differences due to place of articulation within a language — if we only knew enough about the exact articulatory movements involved. There is no real evidence that speakers try to produce different VOTs for different places of articulation. It is more likely that speakers aim for a certain timing difference between articulatory and glottal gestures irrespective of the articulatory gesture involved. The observed VOT is just the inevitable consequence of the physiological movements and the aerodynamic forces. Given enough knowledge about the gestures involved, the differences due to so-called place of articulation may be as determined as the formant frequency changes that occur with particular vocalic gestures.

It also appears that there is a continuum of possible VOTs from which languages may choose. The values chosen by a language reflect the complex interplay of all the various constraints — articulatory ease, auditory distinctiveness, gestural economy, and sociolinguistic pressures — that occur in the development of sound systems (Maddieson 1997, Ladefoged, in press). These forces produce similar VOT patterns in a wide variety of languages. But there is still a great deal of variation between languages that has to be taken into account in the phonetic description — the phonetic interpretation rules — required for each language.

In trying to account for what was known about VOT in different languages at the time, Keating (1984:289) proposed a model in which there are "only as many phonetic categories given by the phonetic features as there are contrasting phonetic types in languages." As necessary evidence, she shows that in order to achieve not only phonological generalization but also the contrasting phonetic differences between languages such as English and Polish stops /p, t, k, b, d, g/, there are two different levels of representations in the grammar. At the first level, various phonetic kinds of /b, d, g/ are defined by the feature [+voice] in both languages. At the second level, the phonetic features further distinguish stops in English from those in Polish by the use of three phonetic categories {voiced}, {voiceless unaspirated} and {voiceless aspirated}. In Polish, as in other languages without aspiration such as French, the phonological features [+voice] and [-voice] are realized as {voiced} and {voiceless unaspirated} respectively, whereas the phonological feature [+voice] in English is usually realized as {voiced}, but can be sometime realized as {voiceless unaspirated} (e.g. word-initially); similarly English [-voice] can be either {voiceless unaspirated} or {voiceless aspirated}, depending on the context. Keating notes that the implementation of the phonologically identical feature Voice is different in different languages, but the categories are chosen from a "fixed and universally specified set" which allows only three discrete phonetic categories {voice},{voiceless unaspirated}, and {voiceless aspirated} without "fuzzy areas of the continuum." In Keating (1990), these three discrete phonetic categories are represented under Aperture Theory (cf. Steriade 1989; 1993):



Keating's approach has many similarities with that of Ladefoged and Maddieson (1996; see also Ladefoged 1997), and that proposed here. We differ from Keating in that we consider what might appear to be phonetic categories as at best modal values within the continua formed by the physical scales — the parameters — that define each feature. It is not at all clear that there are three, and only three, distinct modal values of VOT. As we saw in Figure 9, the aspirated stops in Navajo and Tlingit have values that are very different from those in other languages, and the unaspirated stops in Hupa and Angami differ considerably.

We suggest that there is a feature VOT, definable in terms of the difference in time between the initiation of the articulatory gesture and the initiation of the laryngeal gesture. This is a somewhat different definition of VOT than the traditional definition that has been used throughout this paper, in which VOT is considered to be the interval between the release of an articulatory gesture, usually (always in this paper) a stop, and the beginning of vocal fold vibration. If we redefine VOT in terms of the interval between the gestures involved, then the values become largely unmeasurable without involving the theory of articulatory phonology as described by Browman and Goldstein (1990, 1992). This theory regards gestures as being realized by a task dynamic model (Saltzman 1986; Saltzman and Munhall 1989; see also Hawkins 1992) that would, when fully worked out, take care of the physiological and aerodynamic influences on voicing lag that we have been discussing. The data we have been discussing seem fully compatible with this possibility.

The modal values of VOT (however defined) might well be [voiced], [voiceless unaspirated], and [voiceless aspirated]. As languages never contrast more than three values of VOT, these three values are appropriate for making general statements. But our data show that even if we could measure for each language the VOT in terms of the difference in time between the initiation of the articulatory gesture and the initiation of the laryngeal gesture, there are still fuzzy areas within the VOT continuum. The measured VOTs in Figure 9 are all for virtually the same articulatory gesture, and should therefore reflect comparable intervals between the initiation of the gestures. Nevertheless they show unpredictable variations between languages. In a full description of a language there must be a phonetic interpretation process, in which particular values of the features are assigned for each language. Modal values are not enough.

References

Abramson, A. and L. Lisker. 1971. Voice timing in Korean stops. *Proceedings of The Seventh International Congress of Phonetic Sciences.* 439-446.

Abramson, A. and L. Lisker. 1972. Word-final stops in Thai. *Thai phonetics and phonology,* ed. by Jimmy G. Harris & Richard B. Noss, 1-7. Bangkok: Central Institute of English language.

Anderson, B. B. and I. Maddieson. 1994. Acoustic characteristics of Tiwi coronal stops.*UCLA Working Papers in Phonetics* 87: 131-162.

van den Berg, Jw. 1958. Myoelastic theory of voice production. *J. Speech Hear. Res.* 1: 227-244.

Blankenship, B., P. Ladefoged, P. Bhaskararao and N. Chase. 1993. *UCLA Working Papers in Phonetics* 84: 127-142.

Browman, C. P. and L. Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. Beckman (eds), *Papaers in Laboratory Phonology I: Between the Grammar and Physics of Speech,* Cambridge University Press, pp. 341-376.

Browman, C. P. and L. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica 49*: 155-180.

Chen, M. 1970. Vowel length variation as a function of the voicing of consonant environment. *Phonetica* 22:129-159.

Cho, T. 1996. *Vowel correlates to consonant phonation: An acoustic-perceptual study of Korean obstruents.* The University of Texas at Arlington M.A. Thesis.

Cho, T., P. Ladefoged, M. Dirks and A. Taff (forthcoming) Phonetic structures of Aleut.

Chomsky, N. and M. Halle. 1968. *The sound pattern of English.* Harper & Row, New York.

Cohn, A. 1993. Nasalization in English: phonology or phonetics. *Phonology* 10:43-81.

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

Everett, D. L., and B. Kern. In Press. Wari' (The Pacaas Novos languages of Western Brazil). London: Routledge Publishers, Descriptive Grammar Series.

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

Flemming, E., P. Ladefoged, and S. Thomason. 1994. Phonetic structures of Montana Salish. UCLA Working Papers in Phonetics 87: 1-34.

Grimes, B. F. 1988. Ethnologue languages of the world. Dallas, Texas: Summer Institute of Linguistics, Inc.

Gordon, M. 1996. The phonetic structures of Hupa. UCLA Working Papers in Phonetics 93:164-187.

Halle, M. and K.N. Stevens. 1967. On the mechanism of glottal vibration for vowels and consonants. MIT QPR 85:267-271.

Hardcastle, W.J. 1973. Some observations on the Tense-Lax distinction in initial stops in Korean. Journal of Phonetics 1:263-271

Hawkins, S. 1992. An introduction to task dynamics. In G.J. Docherty and D.R. Ladd (eds.), *Papers in Laboratory Phonology II: Gestures, Segments, Prosody,* Cambridge University Press, pp. 9-25.

Keating, P.A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60: 286-319.

Keating, P.A. 1985. Universal phonetics and the organization of grammars. In *Phonetics linguistics: essays in honor of Peter Ladefoged* (V. Formkin, editor):115-132. Orlando: Academic Press.

Keating, P.A. 1990. Phonetic representations in a generative grammar. *Journal of Phonetics 18:* 321-334.

Kuehn, D.P. and K. Moll 1976. A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics* 4:303-320.

Ladefoged, P. 1993. *A course in phonetics.* 3rd edn. Harcourt Brace College Publishers, Fort Worth.

Ladefoged, P. 1997. Linguistic phonetic descriptions. In *The Handbook of the Phonetic Sciences,* (W. Hardcastle and J, Laver, eds.), Blackwells, Oxford.

Ladefoged, P. In press. The sounds of languages. In *Sound* (the 1997 Darwin Lecture Series), Cambridge University Press, Cambridge.

Ladefoged, P. and I. Maddieson. 1996. *Sounds of the World's Languages.* Blackwells, Oxford.

Ladefoged, P., J. Ladefoged and D. Everett. 1997. Phonetic structures of Banawá, an endangered languge. *Phonetica* (in press).

Lehiste I.. 1970. *Suprasegmentals.* M.I.T. Press, Cambridge, Massachusetts.

Lindblom, B. 1967. Vowel duration and a model of lip mandible coordination. *Royal Institute of Technology, Speech Transmission Laboratory, Stockholm. QPRS* 4/1967.

Lisker, L. 1974. On 'explaining' vowel duration variation. *Glossa* 8: 2:233-246.

Lisker, L. and A. Abramson. 1964. Cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20:384-422.

MacEacheern, M.R, B. Kern and P. Ladefoged. 1997. Wari' phonetic structures. *Journal of Amazonian Languages* 1: 5-25.

McDonough J. and P. Ladefoged. 1993. Navajo Stops. *UCLA Working Papers in Phonetics* 84: 151-164.

Maddieson, I. 1997. Phonetic Universals. In the *Handbook of Phonetic Sciences* , ed by John Laver and William J. Hardcastle, Blackwells Publishers, Oxford: 619-639.

Maddieson, I. and J. Gandour. 1977. Vowel length before aspirated consonants. *Indiana Linguistics* 38:6-11.

Maddieson, I, S. Spajić, B. Sands and P. Ladefoged. 1993. Phonetic structures of Dahalo. *UCLA Working Papers in Phonetics* 84: 25-65.

Maddieson, I. and V. B. Anderson. 1994. Phonetic structures of Iaai. *UCLA Working Papers in Phonetics* 87: 163-182.

Maddieson, I., N. Bessel, and C. Smith. 1996. A preliminary phonetic report on the sounds of Tlingit. *UCLA Working Papers in Phonetics* 93:125-148.

Ohala, J.J. and C. Riordan. 1979. Passive vocal tracct enlargement during voiced stops. In *Speech Communication Papers,* ed by J.J. Wolf and D.H. Klatt. Acoustical Society of America, New York: 89-92.

Peterson, G.E. and I. Lehiste. 1960. Duration of syllable nuclei in English, *J. Acoust. Soc. Amer.* 32:693-703

Pierrehumbert, J. B. 1980. The phonology and phonetics of English intonation. MIT dissertation.

Pierrehumbert, J. 1990. Phonological and phonetic representation. *Journal of Phonetics 18:* 375-394.

Ross, M. 1995. Is Yapese Oceanic? Paper presented to the Second International Conference on Oceanic Linguistics, Suva, Fiji.

Saltzman, E. 1986. Task dynamic coordination of the speech articulators: a preliminary model. In H. Heuer and C. Fromm (eds.), *Generation and Modulation of Action Patterns* (Experimental Brain research Series 15). New York: Springer, pp. 129-144.

Saltzman, E. and K.G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1:333-82.

Shalev, M., P. Ladefoged and P. Bhaskararao. 1993. Phonetics of Toda. *UCLA Working Papers in Phonetics* 84: 89-126.

Shryock, A., P. Ladefoged and K. Williamson. 1995. Phonetic structures of Defaka.*UCLA Working Papers in Phonetics* 91: 89-109.

Silva, J. D. 1992. *The phonetics and phonology of stop lenition in Korean.* Cornell University Ph.D. dissertation.

Silverman, D., B. Blankenship, P. Kirk and P. Ladefoged. 1994. Phonetic structures in Jalapa Mazatec. *UCLA Working Papers in Phonetics* 87: 113-130.

Steriade, D. 1989. Affricates and release. Paper presented at the conference on features, MIT. October.

Steriade, D. 1993. Closure, release, and nasal contours. *Phonetics and Phonology,* vol 5:401- 470.
Stevens, K.N. In press. *Acoustic Phonetics.* Cambridge: MIT Press.
Stevens, K.N., S.J. Keyser and H. Kawasaki. 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.
Tung, T. 1964. *A Descriptive Study of the Tsou Language.* Taipei, Institute of History and Philology, Academia Sinica Special Publication.
Umeda, N. 1977. Consonant duration in American English. *Journal of the Acoustical Society of America* 61: 846-858.
Vatikotis-Bateson, E. and D.J. Ostry. 1995 An analysis of the dimentionality of jaw motion in speech. *Journal of Phonetics* 23: 101-117.
Weismer, G. 1980. Control of the coicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations. *Journal of Phonetics* 8: 427-438.
Wright, R. and P. Ladefoged. 1994. A phonetic study of Tsou. *UCLA Working Papers in Phonetics* 87: 67-92.

The Phonetic Structures of Chickasaw

Matthew Gordon, Pamela Munro and Peter Ladefoged

1. Introduction.

Chickasaw is an American Indian language spoken mainly in the Chickasaw nation of south-central Oklahoma by no more than a few hundred speakers. The Chickasaw speaking area of Oklahoma is indicated in Figure 1. Originally residents of the southeastern United States, the Chickasaw were one of the "Five Civilized Tribes" (along with Choctaw, Creek, Seminole, and Cherokee) forced during the 1830s into the present state of Oklahoma. Chickasaw belongs to the western branch of the Muskogean language family, along with closely related Choctaw. Although the two languages differ markedly in certain respects, Chickasaw and Choctaw share many features (Munro 1987). Perhaps because they are familiar with the Choctaw Bible and Hymnal, most Chickasaw speakers have some knowledge of Choctaw (Munro 1996).

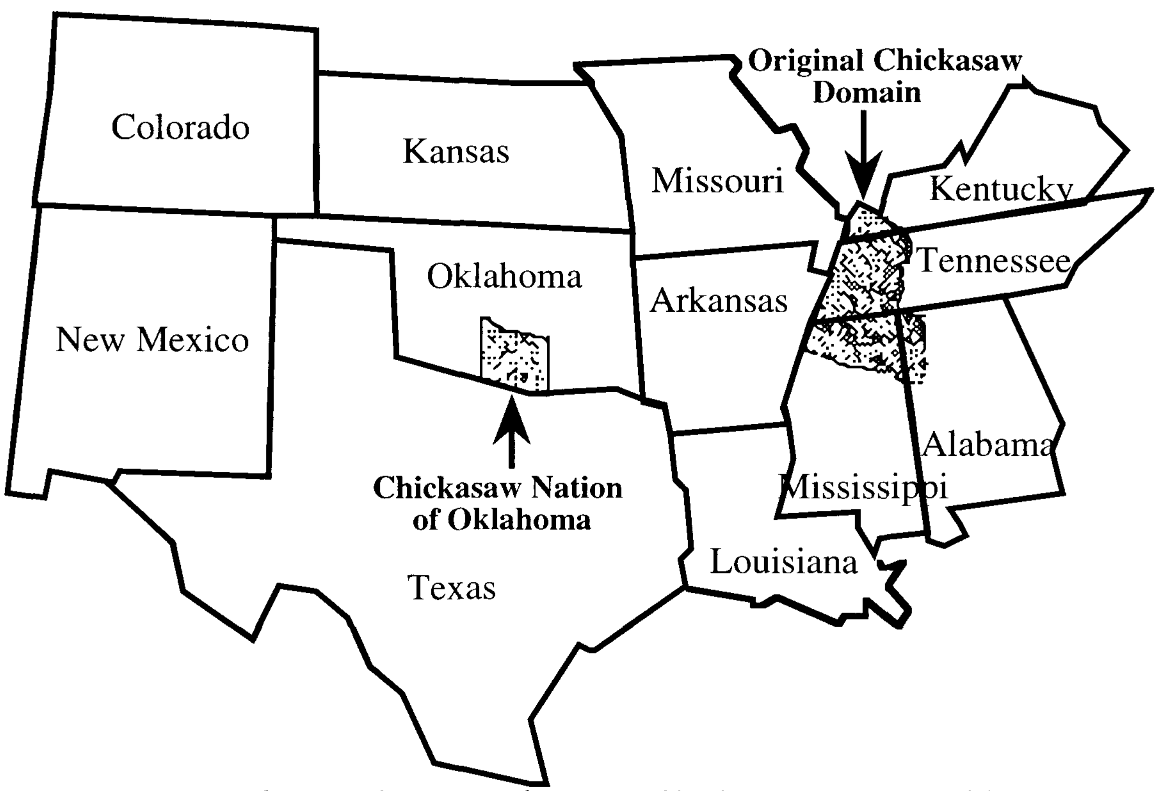


Figure 1: Map indicating location of current Chickasaw nation and location of original Chickasaw homeland (based on map in Gibson 1971) prior to relocation in Oklahoma

Chickasaw has been the subject of less linguistic research than Choctaw. Munro (in press) provides a grammatical overview of Chickasaw and includes an analyzed text of a traditional Chickasaw story. Munro and Willmond (1994) is a dictionary that also contains a thorough description of Chickasaw grammar. These works also cite other literature on Chickasaw (and Choctaw) that the interested reader is urged to consult. This paper relies on Munro and Willmond (1994), Munro (1996) and Munro (in press) for phonological descriptions as well as many qualitative phonetic observations.

The present paper represents the first instrumental study of Chickasaw phonetics. We provide a detailed description of the phonetic properties of Chickasaw vowels and consonants, in order that linguists, as well as speakers of Chickasaw and those who wish to learn the language may gain a better understanding of Chickasaw. Basic acoustic properties such as voice onset time, closure duration and vowel formants are discussed, along with typologically more unusual aspects of Chickasaw, such as the pervasive length contrast in consonants, and acoustic correlates of the three-way distinction in vowel length.

2. The present study

The present study is based on a word list of approximately 150 words that was designed to illustrate the principal phonetic features of Chickasaw. This word list was checked and recorded by one female speaker, F2, in Los Angeles. The remainder of the fieldwork was conducted in Oklahoma, where the word list was elicited from a further thirteen speakers, seven women and six men. Recordings were made in the speakers' homes, using a headmounted noise-canceling microphone, which ensured an approximately 45 dB signal to noise ratio. Speakers repeated each word twice while recordings were made on a Sony Digital Audio Tape Recorder at a sampling rate of 48 kHz. Upon return to the UCLA Phonetics Laboratory, the data that were to be used for acoustic analyses was transferred to the Kay Computerized Speech Lab and down sampled to 8 kHz so as to provide greater accuracy for analyses in the frequency domain (cf. Ladefoged 1996:177). Data from one of the male speakers could not be used; thus, this paper presents data from eight female and five male speakers.

3. Vowels

Chickasaw has three different vowel qualities, which occur short, long and nasalized. The nasalized vowels behave as long vowels, both phonetically in terms of duration and phonologically with respect to rules that are sensitive to syllable weight (see Munro and Willmond 1994 for discussion). The Chickasaw vowels are represented in Table 1 in a broad transcription. The phonetic realization of vowel quality is discussed later.

Table 1: Chickasaw vowels

Short	Long	Nasalized
i	iː	ĩː
o	oː	õː
a	aː	ãː

In addition to the vowels in Table 1, a third set of vowels arises from a phonological process of phonetic lengthening in alternate open syllables. This process of "rhythmic lengthening" lengthens the second in a series of two consecutive phonemically short vowels in open syllables that are not word final (for further discussion, see Munro and Ulrich 1984, Munro and Willmond 1994 and Munro et al. in preparation). Henceforth, the rhythmically lengthened vowels will be denoted in this paper with a half-long symbol; e.g., [**aˑ**] in order to differentiate them from the phonemic long vowels such as [**aː**].

3.1. Vowel quality

The frequencies of the first three formants for short, rhythmically lengthened and long vowels were measured using an LPC display calculated over a 30 millisecond window in the center of the vowel superimposed over an FFT spectrum calculated over a 25.6 millisecond window. In most cases, the LPC window was calculated using 14 coefficients, though more coefficients were sometimes used where resolution of the formants was particularly difficult. The targeted vowels all appeared in the third syllable of words with four syllables. In all cases, the target vowel was surrounded on both sides by /**l**/. Two tokens were measured for each speaker. The examined words appear in Table 2.

Table 2: Words used to examine the quality of Chickasaw vowels.

V	Short	Rhythmically Lengthened	Long
a	**iʃolːali**	**iʃtalaˑli**	**iʃtalaːli**
	'you laugh'	'I bring it here'	'you set it upright'
i	**impihlili**	**toksaliˑli**	**hatʃĩːliːli**
	'I sweep for him'	'I work'	'he hoes for you all'
o	**tʃihagloli**	**ĩːholoˑli**	**hatʃĩːloːloʔ**
	'I listen to you'	'I put on (shoes) for him'	'your (pl) doodlebug'

Average formant values for all nine vowels were taken for the fourteen speakers. Values appear in Table 3. Speakers are indicated in the top left hand corner of each box delimited by thick lines by a combination of a letter that represents the sex of the speaker (F for female, and M for male) and a number. Formant values for the first three formants are listed in columns.

Mean formant values for the first two formants are plotted in Figure 2 for the female speakers. Ellipses indicate two standard deviations from the mean, which is plotted as a larger version of the vowel that it represents. Because Fisher PLSD tests showed no significant differences between the lengthened and long vowels for F1 or F2 for the female speakers, they are collapsed as long vowels in Figure 2.

Table 3: Formant values for Chickasaw short, lengthened and long vowels from 13 speakers

Speaker &Vowel	Length	F1	F2	F3	Speaker &Vowel	Length	F1	F2	F3
F1					**F8**				
	Short	448	1993	2606		Short	476	2296	3138
i	Length	427	2406	2937	**i**	Length	413	2662	3069
	Long	399	2517	3000		Long	400	2358	3027
	Short	696	1545	2703		Short	676	1848	2813
a	Length	883	1427	2683	**a**	Length	869	1579	2641
	Long	820	1345	2676		Long	889	1441	2697
	Short	469	1531	2062		Short	496	1193	2351
o	Length	600	1344	2062	**o**	Length	607	1269	2813
	Long	627	1276	1938		Long	679	1289	2827
F2					**M1**				
	Short	462	2241	3152		Short	427	1745	2475
i	Length	448	2613	3179	**i**	Length	420	1848	2331
	Long	413	2738	3131		Long	324	1882	2186
	Short	537	1855	2807		Short	620	1448	2579
a	Length	882	1814	3614	**a**	Length	655	1344	2455
	Long	889	1931	3606		Long	669	1275	2386
	Short	386	1425	2938		Short	489	1200	2220
o	Length	448	1041	3007	**o**	Length	427	938	2344
	Long	545	1195	3290		Long	427	1055	2331
F3					**M2**				
	Short	351	1848	2379		Short	413	1883	2331
i	Length	379	2048	2200	**i**	Length	434	1965	2345
	Long	379	2241	3152		Long	386	2006	2407
	Short	434	1317	2241		Short	710	1524	2317
a	Length	662	1268	2206	**a**	Length	620	1420	2537
	Long	572	1227	2103		Long	641	1317	2517
	Short	427	1379	2193		Short	482	1289	2110
o	Length	393	1069	2220	**o**	Length	496	1048	2848
	Long	400	1021	2262		Long	413	1062	2124
F4					**M3**				
	Short	427	1999	2275		Short	338	1834	2179
i	Length	372	2455	2827	**i**	Length	331	1820	2276
	Long	317	2469	2821		Long	331	1896	2110
	Short	565	1572	2414		Short	565	1503	2207
a	Length	648	1289	2883	**a**	Length	655	1365	2489
	Long	751	1200	2924		Long	662	1296	2351
	Short	434	1131	2241		Short	413	1324	1903
o	Length	407	1014	2483	**o**	Length	365	875	2200
	Long	400	1028	2200		Long	427	875	2241

Table 3: Formant values for Chickasaw short, lengthened and long vowels from 13 speakers (continued)

F5					**M4**				
	Short	393	2448	2807		Short	399	1931	2531
i	Length	413	2579	2986	**i**	Length	455	1917	2496
	Long	441	2600	2848		Long	393	2000	2379
	Short	696	1903	2717		Short	538	1489	2551
a	Length	855	1496	2745	**a**	Length	648	1386	2586
	Long	875	1476	2772		Long	648	1386	2634
	Short	469	1696	2503		Short	455	1413	2386
o	Length	434	993	2717	**o**	Length	469	1062	2475
	Long	427	951	2648		Long	448	1028	2441
F6					**M5**				
	Short	503	2207	2758		Short	427	2007	2572
i	Length	510	2234	2869	**i**	Length	413	2248	2724
	Long	489	2151	2903		Long	406	2262	2565
	Short	765	1710	2917		Short	503	1538	2606
a	Length	896	1627	2765	**a**	Length	800	1462	2800
	Long	738	1634	2131		Long	759	1400	2710
	Short	572	1365	2496		Short	462	1345	2400
o	Length	565	1510	2675	**o**	Length	434	1034	2469
	Long	545	1255	2806		Long	393	1069	2317
F7									
	Short	365	1972	2827					
i	Length	393	1744	2676					
	Long	310	2200	2896					
	Short	628	1538	2669					
a	Length	875	1420	2724					
	Long	931	1393	2249					
	Short	448	1234	2483					
o	Length	545	1097	2289					
	Long	621	1131	2207					

The female speakers generally show fairly large variability for any single vowel as reflected in the relatively large ellipses in Figure 2. Examination of formant values for individual female speakers in Table 3 shows that this greater variability is a function of both greater variability between tokens for a single speaker and differences between speakers.

The most robust pattern apparent is for the short vowels to be centralized relative to their lengthened and long counterparts. This centralization is most apparent for the high vowels in the F2 (front-back) dimension, where short vowels have significantly different F2 values from both their phonemically long and lengthened counterparts ($p<.01$). In the case of the low vowel, the centralization is a matter of vowel height, which is reflected in the significantly decreased F1 of the short vowels ($p<.0001$). On an individual level, virtually all speakers display centralization of at least one of the two vowels /**a**, **o**/, and typically centralize both. Most speakers also centralize short /**i**/ in the front-back dimension. Speaker F6 was atypical in that none of her short vowels were substantially centralized relative to their lengthened or long counterparts.

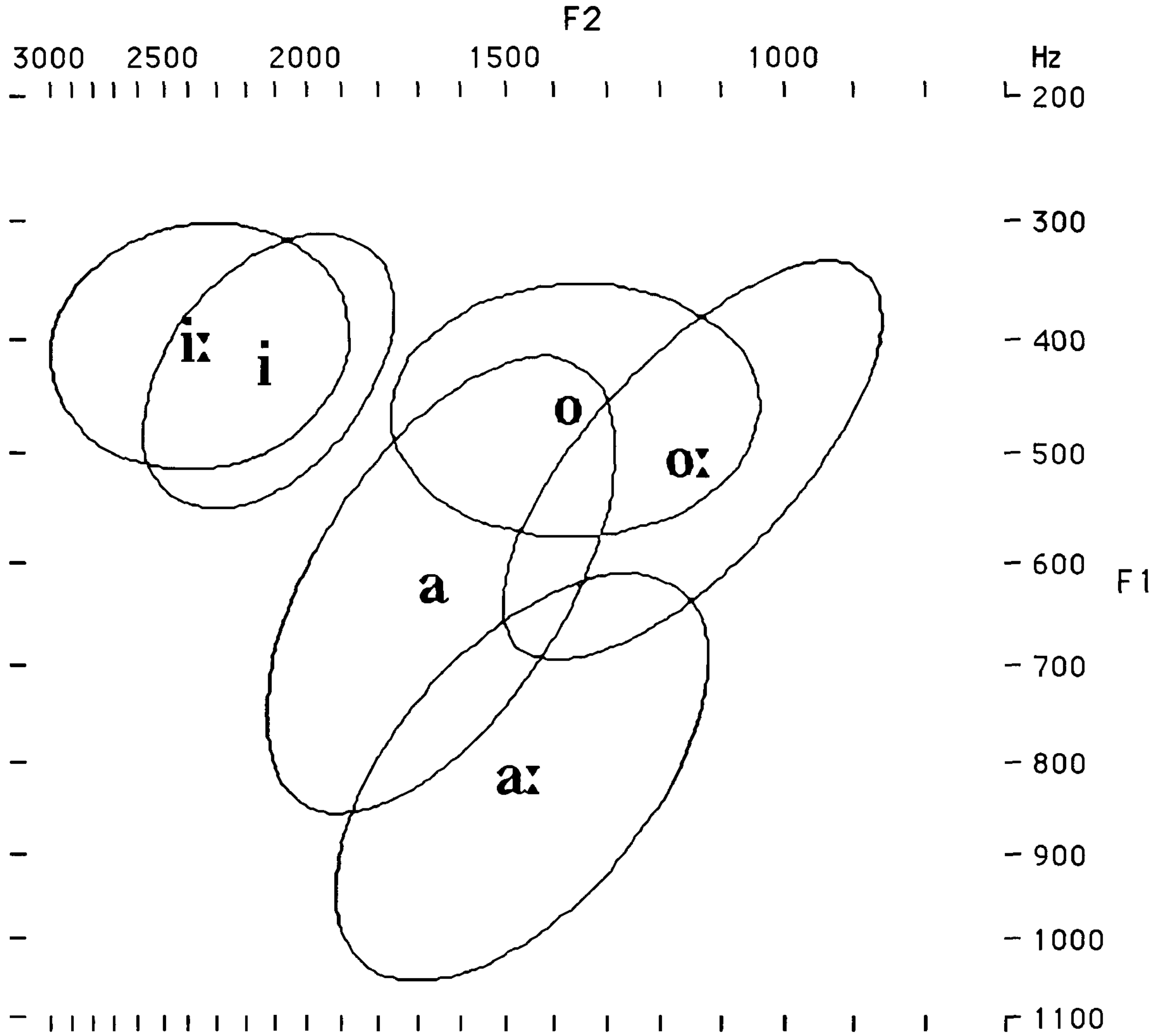


Figure 2: Plot of the first two formants for eight female speakers (phonemically long and lengthened vowels are collapsed)

Values for the first two formants for phonemically short and long vowels for male speakers are plotted in Figure 3. The male speakers differ from the female speakers, in that phonemically long /**a**ː/ and /**i**ː/ are not qualitatively neutralized with their lengthened counterparts. Only long and lengthened /**o**/ are neutralized in terms of formants one and two. Consequently phonetically lengthened /**a**ˑ/ and /**i**ˑ/ are omitted from Figure 3 and are shown in comparison with their phonemically long counterparts in Figure 4.

The male speakers show a similar tendency for the short vowels to be centralized relative to their long and lengthened counterparts. The long and short high vowels differ in the F2 (front-back) dimension (p<.01), and the low vowels differ in both the F1 (height) dimension, and the F2 (front-back) dimension (p<.01).

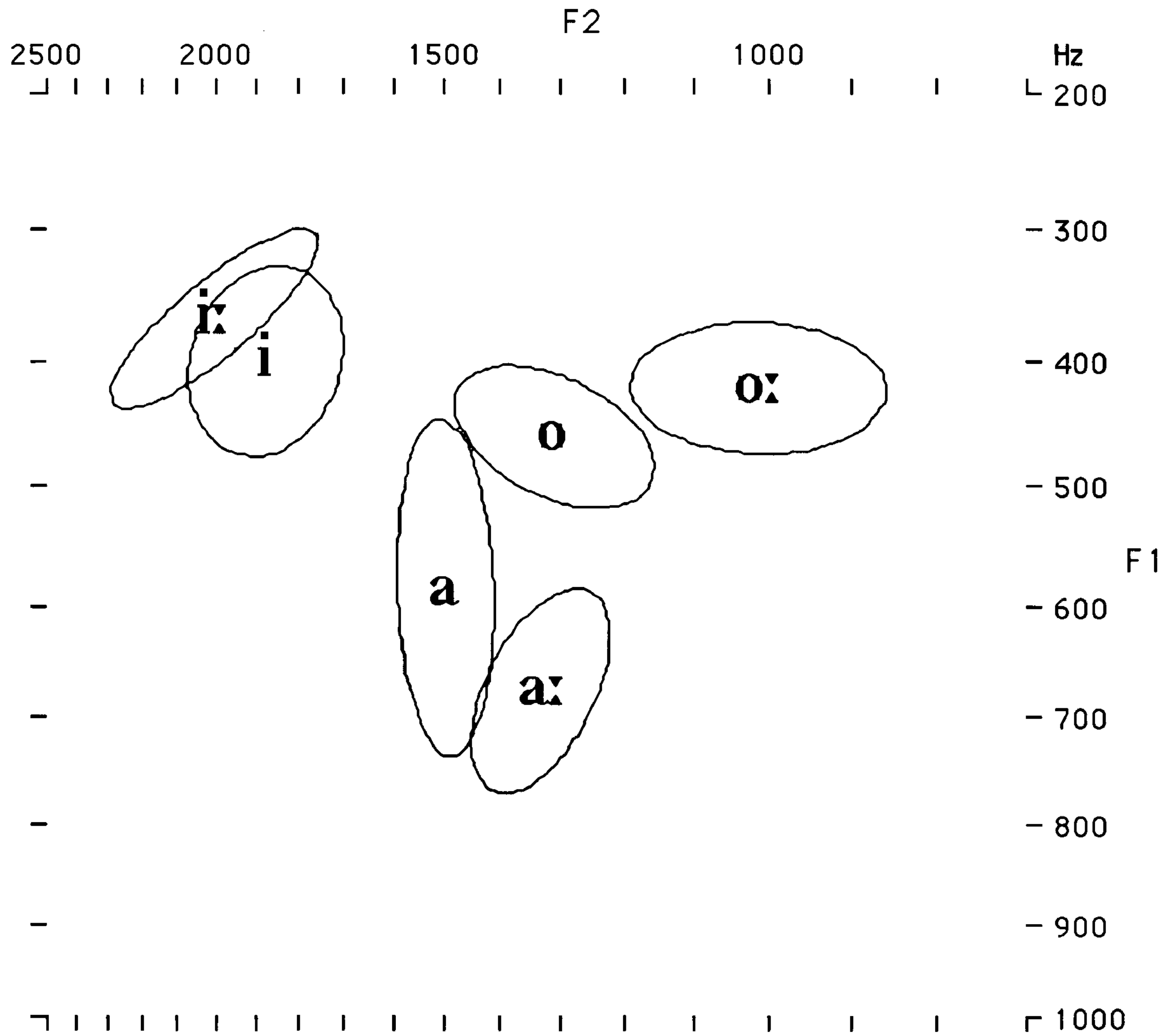


Figure 3: Plot of the first two formants for five male speakers (phonemically short vs. phonemically long vowels)

Looking at Figure 4, F2 for lengthened /**a**ˑ/ is significantly higher than for phonemically long /**a**ː/, and F1 for lengthened /**i**ˑ/ is significantly higher than for phonemically long /**i**ː/. Thus, the lengthened variants of /**i**/ and /**a**/ tend to be somewhat intermediate in quality between phonemically short and phonemically long vowels for the male speakers.

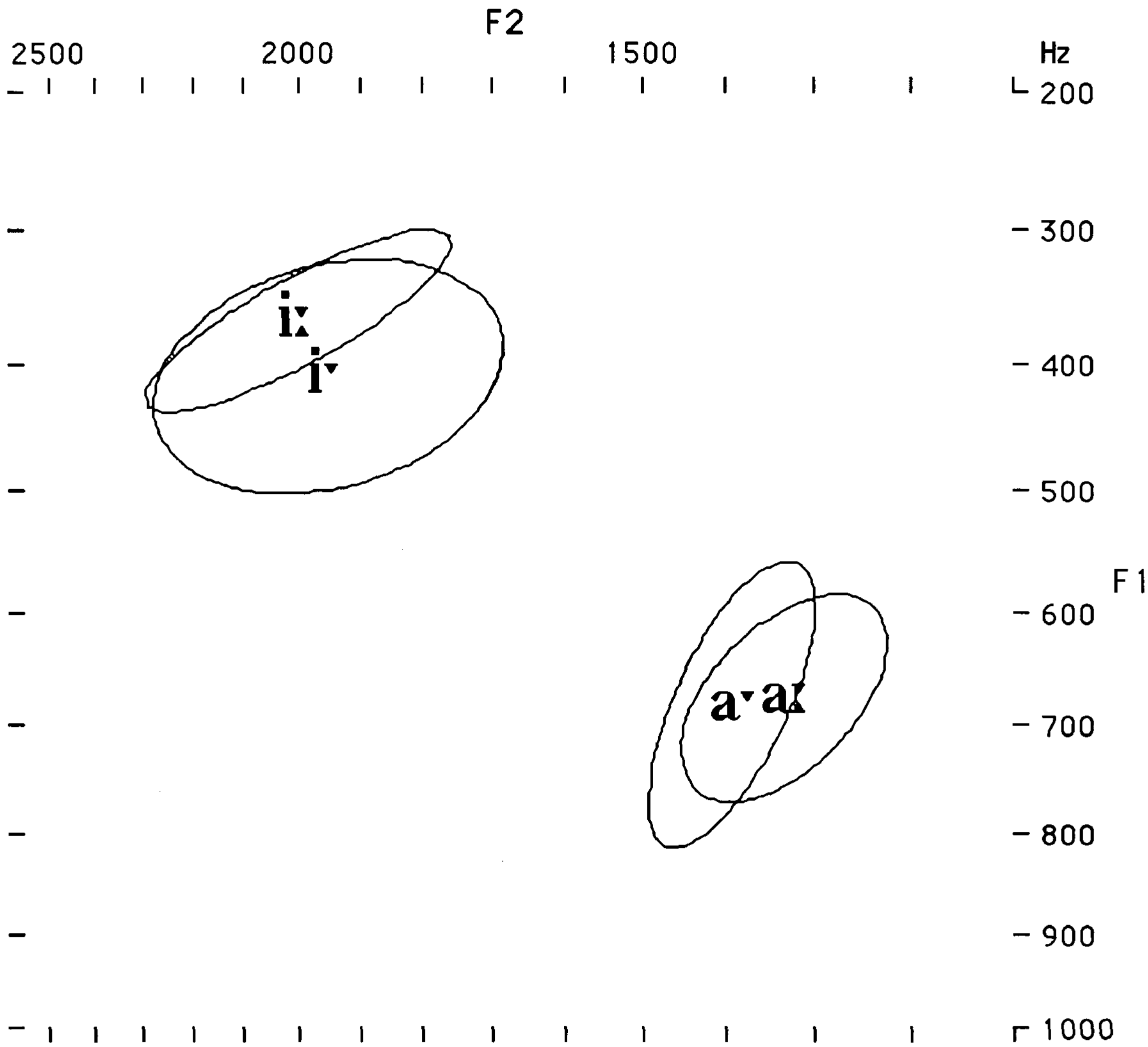


Figure 4: Plot of the first two formants for long and lengthened /**a**/ and /**i**/ for five male speakers

One of the interesting properties of Chickasaw vowels is that the high back vowel is generally not as high as the corresponding front vowel. This tendency reaches significance only in the case of the female long vowels, but is apparent for the other vowels in both Figures 2 and 3. This situation is fairly common in languages with three vowels. A similar system is found for example in Mura-Pirahã (Sheldon 1974) and Hupa (Golla 1970, Gordon 1996), both of which have /**o**/ rather than /**u**/ as the non-low back vowel. In Hupa the situation is further complicated in that the long counterpart to the high front vowel is a mid front vowel. The lack of a phonetic high back round vowel in languages that have a high front vowel is also attested in languages with a greater number of contrastive vowel qualities, e.g. Navajo (McDonough and Austin-Garrison 1994) and Banawá (Ladefoged and Everett 1997), which have four phonemic vowels, Nootka (Sapir and Swadesh 1939), which has five vowel qualities, as well as Wari' (MacEachern, Kern, and Ladefoged 1996), which has six vowel qualities. Based on Crothers' (1978) typology of vowel systems, there is a particularly strong tendency for back vowels to have a relatively low articulation (i.e. more like [**o**]) in three vowel systems compared to systems with a greater number of contrastive vowel qualities. Furthermore, this tendency toward phonetic lowering is more prevalent for back vowels than for front vowels in three vowel systems.

3.2. Vowel duration

The durations of short, lengthened and long vowels were also measured from a waveform using the same corpus used to measure vowel quality in Table 2. Fisher's PLSD post-hoc tests grouping together all speakers found that the three surface vowel lengths were significantly different from one another at the $p<.01$ level. Phonemically short vowels are shortest, and phonemically long vowels are longest, with rhythmically lengthened vowels falling between the two phonemic vowel lengths. The smallest difference in the three categories was between phonemically long vowels and rhythmically lengthened vowels, as shown in Figure 5.

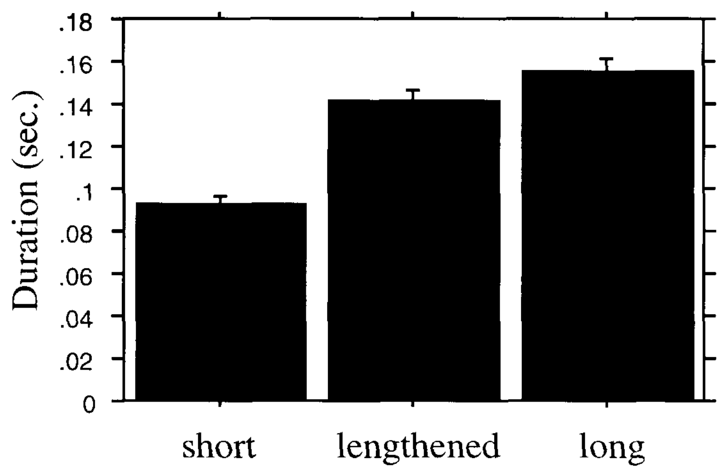


Figure 5: Duration of short, long, and rhythmically lengthened vowels over 13 speakers

Results for individual speakers appear in Table 4.

For all vowel qualities for all speakers, the short vowels are substantially shorter than both the long and the lengthened vowels. However, there is some variation in whether the phonemically long vowels are longer than the rhythmically lengthened vowels as a function of both speaker and vowel quality. For the low vowel, 8 of 10 speakers for whom the comparison could be made show a length difference between long and lengthened /**a**/. For one of these speakers (F4), the difference is only one millisecond. Another speaker (F1) shows no difference between lengthened and long /**a**/, while M1 actually displays a reversal: his long /**a:**/ is actually shorter than his lengthened /**a·**/. For the high front vowel, all speakers observe the same pattern for lengthened /**i·**/ to be shorter than phonemically long /**i:**/, though the difference is negligible in the case of speaker F3. The high back vowel displays the greatest interspeaker variation. For six speakers (F1, F4, F6, M3, M4, M5), long /**oː**/ is more than negligibly longer than lengthened /**oˑ**/. Speakers F3, F5, F6, and F7 essentially show no difference between long and lengthened /**o**/. Speakers F2, M1 and M2 have a longer lengthened /**oˑ**/ than long /**oː**/. In summary, the overall tendency is for lengthened vowels to be slightly shorter than phonemically long vowels.

Table 4: Duration of short, long, and rhythmically lengthened vowels by speaker

	a			**i**			**o**		
Speaker	**Short**	**Length**	**Long**	**Short**	**Length**	**Long**	**Short**	**Length**	**Long**
F1	.088	.179	.179	.073	.126	.186	.089	.173	.197
F2	.070	.117	.141	.073	.079	.096	.042	.120	.105
F3	.090	.156	----	.113	.146	.148	.088	.144	.146
F4	.090	.175	.176	.085	.128	.170	.075	.148	.163
F5	.090	.148	.158	.083	.127	.146	.078	.156	.157
F6	----	----	----	.096	.105	.126	.096	.106	.159
F7	.091	.167	.175	.103	.132	.157	.110	.156	.160
F8	.078	.146	.196	.074	.126	.159	.093	.150	.145
M1	.093	.170	.131	.086	.106	.118	.072	.133	.116
M2	.130	.177	.188	----	----	----	.093	.129	.121
M3	.101	.129	----	.106	.126	.146	.087	.121	.148
M4	.111	.163	.178	.109	.149	.167	.113	.160	.170
M5	.101	.149	.174	.141	.160	.168	.103	.146	.166

4. Consonants

Chickasaw has an inventory of sixteen consonant phonemes, with an additional velar nasal /**ŋ**/, which occurs predictably on the surface before an underlying velar stop in certain morphologically derived forms, called the n- and hn- grades (see Munro and Willmond 1994 for discussion of Chickasaw grades). The consonants of Chickasaw appear in Table 5 below.

Table 5: The consonants of Chickasaw

	Bilabial	Labio-dental	Dental/ alveolar	Postalv-eolar	Labial-velar	Velar	Glottal
Stops	**p b**		**t**			**k**	**ʔ**
Affricates				**tʃ**			
Fricatives		**f**	**s, ɬ**	**ʃ**			**h**
Resonants	**m**		**l, n**				
Glides			**j**		**w**		

The bilabial stop /**b**/ is voiced throughout its entire closure and is often lenited to a voiced fricative or approximant for some speakers. Geminate /**bː**/ more rarely undergoes lenition as well. Singleton but not geminate /**k**/ is sometimes lenited to a fricative; it also may be voiced, particularly before /l/ (see next paragraph). It also appears to be occasionally realized with a more backed articulation than velar stops in English, though it is clearly phonetically distinct from phonemic uvular stops in other languages. Some younger speakers, including one of the speakers examined for this study (F8), produce the lateral fricative as an interdental fricative /θ/. The voiceless stops are basically unaspirated. However, there is a longer aspiration phase in word-initial than word-medial position, though this observation has not been quantitatively tested (see section 3.1. for discussion of voice onset time). All consonants except glottal stop appear as both singletons and geminates.

Most possible combinations of consonants are attested in bisegmental clusters. However, glottal stop may not be the second member of a cluster (Munro 1996). For many speakers, an epenthetic vowel is inserted in clusters consisting of a stop followed by a sonorant (Munro 1996). This vowel is typically realized with a schwa-like quality. The stop is also often voiced in such clusters (Munro 1996). An example of an epenthetic vowel as spoken by speaker F6 is found in the spectrogram in Figure 6.

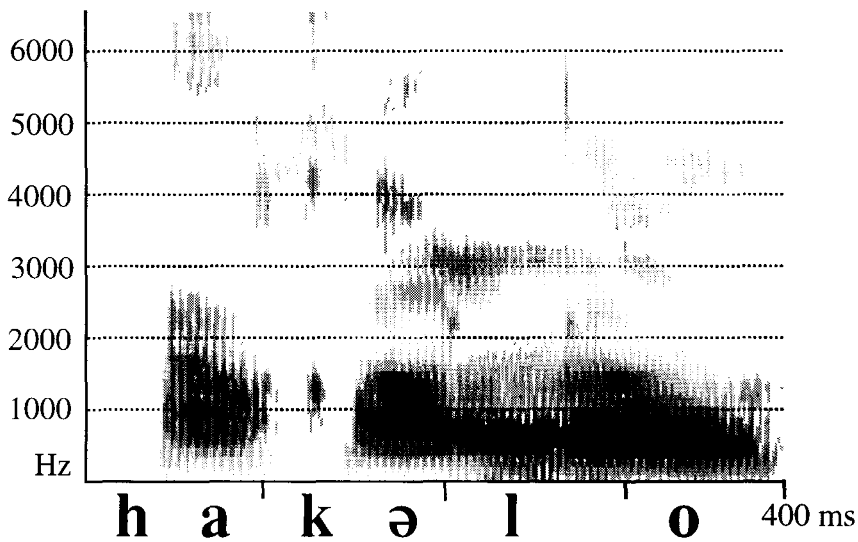


Figure 6: Example of an epenthetic vowel breaking up a stop + sonorant cluster in the word /**haklo**/ 'he hears' (speaker F6)

4.1. Voice onset time (VOT) and closure duration for voiceless stops

Voice onset time (the period from consonant release to the onset of voicing of the following vowel, abbreviated VOT) was measured for the three voiceless unaffricated stops before the vowel /**a**/. Measurements were made from a waveform in conjunction with a spectrogram for the bold-face consonants in the three words: /**ɬipa**/ 'worn out, ragged', /**ayi:mita**/ 'he's excited', /**hika**/ 'stand up!'. Two tokens of each consonant were recorded in most cases for each speaker, except where the consonant underwent lenition. Because of lenition, there are no measurements for /**k**/ for speaker M1.

Results of the VOT measurements pooled over the 13 speakers appear in Figure 7. Results for individual speakers appear in Table 6.

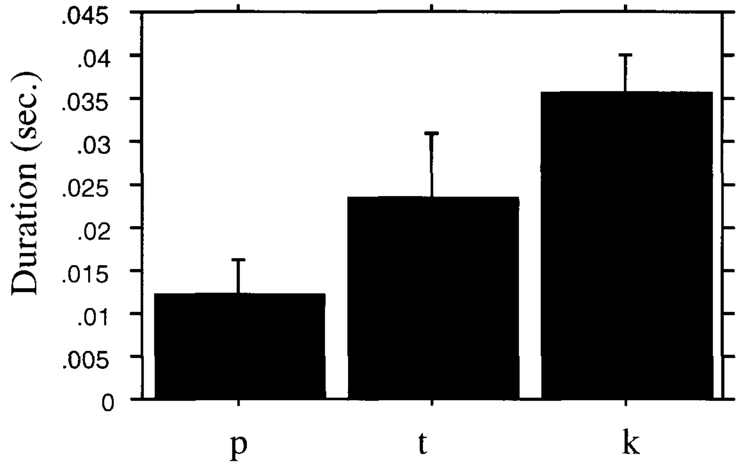


Figure 7: VOT for voiceless unaffricated stops before /**a**/.

Table 6: VOT for unaffricated stops (in seconds)

		Consonant	
Speaker	**p**	**t**	**k**
F1	.030	.027	.037
F2	.014	-.008	.025
F3	.013	.028	.039
F4	.007	.012	.023
F5	.014	.011	.029
F6	.005	.008	.022
F7	.030	.016	.041
F8	.017	.037	.044
M1	.013	.020	lenited
M2	.010	.025	.034
M3	.008	.014	.051
M4	.015	.074	.042
M5	-.008	.022	.040

The VOT of Chickasaw stops is comparable with that of stops in other languages of the world that have been classified as voiceless unaspirated (Cho and Ladefoged 1997, this volume). All three voiceless stops were significantly different from one another in terms of VOT, with the bilabial displaying the shortest average VOT, followed by the dental/alveolar, and finally the velar stop. A few female speakers (F1, F2, F3, F5) did not display the trend for bilabials to have a shorter VOT than alveolars. For all speakers except M4, alveolars had shorter VOTs than velars. For none of the speakers was VOT longer for bilabials than for velars. The differences in VOT associated with the different places of articulation are also similar to those found elsewhere. Cho and Ladefoged (1997) have suggested that "the observed VOT differences are the inevitable consequence of the physiological movements and the aerodynamic forces. Given enough knowledge about the gestures involved, VOT differences due to place of articulation are fully determined."

Stops made at different places of articulation differ not only in VOT but also in the duration of the articulatory closure. Across speakers, closure duration for the voiceless bilabial stop is longest followed by the dental/alveolar stop followed by the velar stop, as shown in Figure 8. Measurements of closure duration for individual speakers appear in Table 7. One speaker (M4) shows a longer closure duration for /**k**/ than for /**t**/, an exception to the general pattern. An additional three speakers (F4, F5 and F8) do not display a substantial difference in closure duration between /**t**/ and /**k**/. Closure duration for /**p**/ is longer than for either /**t**/ and /**k**/ for all speakers.

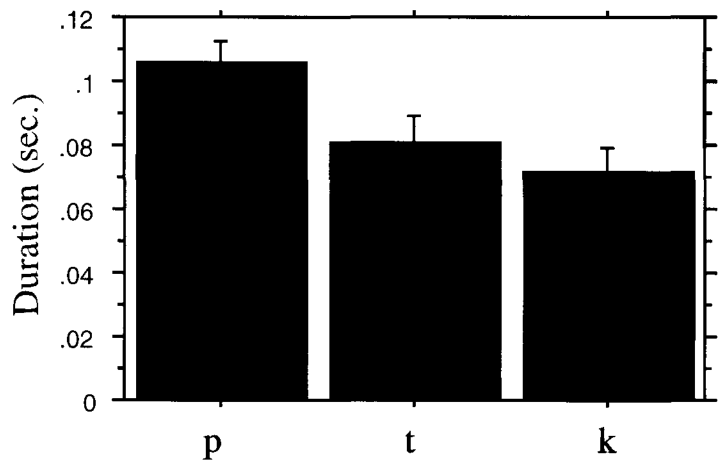


Figure 8. Closure duration for voiceless unaffricated stops before /**a**/

Table 7: Closure duration for unaffricated stops before /**a**/ (in milliseconds)

		Consonant	
Speaker	**p**	**t**	**k**
F1	.110	.097	.075
F2	.084	.071	.064
F3	.113	.101	.083
F4	.124	.095	.096
F5	.114	.100	.098
F6	.112	.101	.082
F7	.115	.089	.067
F8	.111	.060	.059
M1	.080	.058	lenited
M2	.083	.073	.052
M3	.117	.093	.066
M4	.092	.044	.060
M5	.119	.071	.044

A possible reason for the shorter closure duration for the more posterior stops is that the seal for these stops may be more difficult to hold in the face of increased air pressure (Maddieson 1997). When the oral closure is further back, the cavity behind the closure is smaller, and the air pressure reaches a maximum more quickly. The faster the increase in pressure, the shorter the closure is held, perhaps due to a biomechanical feedback mechanism. Following this logic, the seal for a velar will be maintained for a shorter period of time than the closure for an alveolar stop, which should, all else being equal, have shorter a closure than a bilabial stop. The Chickasaw data are compatible with this account, as velars show the shortest closures, whereas bilabials have the longest closures.

The duration hierarchy for VOT is the opposite of the pattern seen for closure duration. For the dental and velar stops the difference in closure duration is almost exactly compensated by the difference in VOT, so that the sum of the closure plus VOT is the same for both places of articulation. However, in the case of the bilabial stops, the compensation is not so exact. The duration of the closure plus VOT is slightly less than 20 ms longer than the corresponding duration for the other two stops. The correlation between VOT and closure duration for stops at different places of articulation appears to be another general cross-linguistic property attested in a large number of languages. In discussing this, Maddieson (1997) offers a possible explanation of the place-dependent nature of both the closure duration and VOT suggesting that the duration of the vocal fold opening may be considered to be fixed, so that when the closure duration is relatively longer, the following VOT becomes relatively shorter (and vice versa). As he puts it: "There is an abduction-adduction cycle of the vocal cords for voiceless stops which is longer in duration than the closure and has a constant time course, anchored to the onset of closure" (Maddieson 1997: 621).

VOT for geminate stops was also measured and compared to that of the corresponding single stops. Relatively little literature has examined differences in VOT between single and geminate consonants, and the existing work shows different results depending on the language. Lahiri and Hankamer (1988) found that VOT is actually significantly longer for single /**t**/ and /**k**/ than for geminate /**t**ː/ and /**k**ː/ in Turkish, though not all speakers show a significant difference. A similar pattern is evident in Han's (1994) data on Japanese VOT for /**p**, **t**, **k**/. On the other hand, Lahiri and Hankamer's study found no significant difference in VOT between single and geminate /t/ in Bengali. A third pattern is attested in Estonian. An unpublished study of VOT in Estonian conducted by Gordon (1995) showed significantly longer VOT values for geminate /**t**ː/ than for single /**t**/ for one speaker. The other speaker in the study did not distinguish between single and geminate /**t**/ in terms of VOT.

In the present study of Chickasaw, single and geminate /**p**/ and /**t**/ were measured before the vowel /**a**/ , and single and geminate /**t**/ and /**k**/ were measured before the vowel /**i**/. The words from which the targeted consonants were measured appear in Table 8.

Table 8: Words used to measure VOT for singleton and geminate stops
(Note the pitch accent, phonetically a high tone, on the first vowel in /**hîk:iʔya**/)
Before /**a**/

p	/**ɬipa**/ 'worn out, ragged'	/**ʃip:a**/ 'it dries up'
t	/**ayi:mita**/ 'he's excited'	/**okʃit:a**/ 'he closes it'

Before /**i**/

t	/**iti**/ 'mouth'	/**it:iʔ**/ 'tree'
k	/**tʃi:ki**/ 'early'	/**hîk:iʔya**/ 'he's standing up'

Results of the measurements pooled over all speakers appear in Figure 9. Results for individual speakers appear in Table 9. Speaker F4 produced a cluster of /**ph**/ rather than a geminate, hence these values were not included.

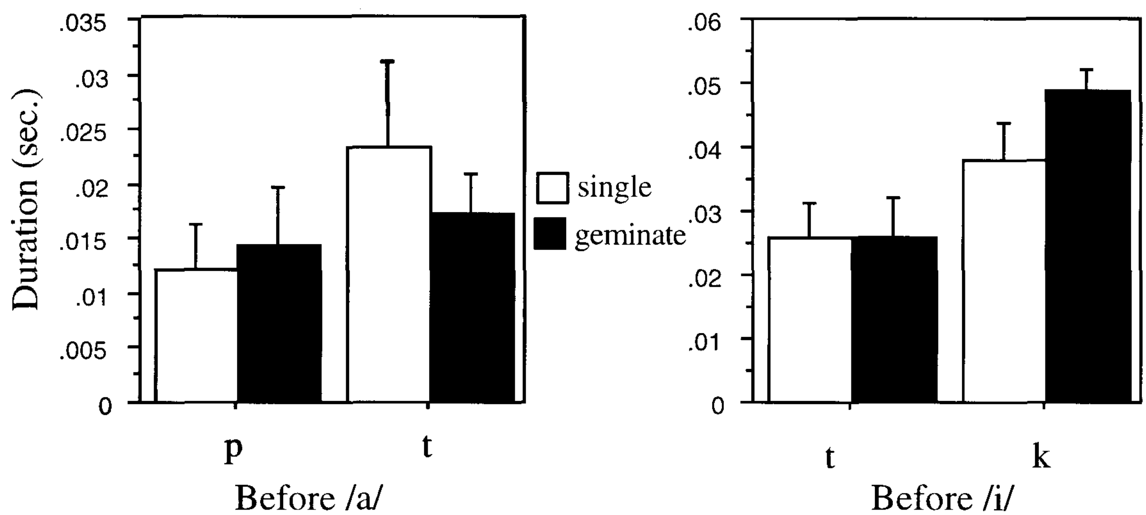


Figure 9. VOT for single and geminate stops

Table 9: VOT for single vs. geminate unaffricated stops (in milliseconds)

	Before /a/				Before /i/			
	p		t		t		k	
Speaker	**Single**	**Gemin**	**Single**	**Gemin**	**Single**	**Gemin**	**Single**	**Gemin**
F1	.030	.010	.027	.027	.030	.025	.036	.042
F2	.014	.013	.011	.018	.014	.011	.021	.047
F3	.013	.019	.028	.021	.028	.021	.057	.069
F4	.007	--	.012	.005	.043	.019	.038	.046
F5	.014	.013	.011	.011	.021	.016	.035	.049
F6	.005	.006	.008	.010	.011	.017	.018	.054
F7	.030	.010	.016	.010	.016	.018	.033	.040
F8	.017	.053	.037	.035	.046	.071	.030	.057
M1	.013	.013	.020	.021	.019	.016	.029	.045
M2	.010	.009	.025	.018	.041	.030	.042	.054
M3	.008	.006	.014	.014	.013	.030	.048	.047
M4	.08	.006	.074	.028	.030	.042	.060	.045
M5	.015	.020	.022	.007	.025	.020	.040	.047

When all speakers are pooled, unpaired t-tests show that the only significant difference in VOT for any pair of single and geminate consonants is that geminate /**kː**/ has a significantly longer VOT than single /**k**/ (p<.001). It is worth noting, however, that the difference between single and geminate /**t**/ before /**a**/, a difference which goes in the opposite direction from /**k**/, approaches significance (p=.0741) according to an unpaired t-test. Chickasaw thus shows traces of the Turkish and Japanese pattern whereby VOT for singletons is longer than VOT for geminates, as well as the Estonian pattern whereby VOT for geminates is longer than VOT for single consonants. Finally, as with Bengali /**t**/, Chickasaw does not show VOT differences for single and geminate /**t**/, at least before /**i**/. The conflicting data between languages and within Chickasaw suggest that VOT differences between single and geminate consonants are a function of multiple properties which are not completely understood.

4.2. Closure duration for single and geminate pairs

In order to examine the durational relationship between single and geminate consonants, durations of several singleton and geminate pairs were measured from a waveform in conjunction with a spectrogram and an energy display. The measured words appear in Table 9.

The consonants /**j**/, /**w**/ and /**h**/ were not measured due to the difficulty of determining precisely where in the waveform these segments began and end. Wherever the relevant examples occurred, consonant pairs were measured in two environments: between /**i**/ and /**i**/ and between /**i**/ and /**a**/. An attempt was made to measure both members of each singleton and geminate pair in prosodically similar environments, ideally in disyllabic words between two short vowels. Occasionally, however, words that either had more than two syllables or contained the target consonant in a non-final syllable were used. In such cases, the target consonant followed a short vowel whenever possible. Results of the duration measurements are found in Figure 10 (before /**a**/) and Figure 11 (before /**i**/) pooling together all fourteen speakers. Results for individual speakers appear in Table 11. Omissions in Table 11 indicate that the speaker did not produce the targeted word. Note that /**m**/ before /**a**/ could not reliably be measured for speaker M2.

In considering the data on geminates, note that for some consonants, the duration data may be confounded by other factors. We now consider these potential confounds. For some pairs, the word in which the geminate appears had more syllables than the word in which the corresponding singleton appears. This would reduce the intrinsic duration difference between single and geminate consonants, under the assumption that segmental duration is inversely correlated with net duration of the word (Lehiste 1970). Furthermore, for certain pairs there are morphological factors that could have confounded the results. The single /**ʃ**/ in /**biʃaɬtʃi**/ falls at the end of a large morphological boundary, which could plausibly have induced lengthening of the /**ʃ**/. The word /**hîkːiʔya**/ has a pitch accent (phonetically a high tone) on the vowel preceding the geminate, which could have influenced the duration of the stop. Finally, in a few words, the target geminate consonant is immediately followed by a syllable ending in glottal stop, which could potentially have influenced the duration of the geminate.

Table 10: Words used for comparing the duration difference between single and geminate consonants in Chickasaw

Environment/ i __ i

Cons	Singleton	Geminate
b	/**itːibi**/ 'they fight'	/**tʃikibːi**/ 'it's all piled up'
t	/**iti**/ 'mouth'	/**itːiʔ**/ 'tree'
k	/**tʃiːki**/ 'early'	/**hȋkːiya**/ 'he's standing up'
f	/**ɬifili**/ 'he drags himself around'	/**pitʃifːi**/ 'he crushes it'
s	/**kisili**/ 'he bites more than once'	/**isːiʔ**/ 'deer'
ɬ	/**yiɬibli**/ 'he demolishes it'	/**kiɬːi**/ 'he gnaws it'
ʃ	/**iʃi**/ 'he takes it'	/**aɬpiʃːiʔ**/ 'pillow'
m	/**hatiːmɪmiʔ**/ 'type of insect'	/**yimːi**/ 'he believes it'
l	/**pihlili**/ 'I sweep it'	/**ilːi**/ 'he dies'
n	/**kostini**/ 'he sobers up'	/**lowak inːi**/ 'he warms himself by the fire'

Environment i __ a

Cons	Singleton	Geminate
p	/**ɬipa**/ 'it's worn out'	/**ʃipːa**/ 'it dries up'
b	/**aɬtʃiba**/ 'it's late'	/**koʃibːaʔ**/ 'poke salad'
t	/**ayiːmita**/ 'he's excited'	/**okʃitːa**/ 'he closes it'
tʃ	/**malːitʃa woːtʃi**/ 'he jumps and barks'	/**ibitʃːalaʔ**/ 'nose'
s	/**pisa**/ 'look!'	/**impat isːa**/ 'he finishes eating'
ʃ	/**biʃaɬtʃi**/ 'it's milked'	/**ibiʃːano**/ 'he has a cold'

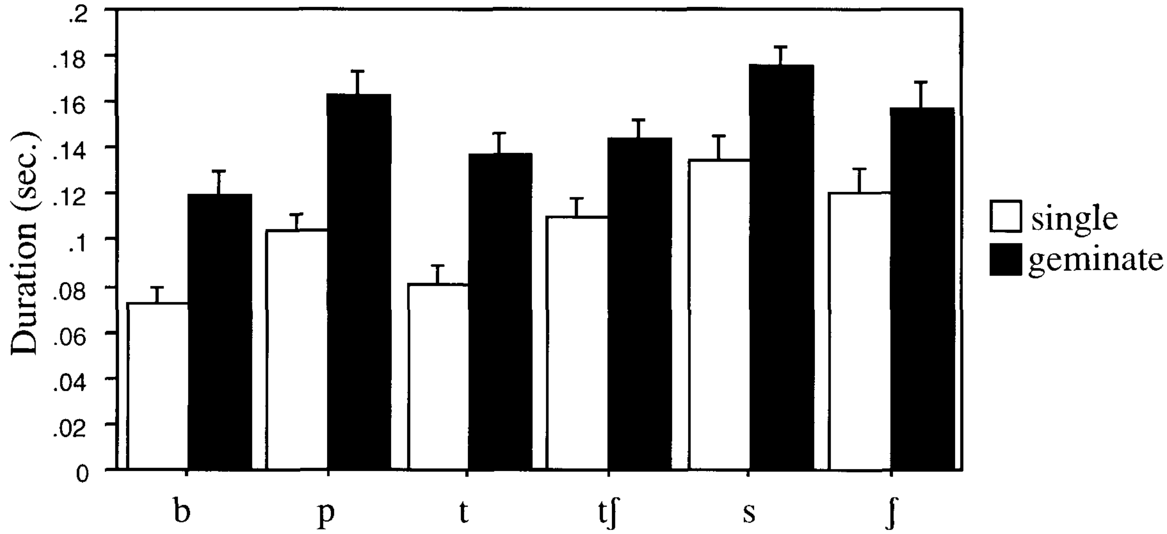


Figure 10: Duration of single and geminate consonants before /a/ (all speakers together)

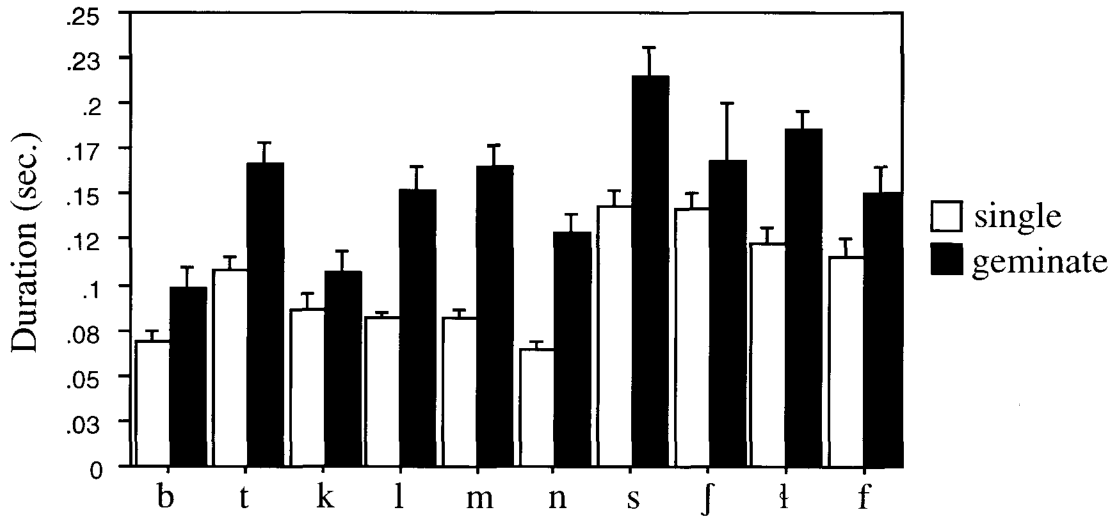


Figure 11: Duration of single and geminate consonants before /i/ (all speakers together)

Interestingly, if results for all speakers are pooled, for very few of the consonants is the geminate twice as long as its corresponding single consonant in the same vowel environment. Only the sonorant geminates /**m**, **n**, **l**/ are twice as long as their corresponding single consonants. For all other single and geminate pairs, the duration ratio between single and geminates is much smaller than 2:1. In some, the small magnitude of the duration difference may be due to independent factors mentioned above, such as number of syllables, morphology, and pitch accents. However, relatively small duration differences between single and geminate consonants are found even for pairs in which these factors are controlled for, suggesting that the length contrast in consonants does not reside only in the duration of the consonant constriction itself. For example, the contrast between single and geminate /**b**/ also appears to be associated with differences in amplitude of voicing: geminate /**bː**/ has stronger voicing than single /**b**/. This difference in amplitude appears to be the principal acoustic cue to the single vs. geminate contrast for speaker F2, F7 and M1 who show virtually no length distinction between single and geminate /**b**/. In fact, geminate /**b**ː/ is shorter than single /**b**/ for speaker M1. Similarly, the contrast between single and geminate /**k**/ appears to involve differences in VOT (see Figure 9) in addition to length differences: geminate /kː/ has a consistently longer VOT than single /**k**/. Additional prosodic cues considered later in this section also help to distinguish single and geminate consonants.

It appears to be a general property of Chickasaw for the ratio in closure duration between geminates and singletons to be smaller than in many other languages. In standard Finnish, most geminates are at least twice as long as their corresponding single consonants (Lehtonen 1970). Similar durational ratios between geminates and singletons are reported for Estonian (Lehiste 1966, Ojamaa 1976) and Bengali (Lahiri and Hankamer 1988). In Japanese (Han 1994), Italian (Farnetani and Kori 1984) and Levantine Arabic (Miller 1987), geminates are for the most part longer than twice the duration of single consonants. In Turkish (Lahiri and Hankamer 1988) and Hungarian (Meyer and Gombocz 1908), geminates are almost three times as long as their corresponding single consonants.

Chickasaw is not alone, however, in showing less than a 2:1 ratio in the duration of geminates relative to single consonants. Other languages display smaller differences between single and geminate pairs at least for certain consonants, e.g. Dogri (Ghai 1980), Icelandic (Orešnik and Pétursson 1977), Norwegian (Fintoft 1961), Sinhala (Letterman 1994). Of these languages, Sinhala and Dogri come the closest to having a 2:1 ratio between the duration of geminates vs. singletons. In Sinhala, geminates are 1.8 times as long as singletons after short vowels, and 1.69 times as long after long vowels. Dogri displays great variability in the relative duration of geminates; single consonants range from being 38% to 66% as long as corresponding geminates depending on the consonant and the speaker.

Table 11: Duration measurements for single and geminate consonants for individual speakers (all measurements in seconds) S = Single, G = Geminate

	Before /a/			**Before /i/**				**Before /a/**			**Before /i/**		
F1	S	G	Ratio	S	G	Ratio	**F2**	S	G	Ratio	S	G	Ratio
p	.110	.193	**1.75**				**p**	.084	.150	**1.79**			
t	.097	.128	**1.32**	.105	.173	**1.65**	**t**	.071	.110	**1.55**	.088	.145	**1.65**
tʃ	.116	.142	**1.22**				**tʃ**	.113	.171	**1.51**			
k				.088	.102	**1.16**	**k**				.077	.082	**1.06**
f				.102	.132	**1.29**	**f**				.114	.132	**1.16**
s	.139	.165	**1.19**	.170	.198	**1.16**	**s**	.136	.149	**1.10**	.108	.150	**1.39**
ʃ	.115	.139	**1.21**	.171	.138	**0.81**	**ʃ**	.103	.128	**1.24**	.110	.121	**1.10**
ɬ				.097	.166	**1.71**	**ɬ**				.108	.155	**1.44**
b	.069	.106	**1.54**	.066	.124	**1.88**	**b**	.071	.073	**1.03**			
m				.080	.147	**1.84**	**m**				.093	.143	**1.54**
n				.075	.155	**2.07**	**n**				.055	.099	**1.80**
l				.082	.114	**1.39**	**l**				.068	.095	**1.40**

Table 11 (continued): Duration measurements for single and geminate consonants for individual speakers (all measurements in seconds) S = Single, G = Geminate

F3	S	G	Ratio	S	G	Ratio	**F4**	S	G	Ratio	S	G	Ratio
p	.113	.189	**1.67**				**p**						
t	.101	.149	**1.48**	.119	.209	**1.76**	**t**	.095	.166	**1.75**	.141	.197	**1.40**
tʃ	.122	.159	**1.3**				**tʃ**	.138	.142	**1.03**			
k				.109	.130	**1.19**	**k**				.089	.148	**1.66**
f							**f**				.160	.214	**1.34**
s	.174	.189	**1.09**	.157	.245	**1.56**	**s**	.173	.195	**1.13**	.166	.256	**1.54**
ʃ	.146	.196	**1.34**	.152	.220	**1.45**	**ʃ**	.132	.184	**1.39**			
ɬ				.132	.232	**1.76**	**ɬ**				.121	.199	**1.64**
b	.076	.144	**1.89**	.073	.106	**1.45**	**b**	.096	.134	**1.40**			
m				.093	.194	**2.09**	**m**				.074	.206	**2.78**
n							**n**				.072	.177	**2.46**
l				.095	.177	**1.86**	**l**				.080	.212	**2.65**
F5	S	G	Ratio	S	G	Ratio	**F6**	S	G	Ratio	S	G	Ratio
p	.114	.167	**1.46**				**p**	.112	.194	**1.73**			
t	.100	.167	**1.67**	.130	.201	**1.55**	**t**	.101	.151	**1.50**	.125	.170	**1.36**
tʃ	.118	.156	**1.32**				**tʃ**	.115	.147	**1.28**			
k				.121	.120	**0.99**	**k**				.115	.129	**1.12**
f				.153	.186	**1.22**	**f**				.109	.139	**1.28**
s	.152	.191	**1.26**	.164	.224	**1.37**	**s**	.130	.195	**1.50**	.140	.242	**1.73**
ʃ	.141	.175	**1.24**	.169	.220	**1.3**	**ʃ**	.104	.142	**1.37**	.132	.116	**0.88**
ɬ				.128	.200	**1.56**	**ɬ**				.131	.190	**1.45**
b	.076	.147	**1.93**				**b**	.060	.105	**1.75**	.073	.093	**1.27**
m				.093	.180	**1.94**	**m**				.084	.154	**1.83**
n				.070	.140	**2**	**n**				.046	.119	**2.59**
l				.087	.171	**1.97**	**l**				.083	.140	**1.69**
F7	S	G	Ratio	S	G	Ratio	**F8**	S	G	Ratio	S	G	Ratio
p	.115	.158	**1.37**				**p**	.111	.139	**1.25**			
t	.089	.139	**1.56**	.116	.196	**1.69**	**t**	.060	.109	**1.82**	.101	.113	**1.12**
tʃ	.111	.143	**1.29**				**tʃ**	.124	.156	**1.26**			
k				.092	.079	**0.86**	**k**				.080	.090	**1.13**
f				.110	.155	**1.41**	**f**						
s	.141	.178	**1.26**	.166	.284	**1.71**	**s**				.140	.183	**1.31**
ʃ	.153	.171	**1.12**	.173	.124	**0.72**	**ʃ**	.127	.148	**1.17**	.158	.172	**1.09**
ɬ				.144	.172	**1.19**	**ɬ**				.151	.154	**1.02**
b	.082	.119	**1.45**	.075	.080	**1.07**	**b**				.069	.090	**1.30**
m				.068	.147	**2.16**	**m**				.067	.122	**1.82**
n				.072	.120	**1.67**	**n**				.063	.112	**1.78**
l				.078	.159	**2.04**	**l**				.076	.118	**1.55**

Table 11 (continued): Duration measurements for single and geminate consonants for individual speakers (all measurements in seconds) S = Single, G = Geminate

M1	S	G	Ratio	S	G	Ratio	**M2**	S	G	Ratio	S	G	Ratio
p	.080	.133	**1.66**				**p**	.083	.135	**1.63**			
t	.058	.154	**2.66**	.088	.175	**1.99**	**t**	.073	.107	**1.47**	.079	.135	**1.71**
tʃ	.089	.123	**1.38**				**tʃ**	.087	.108	**1.24**			
k				.056	.119	**2.13**	**k**						
f				.089	.120	**1.35**	**f**				.117	.138	**1.18**
s				.128	.199	**1.55**	**s**	.110	.157	**1.43**	.146	.198	**1.36**
ʃ							**ʃ**	.084	.127	**1.51**	.136	.116	**0.85**
ɬ				.126	.166	**1.32**	**ɬ**				.128	.198	**1.55**
b	.077	.111	**1.44**	.081	.076	**0.94**	**b**	.051	.107	**2.1**	.048	.079	**1.65**
m				.079	.179	**2.27**	**m**						
n				.068	.123	**1.81**	**n**				.051	.132	**2.59**
l				.083	.133	**1.6**	**l**				.066	.157	**2.38**
M3	S	G	Ratio	S	G	Ratio	**M4**	S	G	Ratio	S	G	Ratio
p	.117	.188	**1.61**				**p**	.092	.149	**1.62**			
t	.093	.145	**1.56**	.118	.154	**1.31**	**t**	.044	.118	**2.68**	.096	.151	**1.57**
tʃ	.105	.140	**1.33**				**tʃ**	.095	.135	**1.42**			
k				.073	.096	**1.32**	**k**				.071	.071	**1**
f				.116	.150	**1.29**	**f**				.099	.130	**1.31**
s	.132	.172	**1.30**	.141	.190	**1.35**	**s**				.124	.210	**1.69**
ʃ	.123	.144	**1.17**	.134	.163	**1.22**	**ʃ**	.135	.138	**1.02**	.132	.108	**0.82**
ɬ				.097	.187	**1.93**	**ɬ**				.119	.203	**1.71**
b	.068	.139	**2.04**				**b**	.081	.099	**1.22**	.070	.097	**1.39**
m				.063	.173	**2.75**	**m**				.090	.182	**2.02**
n				.059	.135	**2.29**	**n**				.073	.126	**1.73**
l				.087	.155	**1.78**	**l**				.085	.192	**2.26**
M5	S	G	Ratio	S	G	Ratio							
p	.119	.158	**1.33**										
t	.071	.135	**1.90**	.103	.145	**1.41**							
tʃ	.099	.143	**1.44**										
k				.049	.148	**3.02**							
f				.104	.156	**1.50**							
s	.096	.156	**1.63**	.119	.219	**1.84**							
ʃ	.086	.196	**2.28**	.121	.272	**2.25**							
ɬ													
b	.059	.122	**2.07**	.065	.146	**2.25**							
m				.085	.162	**1.91**							
n				.068	.113	**1.66**							
l				.093	.157	**1.69**							

Interestingly, in most of the languages (except Sinhala) that display smaller durational differences between single and geminate consonants (less than 2:1), consonant duration appears to be closely related to the duration of neighboring vowels, in particular the preceding vowel. In most of these languages, vowels appear to be substantially shorter before geminates than before single consonants. In Dogri, short vowels preceding single consonants are 1.72 times longer than short vowels preceding geminates. In Icelandic, particularly in southern dialects, geminates are less than 1.5 times longer than single consonants. However, it is unclear in Icelandic whether the length contrast is a property of the consonant itself or is a property of the proceding vowel (see Orešnik and Pétursson 1977 for discussion): vowels are long before short consonants and short before long consonants. In Norwegian, another language in which consonant and vowel length stand in a close trading relationship, geminates are only slightly longer than their singleton counterparts, 1.10 to 1.38 times as long.

In languages that display greater durational differences between geminates and singletons, vowels tend not to differ greatly as a function of whether they precede a geminate or a single consonant. This appears to be true of most of the languages discussed in the preceding paragraph for which measurements were available. For example, in Turkish, where geminates are almost three times as long as single consonants, vowels preceding single consonants are only 3.5% longer than vowels before geminates. In standard Estonian with a similar but slightly greater duration ratio between vowels before geminates and singletons, vowels before geminate /**tː**/ are 5.7% longer than vowels before single /**t**/. In standard Finnish, vowels are 20.3% longer before geminate /**pː** / and /**s**ː/ than before single /**p**/ and /**s**/ (results pooled over both pairs). In Japanese, vowels are between 1.1% and 8.7% longer before geminates than before singletons depending on the consonant. In Sinhala, vowels are also slightly longer before geminates (19.6% for short vowels, 4.4% for long vowels). Italian seems to be somewhat exceptional in manifesting the single vs. geminate contrast in both the consonant itself and the preceding vowel: vowels are 30%-50% shorter before geminates than before single consonants (Farnetani and Kori 1984, 1986).

Like other languages with relatively small duration differences between geminates and singletons, there is evidence that the contrast in consonant length in Chickasaw may be cued by other properties in addition to the consonant durations themselves. Recall from section 3 that Chickasaw displays a process of rhythmic lengthening, which lengthens the second in a series of two consecutive phonemically short vowels in non-word-final open syllables. Vowels fail to lengthen in closed syllables, including syllables closed by a geminate, even if other conditions for the process are satisfied. Thus, in certain positions, the absence of rhythmic lengthening where it would otherwise be expected to occur signals that the following consonant is a geminate. Similarly, the presence or absence of rhythmic lengthening also provides information about the preceding consonant, since rhythmic lengthening only occurs following an open syllable and thus not following a geminate. Rhythmic lengthening patterns thus can provide valuable aid in perceiving contrasts in consonant duration in many environments.

Rhythmic lengthening, however, is not available as a potential cue to consonant duration in all environments. For example, in a two syllable word, rhythmic lengthening is independently blocked from occurring due to the restriction against word-final lengthened vowels. Thus, rhythmic lengthening provides little information about the status of intervocalic consonants in two syllable words. In order to test whether phonetic differences in duration between vowels before single consonants and vowels before geminates also exist in Chickasaw, vowels were measured before single and geminate consonants in the disyllabic near minimal pairs /**iti**/ vs. /**itːiʔ**/ and /**ɬipa**/ vs. /**ʃipːa**/. Recall from Figures 9 and 10 that the difference in closure duration between singletons and geminates in both of these pairs is relatively small. Thus, we might expect there to be large differences in vowel duration as a function of the phonemic length of the following consonant.

Measurements show fairly large durational differences between vowels preceding geminates and vowels preceding singletons, particularly before the bilabial stop. Vowels preceding single consonants are generally longer than those preceding geminates as seen in Figure 12, which pools together results for all speakers. Results for individual speakers appear in Table 12.

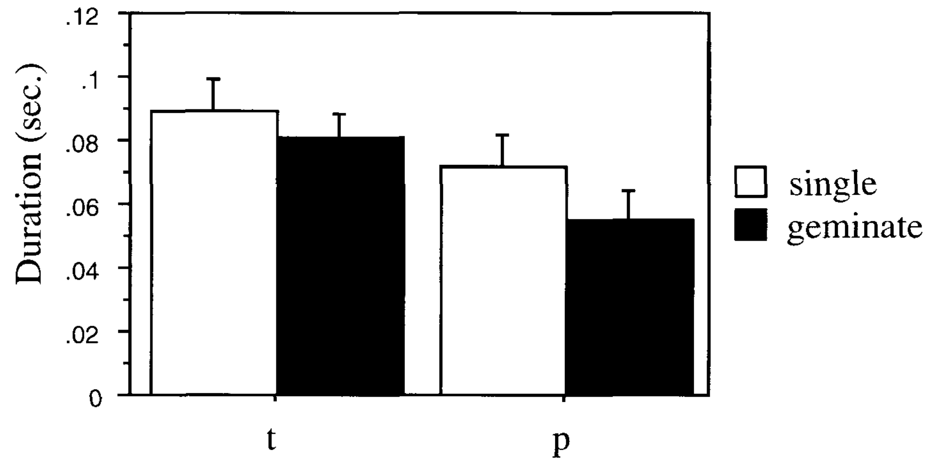


Figure 12: Duration of vowel preceding single and geminate consonants

Not all speakers show substantial differences in vowel length as a function of the duration of the following consonant; however, there is a general trend apparent in Figure 12 for vowels before single consonants to be longer than vowels before geminate consonants. This result is particularly robust in the case of /**p**/. Taken over all speakers (M4 excluded; see below), /**i**/ before single /**t**/ is 8% longer than before geminate /**tː**/, while /**i**/ before single /**p**/ is 27% longer than before geminate /**pː**/. Vowels before single and geminate /**t**/ for one speaker (M4) were not included in the measurements presented in Figure 12, since /**i**/ was substantially longer before the geminate than before the single /**t**/ in his speech. Smaller reversals of the dominant trend for vowels to be longer before singletons are also seen for speaker F6 for /**t**/, speaker M2 for /**t**/, speaker F4 for /**t**/ and speaker F5 for /**p**/, though these results are included in Figures 12. On an individual speaker basis, there does not seem to be a relationship between the magnitude of vowel length differences as a function of the following consonant and the duration difference between single and geminate consonants.

Table 12: Duration (in seconds) of vowel preceding single and geminate consonants

Speaker	p		t	
	Before singleton	Before geminate	Before singleton	Before geminate
F1	.070	.042	.101	.078
F2	.030	.021	.064	.055
F3	.074	.054	.082	.081
F4	.081	.052	.112	.116
F5	.059	070	.072	.073
F6	.065	.056	.049	.068
F7	.072	.049	.092	.072
F8	.062	.031	.073	.076
M1	.062	.033	.110	.075
M2	.117	.090	.090	.096
M3	.071	.055	.091	.086
M4	.105	.094	.082	.124
M5	.070	.070	.132	.086

In summary, at least for /**t**/ in the present corpus, the duration of the preceding vowel does not provide particularly salient cues to the phonemic length of the following consonant. This suggests that the phonemic contrast in consonant duration might rely on other properties in addition to the duration of the consonant and the preceding vowel. For example, VOT appears to provide a cue to the length contrast in /**k**/ and /**t**/ in certain environments. Likewise, voicing is stronger in geminate /**b:**/ than in single /**b**/. Perhaps consonant duration may also be cued by acoustic properties other than those measured in this paper, such as vowel quality or burst amplitude. In addition, the rhythmic lengthening patterns discussed earlier provide useful cues to consonant duration in many environments. Finally, the single vs. geminate contrast is also signaled in many cases through the morphology in Chickasaw, for example, in the choice of suffixes or through grade formation (see Munro and Willmond 1994, Munro et al. forthcoming for discussion).

5. Conclusion

This paper describes some of the basic phonetic properties of Chickasaw, including measurements of closure duration and VOT for single and geminate consonants, and formant and duration measurements for short, long and rhythmically lengthened vowels, and attempts to relate them to common cross-linguistic phonetic trends. Some of the important results are as follows. The three-way length contrast in vowels is phonetically manifested principally in terms of duration. Phonemically long vowels are longer than rhythmically lengthened vowels, which, in turn, are longer than phonemically short vowels. Phonemically long vowels are in general more peripheral than their short, but not their lengthened, counterparts. However, lengthened /**a·**/ and /**i·**/ are intermediate in quality between their phonemically short and long counterparts. Duration differences between single and geminate consonants are relatively small from a typologically perspective. We conjecture that other properties of Chickasaw (e.g. rhythmic lengthening, duration of the preceding vowel, VOT duration in the case of /**k**/, strength of voicing in the case of /**b**/) provide important cues to consonant duration in the face of the less pronounced differences in the consonants themselves.

Acknowledgments

Many thanks are owed to all the kind speakers of Chickasaw who volunteered their time and effort (and voices) to make this study possible: Frankie Alberson, Adeline Brown, Willie Byars, Onita Carnes, Thomas and Lizzie Frazier, Jerry Imotichey, Mary James, William Pettigrew, Lee Fannie Roberts, Mary Ella Russell, Thomas Underwood, Jimmie Walker, and Catherine Willmond. Also, thanks to Catherine Willmond for her insightful linguistic advice and suggestions. Last but certainly not least, special thanks to Catherine Willmond and Onita Carnes for assisting in the preparation of the word lists and for providing invaluable logistics assistance. This work was supported by NSF grant SBR 9511118.

References

Cho, Taehong and Peter Ladefoged
1997 Variations and universals in VOT: Evidence from 17 endangered languages. UCLA Working Papers in Phonetics 95: 18-40

Crothers, John
1978 Typology and universals of vowel systems. *In* Universals of Human Language, volume 2: Phonology, 93-152, edited by Joseph Greenberg. Stanford: Stanford University Press.

Disner, Sandy
1980 Insights on vowel spacing: results of a language survey. UCLA Working Papers in Phonetics 50: 70-92.

Farnetani, Edda and Shiro Kori
1984 Effects of syllable and word structure on segmental durations in spoken Italian. *Quaderni del Centro di Studio per le Ricerche di Fonetica* 3: 143-88.
1986 Effects of syllable and word structure on segmental durations in spoken Italian. Speech Communication 5: 17-34.

Fintoft, Knut
1961 The duration of some Norwegian speech sounds. Phonetica 7: 19-39.

Ghai, Ved Kumari
1980 Contributions to Dogri phonetics and phonology. Annual Report of the Institute of Phonetics 14. University of Copenhagen: 31-94.

Gibson, Arrell
1971 The Chickasaws. Norman: University of Oklahoma Press.

Golla, Victor
1970 Hupa Grammar. Ph.D. Dissertation, University of California, Berkeley.

Gordon, Matthew
1995 Acoustic correlates of stress in Estonian. Ms. UCLA.

Gordon, Matthew
1996 The phonetic structures of Hupa. UCLA Working Papers in Phonetics 93: 164-187.

Han, Mieko
1994 Acoustic manifestations of mora timing in Japanese. Journal of the Acoustical Society of America 96: 73-82.

Ladefoged, Peter and Daniel Everett
1997 Phonetic structures of Banawá, an endangered language. Phonetica 54: 94-111.

Lahiri, Aditi and Jorge Hankamer
1988 The timing of geminate consonants. Journal of Phonetics 16(3): 327-38.

Lehiste, Ilse
1966 Consonant Quantity and Phonological Units in Estonian. Indiana University Press: Bloomington

Lehiste, Ilse
1970 Suprasegmentals. Cambridge: MIT Press.

Lehtonen, Jaakko
1970 Aspects of Quantity in Standard Finnish. Jyväskylä: K.J. Gummerus Osakeyhtiön Kirjapaino.

MacEachern, Margaret, Barbara Kern and Peter Ladefoged.
1996 Wari' phonetic structures. Journal of Amazonian: 3-28.

Maddieson, Ian
1997 Phonetic universals. In Handbook of Phonetic Sciences, edited by William Hardcastle and John Laver, 619-39. Cambridge, MA: Blackwells.

McDonough, Joyce and Martha Austin-Garrison
1994 Vowel enhancement and dispersion in the vowel space of Western Navajo: A study of traditional Navajo speakers. UCLA Working Papers in Phonetics 87: 93-104.

Meyer, E. A. and Gombocz, Z.
1908 Zur phonetik der ungarischen sprache. Le Monde Oriental 2(2): 122-87.

Miller, Ann
1987 Phonetic characteristics of Levantine Arabic geminates with differing morphemes and syllable structures. Ohio State University Working Papers in Linguistics 36: 120-40.

Munro, Pamela
1987 Some morphological differences between Chickasaw and Choctaw. *In* Muskogean Linguistics, edited by Pamela Munro, 119-133. UCLA Occasional Papers in Linguistics 6.

1996 The Chickasaw sound system. Ms. UCLA.

in press Chickasaw. In Native Languages of the Southeastern United States, edited by Heather Hardy and Janine Scancarelli. Lincoln: University of Nebraska Press.

Munro, Pamela and Charles Ulrich.
1984 Structure-preservation and Western Muskogean rhythmic lengthening. West Coast Conference on Formal Linguistics 3: 191-202.
Munro, Pamela and Catherine Willmond
1994 Chickasaw: An Analytical Dictionary. Norman: University of Oklahoma Press.
Munro, Pamela, Matthew Gordon and Peter Ladefoged.
in prep. The phonetics and phonology of Chickasaw Rhythmic Lengthening.
Ojamaa, Koit
1976 Temporal Aspects of Phonological Quantity in Estonian. Ph.D. Dissertation, University of Connecticut.
Orešnik, J. and M. Pétursson
1977 Quantity in Modern Icelandic. Arkiv för Nordisk Filologi 92: 155-171.
Sapir, Edward and Morris Swadesh
1939 Nootka Texts. Philadelphia: Linguistic Society of America.
Sheldon, Steven
1974 Some morphonemic and tone perturbation rules in Mura-Pirahã. International Journal of American Linguistics 40: 279-82.

Some phonetic structures of Aleut

Taehong Cho, Alice Taff, Moses Dirks, and Peter Ladefoged

1. Introduction.

Aleut, the language of the peoples of the Aleutian and Pribilof Islands, Alaska, has two main dialects, Eastern Aleut (own name Unangan) and Western Aleut (own name Unangas). Bergsland (1994) notes that: 'Today Eastern Aleut is represented in six villages by perhaps 400 active speakers. All of these are forty or more years old and all are fluen8akers of Eastern Aleut were recorded in St. Paul and St. George, in the Pribilof Islands (see figure 1). Western Aleut 'is spoken by sixty to eighty Atkan people, including some children, who are all also fluent speakers of English.' (Bergsland 1994, xvi). Our speakers of this dialect were recorded in Atka.

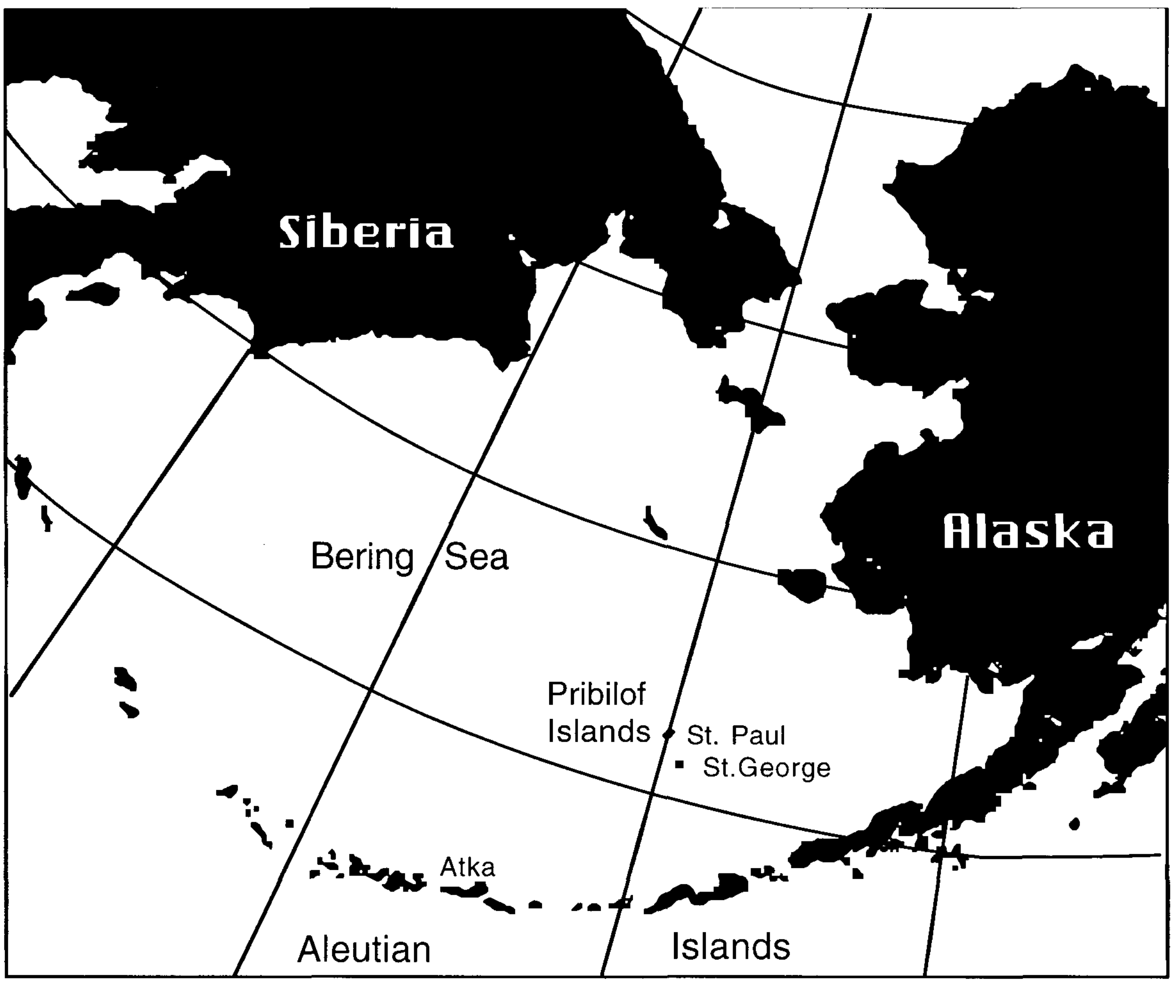


Figure 1, The location of speakers of the two dialects of Aleut. Speakers of Eastern Aleut were recorded in St. Paul and St. George, and of Western Aleut in Atka.

The research reported here considered only the original sounds of Aleut. In the second half of the eighteenth century Russia conquered and exploited the Aleutian islands, and from that time on many Russian words have entered the language. From around 1867, when the United States purchased Alaska (largely for the Pribilof fur trade), English has also been a source of loan words. Loan words were used in our investigation of the sounds of Aleut, but only when they contained no sounds that were not in the original Aleut inventory. Thus we have used words such as **tu:tkaX** *tuutkax̂* 'aunt' from Russian *tyótka,* but not **su:paX**, *suupax̂* 'soup' from English *soup,* because there are no bilabial plosives in the original Aleut inventory.

Both dialects of Aleut have three vowels, **i**, **a**, **u**, each of which can be long or short. In Western Aleut there are 23 consonants which can be roughly categorized as shown in the following chart. Further details are given later in this paper in the consonant section. Eastern Aleut has a simpler system in that the contrasts between voiced and voiceless nasals, sibilants and approximants have been largely lost.

	BILABIAL	DENTAL	ALVEOLAR	PALATO-ALVEOLAR	VELAR	UVULAR	LABIAL VELAR
STOP		**t**		**tʃ**	**k**	**q**	
NASAL	**m̥ m**	**n̥ n**			**ŋ̊ ŋ**		
FRICATIVE		**ð**	**s z**		**x ɣ**	**χ ʁ**	
LATERAL		**l̥ l**					
APPROXIMANT				**ç j**			**ʍ w**

In the Aleut orthography, these consonants are as follows.

	BILABIAL	DENTAL	ALVEOLAR	PALATO-ALVEOLAR	VELAR	UVULAR	LABIAL VELAR
STOP		t		ch	k	q	
NASAL	hm m	hn n			hng ng		
FRICATIVE		d	s z		x g	x̂ ĝ	
LATERAL		hl l					
APPROXIMANT				hy y			hw w

1.1. Data collection

A list of words was developed that illustrated each consonant followed by each vowel, as shown in Appendix A. As far as possible, disyllabic, or, occasionally, monosyllabic words were used. Stress in Aleut is usually on the penultimate syllable, so this ensured that the consonant in question was in initial position in a stressed syllable. For each of the voiceless stops except **q** both long and short vowels were included. The sibilants **s** and **z**, contrast in comparatively few words in Western Aleut (many of the words with **z** being Russian loans) and virtually not at all for most speakers of Eastern Aleut. Some of the non-sibilant fricatives do not contrast in initial position, and accordingly the list included words showing these contrasts in medial and final position, special attention being paid to the contrast between velar and uvular fricatives. The semivowels and their voiceless counterparts have restricted distributions; **w** and **ʍ** occur only before **a**, and **j** and **ç** occur only before **a** and **i**. These sounds were included in medial position. For the vowels, pairs of words were chosen so as to show differences in vowel length, with position in word, stress, and consonantal context being kept constant.

In addition to the wordlist, material was included to demonstrate further aspects of the Aleut stress and intonation system. Speakers were asked to produce sentences as statements, and then to repeat them as questions. There were also sentences differing mainly in the location of the clause breaks, and in the degree of emphasis expected on each word. This material is discussed by Taff (1997).

At the end of each recording the speaker was asked to talk for a few minutes in Aleut, and afterwards to say the same thing in English. Some speakers told short stories and others gave only a couple of sentences. These unrehearsed pieces were transcribed and included as part of each speaker's recorded data.

A total of 24 speakers were recorded, 6 men and 6 women being speakers of Western Aleut, and 4 men and 8 women being speakers of Eastern Aleut. With the exception of three of the Western Aleut women, all the speakers were fluent native speakers. The three younger Western Aleut women had learned Aleut as children in school, but did not speak it often. Each speaker was recorded individually, usually using a DAT recorder and a noise-canceling close-talking microphone. Five of the speakers of Western Aleut had to be recorded with a back up tape recorder. The frequency response was flat (± 2dB) from 70 to 10,000 Hz, but the achieved Signal/Noise ratio was estimated to be only 35 dB.

The general procedure was for speakers to be told the word in English and then asked to repeat it twice in Aleut. When speakers did not know the word required, they were prompted in Aleut. If they were completely unfamiliar with it, the word was omitted. Speakers of Eastern Aleut were not asked to say the words designed to illustrate contrasts not present in their language (e.g. those with orthographic *hm, hn,* etc.).

1.2. Segment frequency

As an aid in the production of the word list, the main entries in a small dictionary of Eastern Aleut designed for use in schools (Bergsland and Dirks, 1978) were entered into a computer. Only words that contained segments that were in the original Aleut phonological inventory were included. In all, there were 8744 segments in 1194 words. In addition to providing a useful means of finding appropriate minimal contrasts, this procedure provided useful information on the relative frequency of the segments in Eastern Aleut.

The relative frequencies of each of the vowels is shown in Table 1. Overall, the vowels are clearly ranked in frequency, with **a** and **i** being much more common than **u**. Long vowels are comparatively uncommon, constituting only 10% of all vowels.

Table 1. The percentage frequencies of the vowels in Eastern Aleut.

SEGMENT	INITIAL	ALL
a	48	39
aa	6	6
i	23	34
ii	1	2
u	21	17
uu	1	2
TOTAL	100	100

The results for the consonants are shown in Table 2. The total numbers for **x**, **χ** and (to a lesser extent) **ŋ** and **n** are inflated by the fact that theseare the only segments that occur in final position in Eastern Aleut; they are part of suffixes such as **-lix** indicating the infinitive forms of the verbs, and the suffixes **-χ**, **-x** that are part of the singular and dual morphemes used for nouns in dictionary citations. The total for **l** is also augmented by these occurrences. In the percentage counts, the examples of final **x**, **χ** and **ŋ**, and **l** in a suffix have been disregarded, making the total number of consonants considered for this purpose 3464 instead of 4900.

Table 2. Relative frequencies of the consonants of Eastern Aleut. The parenthesized figure for **l** indicates the number of occasions this segment was part of the suffix **-lix**,

SEGMENT		INITIAL	FINAL	TOTAL	% INITIAL	% TOTAL
n	n	16		393	2.0	11.3
s	s	112		389	16.0	11.3
g	g	3		346	.5	10.0
l	l	28	(396)	689	4.0	8.5
t	t	82		274	12.0	7.9
k	k	97		250	14.0	7.2
q	q	163		214	23.0	6.2
tʃ	ch	125		203	18.0	5.9
ð	d	13		197	2.0	5.7
m	m	38		194	5.0	5.6
x	x	1	543	692	.1	4.3
ɣ	ĝ	0		142	0.0	4.1
χ	x̂	5	488	615	.5	3.7
j	y	11		116	1.0	3.3
ŋ	ng	2	9	115	.3	3.1
n̥	hn	0		25	0.0	.7
m̥	hm	0		18	0.0	.5
l̥	hl	0		15	0.0	.4
w	w	7		13	1.0	.4
TOTAL		703	1436	4900	99.4	100.1

It is apparent from this dictionary count that some consonants have a more marginal status than others. In fact, none of the words listed as containing **m̥**, **n̥**, **l̥** 'hm, hn, hl′ were pronounced in that way by all our speakers of Eastern Aleut. (Some speakers occasionally used a slight breathy voice in some of these words.) The segment **w** is also noticeably rare in a dictionary count, but the words it occurs in are all in frequent use. The segment **j**, although slightly more common than **w**, occurs almost only before **a**.

2. Vowels

As we have noted, Aleut contains a basic three vowel system, including a front high vowel **i**, a mid low vowel **a**, and a back high vowel **u**, each having its long counterpart as below. Each vowel has a long and short contrast, thus yielding six contrastive phonemes in the system.

i	**u**	**iː**	**uː**
a		**aː**	

There are no vowel sequences within a syllable (other than long vowels, which could potentially be considered as geminates, and are in fact treated as doubled vowels in the orthography; with this in mind we have transcribed long vowels as doubled vowels). The vowels of Eastern Aleut are illustrated in Table 3. As we have noted, some of the vowels are much rarer than others, making it difficult to find good minimally contrasting sets.

Table 3. Words exemplifying contrasts among Eastern Aleut vowels after voiceless stops at the three places of articulation.

DENTAL			VELAR			UVULAR		
ti	**tiχlaχ**	bald eagle	**ki**	**kiɣiχ**	to bite	**qi**	**qiqiχ**	slime
tii	**tiistaχ**	dough	**kii**	**kiikaχ**	cranberry bush, cake	**qii**	**qiiɣaχ**	grass
ta	**taniχ**	forehead	**ka**	**kaðan**	in front of the speaker	**qa**	**qaqaχ**	food
taa	**taaŋaχ**	water	**kaa**	**kaaŋuχ**	healthy	**qaa**	**qaaðan**	little fish
tu	**tukuχ**	chief, boss	**ku**	**kukaχ**	grandmother	**qu**	**qumaχ**	white
tuu	**tuutkaχ**	aunt	**kuu**	**kuuskaχ**	cat			

The phonetic qualities of the vowels were examined, based upon values for the first three formants which were measured using the Kay Elemetrics Computerized Speech Lab (CSL). Two repetitions of each of the words listed in Table 4 were recorded by five female and four male speakers of Eastern Aleut. (Three other female speakers were recorded but did not produce all the required examples.) Data were transferred into the computer at a sampling rate of 10 kHz. A steady state portion of the vowel for each token was centered at about the mid point of its duration from which superimposed LPC and FFT spectra were calculated with 30 ms and 25.6 ms frames, respectively. The formant values were usually determined from the LPC spectra (with 12 or 14 coefficients), using the FFT spectra (and sometimes formant history tracking) as supplementary checks.

The best words in Table 3 for comparing the quality of the vowels, and noting any differences between long and short vowels, are those in the velar column, where the vowels are preceded by **k** and followed by a velar or coronal consonant. Figure 2 shows the relative positions of the first two formants of these vowels. These plots were drawn with the UCLA Plot Formants software. The scales are in accordance with the Bark scale, and the ellipses are drawn with radii of two standard deviation along the axes of the first two principal components of the distribution.

The distribution of the vowels in Figure 2 supports the form of dispersion theory (Liljencrants and Lindblom 1972) which predicts that, in a given system, contrastive vowels are spaced with a sufficient contrast (Lindblom 1986; Lindblom and Maddieson 1988). The vowels in Figure 2 are fairly well in the corners of the vowel space. There is, however, an interesting difference in the height of the front and back vowels. An analysis of variance, with F1 as the dependent variable and vowel quality as the dependent variable, showed that F1 was significantly lower for **i** than for **u** ($p = .0052$ for the male speakers, $p < .0001$ for the female speakers)

An analysis of variance also showed that the only differences in vowel quality attributable to length were for the vowel **a**. There is a significant difference in F2 for long and short **a**, for both men and women, and in the F1 for men (in all cases $p < .01$). The long and short high vowels were not significantly different in any way. It might be thought that some of the difference between long and short **a** could be attributable to the particular words used in the analysis, the short **a** having a following dental consonant, and the long **aa** a following velar consonant. Vowels tend to become fronted after laminal dental consonants in which the tongue blade is raised. However, our auditory impression of Aleut in general is that **a** has a variable but generally mid central quality that could be symbolized by **ə**, whereas long **aa** always has a low central **a** quality. Taff (1992:35) suggests that **a** should be regarded as a featureless vowel.

Taff (1992) has discussed variations in Aleut vowel quality due to context. Our findings are in substantial agreement with hers. There are some differences in vowel quality due to the place of articulation of the preceding consonant. These differences can be demonstrated by reference to the long vowels **i** and **a**, and the short vowel **u** (there are no long **u** vowels after uvular consonants). Figure 3 shows the formant plots for these vowels preceded by stops at each of the three places of articulation. The syllable in which each vowel occurs is also shown, so that possible effects of the following consonant may also be taken into account.

Analyses of variance with F1 or F2 as the dependent variable and place of articulation as the independent variable produced the following results, using Fisher's PLST for post hoc tests. For the male speakers there was no difference in the mean position of the vowel **i** when it is preceded by **t** or **k** (and the following consonant in these words also has no effect), but when **i** is preceded by **q** there is a significant raising of F1, so that the vowel appears lower in the vowel space ($p < .0001$). The high back vowel **u** after **q** also tends to be lower (F1 is raised) in comparison with when it occurs after **t** ($p = .0362$). More strikingly, **u** is more front (F2 is raised) when preceded by **t** than when preceded by **k** ($p < .001$), and there is a strong tendency for it to be more back when preceded by **q** in comparison with when preceded by **k** ($p = .0373$). The low vowel **a** is unaffected by the preceding consonant. The female speakers behaved in a similar way. There were no significant differences in F1 for **i** , but the tendency for this vowel to be lower after **q** can be seen. In the case of the back vowel **u** all three variants differed significantly from each other (p < > .01).

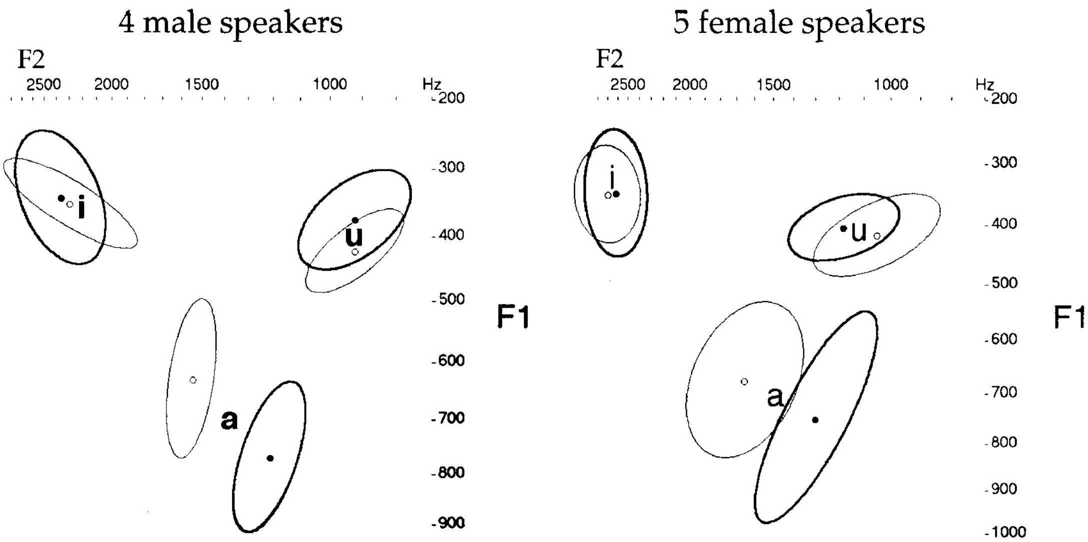


Figure 2. Formant plots of long and short vowels after **k** in the words in Table 3 as produced by four male speakers and five female speakers. The light ellipses enclose short vowels, the centers being marked with open circles, and the heavier ellipses enclose long vowels, the centers being marked with solid points.

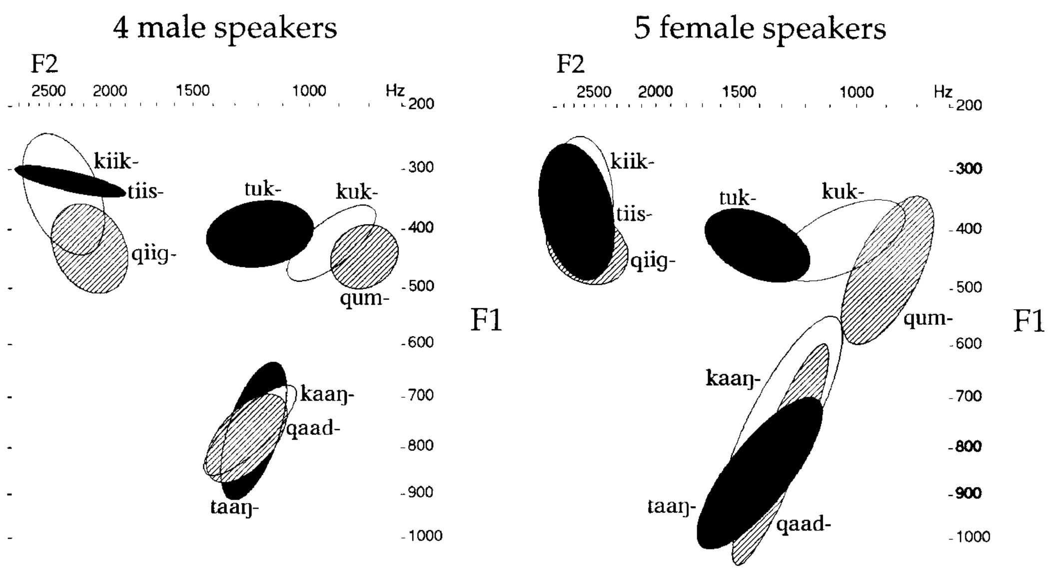


Figure 3. The effect of the preceding consonant on vowel quality. So that the influence of the final consonant may also be considered, the syllables in which the vowels occurred are shown.

We have only a limited amount of data enabling us to test the influence of the following consonant, but we can note the effect of a uvular consonant after **i** and **a** by considering another set of words from Table 3. As can be seen in Figure 4, the vowel **i** in **tiχlaχ** (bald eagle) is very considerably lower than the corresponding vowel in **kiɣiχ** (to bite); and the **a** in **qaqaχ** (food) is further back than corresponding vowel in **kaðan** (in front of the speaker). For both male and female speakers all these differences are significant ($p < .01$). In general, uvular consonants have the greatest affect on the quality of adjacent vowels, and following uvulars probably have a greater effect than preceding uvulars (but we do not have appropriate data to verify this statistically).

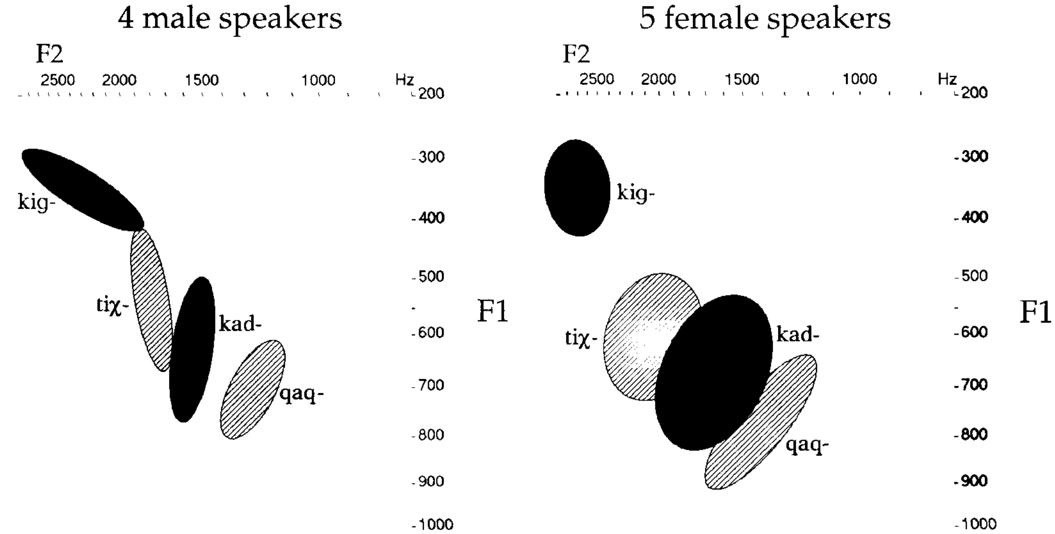


Figure 4. Further effects of consonantal context on vowel quality. So that the influence of the final consonant may also be considered, the syllables in which the vowels occurred are shown.

3. Consonants

The five places of articulation used by Aleut consonants have a somewhat unequal functional load. The only bilabial consonant are the nasals, **m̥** and **m**. The consonants listed in the dental column (see above) actually have varied places of articulation. The stop, nasal and lateral, **t**, **n**, **l**, usually have a laminal articulation, although sometimes, particularly with **l** before **u**, as in **lulix**, 'to believe', the articulation is more apt to be apical and a little further back. The segment **ð** was usually an interdental fricative, but in initial position, where it occurs mainly in Russian loan words, it is often a stop. It is also a stop before a nasal, as in **ajəɣa:dn** '***'. There is no contrast between **s** and **ʃ** in Aleut (some speakers commented that they found this distinction hard in English), and the Aleut sibilant **s** varied from alveolar to almost palato-alveolar as in English. The uvular consonants were made very much at the back of the oral cavity. The voiceless velar and uvular fricatives occur only in medial and final position. In Eastern Aleut they are the only fricatives that can occur in final position. In Western Aleut **s** occurs as a plural marker (instead of **n**), so that it, too, can occur finally. In this section, our phonetic investigations on consonants will be limited to Voice Onset Time (VOT) and characteristics of burst spectra in both Eastern and Western dialects of Aleut.

3.1. Voice Onset Time

This section describes the VOT of stops in word initial position in an effort to investigate how VOT varies according to (1) different dialect (Eastern vs. Western), (2) context of following vowel (short vs. long) and (3) gender difference. In addition, special attention is paid to the variability of VOT in terms of place of articulation. It has been agreed that there is a general tendency for VOT to be longer when the closure for a stop is made further back in the vocal tract (Fischer-Jørgensen 1954, Peterson and Lehiste 1960). Our working hypothesis is that if the variability of VOT is accounted for only by the distance between the open end of the vocal tract and the source of compression, the VOT for a uvular stop will tend to be longer than one for a velar stop.

Table 6. A list of words for Eastern and Western Aleut.

	EASTERN ALEUT		WESTERN ALEUT	
ti	**tiχlaχ**	bald eagle	**tiglaχ**	bald eagle
tii	**tiistaχ**	dough	**tiistaχ**	dough, piecrust
ta	**taniχ**	forehead	**taniχ**	forehead
taa	**taangaχ**	water	**taangaχ**	water
tu	**tukuχ**	chief, boss, rich	**tukuχ**	chief
tuu	**tuutkaχ**	aunt	**tuutχaχ**	stalk of cow parsnip
ki	**kiɣiχ**	to bite	**kiɣil**	to bite
kii	**kiikaχ**	cranberry bush, cake	**kiin**	who
ka	**kaðan**	in front of the speaker	**kaðan**	in front of the speaker
kaa	**kaanguχ**	healthy	**kaanguχ**	healthy
ku	**kukaχ**	grandmother	**kukaχ**	grandmother
kuu	**kuuskaχ**	cat	**kuusxiχ**	cat
qi	**qiqiχ**	slime	**qiqiχ**	slime
qii	**qiiɣaχ**	grass	**qiiɣaχ**	grass
qa	**qaqaχ**	food	**qalɣadaχ**	food
qaa	**qaaðan**	little fish	**qaaðas**	dolly varden (fish)
qu	**qumaχ**	white	**qum̥aχ**	white

The data recorded in DAT tapes were down-sampled to 20,000 Hz, being quantized through the Kay Elemetrics Computerized Speech Lab (CSL). A total of 816 tokens (24 speakers x 17 words x 2 repetitions) were digitized, 775 of which were measured. Some tokens could not be analyzed due to a sudden background noise or speech errors. The VOT for each token was measured from the release of the consonant to the onset of the first formant of the following vowel. The data were statistically analyzed by T-tests and one-factor or two-factor ANOVAs.

VOT Variability . As shown in Table 7 and Figure 5, the VOT's for alveolar stops tend to be longer than the other two stops in both dialects. One-factor analyses of variance of each dialect reveal that the effect of place was significant ($F[2, 359]=34.358$, $p <.0001$ for Eastern Aleut and $F[2,390]=23.871$, $p <.0001$ for Western Aleut). In post-hoc analyses, while the alveolar plosives were distinct from the other two at $p <.0001$ in both dialects, there was no significant between velar and uvular plosives in either dialect. The results for the alveolar stops are in accord with the general tendency that the further forward a stop is made, the shorter the VOT (Fischer-Jørgensen, 1954). However, this tendency does not hold between velar and uvular stops: in the case of Western Aleut, it is reversed in that the velar plosive has a longer mean VOT.

Table 7. Mean VOT of Eastern and Western Aleut Stops in ms.

	Eastern Aleut				Western Aleut			
	Alveolar	Velar	Uvular	Mean	Alveolar	Velar	Uvular	Mean
Speaker	11	11	11		12	12	12	
Token	132	130	100		139	136	118	
VOT	**59**	**75**	**78**	70.7	**76**	**95**	**92**	87.7
Std. Dev...	19	20	24		23	24	27	

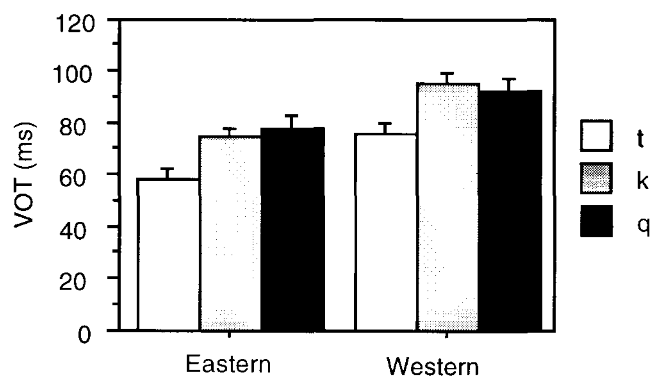


Figure 5. Mean Voice Onset Time (s) by categories of stop and dialect for 24 speakers.

With regard to dialect differences, stops in Western Aleut have about 124% longer VOT's than those in Eastern Aleut. The data were submitted to the Analysis of Variance with the factor of Dialect (Eastern vs. Western). The result of the one-way ANOVA reveals that the difference in VOT between these two dialects is highly significant at $p <.0001$.

One of the interesting acoustic events which may group the velar and uvular plosives together, distinguishing them from the alveolar plosives, is that both velar and uvular plosives frequently have multiple bursts at the release. Quite a few triple bursts as well as double bursts were found in the spectrograms of both the velar and uvular stops. In addition, the mean VOTs for all three stops are quite long, falling into the category of voiceless aspirated stops (Cho and Ladefoged, 1997). The VOTs for alveolar and velar stops are close to the VOTs reported for aspirated stops in English and Cantonese by Lisker and Abramson (1964).

Effect of the Vowel Context (short vs. long vowel) on VOT: The mean VOT of the plosives before a long vowel is about 120 % longer than that before a short vowel both in Eastern and Western Aleuts. The summary is given in Table 8 and Figure 6. The result of unpaired t-tests for VOT with a grouping variable of vowel context (long vs. short) shows a significant main effect of the following vowel context (p <..001 for Eastern Aleut , p<.0001 for Western Aleut). This suggests that the VOT may be considered as some percent of the vowel, so that when the vowel is long the VOT becomes long (Port and Rotunno, 1979).

Table 8. The Voice Onset Time (ms) affected by the following vowel length. The mean values are averaged across the place of articulation and gender. Values in parenthesis are standard deviations.

	EASTERN		WESTERN	
	LONG	SHORT	LONG	SHORT
COUNT	175	187	188	205
MEAN VOT	74 (22)	66 (22)	96 (25)	80 (25)

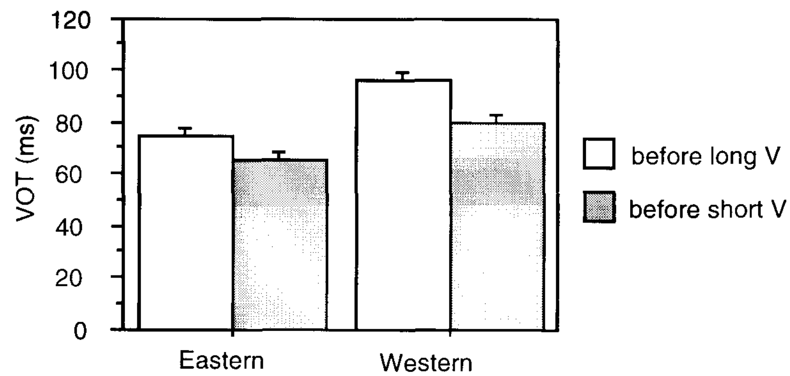


Figure 6. VOT affected by the following phonemic vowel length.

This effect holds irrespective of the place of articulation. Two-way ANOVA's with independent variables of vowel length and place of articulation were conducted to examine the interaction between the two factors. The result shows that there is no significant interaction (F [2,356]=1.131, p=.3239 for Eastern Aleut and F[2,387]=.282, p =.754 for Western Aleut), suggesting that the duration of VOT before a long vowel is longer than that before a short vowel regardless of place of articulation. (Note that the effect of vowel length appears to be greater for Western Aleut.)

Effect of Gender: Female speakers in Eastern Aleut produce stops with about 130% longer VOT than male speakers. However no effect of gender is observed in Western Aleut. The result of an unpaired t-test for VOT with a grouping variable of gender as shown in Table 9 indicates that the difference in Eastern Aleut is significant at p<.0001 while it is not in Western Aleut (p>.05). A three-way ANOVA showed that the effect of gender on VOT in Eastern Aleut is independent of the place of articulation and vowel length. In other words, in Eastern Aleut, female speakers have longer VOT's than male speakers, even when the data are separated by either place of articulation or vowel context (long vs. short) or both.

Table 9. Mean VOT (ms) separated by gender.

	EASTERN		WESTERN	
	MALE	FEMALE	MALE	FEMALE
COUNT	132	230	198	195
MEAN VOT	61 (19)	75 (23)	88 (26)	87 (26)
T-TEST	$t(360) = -6.162$, $p < .0001$		$t(391) = .748$, $p > .05$	

3.2. Burst Spectra

In examining the burst spectra for consonants, there are several suggestions in the literature. Jakobson, Fant, and Halle (1963) suggest that "the essential articulatory difference between the compact and diffuse phonemes lies in the relation between the volume of the resonating cavities in front of the narrowest stricture and those behind this stricture (1963:27)." In their theory, the ratio of the volume of the front cavity to that of the back cavity is higher for the compact spectrum than for the diffuse spectrum.

According to Stevens (in press), the spectrum of velar stops is less flat (more compact) compared with that of bilabial or alveolar stops, in that when the rate of change of cross-sectional area is smaller and the length of the constriction is greater (as it is in the production of a velar stop compared with alveolar and bilabial stops) the formants are moving relatively slower, thus leading to more compact spectral peaks. Stevens suggests that in the production of velar stops, the spectral peak of the burst is usually in the second-formant range, around 1500 Hz.

Keating and Lahiri 1993 note that the spectral peak frequency location for a velar stop largely depends on (1) the resonance frequency of the cavity in front of the constriction and (2) the following vowel context (cf. Fant 1960). For example, the velar stop before the front vowel **i** is associated with a higher frequency peak than before other vowels, because the tongue fronting in the production of **i** will reduce the size of the front cavity, leading to a higher resonance frequency.

Based upon these studies, we set up two hypotheses in analyzing the data. First, the frequency location of the spectra peak is higher for a velar stop than for a uvular stop because the front cavity for the velar stop is smaller, which leads to a higher-frequency spectral peak (Keating and Lahiri 1993) (Hypothesis I). Second, the burst spectrum is more compact for a uvular stop **q** than a velar stop **k**, because the ratio of the volume of the front cavity to that of the back cavity is higher for a uvular stop than for a velar stop, which results in more compactness (Jakobson et al., 1963) (Hypothesis II). In addition, the effect of the following vowel quality on the relative frequencies of the primary peak and the diffuseness will be discussed.

Method: Measurements were taken from the tokens that were used for the VOT measurements for 4 male and 7 female speakers in Eastern Aleut and 2 male and 4 female speakers in Western Aleut. The tokens recorded on analog backup tapes were not used as the high frequency components were not well captured. Using the Kay Computerized Speech Lab system, a 512 point (25.6 ms) window was centered around the burst transient and the FFT spectrum of this window calculated and smoothed. In cases of multiple bursts, the window was centered around the darkest burst. Then, numerical values in dB at each point were averaged for all tokens, separated by vowel quality **i** vs. **a** vs. **u**), gender (male vs. female) and Dialect (Eastern vs. Western). Long and short vowels were not distinguished when pooling the data.

Results and Discussion: The data were analyzed in terms of (1) the relative frequencies of the primary spectral peak (2) compactness vs. diffuseness. Figures 7 and 8 show that mean burst spectra for **k** and **q** before **i**, **a**, **u**. for Eastern and Western Aleut respectively. In general, the principal spectral peak is higher in frequency for **k** than for **q**, supporting Hypothesis I: that is, the effect can be accounted for by the difference in the cavity size and its resonance in front of the constriction (see Keating and Lahiri 1993).

Spectra in the figures can, in general, be characterized as somewhat 'compact' with a prominent spectral peak at low or mid frequencies (somewhere between 500 Hz and 3000 Hz) depending on the following vowel context (cf. Jakobson, Fant and Halle 1963). In addition, it appears that the spectra for female speakers in both dialect are generally flatter (more diffuse) than those for male speakers. However, no striking differences in the relative degree of diffuseness (or compactness) between **k** and **q** are observed, working against Hypothesis II: the higher the ratio of the volume of the front cavity to that of the back cavity, the more compact the spectrum. One of the possible explanations for this might be that the difference in the ratio between a velar and a uvular is not as great as that between a bilabial (or an alveolar) and a velar. which were the objects of the earlier comparisons. However, if we assume that compactness comes from slower change of cross-sectional area and greater constriction length as suggested by Stevens, we may posit that both the velar and uvular stops are produced with no significant difference of constriction length and rate of the articulatory movement. The bandwidth of the primary spectral peak does not show a systematic difference between **k** and **q** in either dialect.

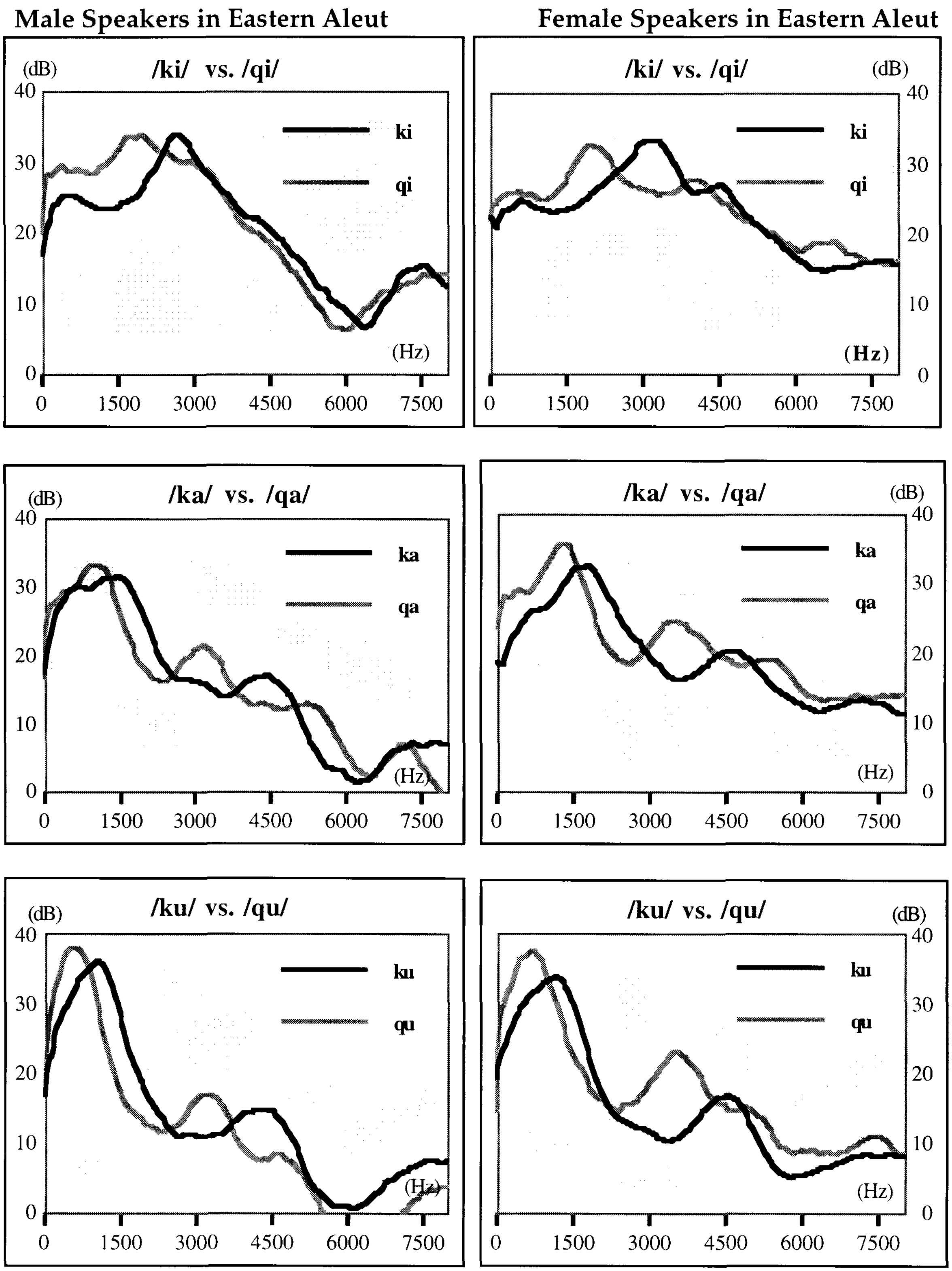


Figure 7. Mean burst spectra of word initial **k** and **q** before, **a**, and **u** in Eastern Aleut. Data from 4 male speakers and 7 female speakers.

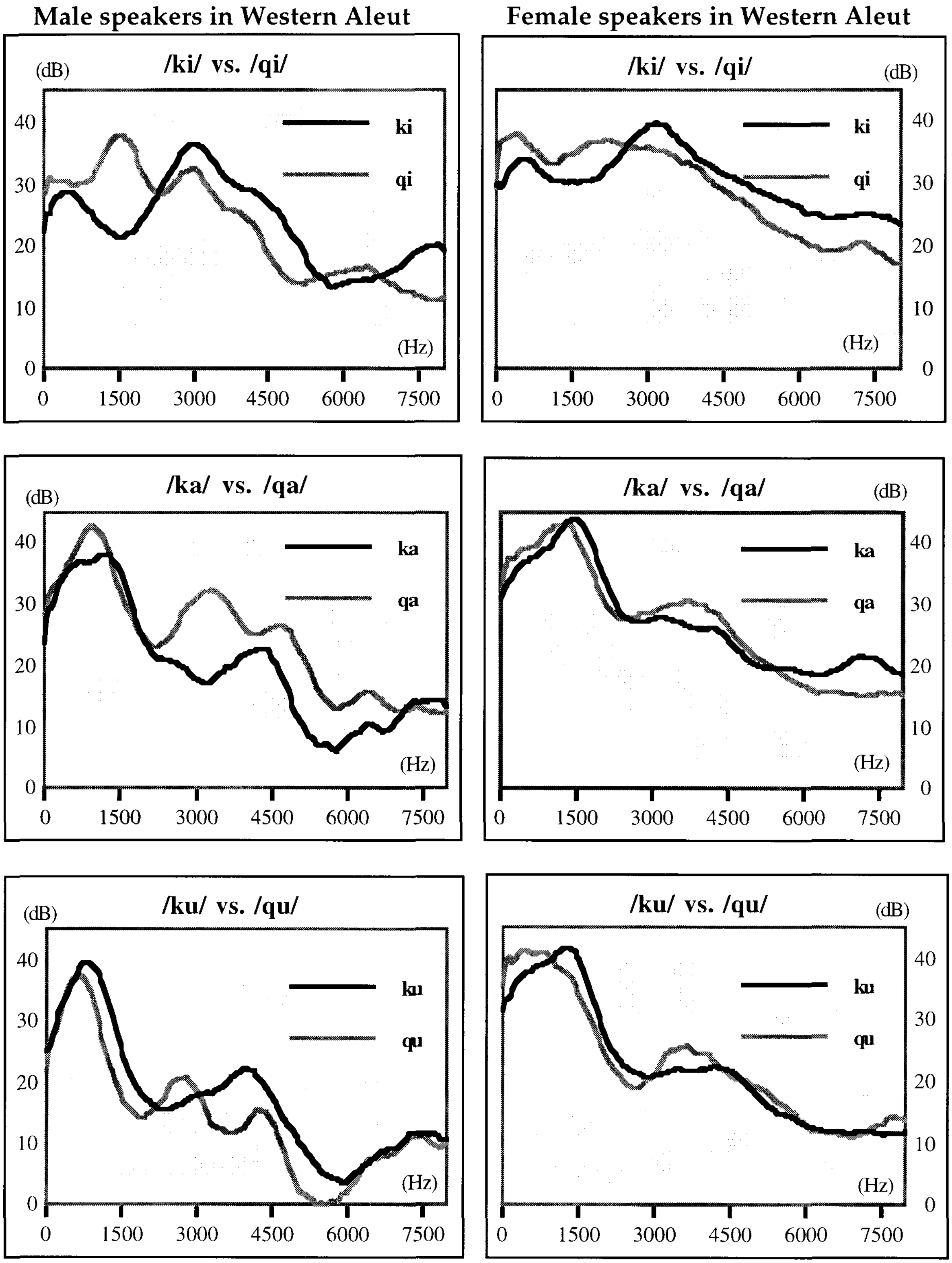


Figure 8. Mean burst spectra of word initial **k** and **q** before, **a**, and **u** in Western Aleut. Data from 4 male speakers and 7 female speakers.

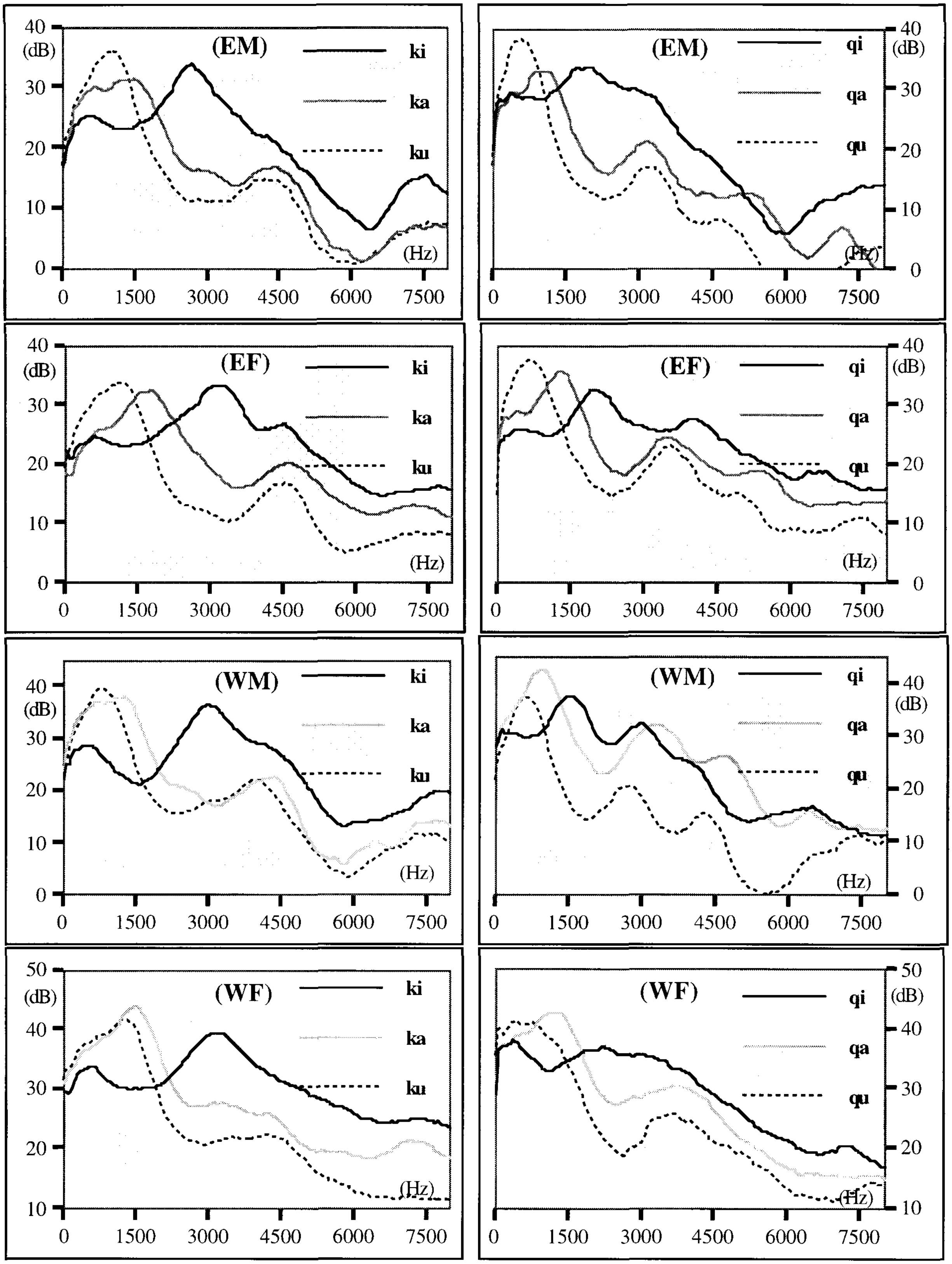


Figure 9. Effect of the following vowel context on Burst Spectra.

As discussed earlier, the frequency location of a prominent spectral peak largely depends on the following vowel. Figure 9 presents data showing that Aleut is in accordance with the earlier studies; that is, the frequency location of the spectral peak is higher before **i** than before **a** or **u**. This effect is mainly accounted for by the assumption that the consonant peak is determined by the front cavity resonance (Keating and Lahiri 1993). The tongue body is raised and somewhat fronted when **k** or **q** is produced followed by the front vowel /i/, so that the front cavity becomes shorter compared with cases for back vowels **a** and **u**. This leads to a higher frequency location of the spectral burst. In addition, the consonant peak is slightly higher before the unrounded vowel **a** than before the rounded vowel **u**, as can be observed in Figure 9. Here again, this effect is presumably due to the fact that the rounded **u** is produced with lip protrusion, which results in a relatively longer front cavity as compared with the unrounded **a**.

Figure 9 also shows that the spectrum is relatively more diffuse (flatter) when associated with the following vowel **i**. This effect can possibly be explained by the analysis of Jakobson et al. (1963) that the ratio of the volume of the front cavity to that of the back cavity is higher for the compact spectrum. As the volume of the front cavity before **i** is smaller than before either of the vowels **a** or **u**, the ratio is lower, leading to the relatively diffuse spectrum. Likewise, we also observe that the spectrum of the burst before **u** is slightly more compact than that before **a**, due to the rounding effect which increases the ratio of the volume of the front cavity to that of the back cavity.

3.3. Summary of stop consonant differences

The VOT analyses show that VOT varies according to (1) place of articulation, (2) dialect, (3) the following vowel duration and (4) gender. In general, VOT is shorter for alveolar than for velar or uvular. However, the difference in VOT between velar and uvular was not statistically significant. With regard to dialect difference, stops in Western Aleut have about 124% longer VOT's than those in Eastern Aleut. The effect of the following vowel duration on VOT is also significant: VOT is longer before the long vowel than before the short vowel, suggesting that the VOT may be considered as the part of the vowel. Finally, we observed a gender difference on VOT in Eastern Aleut but not in Western Aleut. The overall VOT study also suggests that the stops in Aleut belong to the category of voiceless aspirated stops reported in Lisker and Abramson (1964).

Illustrative spectrograms for **k** and **q** are shown in Figure 10. Qualitative comparison of the burst and frication noise between the velar and the uvular suggests that not only more overall noise energy of frication but also more energy in the low frequency range serve as one of the significant cues for native speakers of Aleut to distinguish uvular from velar plosives. The difference between the velar stop **k** and the uvular stop **q** can be mainly characterized by the different frequency location of spectral peak: the velar stop is associated with a higher frequency spectral peak as compared with the uvular stop (supporting Hypothesis I). However, no striking difference in the degree of diffuseness between **k** and **q** was observed (working against Hypothesis II). Finally, the data show that the frequency location of the spectral peak largely depends on the following vowel context: it is higher before the front vowel **i** than before **a** or **u**, and it is higher before the unrounded vowel **a** than before the rounded vowel **u**. These acoustic correlates are mainly accounted for by the resonance frequency of the cavity in front of the constriction.

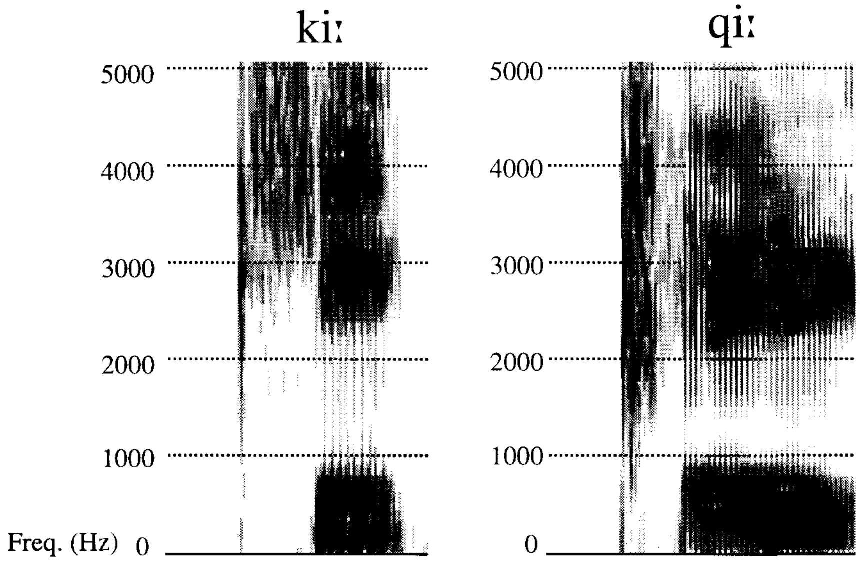


Figure 10. Spectrograms for the velar **k** in **kiikaχ** (cranberry bush) and the uvular **q** in **qiiɣaχ** (grass) as produced by a female speaker of Eastern Aleut. The uvular **q** is associated with greater low frequency energy, and, when considered over the spectrum as awhole, more overall noise energy of frication.

4. Final Remarks.

Aleut has many other interesting properties that we hope to discuss in a subsequent paper. The voiceless nasals and laterals in Western Aleut need further analysis. The fricatives need more careful investigation, particularly with respect to the degree of voicing that may be present, and the contrast between velar and uvular fricatives. Some of the consonant clusters such as **kθ** as in **kθaχ** 'kdax̂' (ice) and **tʃŋ** as in **tʃŋaχ** 'chngax̂' (fur, pelt) need further description. The investigation reported here is only a first step in the acoustic analysis of Aleut. In addition we have tried to contribute not only to Aleut studies, but also to general phonetic knowledge by, for example, detailing the qualities of the vowels in a three vowel system, and adding to the few studies of the contrast between velar and uvular stops.

References

Bergsland, K. 1994. *Aleut Dictionary (Unangam Tunudgusii).* Alaska Native Language Center, University of Alaska Fairbanks.

Bergsland, K. and M. Dirks. 1978. *Qawalangim tunugan kaduĝingin.: Eastern Aleut grammar and lexicon.* National Bilingual Materials Development Center, Rural Education Affairs, University of Alaska.

Cho, T. and P. Ladefoged. 1997. Variations and universals in VOT: evidenced from 17 endangered languages. *UCLA Working Papers in Phonetics, t*his volume.

Jakobson, R., G. Fant, and M. Halle. 1963. *Preliminaries to speech analysis.* 4th ed. Cambridge: MIT Press.

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.

Keating, P. and A. Lahiri. 1993. Fronted velars, palatalized velars, and palatals. *Phonetica* 50:73-101.

Liljencrants, J. and B. Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language 48:* 839-862.

Lindblom, B. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America 35:* 1773-1781.

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

Lindblom, B. and I. Maddieson. 1988. Phonetic universals in consonant systems. In L. Hyman and C. Li (eds.), *Language, Speech and Mind,* Routledge, pp. 62-78.

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

Peterson, G.E. and I. Lehiste. 1960. Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America 32:* 693-703.

Port, R. F. and R. Rotunno. 1979. Relation between voice-onset time and vowel duration. J. Acoust. Soc. Am 66(3), Sept. 1979.

Stevens, K. (in press). *Acoustic Phonetics.* Cambridge: MIT Press.

Taff, A. 1997. Aleut intonation. *UCLA Working Papers in Phonetics,* this volume.

Taff, A. 1992. Uvularization of vowels in Aleut. M.A. thesis, University of Washington.

Acknowledgments

This paper is for the Aleut people. We are specially grateful to our friends in St. Paul, St. George and Atka who gave us so much help in recording the data. In the Pribilof Islands: Edna Floyd, Gregory Fratis, Fanny Galanin, Andronik Kashevarof, Ella Kashevarof, Ludmilla Mandregan, Marva Melovidov, Rufina Merculief, George Rukovishnikoff, Olga Rukovishnikoff, William Shane and Edna Stepetin. In Atka: Oleana Dirks, Dennis Golodoff, Raymond Golodoff, Victor Golodoff, Eva Nevzoroff, Katherine Nevzoroff, Millie Prokopeuff, Judy Zaochney Simmens, Daniel Snigaroff, Mark Snigaroff, Michael Snigaroff and Sally Swetzoff

Work supported by NSF grant 9319705.

Appendix
Aleut wordlist used as a basis for making recordings

The first column shows the contrasts sought, using Aleut orthography as an aid to Aleut specialists. The second column gives the word recorded in Aleut orthography. The third column gives an approximate English gloss. The final column gives an IPA transcription appropriate for most of our speakers. Blanks are left to show where no appropriate word was found.

1. STOPS IN INITIAL POSITION

	ALEUT	ENGLISH	TRANSCRIPTION AND NOTES
ti	tiĝlax̂	bald eagle	**'tiʁlaχ**
tii	tiistax̂	dough, piecrust	**'tiistaχ**
	tiihnax̂	mound	**'tiin̥aχ**
ta	tanix̂	forehead	**'taniχ**
taa	taangax̂	water	**'taaŋaχ**
tu	tukux̂	chief	**'tukuχ**
	tulax̂	forearm	**'tulaχ**
tuu	tuutx̂ax̂	stalk of cow parsnip	**'tuutχaχ**
chi	chis	to stick, adhere, cleave	**'tʃis**
	chisil	scatter	**'tʃisil**
	chigdax̂	Aleut raincoat	**'tʃiʁdaχ**
chii	chiidax̂	baby animal, e.g. seal	**'tʃiiðaχ**
cha	chagix̂	halibut	**'tʃaɣiχ**
chaa	chaasxix̂	cup	**'tʃaasxiχ**
chu	chugux̂	sand	**'tʃuɣuχ**
chuu	chuulkix	socks	**'tʃuulkix**
ki	kigil	to bite	**'kiɣil**
kii	kiin	who	**'kiin**
ka	kadan	in front of the speaker	**'kaðan**
kaa	kaangux̂	healthy	**'kaaŋuχ**
ku	kukax̂	grandmother	**'kukaχ**
kuu	kuusxix̂	cat	**'kuusxiχ**
qi	qiqix̂	slime	**'qiqiχ**
qii	qiigax̂	grass	**'qiiɣaχ**
qa	qalgadax̂	food	**qal'ɣaðaχ**
qaa	qaadas	dolly varden (fish)	**'qaaðas**
qu	quhmax̂	white	**'qum̥aχ**

2. SIBILANTS.

si	sisax̂	hundred	**'sisaχ**
sii	siintax̂	cent	**'siintaχ**
sa	sas	birds	**'sas**

saa	saalax̂	lard, reindeer fat	**'saalaχ**
	saakux̂	king eider	**'saakuχ**
su	susux̂	pus	**'susuχ**
suu	suuskax̂	baby's bottle	**'suuskaχ**
zi			
zii	ziilitax̂	vest	**'ziilitaχ**
za	azax̂	to be (no time marker)	**'azaχ**
zaa	zaavtrikax̂	breakfast	**'zaavtrikaχ**
zu	huzus	to make plenty	**'huzus**
zuu	zuulutax̂	gold	**'zuulutaχ**

3. VOICED FRICATIVES.

Orthographic **d, g** are normally fricatives, but in initial position (where they are comparatively rare) they may be stops, particularly in Russian loan words. Many initial **d, g** , are in loan words. In the transcription here they are given as **d** although many speakers have **ð** for many of these words

di	dix̂	soot	**'iχ**
dii	diikal	to be mischievous	**'iikal**
da	dax̂	eye	**'aχ**
daa	daax̂tux̂	kidney	**'aaχtuχ**
du	duskax̂	washboard, plank	**'uskaχ**
gi	gil	he is envious	**'ɣil**
gu	agul	to make, build	**'aɣul**
ĝaa	x̂aayax̂	steam bath	**'χajaχ**
x̂a	x̂ax̂	(halibut) stomach	**'χaχ**
gaa			
gu	gul	to go through	**'ɣul**
hi	hinus	piece of sod	**'hinus**
ha	halal	to turn the head	**'halal**
hu	hudax̂	dried fish	**'hudaχ**
ĝ	aĝal	to open, ebb tide	**'aʁal**

4. UVULAR FRICATIVES

Voiceless velar and uvular fricatives occur in medial and final position. In Eastern Aleut none of the other fricatives occur in final position. In Western Aleut **s** occurs as a plural marker (instead of **n**, as in Eastern Aleut). (Parenthesized words in the second column have been listed above; they are shown again so that the contrasts can be seen.)

ux	asxinux̂	(one) girl	**as'xinuχ**
ux	alax asxinux	two girls	**'alax as'xinux**
ix	(dix̂)	soot	**'iχ**
ix	(chuulkix)	socks	**'tʃuulkix**
ax	(taangax̂)	water	**'taaŋaχ**
ax	hizax	almost	**'hizax**

5. VOICED AND VOICELESS NASALS AND LATERALS,

mi	miilax̂	soap	**'miilaχ**
ma	mal	to do	**'mal**
mu	mukax̂	flour	**'mukaχ**
ni	Niiĝuĝis	people of Atka	**'niiʁuʁis**
na	naga	it's inside	**'naɣa**
na	anax̂	mother	**'anaχ**
nu	qanul	smells of fish	**'qanul**
ngi	qungix̂	hump (back)	**'quŋiχ**
nga	hangal	to go up	**'haŋal**
ngu	qangul	to go in	**'qaŋul**
hmi	hmiichix̂	ball	**'m̥iitʃiχ**
hma	hmachil	to get stuck	**'m̥atʃil**
hmu	hmuqatix̂	fish's gill cover	**m̥u'qatiχ**
hni	chuhnil	to poke	**'tʃun̥il**
hna	ahnatix̂	marker	**a'n̥atiχ**
hnu	hnul	to reach	**'n̥ul**
hngi	kahngil	to bend	**'kaŋ̊il**
hnga	ahngal	to acknowledge	**'aŋ̊al**
hngu	qahngus	seaweed	**'qaŋ̊us**

6. APPROXIMANTS

wi			
wa	waya	right here, now	**'waja**
wu			
li	lil	to look like, appear	**'lil**
la	lal	to pick up, gather	**'lal**
lu	lul	to believe	**'lul**
yi			
ya	yas	reef	**'jas**
yu	ayul	to fall	**'ajul**
hwi			
hwa	hwaĝix̂	smoke	**'ʍaʁiχ**
hwu			
hli	tahlidax̂	knot in wood	**ta'l̥iðaχ**
hla	hlax̂	boy	**'l̥aχ**
hlu	qihluxs	barking	**'qil̥uxs**
hyi			
hya	hyal	tide	**'çal**
hyu	hyul	to pour, spill	**'çul**

7. VOWEL LENGTH

aa	aalal	being done	**'aalal**
a	alax̂	whale	**'alaχ**
	alal	want	**'alal**

ii	chiidax̂	baby animal. e.g. seal	**'tʃiiðaχ**
	hiilal	done that way	**'hiilal**
i	chidaĝa	border, space beside	**tʃi'ðaʁa**
	inux̂	piece of tobacco	**'inuχ**
	hilal	to read	**'hilal**
uu	huudal	to sound horn	**'huuðal**
u	udax̂	bay	**'uðaχ**
	hudal	to dry fish/meat	**'huðal**

8. STRESSED VS, UNSTRESSED VOWELS

i	(kitax̂)	foot	**'kitaχ**
'i	hlam kitaa	boy's foot	**'l̥am ki'taa**
u	tunux̂	word	**'tunuχ**
'u	Unangam tunuu	Aleut language	**u'naŋam tu'nuu**
a	(tanix̂)	forehead	**'taniχ**
'a	hlam tanii	boy's forehead	**'lam ta'nii**

9. VELAR VS UVULAR INFLUENCE ON THE VOWELS

kak	kakix̂tal	looks up	**'kakiχ'tal**
qaq	qaqaĝix̂	arctic loon	**qa'qaʁiχ**
kix	kixs	to bite	**'kixs**
qiq	(qiqix̂)	slime	**'qiqiχ**
kuk	(kukax(fl	grandmother	**'kukax**
quq	quqdax̂	dirty	**'quqðaχ**
ka	(kadan)	in front (of speaker)	**'kaðan**
qa	qadux̂	scab on skin	**'qaðuχ**
ki	(kitax̂)	foot	**'kitaχ**
qi	qilax̂	morning	**'qilaχ**
ku	kudux̂	leg	**'kuðuχ**
ak	chaknax̂	stinky	**'tʃaknaχ**
ax	(daaxtux̂)	kidney	**'aaχtuχ̊**
ak	akalux̂	way, path	**a'kaluχ**
aq	aqal	stretch	**a'qal**

10. ILLUSTRATIVE CLUSTERS

chng	chngax̂	fur, pelt, yarn	**'tʃŋ̊aχ**
mg	humgix̂	lung	**'humɣiχ**
mx	amxix̂	flesh of fish	**'amxiχ**
mɣ	umĝix̂	bait	**'umʁiχ**
mx			
kd	kdax̂	ice	**'kðaχ**
sd	sdax̂	star	**'sðaχ**

The effect of stress on vowel length in Aleut

Lorna Rozelle
University of Washington

Although duration is typically one of the physical correlates of stress, (Fry, 1955; Crystal and House, 1988; i.a.), Berinstein (1979) claims that languages with contrastive vowel length do not use duration as a correlate of stress. This study addresses three questions about the phonetic correlates of stress in Aleut, a language with contrastive vowel length. Is duration a correlate of stress in Aleut? If so, does the change in duration caused by stress neutralize the phonemic contrast between short and long vowels? Finally, does stress assignment precede or follow the process of final vowel deletion common in Aleut? Measurements of vowel durations from spectrographic analysis show that duration is indeed a correlate of stress, counter to Berinstein's claim. However, despite the change in duration wrought by stress, contrastive ratios are maintained between short and long vowels. It is also clear that stress is assigned before the final syllable is deleted.

1. Introduction

1.1 Aleut

Aleut is the smallest branch of the Eskimo-Aleut language family. It is spoken in the United States on the Aleutian and Pribilof Islands in the Bering Sea by fewer than 1,000 speakers and in Russia on the Commander Islands by fewer than 100 speakers. There are two major, mutually intelligible dialects, Eastern and Western, both of which have subdialects. The Pribilof Island dialect spoken by the subjects for this study is classified as a western subdialect of Eastern Aleut (Bergsland, 1991).

Underlyingly, Aleut has three vowels, / i /, / u /, and / a /, all of which can be both long and short. Examples of length contrasts producing lexical distinctions are shown in (1). Words are written in the standard orthography: long vowels are written as two vowels; stress is not marked. In the transcription, length is indicated by the IPA length mark ː, č is used for IPA **ʧ**, and stress is shown by an acute accent over the vowel instead of the IPA stress mark..

(1)	siching	**síčiŋ**	*four*	chang	**čáŋ**	*my hand*
	sichiing	**sičíːŋ**	*nine*	chaang	**čaːŋ**	*five*

Aleut has a quantity sensitive stress system. Taff (1992) states weight is assigned at the nuclear level so that consonants are weightless. Syllables containing long vowels are heavy, and syllables containing short vowels are light. Stress falls on the ultimate syllable if it is heavy and the penultimate syllable is light, (2a). Otherwise, stress falls on the penultimate syllable, (2b and 2c). Both long and short vowels can be in the stressed penult or in the unstressed ultima.

(2)	a.	sichiing	**sičíːŋ**	*nine*
	b.	siching	**síčiŋ**	*four*
	c.	adaadaa	**aðáːðaː**	*their father*

The situation is still more complex. Bergsland (1994) says, "Stress in Aleut is somewhat indeterminate and difficult to define, depending upon several factors ... Stress, in turn, may affect the length of vowels and consonants." Although length is contrastive, the durational effect of stress appears to neutralize the length distinction. According to Taff (1992), a stressed short vowel can sound long, as in (3a), and an unstressed long vowel can sound short, as in (3b).

(3)	a.	amaxsix	**amáxsix**	*to spend the night*
		amaagaasalix	**amaːɣaːsálix**	*to take over there*
	b.	amaagaasix	**amaːɣáːsix**	*to get there*
		adaadaa	**aðáːðaː**	*their father*

1.2 Previous research

Stress and its phonetic correlates in English have been well-studied. Whether the findings of these studies apply to Aleut is unknown, but they give us some idea of the range of factors that influence stress. Fry (1955) determined that duration is a more effective cue to stress than intensity. ("Duration" is used to refer to the physical measurement of time, while "length" is used to when refer to the phonological categorization of duration.) Fry (1958) determined that pitch is an even more effective cue than duration. Other studies, such as Peterson and Lehiste (1960), Umeda (1975), and Crystal and House (1988) have measured vowel duration in English in various phonological environments, including stress. Adams and Munro (1978) determined that duration is the most frequent cue to stress for their subjects. Their study is distinctive in that it analyzed stress cues for words in complete sentences rather than words in isolation. They ask what "distinguishes stressed from unstressed syllables *in the stream of speech*. ... What does a speaker *do* that causes the listener to receive an impression of stress?"

Since vowel length is not phonemic in English, none of these studies address the issue of neutralization of phonemic vowel length. Bond (1991) measured vowel duration in Latvian, a language with phonemic vowel length, and concluded that contrastive ratios are maintained in spite of adjustments that occur for reasons of stress, morphology and syntax. She suggests that "...adjustments in duration may be universal in the phonetic structure of languages, but these adjustments have language-specific and different implementations."

Beckman (1986) compared the phonetic correlates of stress in English and Japanese. She concluded that these systems differ in that the Japanese stress system uses pitch as a correlate of stress to a greater extent than English, hence accounting for the traditional distinction between these so-called "pitch" accent and "stress" accent systems. However, since it is again the case that neither of these languages uses pitch contrastively, Beckman (1986) does not address the question of the neutralization of a contrast by stress. However, this study is relevant in that it shows that the extent to which one of the phonetic cues to stress is used can vary cross-linguistically.

Berinstein (1979) examined the interaction between stress and vowel length in K'ekchi, a Mayan language with distinctive vowel length for unstressed vowels. In a perception experiment she synthesized four syllable tokens of the form bIbIbIbI. The duration of three of the four vowels was 100 ms, while the duration of one of the vowels was one of six durations: 70, 100, 120, 140, 160, 200 ms. The location of the syllable with varying length could be in first, second, third or fourth position in the token. Subjects were asked to judge which syllable was stressed. Berinstein concluded that English speakers used both position and duration as cues to stress, while K'ekchi speakers used only position. Increases in duration had no influence on the perception of stress for K'ekchi speakers. In a production study, Berinstein elicited twenty words varying in vowel quality, stress and phonemic length and measured vowel duration via spectrographic analysis. She found that short vowels are not lengthened by stress. (She was unable to construct examples with stressed long vowels. Apparently long vowels do not appear in both stressed and unstressed positions in K'ekchi.) Berinstein concluded that duration is not a correlate of stress in K'ekchi, and she hypothesized that all languages with phonemic length do not use duration as a correlate of stress.

2. The experiment

2.1 Equipment and method

Data from subject ML was recorded with a PMD 430 Marantz cassette tape recorder and Electro-voice D054 dynamic omni-directional microphone in the subject's home and in a classroom. Data from the other three subjects, BS, LM, and GF, was recorded in home settings on DAT equipment with a flat frequency response throughout the auditory range, using a close-talking, noise-canceling microphone so that the signal to noise ratio was always better than 40 dB. For all data, vowel duration was measured from the spectral analysis produced on a Kay CSL 4300B speech analysis system.

In measuring vowel length, the section of the spectrogram where formant structure was visually identifiable was listened to. If necessary, the range was narrowed until none of the preceding or following consonants were audible, eliminating on and off glides and measuring only the steady state formant structure in order to mitigate the influence of surrounding consonants on vowel length. Figure 1 provides an example.

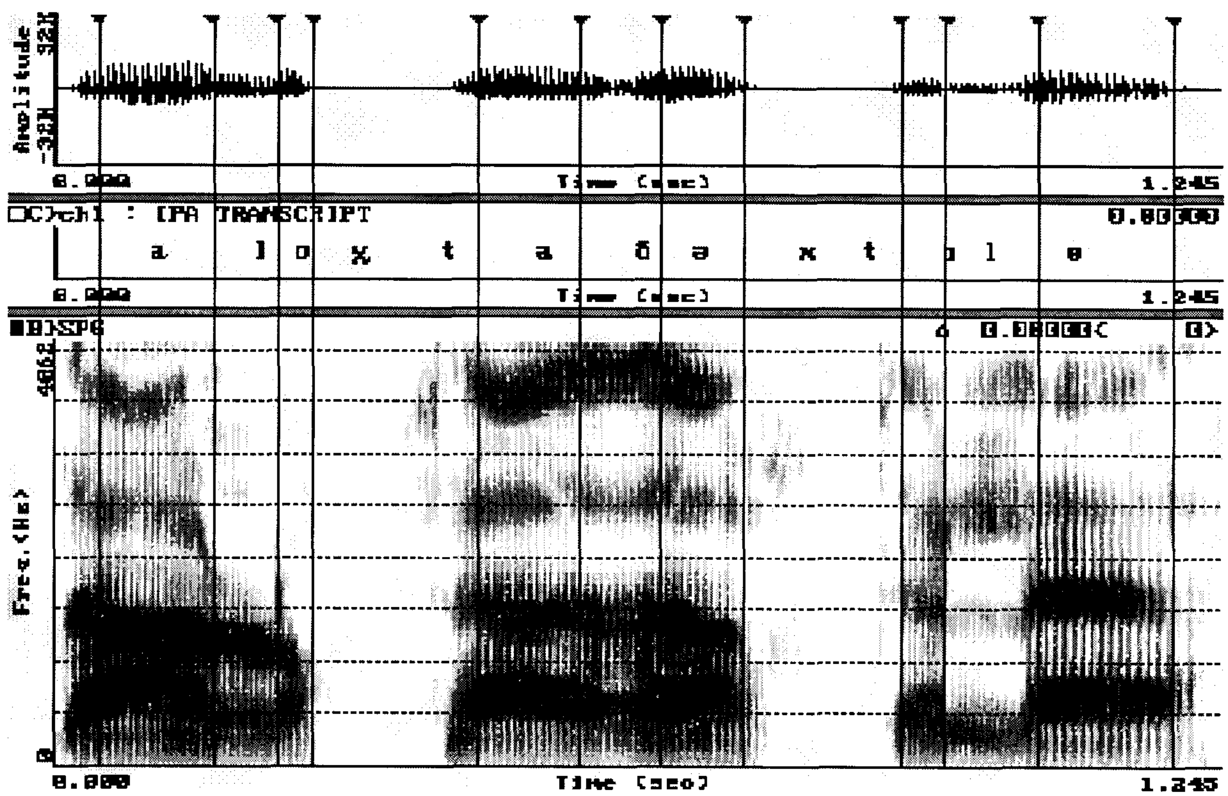


Figure 1. The waveform, transcription and spectrogram of *aalux̂taadax̂til, ii* **aːluχtaːðaχtil iː** 'Did (the girls) laugh?' The intervals that were measured are indicated. The transcription beneath the waveform is slightly narrower than the transcription in the text.

In order to check the accuracy of the measurements, a random sample of vowels was measured by another phonetician. This person was given the instructions in the preceding paragraph but worked independently. Approximately seven percent of the data was checked.

These measurements were compared to the corresponding measurements that had been make previously by the author. The difference between the means of the two sets of measurements was 3 ms. The mean of the differences was 18 ms, with a standard deviation of 13 ms.

2.2 Subjects

This analysis is based upon the speech of four subjects. All subjects are adults who were born on the Pribilof Islands. They learned Pribilof Aleut as their first language and also speak English. BS, a man in his early 70's, speaks Aleut at home with his wife and with his peers. LM, a woman in her late 60's, speaks Aleut daily with her peers. GF, a man in his early 50's, is in the transition generation. The generation before his learned Aleut as their first language, while the generation after learned English. He eagerly speaks Aleut with his peers but mostly uses English at home as his wife understands but does not speak Aleut. BS, LM and GF have lived on the Pribilof Islands all their lives. The fourth subject, ML, is a man in his mid 80's. Although ML has been living off of the Pribilof Islands for several years, he speaks Aleut daily with his wife, who is also a native speaker.

2.3 Materials

Many factors influence vowel duration besides distinctive length and stress. Some of these factors are vowel quality, post-vocalic voicing, post-vocalic place of articulation (Peterson and Lehiste, 1960), speaking rate (Crystal and House, 1982), word length (Lehiste, 1972), morphological structure (Bond, 1991), prepausal position (Umeda, 1975; Crystal and House, 1988), and word prominence (the information load the word carries in the message) (Umeda, 1975). According to Peterson and Lehiste (1960) and Umeda (1975) the *pre*vocalic consonant has no consistent effect on vowel length in English. Another factor, perhaps similar to rate of speech, is whether measurements are made of connected speech or of isolated citation forms. The influence of all but the last of these factors on vowel duration in English has been well-documented; however, their influence in other languages has been studied less, and in Aleut not at all. In order to mitigate these potential confounds all three Aleut vowels were measured in a variety of phonological, morphological and syntactic environments in two different elicitation tasks. Though it is possible that whether the word was part of a connected speech or was an isolated citation form will affect the results, for the purpose of this study data from different sources has been pooled.

The measured vowels have three different sources. The first is a connected narrative of approximately two minutes elicited from subject ML who was instructed to tell a short story about something. The second source of data is a set of tapes from a field methods class in which citation forms were elicited from ML. The stimulus was an English word. The response was an Aleut word. Words were often repeated several times, allowing several measurements of the same word to be made. Multiple measurements of the same word were averaged. Repetitions which seemed unnaturally slow were not used. The third source of data is a set of sentences elicited from each of three subjects, BS, LM and GF, which was collected by Alice Taff and Peter Ladefoged and has been described elsewhere (Taff and Wegelin, 1997; Cho, Taff , Dirks, and Ladefoged, 1997). The stimulus was an English sentence, the response the Aleut translation. The stimuli were designed without regard for vowel length, so include a realistic sampling of long and short vowels in stressed and unstressed positions. Although the responses were not uniform across speakers, no amendments were suggested so that the responses remained as natural as possible. When sentences were repeated, corresponding measurements were averaged.

Every vowel from each source was measured, although not all measurements were included in the analysis. Certain measurements were excluded for two reasons. First, vowels next to glides or voiced velar or uvular fricatives were often excluded because in these cases it was difficult to determine the beginning or end of the vowel. Second, Aleut has an optional process in which final syllables can be deleted. As the interaction between deletion and stress assignment is one of the questions to be addressed, words in which deletion has occurred were not included in the primary data set.

3. Results

The table in (4) summarizes the means and standard deviations for short and long vowels in stressed and unstressed position. Figure 2 shows a chart of the same data.

(4)

	short	**long**	
unstressed	64 ms	130 ms	mean
	22 ms	27 ms	standard deviation
	317	38	number of tokens
stressed	78 ms	151 ms	mean
	29 ms	44 ms	standard deviation
	99	84	number of tokens

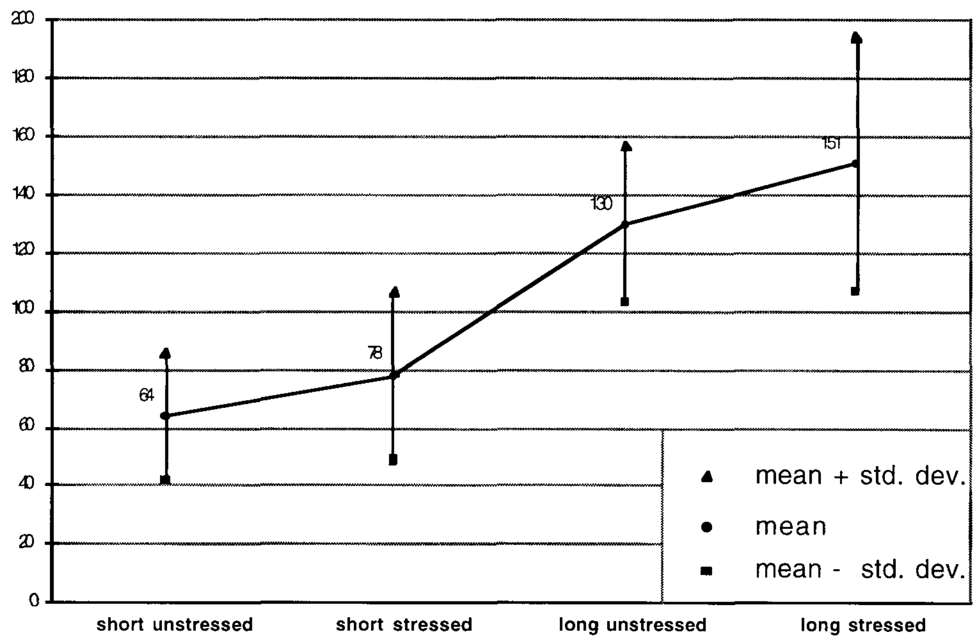


Figure 2. Vowel duration in milliseconds.

First, consider the difference between short and long vowels. When not stressed, short vowels average 64 ms while long vowels average 130 ms. The duration of a long vowel is 2.0 times that of a short vowel. Similarly, when stressed, short vowels average 78 ms, while long vowels average 151 ms. The duration of a long vowel is 1.9 times that of a short vowel. Thus, the ratios of the duration of long to short vowels supports the theoretical notion that a long vowel occupies two timing slots while a short vowel occupies only one (McCarthy, 1979).

Second, consider the effect of stress on duration. The duration of a short vowel is 64 ms when not stressed and 78 ms when stressed. The duration of a long vowel is 130 ms when not stressed and 151 ms when stressed. Given the variation in the data, are these differences significant enough to say that stressed vowels are longer than unstressed vowels?

Since the data sets do not have normal distributions, the non-parametric Mann-Whitney rank sum test was used to compare means (Snedecor and Cochran, 1980). The null hypothesis is that the means of the data sets are the same. Rejection of the null hypothesis implies that one set has a higher mean than the other. Application of the rank sum test on the sets of short stressed and short unstressed vowels as well as on the sets of long stressed and unstressed vowels shows that the null hypothesis is rejected with $p \leq 0.01$ in both cases. The conclusion is that stress does affect vowel duration; hence, duration is a correlate of stress in Aleut.

Third, let us consider whether the effect of stress on duration obscures phonemic length contrast. The duration of a stressed short vowel averages 78 ms while the duration of a long unstressed vowel averages 130 ms. Again, given the variation in the data, are the durations of these two sets of vowels significantly different? Application of the rank sum test allows us to reject the null hypothesis with $p \leq 0.01$ and conclude that the two means are different. Thus, although stress affects duration, it does not destroy the length contrast between a long and a short vowel, even when the short vowel is stressed and the long vowel is not.

To further verify the significance of these results, variation between subjects was examined. The means over all vowels for each subject excluding citation forms are as follows: GF 75 ms, BS 82 ms, LM 84 ms, ML 86 ms. Figure 3 shows that the results for all four subjects are similar. As is evident from the means, subject GF speaks more quickly than the others, so his lines are lower, but they have the same slopes.

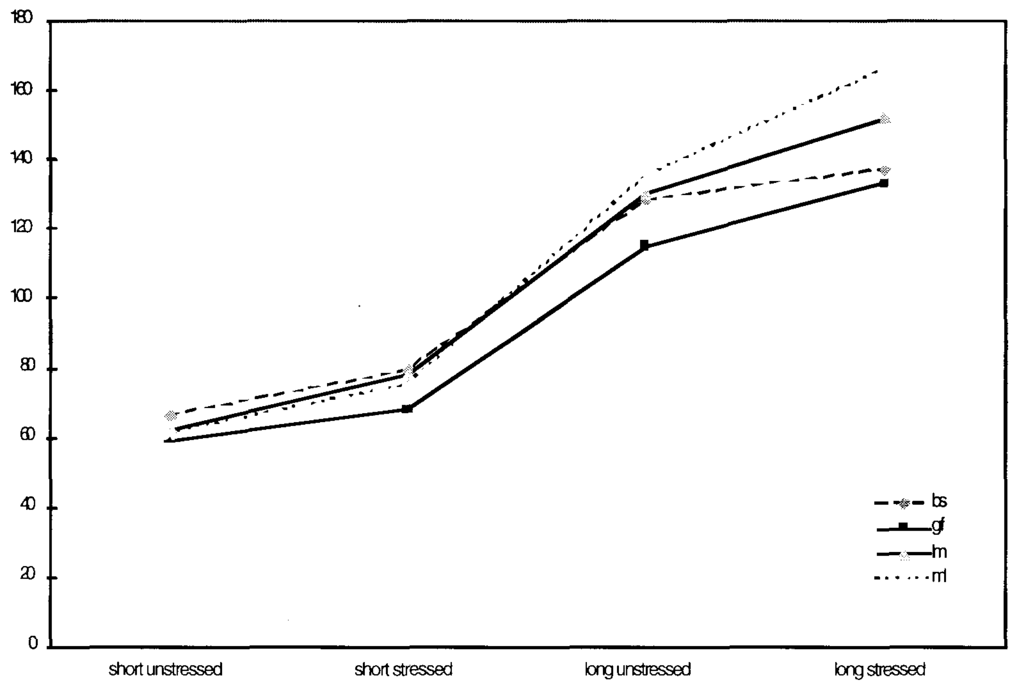


Figure 3. Comparison of the means of subjects' durations in milliseconds.

The durations of short stressed and unstressed vowels for all subjects is very similar. A power test was performed on the means of the durations of short stressed and unstressed for each subject. It showed that in order to detect a twenty percent increase in duration between short unstressed and stressed vowels, only .5 speaker was required, indicating that the intersubject variability was very small. As is clear from the chart in (7), there is more intersubject variability for the long vowels. In particular, while subject BS does lengthen long vowels under stress (128 ms for long unstressed vowels; 137 ms for long stressed vowels), he does not lengthen as much as the other subjects do nor as much as he does for short vowels. Including BS in the power test shows that in order to detect a twenty percent increase in vowel duration, it is necessary to use 4.6 subjects instead of the four actually used in this study. However, if BS is excluded and only the data from the other three subjects are analyzed, only 2 subjects are necessary, indicating that there is little intersubject variability amongst the other three subjects. This is in accordance with Kehoe, et. al. (1995) who report: "Across all age ranges, there was great diversity in how individual subjects employed different phonetic parameters. Some subjects employed certain acoustic parameters over others, that is, employed duration only rather than f_0, or employed f_0 only rather than duration." It is possible that BS employs duration to a lesser extent than the other subjects for marking stress in long vowels. If this is the explanation, it is unclear, however, why he employs duration for marking stress in short vowels.

4. Final syllable deletion

Final syllables are frequently deleted in Aleut. Does the word receive stress before or after the final syllable is deleted? (The stress rule was given in section 1.1.) For some words, the question of order is irrelevant. For example, when the final syllable is deleted, *ayagaadax̂* 'girl' is stressed on the underlying penultimate syllable regardless of the order of application of these two rules; see (5a).

(5) a.		/ ayaɣaːðaχ /		/ ayaɣaːðaχ /
	stress	**ayaɣáːðaχ**	final deletion	**ayaɣaː**
	final deletion	**ayaɣáː**	stress	**ayaɣáː**
		ayaɣāː		**ayaɣāː**
b.		/ iɣaχtanaχ /		/ iɣaχtanaχ /
	stress	**iɣaχtánaχ**	final deletion	**iɣaχta**
	final deletion	**iɣaχtá**	stress	**iɣáχta**
		iɣaχtá		**iɣáχta**

However, for other words, the two orders yield two different outputs. For example, see (5b). In *igax̂anax̂* 'airplane,' if the word receives stress before the final syllable is deleted, the penultimate syllable is stressed: **iɣaχtá**. But, if the final syllable is deleted before the word receives stress, the antepenultimate syllable is stressed: **iɣáχta**. Since duration is a strong correlate of stress in Aleut, examination of vowel duration provides an objective, phonetic way to determine which syllable is stressed. The data collected contained 21 examples of final syllable deletion. Of these, nine were of the *ayagaadax̂* type, which are stressed the same regardless of the order of application of the two processes. Of the remaining twelve, four were repetitions whose measurements were averaged, and one was only two syllables. A summary of vowel durations in milliseconds for each of the seven words is shown in (6). The syllable that was deleted is underlined.

(6)

		antepenult	penult	ultima
igaχtanaχ	66	48	67	∅
iχtada		64	104	∅
kumsiχtadan	x	55	79	∅
kumsiχtakun	x	64	71	∅
kumsiχtalix	43	45	57	∅
qakchiklukuχ	53	36	49	∅
qaxchiklulakaχ	50	46	88	∅

In each of these words, if the word receives stress before the final syllable is deleted, the penultimate syllable will be stressed. If, on the contrary, the final syllable is deleted before the word receives stress, the antepenultimate syllable will be stressed. All vowels are short, and in each case, the vowel with greatest duration is the penultimate, not the antepenultimate, indicating that the penultimate syllable is stressed. Words receive stress before the final syllable is deleted.

Additional support for the conclusion that final syllable deletion occurs later is found in the nature of the process. This process has one of the characteristics of a post-lexical rule: it can have variable output (Kiparsky, 1985). Sometimes the final syllable does not delete, sometimes it deletes entirely, as shown in the upper part of figure 4, and sometimes it only devoices, as shown in the lower part.

5. Discussion

One question that this study raises is why, according to other researchers, are short stressed and long unstressed vowels difficult to distinguish in Aleut, even though they are phonetically distinct? This question is all the more puzzling when one considers the difference between vowel durations shown in (6). The greatest difference in duration occurs between short stressed and long unstressed vowels: a long unstressed vowel is longer than a short stressed vowel. Moreover, upon reexamination, the chart in Figure 2 shows that, despite the variation in the data, the indicated ranges of durations for short unstressed vowels and long unstressed vowels overlap only slightly.

(6) a. short stressed - short unstressed 78 ms - 64 ms = 14 ms
b. long stressed - long unstressed 151 ms - 130 ms = 21 ms
c. long unstressed - short stressed 130 ms - 78 ms = 52 ms

One possible explanation for the difficulty in distinguishing short stressed and long unstressed vowels is that speakers of languages that do not use length as a primary phonemic contrast, such as English, are not adept at perceiving length contrasts, regardless of the role of duration in the manifestation of stress. Yet, Bergsland, the author of the *Aleut Dictionary*, who states that stress is “somewhat indeterminate and difficult to define,” is a native speaker of Norwegian, a language that does use length phonemically. Also, Fox and Lehiste (1989), in a perceptual experiment, tested the abilities of English and Estonian speakers to distinguish three different durations. Although Estonian is controversially analyzed as exhibiting a three-way length contrast, Estonian speakers do no better than English speakers at distinguishing these durational distinctions. Hence, it seems unlikely that Aleut speakers can distinguish four different durations, short unstressed, short stressed, long unstressed and long stressed, and unlikely that English speakers confuse short stressed and long unstressed vowels due to inability to perceive duration.

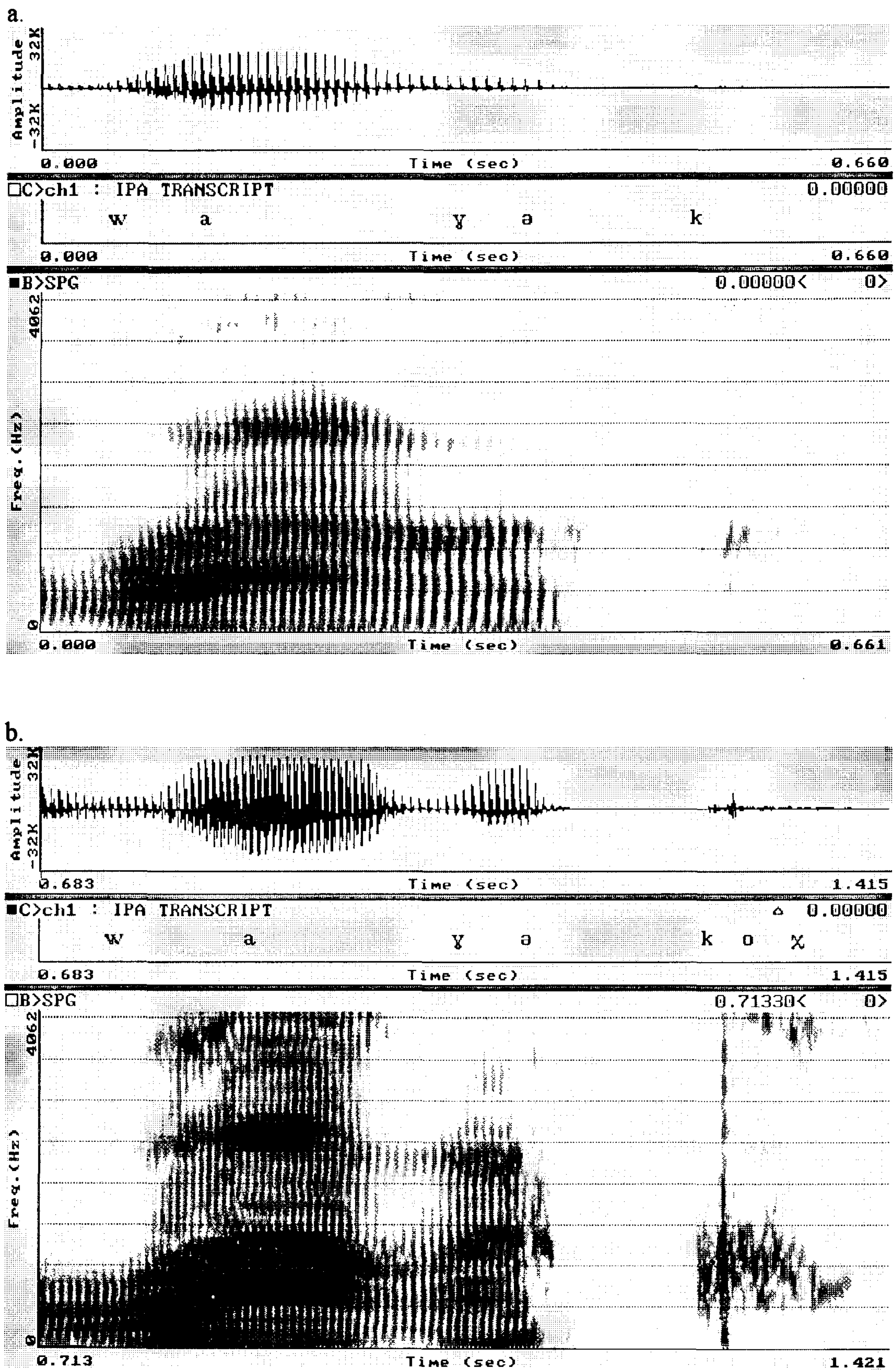


Figure 4. Waveforms and spectrograms showing (top part of the figure) the deletion of the syllable **oχ**, and (lower part) the devoicing of this syllable.

Alternatively, perhaps stress in Aleut does lengthen vowels, but to a lesser degree than it does in other languages. It might be the case that languages that use length contrastively use duration as a weaker correlate of stress than languages that do not use length contrastively. This is a weakening of Berinstein's (1979) claim that languages that use length contrastively do not use duration as correlate of stress. A preliminary search suggests that this hypothesis is correct. In Aleut, the ratio of the durations of stressed to unstressed vowels is 1.2, whether the vowels are long or short. In Latvian, another language with length contrasts, the ratio of stressed to unstressed vowels is 1.3 for both long and short vowels (Bond, 1991). In K'ekchi, the ratio of stressed short to unstressed short vowels is 1.0 (Berinstein, 1979). In contrast, in a language without length contrasts, like English, the ratio is much greater, 1.6 or 1.7 (Fry, 1955 and Crystal and House, 1988, respectively). Further acoustic analysis is necessary to confirm or deny this hypothesis. In addition, it would be illuminating to access native speakers' abilities to distinguish long and short vowels in stressed and unstressed positions in perception experiments to determine their sensitivity to durational distinctions in various languages.

In summary, stress manifests itself in various ways. Not only is stress without consistent phonetic correlates even within a single language (Hayes, 1995), but also the phonetic correlates that have been identified as significant in English are used phonemically in other languages. This acoustic analysis of Aleut shows that duration is indeed a robust correlate of stress, used even in languages that use length phonemically. Moreover, the use of duration as a correlate of stress does not compromise the phonemic use of duration. These measurements also provide phonetic verification of Bergsland's transcriptions of vowel length as well as Taff's stress rule. The vowels transcribed as long in the dictionary are about twice as long as those transcribed as short. Likewise, the vowels that Taff's stress rule assigns main stress are about twenty percent longer than those not assigned main stress.

6. Acknowledgments

I want to thank Alice Taff for all her help, not the least of which was collecting the data. Work supported in part by NSF grant 9319705.

7. References

Adams, Corrine and R. R. Munro (1978) "In search of the acoustic correlates of stress: fundamental frequency, amplitude, and duration in the connected utterance of some native and non-native speakers of English." *Phonetica* **35**, 125-156.

Beckman, Mary E. (1986) *Stress and Non-Stress Accent.* Dordrecht: Foris Publications.

Bergsland, Knut (1994) *Aleut Dictionary.* Fairbanks, Alaska: Alaska Native Language Center, University of Alaska.

Bergsland, Knut (1991) "Eskimo-Aleut languages." *The New Encyclopaedia Britannica*, 15th ed., vol. 22. Chicago: Encyclopaedia Britannica, 778-780.

Berinstein, Ava (1979) "A cross-linguistic study on the perception and production of stress." *UCLA Working Papers in Phonetics* **47**, Department of Linguistics, UCLA.

Bond, Dzintra (1991) "Vowel and word duration in Latvian." *Journal of Baltic Studies* **22**, 133-144.

Cho, T. Taff, A., Dirks, M, and Ladefoged, P. (1997) "Some phonetic structures of Aleut." *UCLA Working Papers in Phonetics* (this volume).

Crystal, T. and A. House (1982) "Segmental durations in connected speech signals: Preliminary results." *Journal of the Acoustical Society of America* **72**, 705-716.

Crystal, T. and A. House (1988) "Segmental durations in connected speech signals: Syllabic stress." *Journal of the Acoustical Society of America* **83**, 1574-1585.

Fox, Robert Allen and Ilse Lehiste (1989) "Discrimination of duration ratios in bisyllabic tokens by native English and Estonian listeners." *Journal of Phonetics* **17**, 167-174.

Fry, D. B. (1955) "Duration and intensity as physical correlates of linguistic stress." *Journal of the Acoustical Society of America* **27**, 155-158.

Fry, D. B. (1958) "Experiments in the perception of stress." *Language and Speech* **1**, 126-152.

Hayes, Bruce (1995) *Metrical stress theory: Principles and case studies.* Chicago: University of Chicago Press.

Kehoe, Margaret, Carol Stoel-Gammon, and Eugene H. Buder (1995) "Acoustic correlates of stress in young children's speech." *Journal of Speech and Hearing Research* **38**, 338-350.

Kiparsky, Paul (1985) "Some consequences of lexical phonology." *Phonology Yearbook* **2**, 85-138.

Lehiste, Ilse (1972) "The timing of utterances and linguistic boundaries." *Journal of the Acoustical Society of America* **51**, 2018-2024.

McCarthy, John (1979) *Formal problems in Semitic phonology and morphology.* Ph.D. dissertation, MIT.

Peterson, George E. and Ilse Lehiste (1960) "Duration of syllable nuclei in English." *Journal of the Acoustical Society of America* **32**, 693-703.

Snedecor, George W. and William G. Cochran (1980) *Statistical Methods.* Iowa: The Iowa State University Press.

Taff, Alice (1992) *Uvularization of Vowels in Aleut.* M. A. thesis, University of Washington.

Taff, Alice and Wegelin, J. (1997) "Intonation of declaratives and questions in Unangan (Eastern Aleut). *UCLA Working Papers in Phonetics* (this volume).

Umeda, N. (1975) "Vowel duration in American English." *Journal of the Acoustical Society of America* **58**, 434-445.

Intonation of declaratives and questions in Unangan (Eastern Aleut)

Alice Taff
University of Washington
Department of Linguistics, Box 354340
Seattle, Washington 98195-4340, USA
taff@u.washington.edu

Jacob Wegelin
University of Washington
Department of Statistics, Box 354322
Seattle, Washington 98195-4322, USA
wegelin@stat.washington.edu

1. INTRODUCTION

This paper provides phonetic evidence in the form of pitch tracks of sentences elicited from eight native speakers of Unangan (Eastern Aleut). Statistical analysis of the pitch tracks supports the following claims about the patterning of intonation in Unangan.

1. Each content word in the language ha2inning and trough at its end.
2. Word contours combine with sentence downtrends to form sentence contours of cascading F0, each word a step in the cascade.
3. Yes/no questions have the same intonation contours as declarative sentences.

This paper is concerned with content words. Preliminary analysis indicates that function words may have different intonation characteristics than content words but this issue requires study beyond the scope of this investigation. Content words make up the bulk of Unangan vocabulary while function words are few. Many sentences have no function words at all.

Unangan is the indigenous language of the Eastern Aleutian Islands and the Pribilof Islands in Alaska, USA. *Unangan* is the language name of self designation by speakers of the eastern dialect of the language known as Aleut. (*Unangas* is the name in the western dialect.) A moribund member of the Eskimo-Aleut language family, Unangan has approximately two hundred fluent native speakers, all over the age of fifty. Pertinent morphology here is that Unangan is agglutinating with the possibility of affixing several suffixes to a lexical (root) morpheme. A single word as in (1)a can, with appropriate affixes, be a sentence as in (1)b.

(1)	a.	*Sadan*				"outside of"
		sada +	n			
		outside	p			
	b.	*Sadaatakun.*				"We were out there."
		sada +	ata +	ku +	n	
		outside	to have	imed pst	p	

Main word stress is usually on the penultimate but can also be on the ultimate syllable (Bergsland 1994, Taff 1992). The phonetic correlate of stress is duration despite the fact that vowel length is contrastive in Unangan (Rozelle 1997).

Yes/no questions are marked morphologically by the conjunctive morpheme, *-lix*, **liX**, which is frequently syncopated to **l** The term 'conjunctive' is derived from the use of this morpheme in conjoining predicates whereby the first verb of a conjoined set has the *-lix* ending and the final verb has the tense/person markers. This conjunctive morpheme is usually followed by a sentence-final question clitic which is *ii*, in the orthography, but which is realized as a different form, **hae**, in the Eastern dialect. In (2) the underlined morphemes show the contrast between declarative vs. yes/no question morphology. The orthography use here is the practical school orthography developed in 1972 (Bergsland 1994). ĝ = a voiced uvular fricative. x̂ = a voiceless uvular fricative. In actual speech, the final rhyme of the final content word in a sentence is nearly always omitted so that the final spoken syllable of (2)a is lak and the final syllable of (2)b is lal (+ ii).

(2)	a.	*Agyuĝum uyuungin aduklakun.*				"Cormorants necks are long."			
		agyuĝu +	m	uyu +	ungin	adu +	kla +	ku +	n"
		cormorant	rlp	neck	rlp	long	clumsy	prs	p"
	b.	*Agyuĝum uyuungin aduklalix, ii.*				"Are cormorants necks long?"			
		agyuĝu +	m	uyu +	ungin	adu +	kla +	lix	ii"
		cormorant	rlp	neck+	rlp	long	clumsy	cnj	Q"

2. METHOD

2.1. Data

Recordings were made of eight native speakers of Unangan, four men and four women. In the recording process the researcher gave a stimulus sentence orally in English, and the speaker then provided an oral translation in Unangan. Each speaker recorded the same list of around 18 sentences. The list included declarative vs. yes/no question pairs and simple vs. complex sentences. The speakers were recorded indoors on DAT equipment with a flat frequency response throughout the auditory range, using a close-talking, noise-canceling microphone. Signal-to-noise ratio was always better than 40 dB.

To avoid gaps in pitch tracks and pitch perturbations caused by voiceless segments, sentences were designed to contain words with mostly sonorant and all voiced segments. However, since obtaining natural intonation contours was the primary objective, speakers were not asked to rephrase their responses if they made different word choices than those anticipated by the researcher. A total of 172 sentences are included in this study. Each sentence was transcribed phonetically and in Unangan orthography. Translations, glosses, and speaker intention (e.g. question vs. declarative) were later confirmed by a native speaker.

2.2 Measurements

Utterances were sampled at 10,000 Hz on a Kay Computerized Speech Lab (CSL) 4300B. Pitch tracks were generated using a frame length of 25 ms. and a frame advance of 20 ms. Numerical results of the entire pitch track for each sentence were entered into a data base for analysis. Measurements were made by hand for cases in which the amplitude of the wave was too low for the pitch tracking function of the CSL. These were calculated from the waveform and entered into the database.

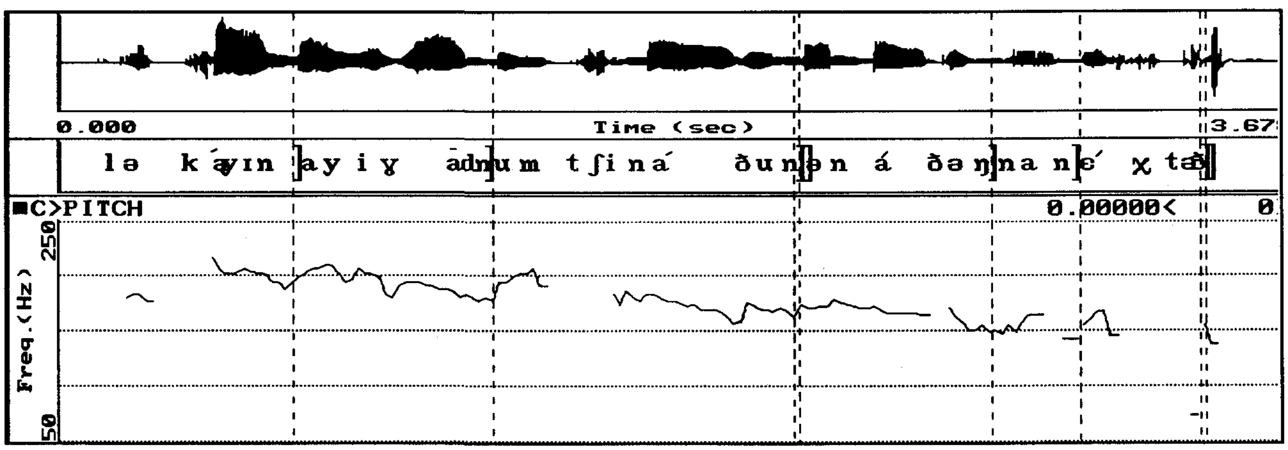
3. RESULTS

3.1 Pitch tracks

Pitch tracks of declarative sentences suggest that each content word is mapped to its own pitch contour as illustrated in examples (3) a-e. Examples a. and b. are responses from two different speakers to the English stimulus, "When boys try to kiss girls, tell their mothers." Example c. is the response from a third speaker to the English stimulus, "When boys try to kiss, girls tell their mothers." Examples d. and e. are from a fourth speaker. d. is the response to "Cormorants' necks are long." e. is the response to "Are cormorants' necks long?" Each example includes an orthographic rendering and gloss, wave form, IPA transcription with main word stress, and pitch track sampled every 20 ms. analyzed and displayed from 50 to 250 Hz. Dotted vertical lines through the displays and square brackets in the transcriptions indicate word boundaries. Double lines and brackets indicate clause boundaries. Each content word has a peak near its beginning and a trough near its end. In (3)a the word *ngan* is a function word as is *naan* in (3)b.

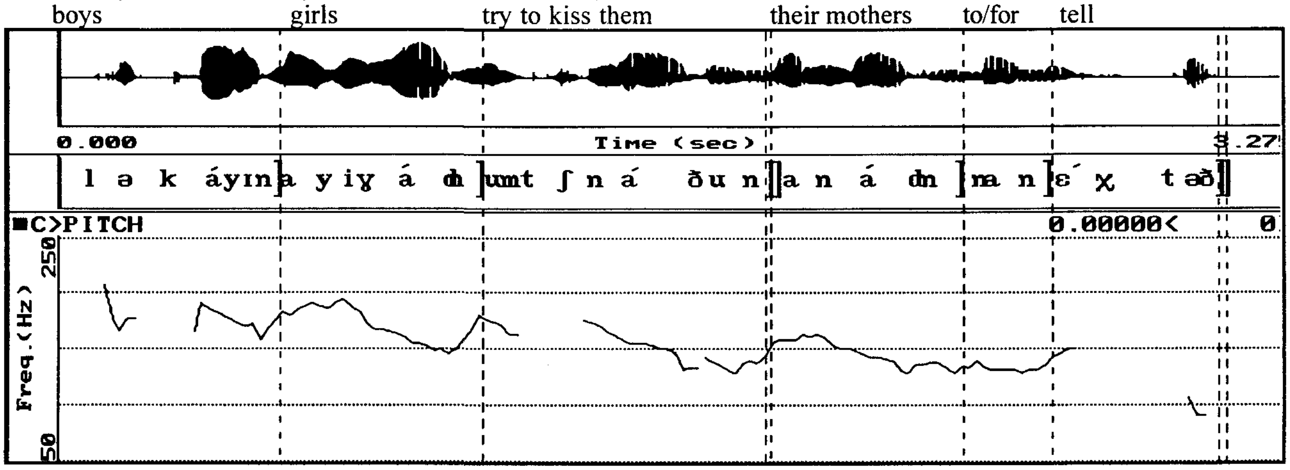
a. Speaker FG

Lakaayan	ayagaadan	umchunaaĝidumchin,	anaadangin	ngan	ix̂ada.
boys	girls	try to kiss them	their mothers	to/for	tell



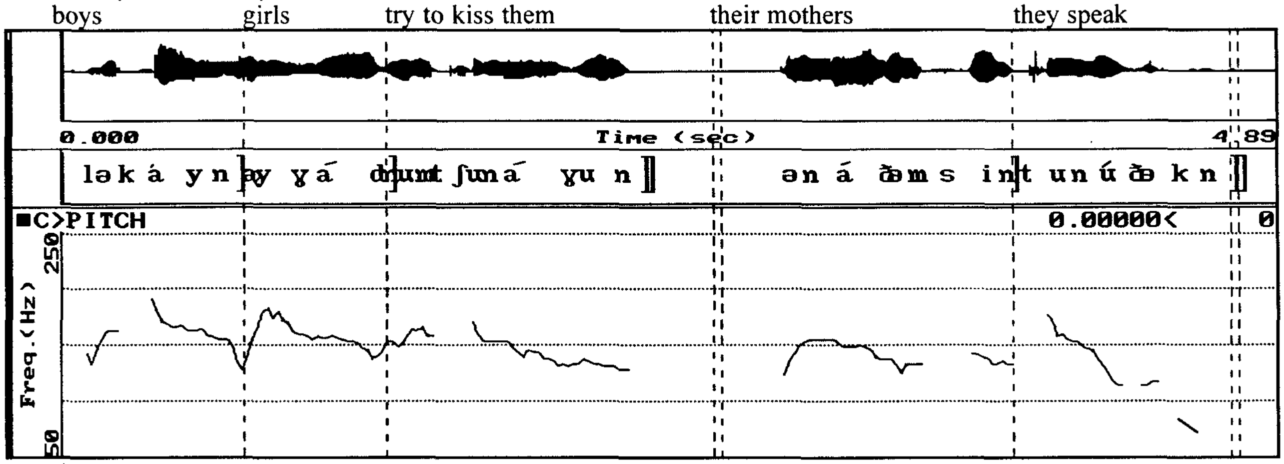
b. Speaker OR

Lakaayan	ayagaadan	umchunaaĝidun,	anaadangin	naan	ix̂ada.
boys	girls	try to kiss them	their mothers	to/for	tell



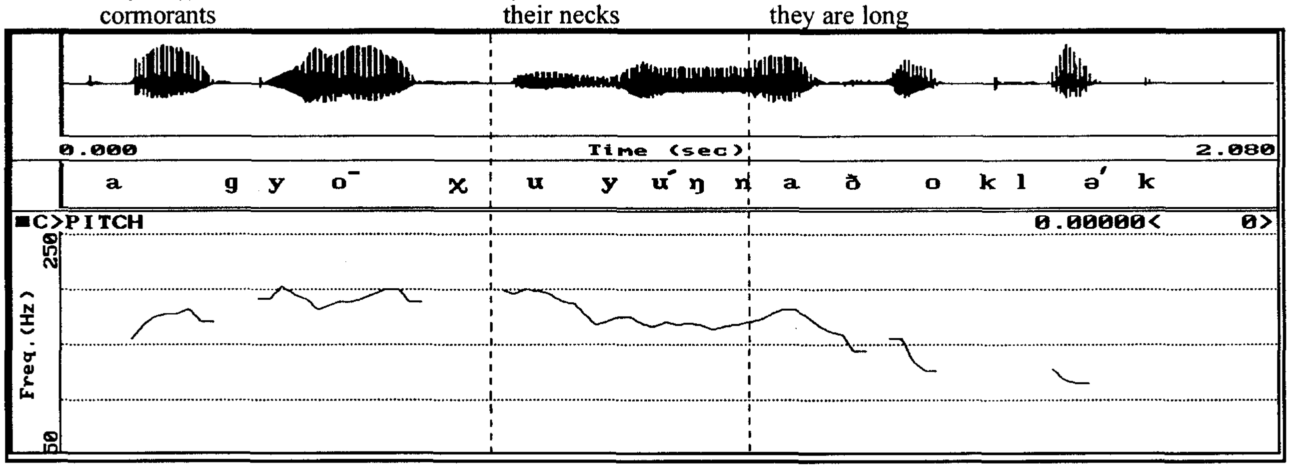
c. Speaker EK

Lakaayan ayagaadan umchunaaguun, anaadamchin tunudakun.
boys girls try to kiss them their mothers they speak



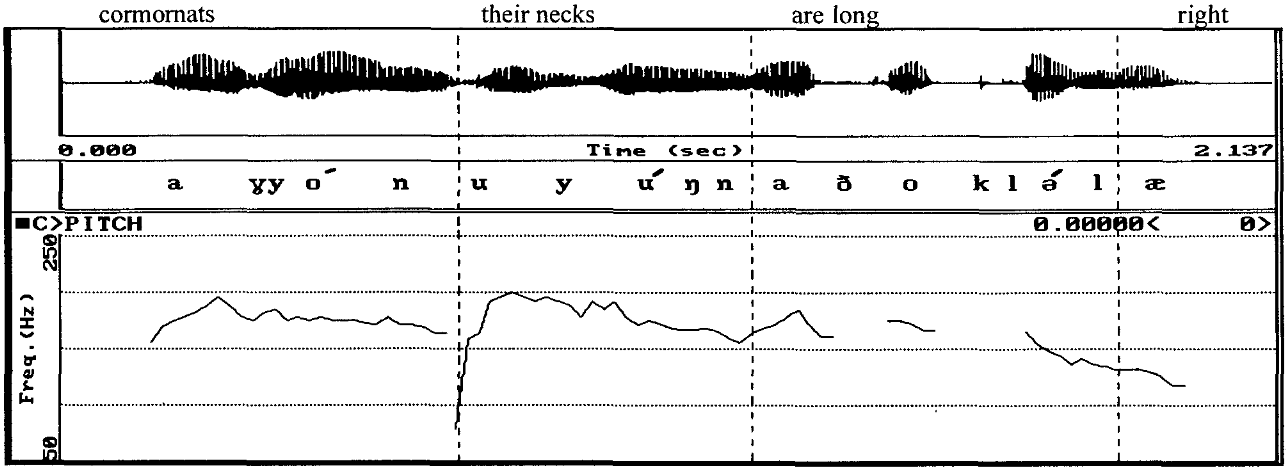
d. Speaker OR

Agyuĝuχ uyungin aduuklakun.
cormorants their necks they are long



e. Speaker OR

Agyuĝun uyuungin aduuklalix ii?
cormornats their necks are long right



3.2 Statistical Analysis

3.2.1 Smoothing

Tokens (sentences) were grouped by speaker, plotted together, and smoothed, to explore the following possibilities:

1. The beginnings and endings of words might be visible in the intonation contours of tokens,
2. Tokens themselves might display a distinctive intonation contour,
3. A difference might be visible between the contours of declaratives and interrogatives.

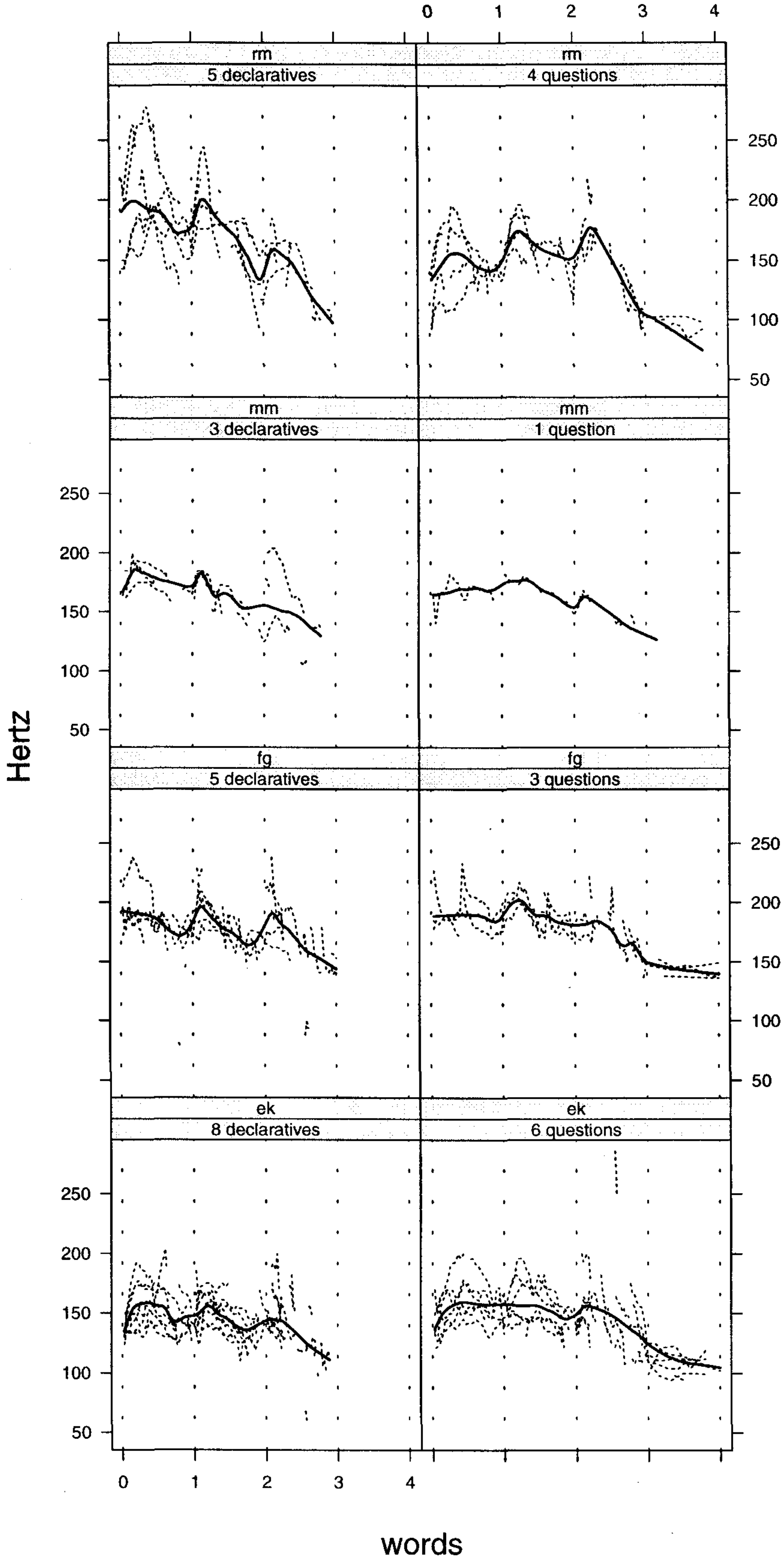
Smoothed pitch tracks of three-word declaratives and yes/no questions are presented in (4) for each speaker. The declaratives appear in the left-hand column and the yes/no questions appear in the right-hand column. Note that in Unangan, the question clitic, *ii*, exemplified in (2)b and (3)e, makes three-word declaratives convert to four-word interrogatives. The first four speakers, FG, RM, EK, and MM, are women; the last four, GR, AK, BS, and GF, are men.

To make words comparable, word durations were converted from time units (i.e., seconds) to word units. Thus, in each of the graphs in (4), the first words start at 0.0 and end at 1.0; the second words start at 1.0 and end at 2.0, and so on. Next, as a first attempt at summarizing intonation patterns, a smoothing function was applied to the data. (The smoother used here is the default setting of the "supsmu" function of the software S-PLUS, version 3.4, release 1 for Silicon Graphics Iris, IRIX 5.3:1996.) This summary is visible as the bold curve on each plot. Given data x_i and y_i (i=1...n), a smoother finds a continuous function y=f(x) under the assumption that the y_i are noisy realizations of $f(x_i)$. The actual pitch track data is given as dotted lines. Statements in this paper about the intonation contours of words and sentences are based on examination of the smoothed contours in (4).

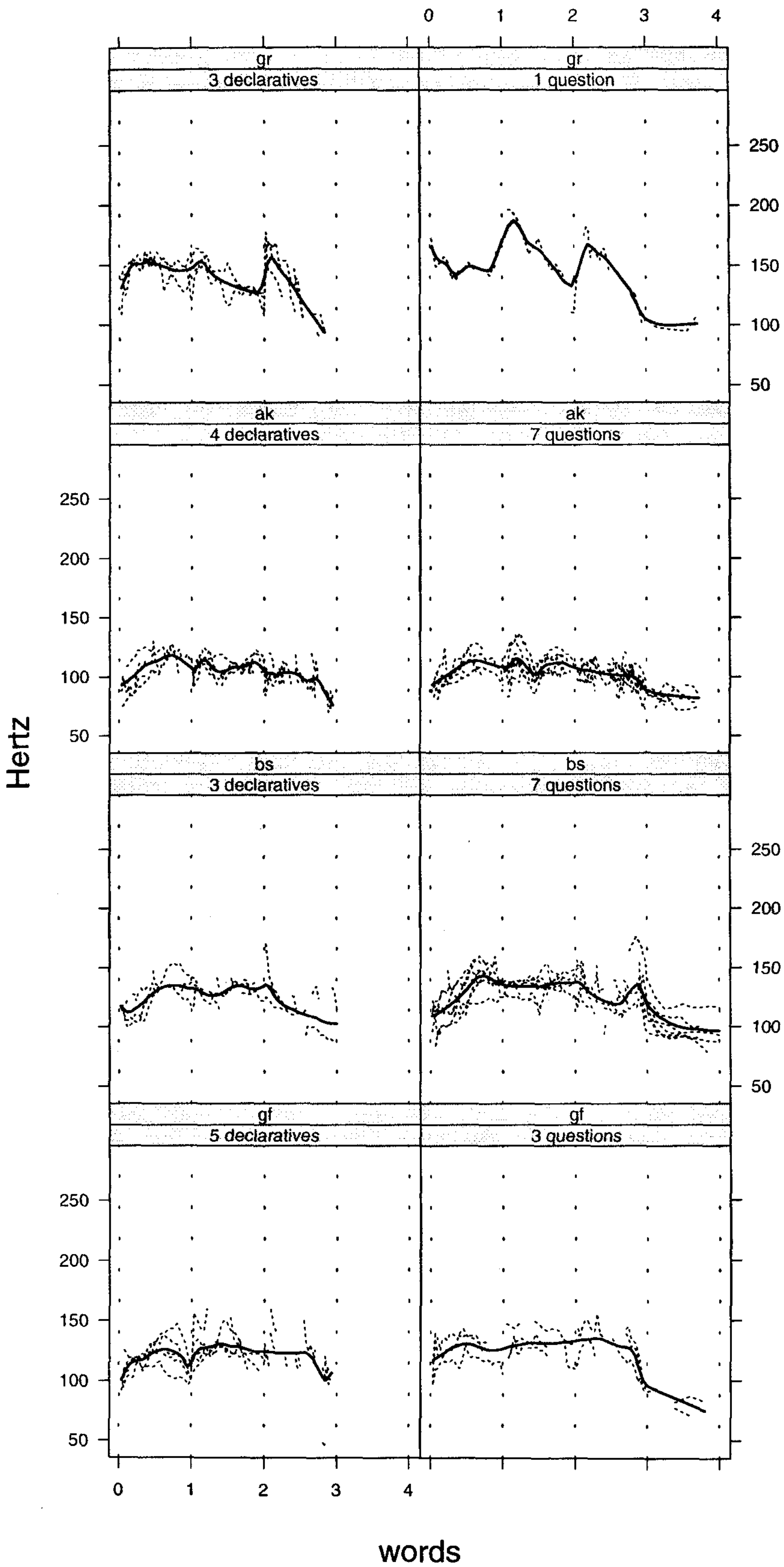
3.2.2 Pitch range medians

In order to determine whether there is a contrast in overall pitch range between declaratives and yes/no questions, the median pitch was computed for each sentence and plotted by speaker and sentence type. Results appear in (5). The plots in (5) include all sentences in the data, whereas those in (4) are just the three word (plus question clitic) tokens. In (5), sentence median pitches are plotted as open circles and compared by speaker and sentence type. Pitch is plotted on the horizontal axis. Vertical scatter is random. Results from female speakers appear at the top half of the figure; results from male speakers appear at the bottom half. 172 sentences are included in the data, i.e. all the sentences in the database, regardless of the number of words.

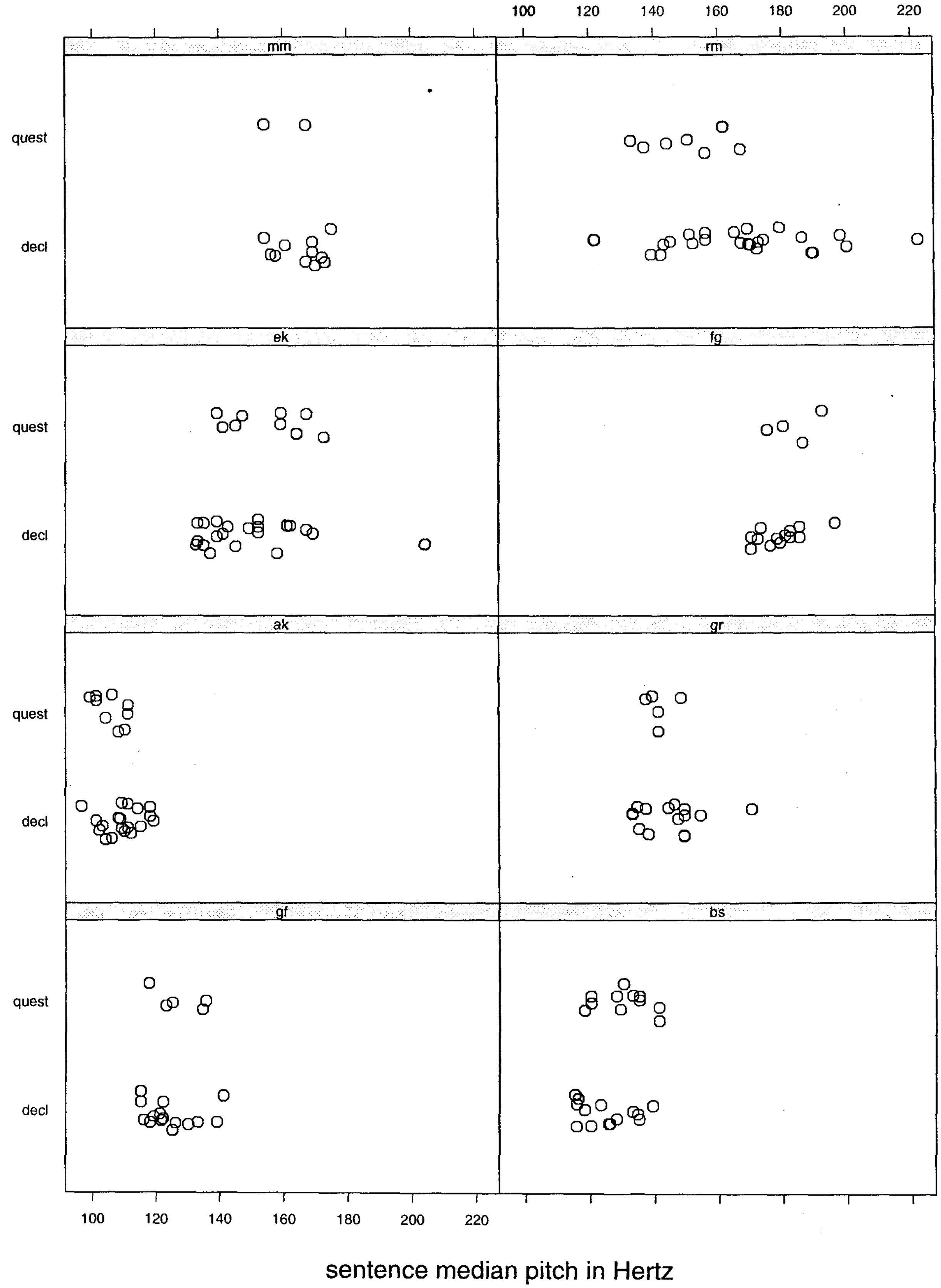
(4) female, 3–word declaratives, 4–word interrogatives



(4 cont.) male, 3–word declaratives, 4–word interrogatives



(5) Sentence median pitches by speaker and sentence type



4. DISCUSSION

4.1 Word contours

Smoothed results of declaratives (in the left-hand column of (4)) from the first five speakers, RM, MM, FG, EK and GR, show a mapping between word boundaries and pitch. There are pitch peaks near word beginnings and troughs near word ends. Downtrends are in evidence across the sentences in that sentences end lower than they start and successive lows (except for some cases of initial low) are successively lower. Results from these speakers suggest that in Unangan, each word is matched to its own intonation contour: peak near the beginning and trough near the end.

Results from declaratives of the last three speakers are not as easy to analyze. AK shows distinct troughs at word boundaries but also has a trough in the middle of the second word and peak near the end of the first, second, and third words. Results from speaker BS make a case for word-length intonation contours but here the contour is the opposite of that described for speakers one through five. BS shows troughs at word beginnings and peaks near word ends (except for the final word). Results from speaker GF show only one trough, at the boundary between the first and second words.

For the most part, word length information chunks are identified intonationally by a particular pitch contour: a peak near a word beginning and a trough near a word end.

4.2 Sentence contours

Pitch tracks in (3) and their smoothings in (4) suggest that sentence contours are concatenations of simple peak-trough word contours. These linked word contours are affected by downtrends to form a cascade effect in which each successive peak is slightly lowered from the beginning to the end of each sentence. In addition, sentences appear to begin by rising from a trough which delays the peak of the word; and then end by dropping steeply into one.

4.3 Declarative vs. yes/no question intonation contrasts

4.3.1 Smoothing

Since yes/no questions in Unangan are marked morphologically, the question asked here is whether they are also marked by intonation that contrasts with the intonation pattern for declaratives. This might seem like a pointless question; there is no 'need' for an intonation contrast between declaratives and yes/no questions in Unangan since morphology is used for this contrast; however, in other languages, e.g. English, declaratives contrast with yes/no interrogatives both syntactically, *She went to the store.* vs. *Did she go to the store?* as well as by intonation contour, final fall vs. final rise.

Smoothing results in (4) do not show that yes/no interrogatives have intonation contours that clearly pattern differently from declaratives. The appearance of a long, flat final word is misleading (and points out a flaw in this word-stretching and squeezing methodology). The fourth and final word in each of the yes/no questions is the clitic, *ii*. This word is very short in duration and its pitch drops steeply in the real time of actual speech just like the final drop described above for declaratives. C.f. (3)d and e. The fourth word question clitic appears as a flat, near horizontal, line in the smoothed examples in (4) because its duration is stretched.

Looking at the smoothed contours of yes/no questions in the right column of (4), results from the first speaker, RM, suggests that downtrending of troughs is not a feature of yes/no questions but this is not the case for the other speakers. The first five speakers show peaks near word beginnings and troughs near word ends for the y/n questions much like their declaratives. The sixth speaker, AK, has less dip at the second word boundary for interrogatives than for declaratives, otherwise there is not much contrast between his declaratives and y/n question contours. Speaker BS has less dip at the beginning of the second word for interrogatives than for declaratives and a noticeable peak at the end of the third word. Speaker GF has a less noticeable trough at the end of the first and word of interrogatives; otherwise, both his declaratives and questions can be said to be fairly flat.

There may be an intonational contour difference between declaratives and yes/no questions in the small peak on the final syllable of the final content word of the questions, just before the question clitic. This is most pronounced for speakers FG, AK, BS, and GF, showing up near the end of the third word. This small peak appears as a shoulder at the end of GR's question (note that there is only one question smoothed for GR.) The peak at this position does not show up in the smoothed questions contours for RM, MM or EK but the actual pitch tracks (which appear as dotted lines on the graphs in (4)) suggest that such peaks are produced. However, this final syllable peak also appears sometimes in the declaratives. This is most pronounced in the smoothed declaratives contour for AK but also appears in the dotted data from the other speakers. Further work will establish whether there is a correlation between this final peak and a declarative/question intonation contrast.

The conclusion here is that there is no intonation contour contrast between declaratives and yes/no questions with the possible exception of a small final peak on the last content word.

4.3.2 Median pitch comparison of declaratives and yes/no questions

Since intonation contours do not seem to contrast between declaratives and yes/no questions, we wondered whether questions are higher overall in pitch than declaratives. Median pitches per sentence plotted in (5) show that the range of the medians of yes/no questions is within that of declaratives; the center points are in about the same area for each speaker.

For each speaker, pitch medians for declaratives are dispersed over a wider range than that of questions but this may be an artifact of the data; there are more declarative sentences than yes/no questions. The speakers with the most questions show the least difference between the range of declaratives and questions.

5. CONCLUSION

Each content word in Unangan is defined intonationally by its own contour, a peak near the beginning and a trough near the end. The fact that a word-sized domain functions as an intonational domain may be related to the morphological structure mentioned in §1. Perhaps each word is treated intonationally as an individual phrase within the sentence since it must have inflectional suffixes and may have numerous derivational suffixes which provide it the potential for standing alone as a sentence. This kind of one-to-one isomorphy between intonation contours and each (content) word in a sentence has not been found in Japanese (Pierrehumbert and Beckman 1988), English (Pike 1945 and numerous others), and Bengali (Hayes and Lahiri 1991). However, in Cup'ik (Central Alaskan Yupik), geographically proximate to and distantly related to Unangan, content words have characteristic intonation contours (Woodbury 1993) but in Cup'ik such words begin with troughs and end with peaks .

Unangan sentence contours are serial peak-trough word contours. These are affected by downtrends so that sentences appear as a set of peak-trough cascades, each peak-trough lower than the preceding one. Sentences start with a trough that may delay the peak of the first word. Sentences end with a sharp drop. Yes/no question contours are not markedly different from declaratives. One contrast may be in the height of a small peak on the final syllable of the final content word of questions. The range of median pitch for questions is within that of the range of median pitch for declaratives.

The data analysis presented here is exploratory and incomplete. Tokens representing many different sentences were grouped together in an attempt to detect an overall trend. This explicitly ignores the possibility that different sentences might have different intonation contours, and that utterances of the same sentence by different speakers might display a common intonational pattern.

In the next phase of exploratory work, we remedy this by grouping tokens by sentence, rather than by speaker. We are plotting utterances of the same sentence by several different speakers together on the same plot and again applying a smoother.

6, ACKNOWLEDGEMENTS

The fieldwork was supported by a National Science Foundation grant SBR9511118 to Peter Ladefoged and Ian Maddieson, "Phonetic Structures of Endangered Languages". Many thanks to Moses Dirks who, as a native speaker, went over the translations, and glosses. Any remaining errors are entirely the responsibility of the first author. Funding for data corroboration was provided by the Jacobs Research Funds, 1997, "Aleut Intonation".

7. REFERENCES

Bergsland, K. (1994). *Aleut Dictionary.* Fairbanks, AK: Alaska Native Language Center.

Bolinger, D. (1978). "Intonation across Languages". In (ed. J. Greenberg). *Universals of Human Language*, Vol.II. Stanford: Stanford University Press. 471-524.

B. Hayes and A. Lahiri, "Bengali Intonational Phonology", *Natural Language and Linguistic Theory* Vol. 9, pp. 47-96, 1991.

Pierrehumbert, J. and M. Beckman (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.

Pike, K. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.

Rozelle, L. (1997). "The Effect of Stress on Vowel Length in Aleut", *University of Washington Working Papers in Linguistics* 15: 148-160.

Taff, A. (1992). "Uvularization of Vowels in Aleut", M.A. thesis, University of Washington.

Woodbury, A. (1993). "Against Intonational Phrases in Central Alaskan Yupik Eskimo", Presentation at LSA, Los Angeles, January, 1993.

Phonetic Structures of Scottish Gaelic

Peter and Jenny Ladefoged
University of California, Los Angeles
Alice Turk, Kevin Hind, and St. John Skilton,
University of Edinburgh

1. 1. Introduction

There are probably no monoglot speakers of Scottish Gaelic (henceforth Gaelic, pronounced **gælɪk**, as opposed to **gelɪk**, which is the form of the language spoken in Ireland). There are certainly very few fluent speakers under 60. Very little Gaelic is spoken on the Scottish mainland, but it is more widely spoken in the Outer Hebrides, the group of islands to the North West of the Scottish mainland.

Gaelic has a number of features of phonetic interest. The vowel system has some unusual back unrounded vowels that are found together in less than 1% of the world's languages (Maddieson 1984). There are contrasts in vowel length, oral and nasal vowels, and a large number of diphthongs. The stop system contrasts voiceless and voiceless aspirated stops in initial position, and has voiceless stops contrasting with pre-aspirated stops in medial position; there are no voiced stops. There are two series of stops which can loosely be called palatalized and velarized. (In some analyses the palatalized series are regarded as sequences of **p+j**, **b+j**, etc.) The same kind of contrast appears among the sonorants, with the additional possibility of there being plain sonorants which are neither velarized nor palatalized. The syllable structure is of special interest as there is a contrast between words that are clearly two syllables and those that have two vowels with an inter-vocalic consonant, but which might be considered phonologically monosyllabic and certainly are so in the opinions of speakers. These and other aspects of the phonetics will be discussed in this paper.

The data for this study were collected in the neighborhood of Greater Bernera, Lewis, in the Western Isles of Scotland. This is one of the forms of Gaelic described by Borgstrøm (1940) in his classic work on the phonology of the language. Recordings were made of 11 speakers, 7 men and 4 women. The youngest speaker was 63, the oldest was 88. Five of the men were between 70 and 74, the other two were in their sixties. Aerodynamic and palatographic data were collected from one 70 year old male speaker. In addition we had access to a major collection of palatographic data made in 1955 by Frederick Macaulay, a Gaelic speaker from South Uist, another island in the Hebrides. The South Uist dialect is different from that Bernera, but as Borgstrøm notes, the differences are mainly in "the phonemes chosen to make up words" (Borgstrøm 1940:9).

Each of the 11 speakers recorded the word list given in Appendix 1. Some of the speakers failed to produce a few of the words that were unfamiliar to them, as the list contained some words that were no longer in common use, These words were included because, for those who knew them, they provided the best possible examples of certain phonological contrasts. The transcription in the second column is that typical of most of the speakers, but there were a large number of small differences in pronunciation, some of them being due to the speaker producing a different form of the word. Further details, together with copies of the recordings, are available on request.

The general procedure was for KH, who is a fluent but non-native speaker of Gaelic, to give an English gloss of the required Gaelic word, and, on some occasions, to say the Gaelic word. The speakers then repeated the Gaelic word twice. Some of the speakers had a copy of the word list, so that they could read the words. In nearly all cases the style of speech was a formal pronunciation appropriate for citation forms.

Each speaker was recorded individually on a Sony DAT recorder with a Shure noise-canceling, close-talking microphone. All of the recordings were made in private homes, with little background noise, except for the gale that was blowing most of the time. Despite the outside wind noise, the signal to noise ratio was always greater than 45 dB.

1.2. Phonological inventory

Adopting a somewhat phonological stance, we can categorize the principal oppositions among Gaelic consonants as shown in Table 1. 1

Table 1. 1. The consonants of Gaelic.

	PLAIN			PALATALIZED			
	LABIAL	CORONAL	VELAR	LABIAL	CORONAL	VELAR	GLOTTAL
STOPS	**pʰ p**	**tʰ t**	**kʰ k**	**pʰʲ pʲ**	**tʰʲ tʲ**	**kʰʲ kʲ**	
NASALS	**m**	**n**			**nʲ**		
		(nˠ)					
FRICATIVES	**f v**	**s**	**x ɣ**	**fʲ**	**ʃ**	**ç ɣʲ**	
RHOTICS		**ɾ**			**ðʲ**		
		ɾˠ					
APPROXIMANTS		**j**					**h**
LATERALS		**l**			**lʲ**		
		lˠ					

The plain vs. palatalized contrast is pervasive in all forms of Gaelic. Among the coronals **n**, **r**, **l** there may be a three way opposition, in that a velarized form of each of these consonants contrasts potentially with both a plain and a palatalized form. (We did not find this contrast in our data for the coronal nasal.) In this table, the palatalized counterpart of **rˠ** is shown as **ðʲ**, a variant which is found in a number of Hebridean dialects of Gaelic (Borgstrøm 1940). Some speakers of the Bernera dialect (and of the dialects on the mainland opposite Bernera) have **rʲ**, a trilled **r** with a palatal fricative release. The approximant **j** is sometimes pronounced with a very close approximation that may be fricative. Velar nasals are not shown in the chart as they do not occur in underlying forms. They are found only before plain and palatalized velar consonants, as in *banca* **baŋkə** (bank) and *taingeil* **taŋʲkʲal** (thankful) and in derived forms such as *an gunna* **əŋunˠə** (the gun) and *an cat* **əŋʰatʰ** (the cat). Also not shown are the retroflex variants of the coronals **t**, **d**. **s**, **n** which occur before an underlying **r**. as in *freagairt* **frɤkəʈ** (answer), *bàrdachd* **paɖəxk** (poetry), *arsa* **aʂə** (said), *beàrn* **pʲaːɳ** (gap).

Gaelic has 9 vowels, each of which can be short or long. Seven of these vowels have the qualities associated with the seven vowel systems common in many languages. To these Gaelic adds the high and mid back unrounded vowels **ɯ** and **ɤ**, so that the complete set can be categorized as shown in Table 1.2, The qualities of these vowels will be discussed in section 2 of this paper.

Table 1.2, An overview of the vowels of Gaelic,

	FRONT	CENTRAL	BACK	
			ROUNDED	UNROUNDED
HIGH	**i iː**		**u uː**	**ɯ ɯː**
MID HIGH	**e eː**		**o oː**	**ɤ ɤː**
MID LOW	**ɛ ɛː**		**ɔ ɔː**	
LOW		**a aː**		

The stress and intonation of Gaelic will be discussed in the third section of this paper.

2. Gaelic consonants

Gaelic has two sets of stop consonants, one, orthographically *b, d, g*, is usually voiceless unaspirated **p**, **t**, **k**, and the other, orthographically *p, t, c*, is usually aspirated **p**^{h}, **t**^{h}, **k**^{h}. Each of these stops has a palatalized counterpart (orthographically *b, d, g* and *p, t, c*, followed by *e* or *i*), so that there are also **p**^{j}, **t**^{j}, **k**^{j} and **p**^{hj}, **t**^{hj}, **k**^{hj}. Intervocalically the aspirated consonants **p**^{h}, **t**^{h}, **k**^{h} and **p**^{hj}, **t**^{hj}, **k**^{hj} are realized as preaspirated, so that they could be transcribed ^{h}**p**, ^{h}**t**, ^{h}**k** and ^{h}**p**^{j}, ^{h}**t**^{j}, ^{h}**k**^{j} in an allophonic transcription. We will begin by considering the voicing differences among these consonants.

2. 1. Voicing

The phonetic nature of the voicing contrasts was determined by considering the onset of voicing for the stops in initial position in the words in Table 2.1, and the offset of voicing, the closure duration, and the onset of voicing for the intervocalic stops in Tables 2.2. The speech of the six male and four female speakers described above was analyzed using the Macquirer computer speech analysis system. The DAT recordings were re digitized at 22,050 Hz, and measurements made on waveform displays supplemented occasionally by wide band (172 Hz bandwidth) spectrograms.

Table 2.1. Words used in the investigation of VOT in stops in word initial position.

p^{h}	pana	**p**^{h}**anə**	pan
t^{h}	tana	**t**^{h}**anə**	thin
k^{h}	cana	**k**^{h}**anə**	can
p^{hj}	peallag	**p**^{hj}**al**^{ɣ}**ak**	rag
t^{hj}	teannadh	**t**^{hj}**an**^{ɣ}**əɣ**	tightening
k^{hj}	ceannach	**k**^{hj}**an**^{ɣ}**əx**	buying
p	baga	**pakə**	bag
t	daga	**takə**	pistol
k	gagach	**kakəx**	stammering
p^{j}	beannachd	**p**^{j}**an**^{ɣ}**əxk**	blessing
t^{j}	dealan	**t**^{j}**al**^{ɣ}**an**	electricity
k^{j}	gearradh	**k**^{j}**ar**^{ɣ}**əɣ**	cutting

Table 2.2. Words used in the investigation of aspiration and closure duration in intervocalic stops.

a^{h}p	apag	**a^{h}pak**	little ape
a^{h}t	atadh	**a^{h}təɣ**	swelling
a^{h}k	aca	**a^{h}kə**	at them
*a^{h}p^{j}	(not found)		
a^{h}t^{j}	aiteal	**a^{h}t^{j}əl^{ɣ}**	glimpse
a^{h}k^{j}	se taice a thug e dha	**t^{h}a^{h}k^{j}ə**	it's support that he gave him
ap	abaich	**apiç**	ripe
at	adag	**atak**	stook
ak	gagach	**kakəx**	stammering
***ap^{j}**	(not found)		
at^{j}	aideachadh	**at^{j}əxəɣ**	admitting
ak^{j}	aigeann	**ak^{j}ən^{ɣ}**	abyss

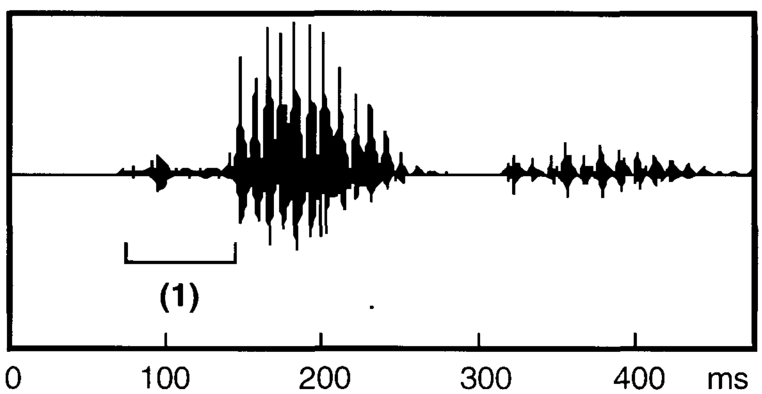


Figure 2.1. Waveform of Speaker 1 saying *cana* , **k^{h}anə**, (can). (1) aspiration interval.

The VOT of initial stops was measured directly from the waveform display. The word as a whole was first isolated as shown in Figure 2.1, and then the appropriate part of the display (slightly more than the duration of presumed aspiration) was expanded. The interval between the first indication of the release of the stop and the first full cycle of vocal fold vibration was measured. Often, as in Figure 2.1, the first part of the release had a somewhat smaller amplitude than a point about 20 ms later. But there was usually little doubt as to the moment of the release of the articulation.

The measurements of the intervocalic stops were also made on the waveform displays, but for these consonants, as exemplified in Figure 2.2, there was more uncertainty concerning the correct points to measure. The preaspiration interval is indicated by (1) in Figure 2.2. The cessation of voicing in this example was fairly easy to determine, but the beginning of the closure here and elsewhere was often less well marked. It may be that on a number of occasions there was incomplete closure. There are certainly some indications of noise on the spectrogram during interval (2), which has been marked as the closure in this stop. The release of the closure was also hard to determine. As the final syllable in this set of words was often said with even less intensity than in the utterance in Figure 2.2, the onset of voicing was sometimes difficult to determine. All these points should be borne in mind when considering the results reported below.

Despite these potential problems, there is no doubt that the phonological contrast between the two sets of Gaelic stops always depends on differences in aspiration and never on differences in voicing. None of the speakers ever had any indication of voicing during the closure in the word initial stops. Even in the intervocalic stops the only voicing that was ever apparent during the closure was a few periods at the beginning in the unaspirated stops. These so-called unaspirated intervocalic stops were in fact slightly aspirated as will emerge in the discussion of the results.

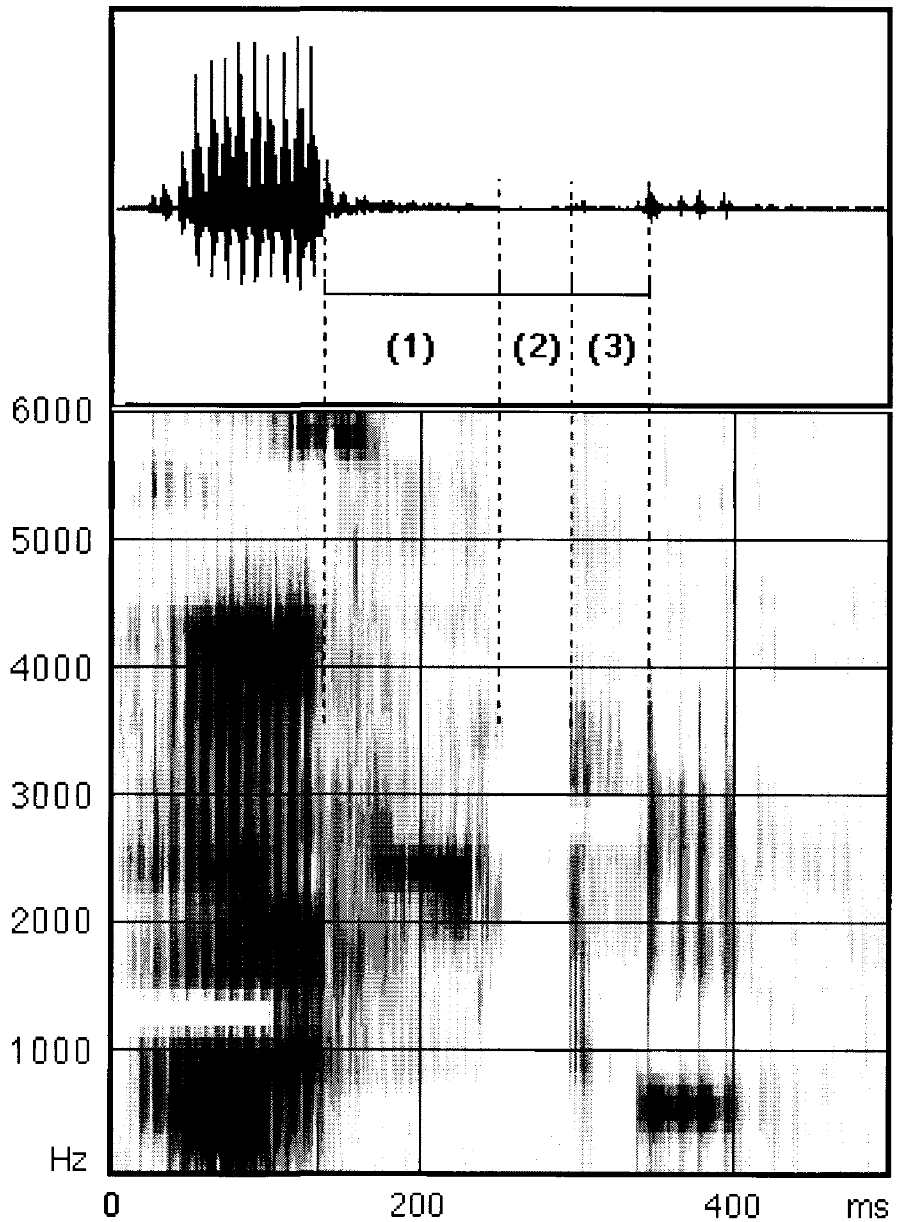


Figure 2.2. Waveform and spectrogram of Speaker 1 saying *aca*; **a^{h}kə**; (at them). (1) pre-aspiration; (2) closure; (3) post-aspiration.

Detailed differences in VOT were investigated by analysis of variance. The word initial stops were analyzed with VOT as the dependent variable, the independent variables being speakers, sex, place of articulation and palatalization. For the intervocalic stops, each of the intervals shown in Figure 2.2 was used as the dependent variable, with the same independent variables as for the initial stops. Post-hoc tests were made using Fisher's PLSD. We will consider as significant only values of $p < .01$, noting, when appropriate, the actual significance level achieved. The error bars in the figures provide further data by showing 5% confidence levels.

An overall view of the VOT of the initial stops as produced by the male speakers is given in Figure 2.3, and a similar view for the female speakers is given in Figure 2.4 The only significant difference between the male and female speakers for the four groups of consonants shown in Figures 2.3 and 2.4 is that the female speakers have a significantly shorter VOT ($p = .0013$) for the voiceless unaspirated stops. It is unlikely that these differences are due to the particular style of speech of the female speakers. Any stylistic difference, such as a greater overall rate of speech, would have affected the aspirated stops as much as the unaspirated stops. It seems more reasonable to presume that this difference has a purely physiological basis. The vocal folds will vibrate only when they are suitably positioned and there is an appropriate airflow between them. The male vocal folds, being heavier, take longer to commence vibrating once these conditions have been established. Because of the difference between the male and female speakers, and the smaller number of female speakers, only the male speakers will be considered in what follows.

There are a number of differences in VOT associated with the different places of articulation. Among the unaspirated stops, **p** and **k** are significantly different ($p < .01$), but **t** is not different from either. These findings are in accord with generally observed differences in VOT due to place of articulation, as summarized by Cho and Ladefoged (1997). There are several reasons why velar stops tend to have longer VOTs. One is that the articulatory closure involves extensive contact between the back of the tongue and the roof of the mouth with the result that the release of this closure is comparatively slow, sometimes involving a re-formation of the closure and a second burst. The slow release causes a delay in the lowering of the air pressure behind the closure, and hence a delay in a sufficient airflow between the vocal folds to enable them to vibrate. Another cause of the longer VOT is the relative sizes of the cavities behind the closure and in front of the closure. In velar stops the cavity behind the closure is smaller than it is for bilabial stops. As a result there may be a higher pressure during the closure, In addition, because the mass of air in front of the closure is larger for velar stops than for the other stops, there will be a greater impedance to the outgoing air. Both these factors will increase the interval before there is a sufficient flow of air between the vocal folds for vibrations to commence.

All the aspirated stops are (not surprisingly) very significantly different from all of the unaspirated stops ($p < .0001$). But there is only a trend in the expected direction of velar stops having a longer VOT than dental or bilabial stops; there are no significant differences between $\mathbf{p^h}$, $\mathbf{t^h}$ and $\mathbf{k^h}$. Apparently, in Gaelic, the aerodynamic effects of differences in place of articulation are not significant after the comparatively long aspiration interval.

Among the unaspirated palatalized stops, $\mathbf{p^j}$ and $\mathbf{k^j}$ are very significantly different, as are $\mathbf{p^j}$ and $\mathbf{t^j}$ ($p < .0001$ in both cases); in the case of the aspirated palatalized stops, only $\mathbf{p^{hj}}$ and $\mathbf{t^{hj}}$ are significantly different from each other ($p < .001$). The palatalized dental stops have long VOTs in comparison with palatalized stops at other places of articulation because their articulation involves the tongue being close to the roof of the mouth for much of the front part of the oral cavity. Thus the outgoing air has to overcome a considerable impediment before aerodynamic conditions suitable for vocal fold vibration can be established.

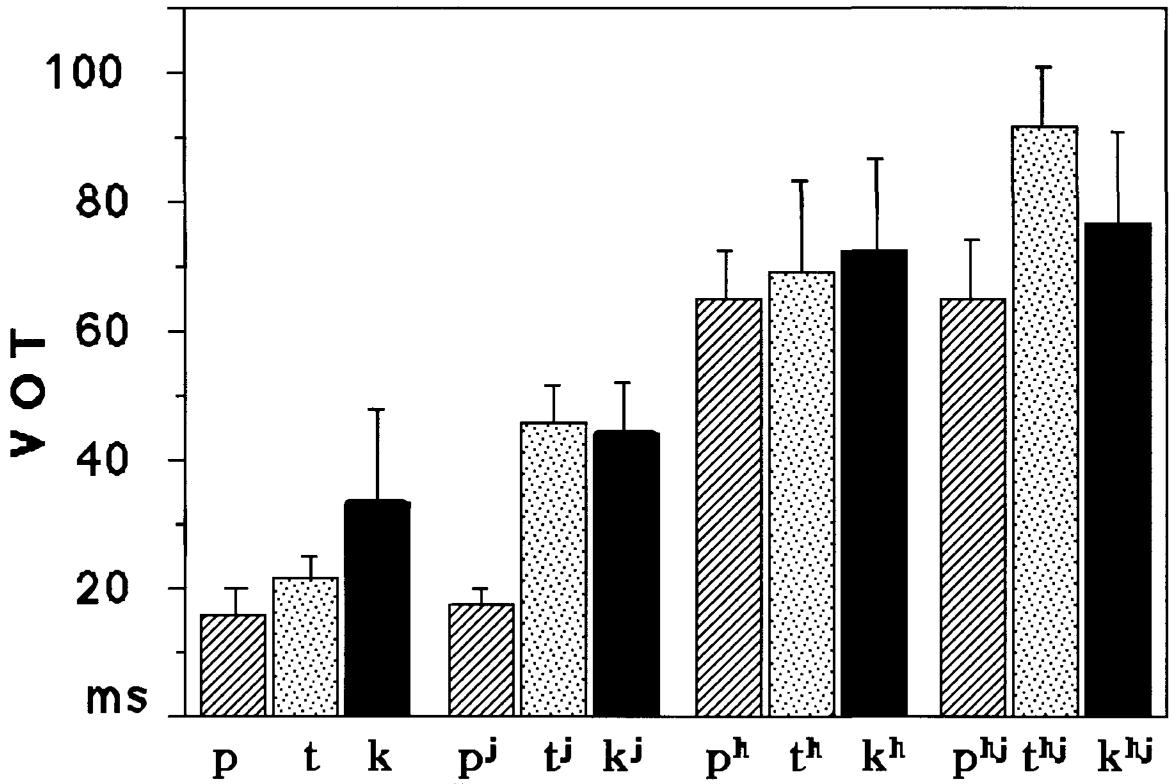


Figure 2.3. Mean VOT in Gaelic word initial stops as produced by 6 male speakers. Error bars indicate a 95% Confidence Interval.

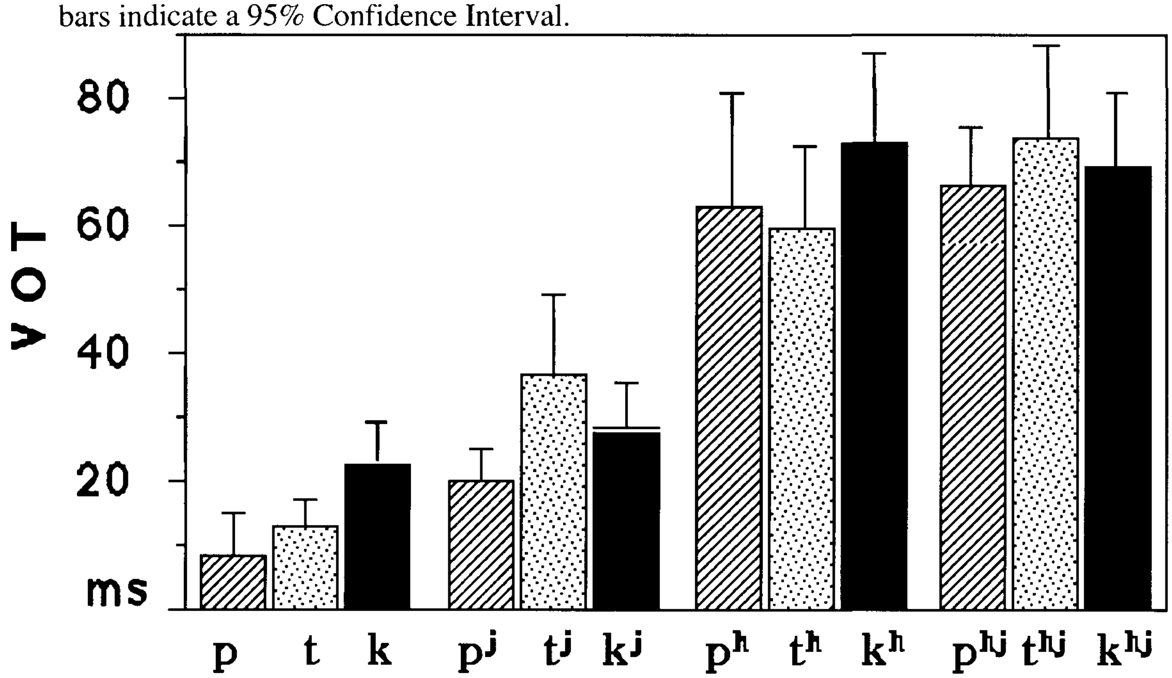


Figure 2.4. Mean VOT in Gaelic word initial stops as produced by 4 female speakers. Error bars indicate a 95% Confidence Interval.

In comparing the plain versus palatalized stops, there are no significant differences between the plain and palatalized pairs, **p** and **p**^{j}, **p**^{h} and **p**^{hj}, **k** and **k**^{j}, and **k**^{h} and **k**^{hj}, but **t** and **t**^{j}, and **t**^{h} and **t**^{hj}, are significantly different ($p < .0001$ and $p = .0015$, respectively). The addition of palatalization to the dental stops adds sufficient impediment to the outgoing air (sometimes even causing fricative noise) so that vocal fold vibration cannot start as early as in the case of the plain articulations.

The aspirated vs. unaspirated contrast among stops in initial position becomes transformed into a pre-aspirated vs. unaspirated contrast in medial position. It is as if the aspiration which occurs after the initial stops was moved earlier so that it occurs before them in intervocalic position. These points can be observed in the aerodynamic records of *Pabach* **pʰapəx** (person from Pabaigh), and *apak* **aʰpak** (little ape), reproduced in Figure 2.5.. The audio signal, oral air flow and oral air pressure were captured on a Macintosh powerBook with the aid of the Macquirer multichannel recording system. The audio recording was made by means of a microphone placed just outside a mask that the speaker held against his face below the nose. This signal, which is slightly overloaded (and is distorted as the sound was muffled by the mask), was digitized at 11,400 Hz. The mask capturing the flow of air from the mouth was connected to the Macquirer transducer system, and digitized at 440 Hz. A thin tube between the speaker's lips recorded the pressure of the air in the front part of the mouth (but not behind the velar articulations which occur at the ends of these words). This pressure was also transduced by the Macquirer system and digitized at 440 Hz.

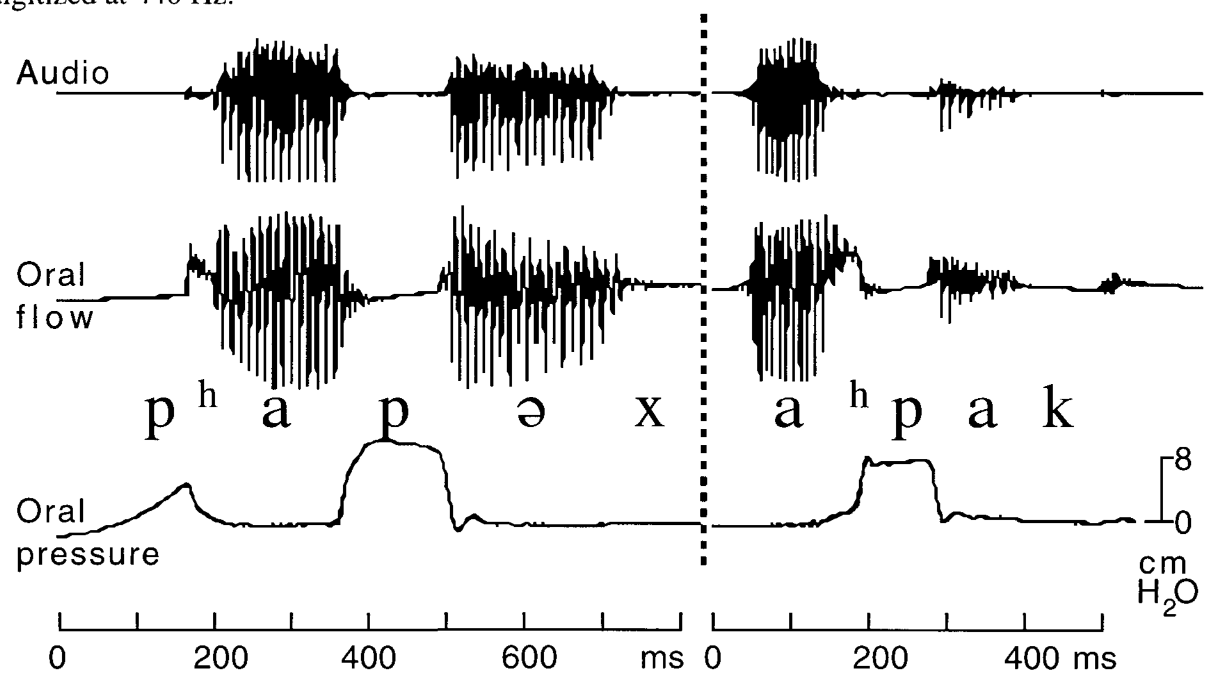


Figure 2.5 Aerodynamic records of speaker 1 saying *Pabach* **pʰapəx** (person from Pabaigh), and *apag* **aʰpak** (little ape). (See text for details of the instrumentation.)

The aerodynamic records show the build up of oral pressure for the first consonant, **pʰ**. The peak pressure is only about 5 cm H_20, considerably lower than the oral pressure for the intervocalic **p** later in the same word. The small initial pressure accounts for the comparatively small burst, which is typical of this speaker (and several of the other Gaelic speakers we recorded). The voiceless air flow (the aspiration) lasts for about 50 ms. Later in this word, the intervocalic voiceless unaspirated consonant, **p**, has two or three periods of voicing during the first part of the closure while the oral pressure is rising. After the release there is a very short interval before regular voicing commences. In the second word, the intervocalic stop, **ʰp**, has an increasing rate of mean oral air flow at the end of the first vowel. During this interval of preaspiration the oral air pressure rises slightly, but it is not until the lips close (at about 200 ms on the time scale) that the oral pressure rises substantially. The release of this consonant is accompanied by a very short interval before regular voicing occurs.

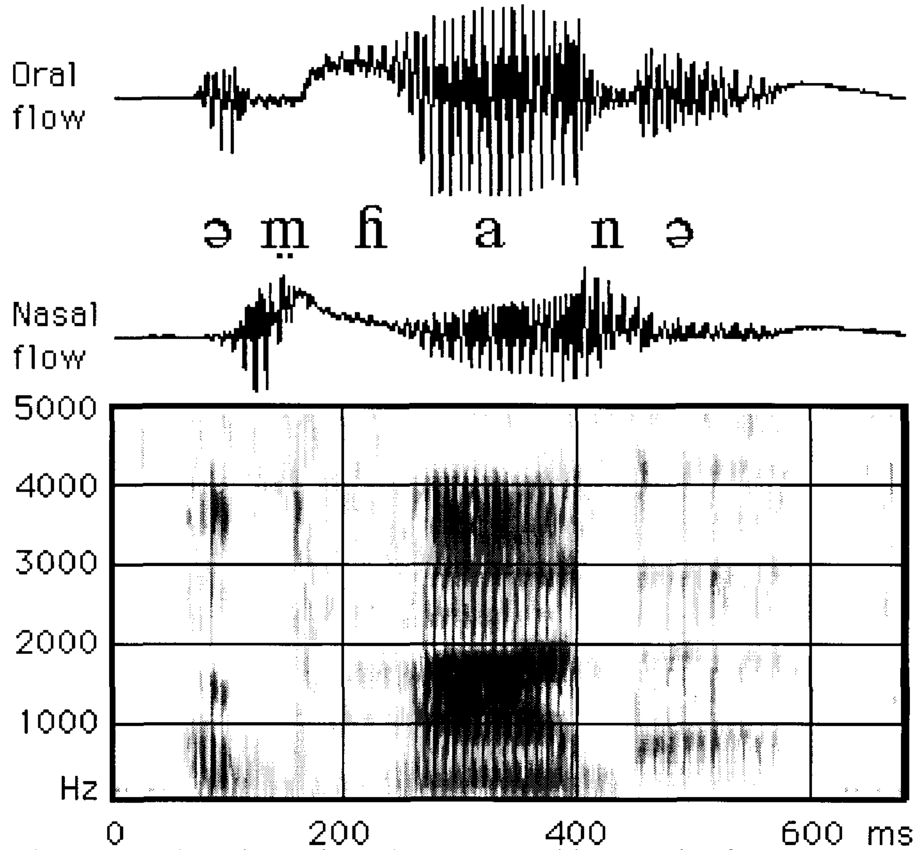


Figure 2.6. Aerodynamic and spectrographic records of speaker 1 saying *am pana* **əm̤ɦanə** (the pan). Note that the spectrogram is made from an audio recording that is somewhat distorted by the airflow mask.

When a word with an initial aspirated consonant occurs after certain closed class words which historically ended in a nasal consonant, the initial aspirated plosive is replaced by a homorganic aspirated nasal. For example, the phrase *am pana* (the pan) is pronounced as **am̤ɦana**. The articulations involved are illustrated in Figure 2.6, which shows aerodynamic records of this phrase, made as described above. This figure also includes a spectrogram made from the audio recording. Again it should be remembered that the audio recording is distorted by the procedures involved in making the aerodynamic records, so the spectrogram does not show the segments as they would normally appear. Nevertheless the combined aerodynamic and spectrographic records make it plain that there is no oral stop in this phrase. After the initial vowel the oral flow ceases while the nasal flow increases. The first part of this nasal flow has some breathy voice vibrations, which cease when the oral flow starts again, somewhat abruptly, as the lips open. This cannot be called a stop release, however, as there is still considerable nasal airflow. As the nasal airflow decreases and the oral airflow increase breathy voice vibrations appear in the oral airflow. The whole sequence is similar to that in the aspirated nasals in the Tibeto-Burman language, Angami, as described by Blankenship et al. (1993).

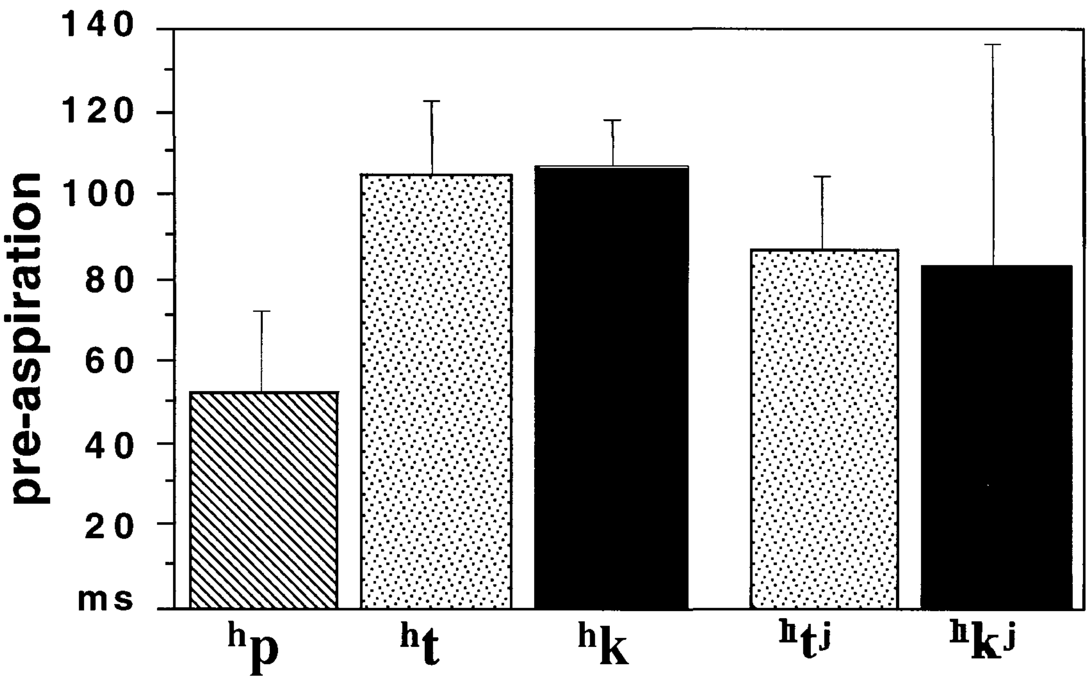


Figure 2. 7. Mean preaspiration for six male speakers in Gaelic intervocalic aspirated stops.

Preaspiration is a very notable feature of Gaelic. Generally speaking, the voiceless interval between the end of the vowel and the start of the consonant closure is considerably longer than the aspiration in initial consonants. The mean duration of the pre-aspiration for the six male speakers saying the words in Table 2.2 is shown in Figure 2.7. There is one difference (for which we have no satisfactory explanation) due to place of articulation. Pre-aspiration is very significantly less before **p** than before **t** or **k** ($p < .0001$). The palatalized stops are not significantly different from each other, nor are the plain stops significantly different from their palatalized counterparts. Intervocalic $^{h}\mathbf{p}^{j}$ does not occur. The large error bars associated with $^{h}\mathbf{k}^{j}$ reflect the fact that several of our speakers had problems recognizing the only example we could find for $^{h}\mathbf{k}^{j}$ after **a**, leaving us with only four valid tokens. Intervocalic $^{h}\mathbf{k}^{j}$ is quite common in conjunction with other vowels, but we wanted to keep the vowel environment the same for all the intervocalic consonants.

As we noted in the discussion of Figure 2.2, there was often some noise in the part of the articulatory gesture that we took to be the stop closure, which caused some measurement problems. Perhaps as a result of these problems, we did not find any noteworthy differences in closure duration. All these stops were completely voiceless, even having a slight delay after the release of the closure, before the onset of the voicing for the following vowel. There were no significant differences due to place of articulation in this post-aspiration, although the trend towards a longer VOT after velars is apparent in the data shown in Figure 2.8. As in the case of the initial stops, only the dental palatalized stops had a strong tendency towards a longer VOT than their counterparts ($p = .0134$). The difference between $^{h}\mathbf{k}$ and $^{h}\mathbf{k}^{j}$ did not reach even that level of significance, perhaps because of the small number of tokens of $^{h}\mathbf{k}^{j}$.

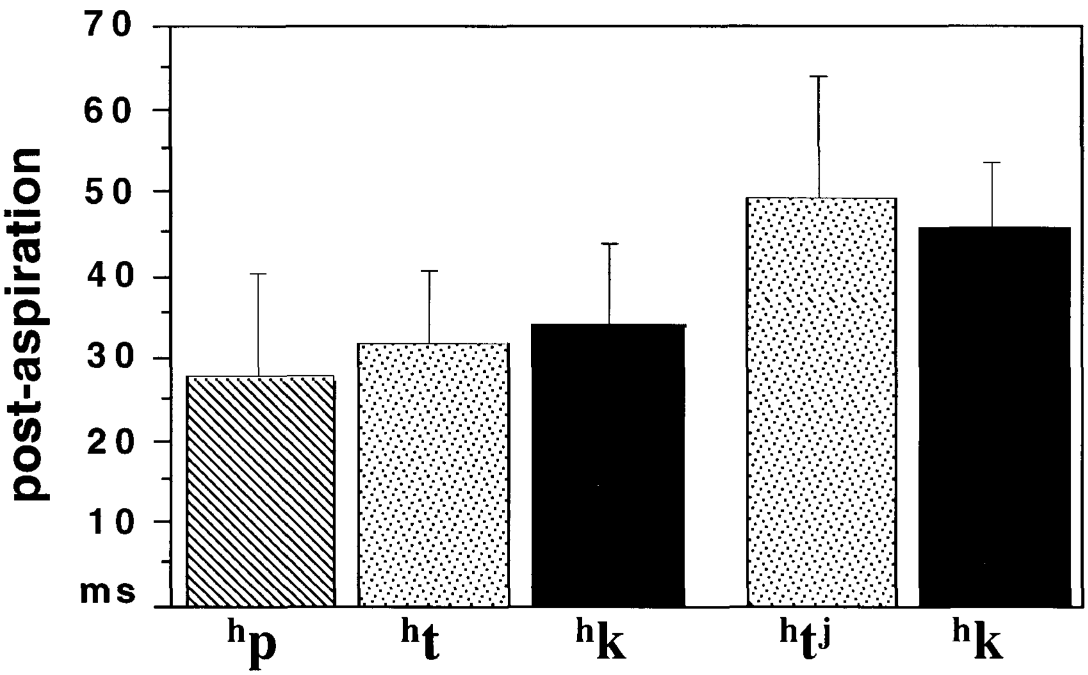


Figure 2.8. Mean post-aspiration in Gaelic intervocalic aspirated stops.

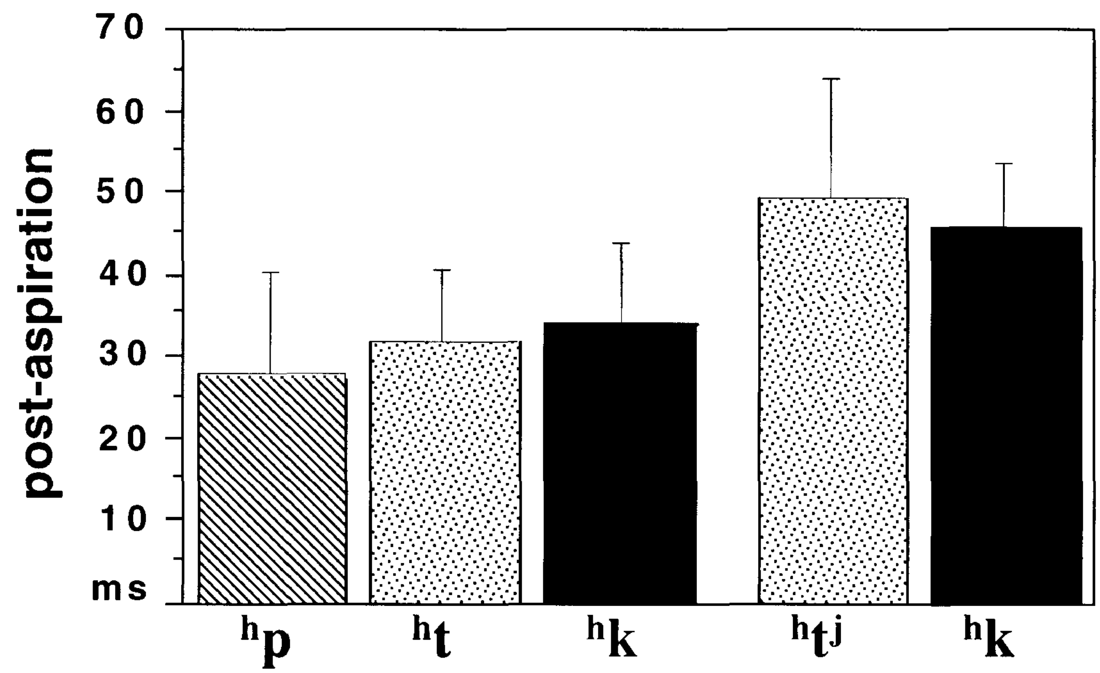


Figure 2.9. Mean closure duration in Gaelic intervocalic unaspirated stops.

In intervocalic position the unaspirated stops were completely voiceless except for a few periods voicing at the start of the closure as illustrated by the first word, *Pabach* **pʰapəx** (person from Pabaigh), in Figure 2.5. The mean closure durations for these stops are shown in Figure 2.9. There is an orderly (but not statistically significant) decrease in duration from the labial to the velar stop, which is counterbalanced by the duration of the VOT that occurs on the release of these stops, shown in Figure 2.10. In these so-called unaspirated stops, **k** probably has more aspiration than **p** ($p = .0132$), and also than **t** ($p = .0298$). Only the palatalized velar stops do not follow these trends. For reasons that are not apparent to us, the closure duration of **k^j** is very significantly longer than that of **t^j** ($p < .0006$), and the VOT of **k^j**, is significantly longer than all the other stops ($p < .01$).

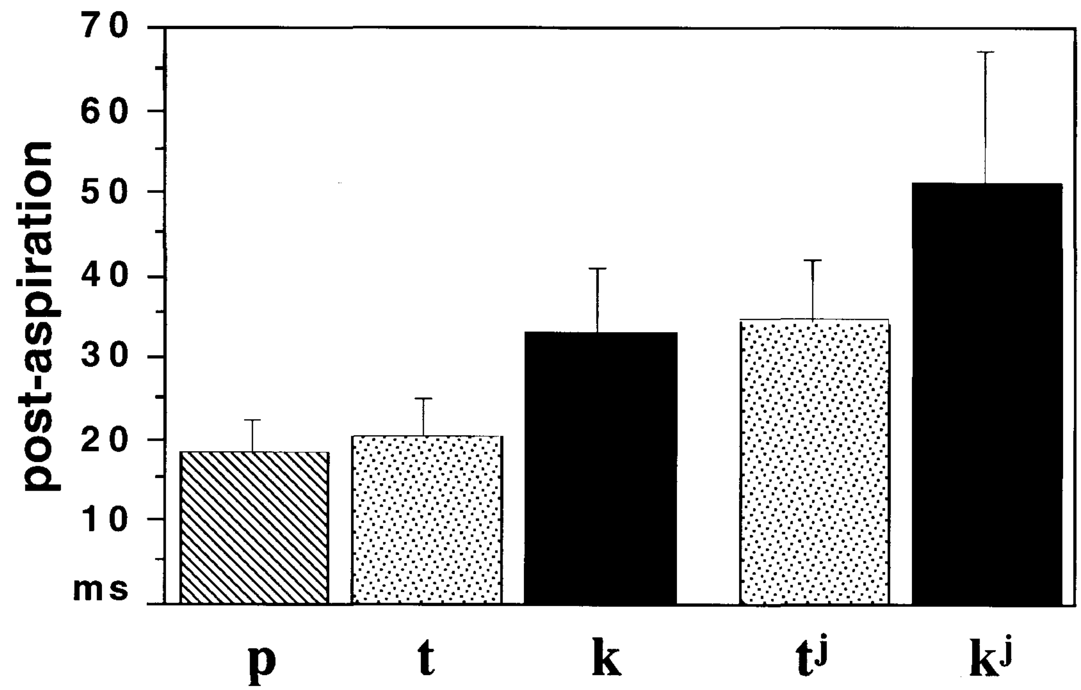


Figure 2.10. Mean post-aspiration in Gaelic intervocalic unaspirated (sic) stops.

2. 2. Palatalization

Spectrograms of contrasts between aspirated plain and palatalized stops are shown in Figure 2.11; the unaspirated consonants are illustrated in Figure 2.12 The acoustic differences in these stops are in the bursts and in the formant transitions. For both the aspirated and unaspirated stops, the burst in the palatalized series have a greater intensity in the region around 3–6000 Hz. However, probably the major cue to palatalization (we have not tested this in auditory perception studies) is the high second formant at the beginning of the release, usually accompanied by a lowered third formant. Vowels after palatal consonants always begin with a palatal onglide.

The lateral, rhotic and nasal coronals **l**, **r**, **n** are interesting in that in the dialect of Bernera Gaelic described by Borgstrøm (1940) each of these sounds is involved in a three way contrast. In Old Irish there was probably a contrast between laminal dental and apical alveolar coronals, each of which could be palatalized or velarized, giving four possibilities. If we use the dental diacritic to mark the laminal dentals, these possibilities for the laterals would be **l̪**$^{\gamma}$, **l**$^{\gamma}$, **l̪**^{j}, **l**^{j}. In present day Bernera Gaelic (and in many other dialects, including a dialect of Irish Gaelic described by Ní Chasaide, 1979), the first two of these, **l̪**$^{\gamma}$ and **l**$^{\gamma}$, have fallen together, and the palatalized apical alveolar **l**^{j} has lost its palatalization. The remaining sounds are thus **l**$^{\gamma}$, **l**, **l**^{j}, which are the symbols we have used in describing our data. There is a similar contrast between **r**$^{\gamma}$ and **r**, but the third possibility is not **r**^{j} but a dental fricative **ð**^{j}. We did not find any evidence of a contrast between **n**$^{\gamma}$ and **n** in our data, even in the words described by Borgstrøm (1940) as having these sounds; there was, however, the contrast between **n** and **n**^{j}.

In our analysis of the laterals we found exactly what Ní Chasaide (1979) had found many years previously in her investigation of Irish. We examined spectrograms of the words in the upper part of Table 2.3, showing the three laterals in initial and medial position. All of our speakers produced these laterals with markedly different second formants. The second formant had the lowest frequency for **l**$^{\gamma}$, an intermediate frequency for **l**, and the highest frequency for **l**^{j}. Spectrograms of one set of tokens from speaker 1 are shown in Figure 2.13. The variations in the second formant frequency are probably associated with a high back position of the tongue for **l**$^{\gamma}$, a neutral position of the body of the tongue for **l**, and a raised position of the front of the tongue for **l**^{j}.

Table 2.3. Words used for spectrographic analysis of laterals and rhotics.

	INITIAL			MEDIAL		
l$^{\gamma}$	latha	**l**$^{\gamma}$**a-ə**	day	balla	**pal**$^{\gamma}$**ə**	wall
l	leat-sa	**la**^{h}**tsə**	with you	baile	**palə**	town
l^{j}	leannan	**l**^{j}**an**$^{\gamma}$**an**	sweetheart	caillidh	**k**^{h}**al**^{j}**i**	lose (future tense)
r$^{\gamma}$	rathad	**r**$^{\gamma}$**a-ət**	road	bearradh	**p**^{j}**ar**$^{\gamma}$**əɣ**	shearing
r	rinn	**rain**^{j}	did	aran	**aran**	bread

A similar analysis was conducted for the rhotics in the words in the lower part of Table 2.5. The third possibility, the palatalized rhotic, was not included in the spectrographic analysis, as it is realized as a palatalized dental fricative. The findings were similar in that the second formant had a lower frequency for **r**$^{\gamma}$ than for **r** for all eight male speakers. The third formant was also significantly lower for initial **r**$^{\gamma}$, but not for medial **r**$^{\gamma}$. Both initial rhotics were approximants with voiceless onsets, and could be symbolized as **ɹ̥**$^{\gamma}$ and **ɹ̥** in a narrow transcription. The medial rhotics were both voiced taps that could be symbolized **ɾ**$^{\gamma}$ and **ɾ**. Spectrograms of one set of tokens from speaker 1 are shown in Figure 2.14.

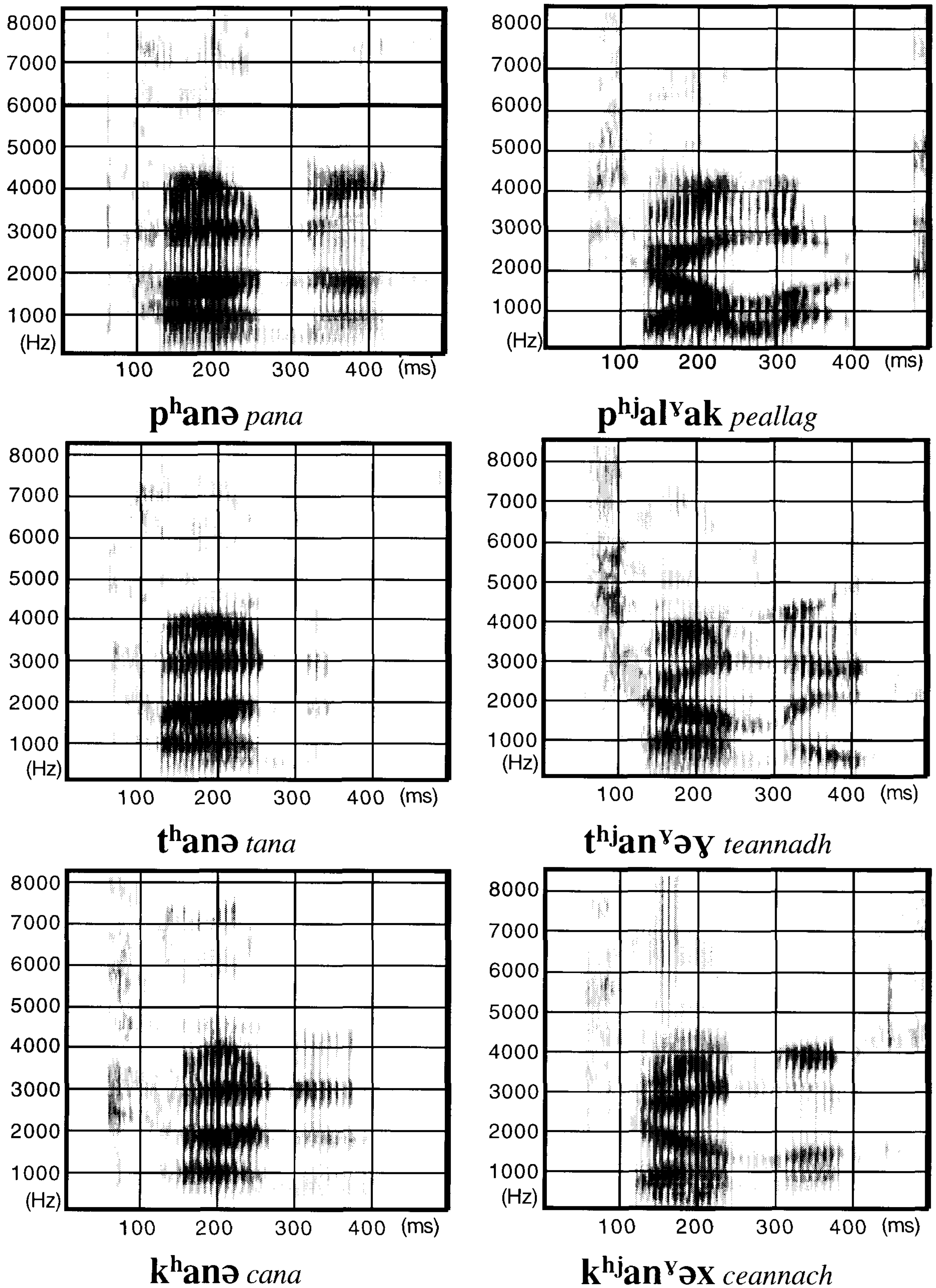


Figure 2.11 Plain and palatalized aspirated consonants as spoken by speaker 1.

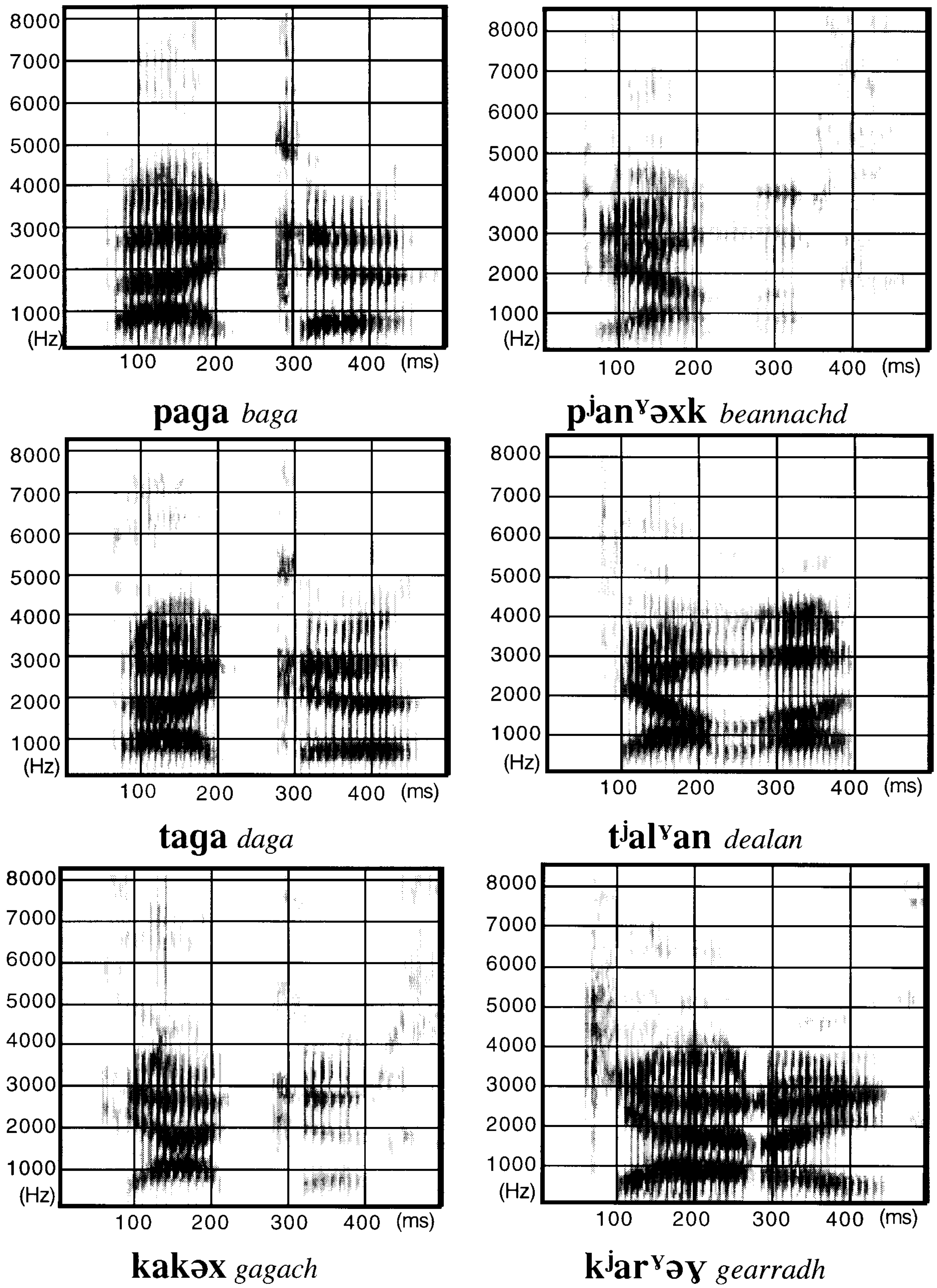


Figure 2.12 Plain and palatalized voiceless unaspirated consonants as spoken by speaker 1.

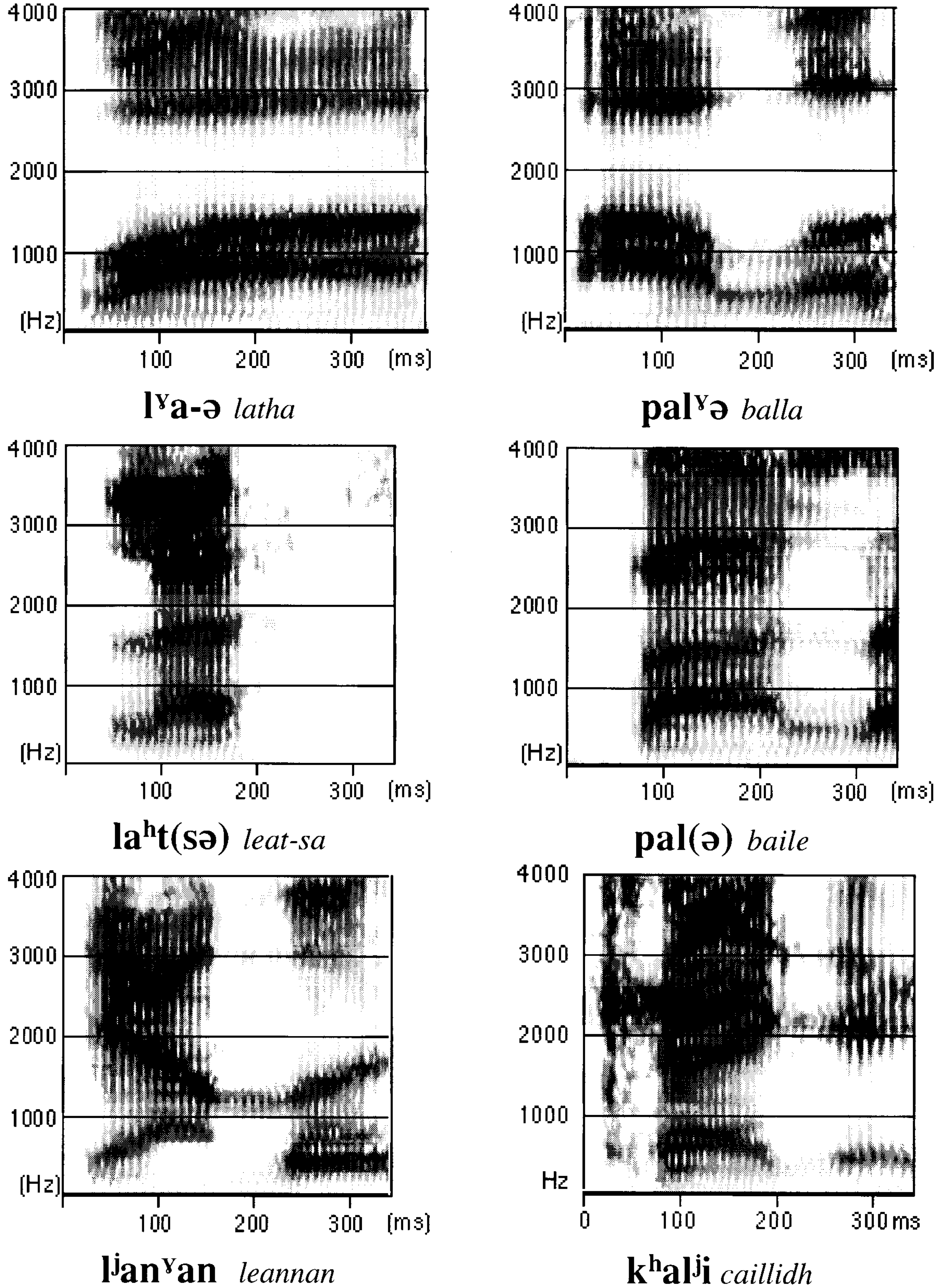


Figure 2.13. Spectrograms illustrating the contrasts among the three laterals in initial position (left-hand column) and medial position (right-hand column).

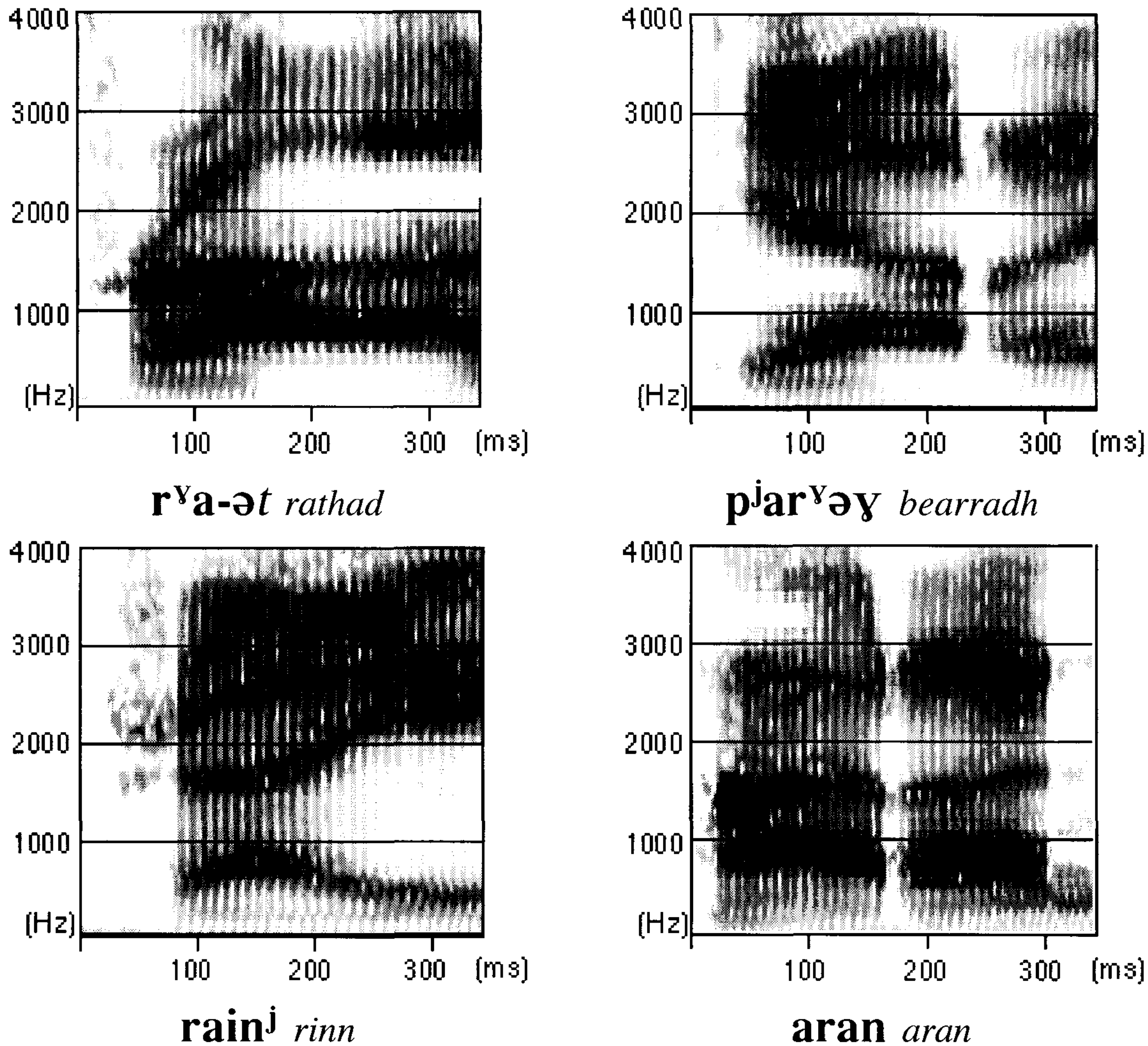


Figure 2.14 Spectrograms illustrating the contrasts among the three laterals in initial position (left-hand column) and medial position (right-hand column).

We have some palatographic data on one of our speakers, enabling us to describe the articulatory differences in some of these sounds. We were also given access to a set of palatograms made by Frederick Macaulay in 1955. These palatograms, which are of very high quality, have not been analyzed to any great extent. Macaulay's Gaelic is slightly different from that spoken in Bernera, where our recordings were made, in that he was a native of South Uist. But judging from the limited number of palatograms we made of one of our speakers it appears that there are no substantial differences in the consonantal articulations. Borgstrøm (1940) indicates that the differences between the two dialects are more likely to be in the distribution of the phonemes (i.e. in which words particular phonemes occur) rather than in their phonetic realization.

Macaulay's palatography database contains no minimal pairs establishing the contrast in palatalization. This is partly because palatalized consonants are associated with front vowels and non-palatalized consonants with back vowels. In addition, Macaulay limited his palatographic investigation to words in which there was only one consonantal contact, so as to ensure a clear view of the articulation of that consonant. This prevented him from using some of the words we used in our acoustic investigation, such as **lʲanˠan** *leannan* (sweetheart). His database has no examples of **lʲ**. Illustrations of the contrasts between **lˠ**, **l** are shown in Figure 2.15. The most apparent difference is in the amount of articulatory contact with the teeth. The teeth have been wiped clean in **l^{y}ã:v** *làmh* (hand), but there was no contact between the tongue and the teeth in **le:m** *leum* (jumped). We may conclude that **l^{y}** has a laminal denti-alveolar articulation and **l** an apical alveolar articulation. There is no sign of the raising of the back of the tongue in **l^{y}**, perhaps because the tongue has been narrowed for the lateral articulation. But in **l** the front of the tongue is raised so that more of the palate alongside the molars has been contacted. All 13 examples of **l^{y}** in Macaulay's palatograms show clear evidence of the tongue having contacted the teeth so as to form a denti-alveolar articulation. In all 15 examples of **l** the teeth remain untouched, confirming the apical alveolar articulation of this sound.

l^{y}ã:v làmh   le:m leum

Figure 2.15. Palatograms (made by F. Macaulay) illustrating the laminal denti-alveolar articulation of **l^{y}** and the apical alveolar articulation of **l**.

3.1 Gaelic vowels

As mentioned section 1, Borgstrøm (1940) describes the Bernera dialect as having nine short and nine long oral vowels as shown in Table 1, He makes the point that the labels 'back' and 'front' are "to be taken in their phonological, and not in their strict phonetic sense". In particular, "the phoneme **u**(:) is realized either as a pure back vowel **u**, or as a back-mixed **u̇**." (Borgstrøm (1940:11). Later he notes: "There are several varieties of **u** sounds in the dialect. The most frequent sound is the one I write **u̇**, **u̇**:; its point of articulation is not strictly back, but rather between back and mixed and very narrow; the lips are rounded though not protruded". Borgstrøm also notes that these vowels exhibit a noticeable amount of contextual variability. He points out that there are more rounded versions of the same vowel occurring "in the neighborhoods of the consonants **x**, **ɫ**, **ɴ**, **ʀ**, and the cerebrals **ḍ**, **ṇ** [IPA **ɖ**, **ɳ**]; ... these two sounds, **u** and **u̇**, are strikingly different from each other, and the difference is perceived by many native speakers." (Borgstrøm (1940:32). He also notes that "The articulation of **λ** [IPA **ɯ**] varies a good deal according to its surroundings. Before palatal consonants most people use a very high and narrow articulation with a tendency to glide from the back-flat into a mixed position; before non-palatal consonants the articulation is often rather open." (Borgstrøm (1940) , p. 35.)

One of the more interesting aspects of Borgstrøm's description is the claimthat this dialect has two non-low back unrounded vowels, **ɯ**(:) and **ɤ**(:), in addition to the high back rounded vowel **u**(:), for both the short and long vowel series. According to Maddieson (1984) non-low back unrounded vowels are quite rare, and are found in less than 1% of the world's languages. Maddieson (1984) lists 20 languages that have unrounded high back vowels. Of these 20, 8 have only an unrounded high back vowel **ɯ**, while 12 have both rounded and unrounded high back vowels, **u** and **ɯ**. Only one language (Apinaye) is classified as having two back unrounded vowels, with an inventory containing the mid back unrounded vowel **ɤ** as well as the high back unrounded **ɯ**.

A primary purpose of this section is to see if there is a clear distinction between the high non-front rounded vowel **u** and the high non-front unrounded vowels **ɯ** and **ɤ**, for both the short and long series. To this end we present plots of formant values from 7 male speakers' vowels, and we discuss statistical analyses of measurements of the non-low back vowels. If, as Borgstrøm indicates, there are indeed nine distinct vowels in this dialect, then we would expect to see nine relatively distinct areas in the vowel space for both long and short vowels.

3.2. Short Vowels

Table 3.1 lists the words used for investigating short vowels. As far as possible, we tried to elicit each vowel in two environments: (1) following an initial alveolar consonant; and (2) following an initial bilabial consonant. The environment following the test vowel was not controlled. We were unable to find a well-known bilabial environment for the short vowel **ɯ** so this word was elicited in word-initial position.

Table 3.1. Word-list used for investigating short oral vowels.

i	(cha) tig	**(xa) tʲikʲ**	won't come
i	binneas	**piɲəs**	harmony
i	biodag	**pitak**	dagger
e	deireadh	**tʲeðʲəɣ**	end (noun)
e	beinne	**peɲə**	mountain
e	beithe	**pehə**	birch tree
ɛ	dearcadh	**tʲɛr̥kəɣ**	looking
ɛ	beatha	**pɛhə**	life
ɛ	bean	**pɛn**	woman
ɛ	beadaidh	**pɛti**	ill-mannered
a	tagsaidh	**tʰaksi**	taxi
a	bannag	**panˠak**	new year's cake
a	badan	**patan**	little sod
ɔ	toddy	**tʰɔdi**	toddy
ɔ	tocasaid	**tʰɔʰksatʲ**	barrel
ɔ	bonnach	**pɔnˠəx**	scone
ɔ	bodach	**pɔtəx**	old man
o	togail	**tʰokal**	lifting
o	bogadh	**pokəɣ**	soaking
ɤ	tagh	**tʰɤɣ**	choose
ɛ	beagan	**pɛkan**	a little
u	tugainn	**tʰukiɲ**	come on!
u	bun	**pun**	foot (of hill)
u	buideal	**putʲəlˠ**	bottle

ɯ	turadh	**tʰɯrəɣ**	dry weather after wet
ɯ	uinneag	**ɯɲak**	window

In general, a token was included in our analyses if it was clear that the speaker had correctly identified the word, and if there was no hesitation over its pronunciation. While we had hoped to elicit each vowel in the context of both a preceding bilabial context and a preceding alveolar context, we were unsuccessful in the case of the vowel **ɤ**, where speakers consistently produced the word 'beagan' as **pɛkan** where we had expected **pɤkan**. This word is shown in the list in Table 3.1 in the order appropriate for an **ɤ** vowel, but the tokens were categorized as containing **ɛ**. There were a few other anomalous pronunciations. Speaker 2 produced both tokens of *dearchadh* as **tʲar̥kəɣ**, where **tʲɛr̥kəɣ** had been expected. The near-homophonous pair *beithe* **peh**- *beatha* **pɛh** led to some speakers pronouncing the wrong word. If they were corrected by KH, and repeated the correct version, only the correction was taken. There were also other occasional confusions, such as in the case of the word *bonnach* which appeared as **pɔnˠəx** or **panˠəx**. Only if it was clear that the speaker was producing the correct word was it included. For many speakers the word beadaidh **peti** did not seem to be well-known and led to a variety of pronunciations, and was thus rejected for four of the speakers. Rejections of this type were more common when speakers were reading the word, rather than repeating it after hearing it said. In general, any token where the speaker was unsure of the pronunciation was rejected. After all dubious tokens had been omitted, there were a total of 349 reliable short vowel tokens for analysis from a possible 382.

Measurements were made from the 7 male speakers' recordings. The DAT recordings were low-pass filtered at 7,800 Hz and re-digitized at a sample rate of 16,000 Hz. The data was examined with the use of XWaves software from Entropics, run on a SUN workstation with a UNIX-based operating system. Waveforms and spectrograms were used to delimit the test vowels. Vowels were judged to begin at the release of the previous consonant (as it appeared on the waveform), and to end at the beginning of the closure for the following consonant. Formant measurements were taken at the midpoint of each vowel, and were estimated automatically from LPC power spectra (with a 25 ms rectangular window and an analysis order of 18). The LPC formant values were checked by eye using the relevant power spectra in conjunction with wide-band spectrograms. One dubious token was rejected because the formant values obtained from the LPC analysis were not relatable to what was observed on the spectrogram.

The mean formant values are given in Table 3.2, and a plot of the overall vowel space in the post-alveolar environment is presented in Figure 3.1. In this and subsequent figures the vowels are plotted with the second formant (F2) on the abscissa and the first formant (F1) on the ordinate, using a Bark scale which reflects perceptual distances. The ellipses enclose all vowels within two standard deviations of the mean. The vowel symbols are at the centers of the ellipses except for **u**, which is placed outside the ellipse as otherwise it would have been on top of **ɯ**.

Table 3. 2 Mean formant values (Hz) and standard deviations (s.d.) for short vowels in the post-alveolar context

VOWEL	F1	S.D.	F2	S.D.
i	346	35	2077	91
e	415	29	1896	41
ɛ	579	39	1690	109
a	652	35	1447	94
ɔ	569	52	1082	124
o	496	44	985	72
u	405	39	1427	92

(ɯ)	405	30	1433	75
ɤ	505	44	1426	45

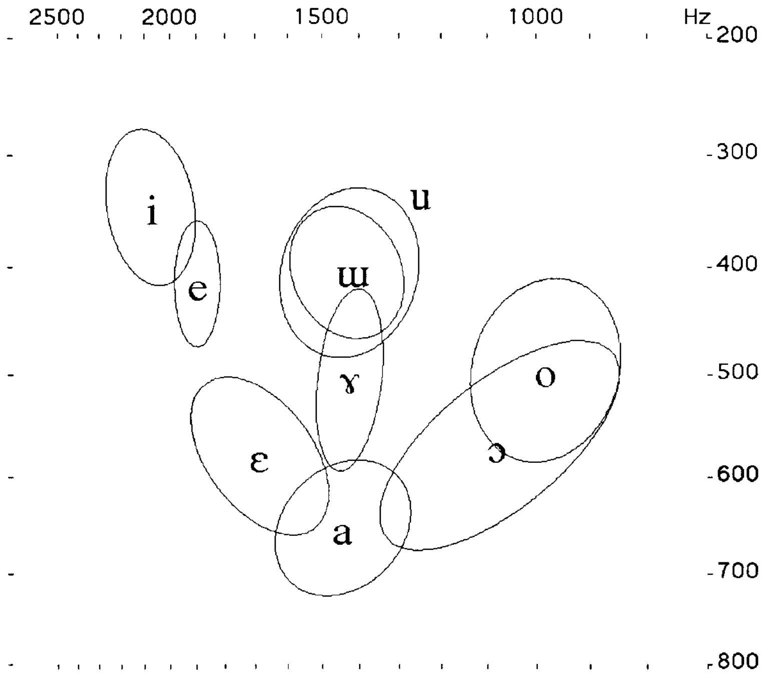


Figure 3.1. The first and second formants of Gaelic vowels following an alveolar consonant as produced by 7 male speakers of Gaelic. See text for a description of the axes and the form of the plots.

For the most part, the pattern for each speaker is well represented by the pattern for the means, as indicated by the large vowel symbols. The overlap between ellipses is largely due to the constant displacements that are a function of individual vocal tract size and shape. While most of the vowels are reasonably distinct, this is far from true for the vowels **u** and **ɯ** which are virtually completely overlapped. The other back unrounded vowel, **ɤ**, is positioned in the middle of the vowel space, and overlaps with some of the tokens labeled as **ɯ** and **u**. The spaces for these vowels will be discussed in greater detail below, but before we do this there are some other cases that are worth commenting on.

The high front vowels are close together. In vowel plots for the speakers individually (not included in this paper), tokens of **i** and **e** overlap for every speaker except speakers 2 and 3. As predicted by Borgstrøm (1940), there is a small amount of overlap in the spaces for **a** and **ɛ**. This is due to the word *dearcadh*. As we have noted, the vowel produced by speaker 2 in this word was classified as an **a**, based on its auditory impression, and confirmed by visual examination of the plot. But although he was the only speaker whose vowel we decided to reclassify, all the other speakers produced their lowest tokens of **ɛ** (highest value of F1) in this word. Borgstrøm (1940) specifically refers to such cases "The articulation and acoustic impression of the *æ* -sounds cover an area between *a* and *ɛ,* which gives room for several varieties (p.27)". According to our analysis, the vowel should be classified with the **ɛ** tokens for all speakers except for speaker 2.

Some of the other overlaps were also due to particular words. The **ɔ** vowel in *bonnach,* **pɔnˠəx** is lower than might be expected. Two speakers originally said this word as **panˠəx** before noticing that a different word had been requested. But even after their original pronunciations had been discarded, and only the corrected tokens included in the analysis, there was still overlap with the **a** tokens in other words. The **ɔ** vowel also varied in another way. The tokens of this vowel in *tocasaid*, **tʰɔʰksatʲ**, were all produced closer to **o**.

We will now consider the high, non-front vowels. Of particular interest is whether there is a clear difference between the **u** and **ɯ** tokens. If a language has two distinct vowels, **ɯ** and **u**, then we would expect their formant values to be statistically different. In this case we might expect that the F2 for **u** should be lower than fo **ɯ**, as lip-rounding effectively increases the length of the vocal tract, thus lowering the formant frequency. Similarly, any difference between **ɯ** and **ɤ** should at least show up as a difference in height, expressed by a difference in F1.

As only the alveolar environment had a full set of tokens for comparison, a repeated measures ANOVA was run for just those tokens presumed to contain **u** and **ɯ** occurring in that environment. The results are shown in Table 3.3.

Table 3.3. Repeated measures ANOVA, for **ɯ** versus **u** with each of the formants as dependent variable.

	u MEAN	**ɯ** MEAN	F(8,48)	p
F1	405	405	0.071	n.s
F2	1427	1433	0.168	n.s.
F3	2110	2232	5.88	<0.01

The difference in F3 may be partly attributed to the difference in the segments following the vowels. The word containing **u** was *tugainn*, **tʰukiɲ**, 'come on', and that for **ɯ** was *turadh*, **tʰɯrəɣ**, 'dry weather after wet'. A velar consonant thus followed **u**, which could lead to a lower F3 by the end of the vowel in comparison to the **ɯ** tokens which had a tap (or alveolar approximant) following the vowel.

The plots and statistical analyses of the non-low vowels in the alveolar environment might be considered to indicate that the tokens which we had originally classified as **ɯ** (following Borgstrom's description) were actually pronounced as **u**. However, a different conclusion emerges when we consider some of the speakers as belonging to different groups, and consider the pronunciations of separate words. An examination of the supposed **ɯ** tokens indicates that all speakers produced the vowel in *turadh*, supposedly **tʰɯrəɣ**, in a similar way to the **u** vowel in *tugainn*, **tʰukiɲ**, apart from the F3 difference mentioned above. We may therefore say that *turadh*, is now pronounced as **tʰurəɣ**. However speakers behaved in different ways when it came to the other tokens originally taken to be examples **ɯ**.

We will discuss first a group of speakers which we will call the separating group, consisting of speakers 1, 4, 6 and 7. This group produced tokens of what had been thought of as **ɯ** in two different ways. Like all the other speakers they produced an **u** vowel in *turadh*, Moreover it was still true for this group that only F3 was statistically different for the first vowels in *tugainn* and turadh ($F=12.8$, $p<0.05$), in keeping with the results for the speakers as a whole. But this group produced all tokens of the vowel in *uinneag* , (supposedly **ɯɲak**), in a similar way to the tokens of **ɤ** (as in *tagh*, **tɤ**). We may therefore say that for these speakers *uinneag* is now pronounced as **ɤɲak**, Figure 3.2 shows a plot of the relevant vowels. The ellipse enclosing the vowels in*uinneag* is close to the ellipse for the tokens of **ɤ** (as in *tagh*, **tɤ**), and the ellipse for the vowels in *turadh* overlaps with that for **u** (as in *tugainn* **tʰukiɲ**).

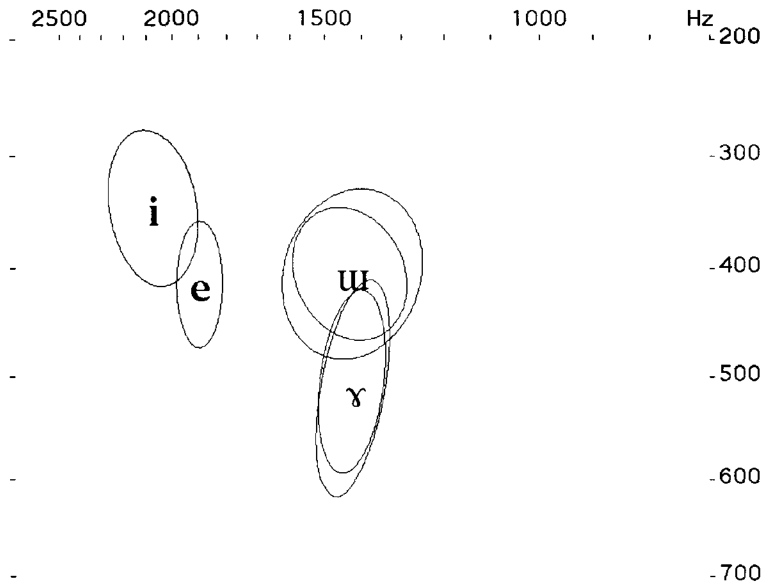


Figure 3.2 Ellipses enclosing the **u** vowels in *tugainn* **t^{h}ukiɲ** and the **ɤ** vowels in *tagh*, **tɤ**, as produced by all speakers. Within each of these ellipses the slightly smaller ellipses enclose the vowels in *turadh* , **t^{h}ɯrəɣ**, and in*uinneag*, **ɤɲak**, as pronounced by speakers 1, 4, 6 and 7. The front vowels **i** and **e** are also shown so as to provide reference points.

There is also what we may call a non-separating group, consisting of speakers 2 and 3 who produced what was expected to be **ɯ** in much the same way as **u** in all cases: the vowels in both *turadh* and *uinneag* were like those in *tugainn*, *bun* and *buideal*. For the alveolar environment, there were no statistically significant differences between what was expected to be **ɯ** and **u** for any of the formants, so it does seem that for these speakers, **ɯ** has completely merged with **u**. Figure 3.3 shows the vowels **u**, **ɤ** and what was expected to be **ɯ** for these two speakers.

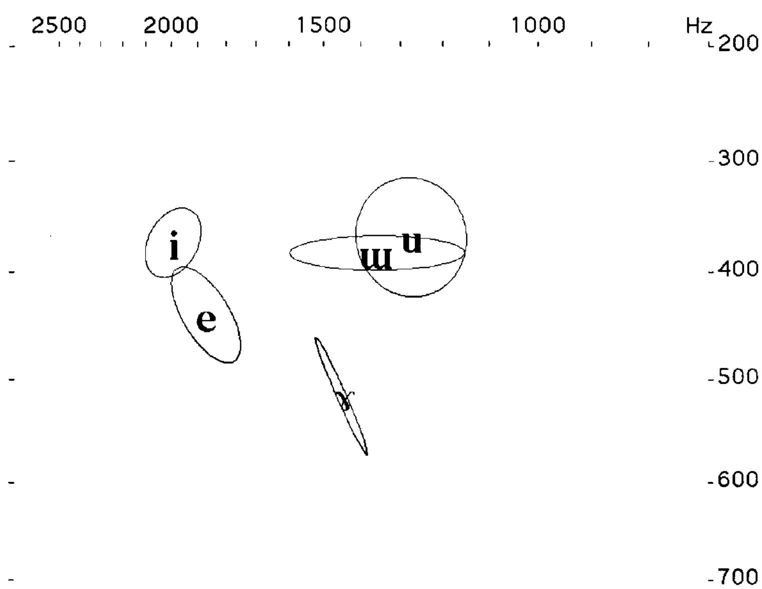


Figure 3.3. Vowel spaces for the vowels **u**, **ɤ** and what was expected to be **ɯ** for speakers 2 and 3. The front vowels **i** and **e** are also shown so as to provide reference points.

Speaker 5 did not fall clearly into either of the two groups. He produced a distinction between **ɯ** and **u**, so could perhaps be classified as a 'separator'. His vowel space for **u**, **ɤ** and what was expected to be **ɯ** is shown in Figure 3.4. Two of the tokens fall within the **u** region, but the tokens from *uinneag* are more fronted towards **ɛ**, rather than lowered towards **ɤ** as in the separators group.

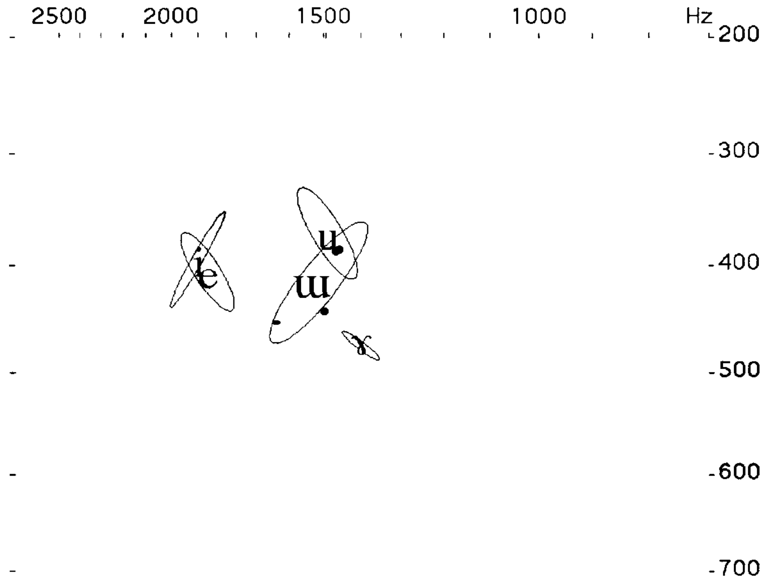


Figure 3.4. Vowel spaces for the vowels **u**, **ɤ** and what was expected to be **ɯ** for speaker 5. The front vowels **i** and **e** are also shown so as to provide reference points. The four tokens of what was expected to be **ɯ** are enclosed within one ellipse. Two of these tokens are within the **u** ellipse.

Summarizing our findings for the high back and central short vowels, the evidence presented here does not support the view that the short vowels **ɯ** and **u** are distinct for these speakers, at least in the alveolar environment. There were two different patterns of behavior, which we loosely labeled 'separators' and 'non-separators'. For the separators, vowels that were initially classified as **ɯ** could be said to have merged with **u** in some words, and with **ɤ** in other words. For the non-separators, nearly all the original **ɯ** vowels have merged with **u**. We should also note that the combined **ɯ** and **u** vowel space has a higher value for F2 than might be expected, indicating that this remaining high back vowel is either further forward or has less rounding than a cardinal **u** vowel. In any case, it is clear that for short oral vowels, there is no evidence for a distinction between two high central or back vowels in this dialect.

The analysis by Borgstrøm (1940) was a detailed one, and various possibilities are available to explain why our findings are different. First and foremost, we might have found a distinction if we had used other vowels. However, assuming our data is representative of the dialect as a whole, the lack of a contrast between **ɯ** and **u** may not be surprising given the relative rarity of back unrounded vowels in the world's languages. But how is one to explain the separate groupings of speakers? Examples that Borgstrøm (1940) described as **ɯ** seem to have different realizations for different speakers. The separators split the **ɯ** tokens by height, patterning tokens from *turadh* with **u**, and tokens from *uinneag* with **ɤ**. The non-separators, on the other hand, merged all tokens along with **u**. On further investigation, it turns out that the non-separators were the speakers that had had most exposure to English, one being born in England, and the other having been away from the area for a substantial period of his life. English does not have a distinction for roundedness in back vowels. The **ɯ** vowel historically present in this dialect has either disappeared, or is in the process of being 'split' so that it merges with nearby vowels. Under this scenario, Separators are the people who split the **ɯ** vowel (which was once in between **u** and **ɤ** in terms of the vowel space) merging it with nearby vowels without the roundedness distinction. The nearby **u** and **ɤ** vowels act as 'magnets' for the new realizations. Support for this notion comes from three areas:

Firstly, with the exception of speaker 5, the vowels that were classified as **ɯ** pattern with one or the other of **u** or **ɤ** rather than falling in between the two areas. The non-separators produce all their tokens like **u**. If they really are more English-oriented in their phonological system, this may not be surprising, as English high back vowels are somewhat rounded by default.

Secondly, speaker 5 also seemed to split the **ɯ** tokens, but along a front-back dimension rather than a height one. Assuming that a **ɯ** vowel would be situated in between **ɛ**, **u** and **ɤ**, it is conceivable that a disappearing **ɯ** vowel could be attracted towards these vowels. In cases where the vowel moved frontwards, the unrounded nature of the vowel would have been maintained, as is typical for front vowels. In other speakers (the separators), the vowel split along the height dimension. This might be further evidence that nearby vowels act as magnets for the **ɯ** vowel.

Thirdly, we must remember that Gaelic is in serious decline. Studies such as that by Dorian (1978) on East Sutherland Gaelic show how English has deeply affected the language. Our Bernera speakers did not have the syntactic and morphological influences reported by Dorian, but they may have been influenced by English in more subtle ways. Perhaps it is no surprise that the English vowel system, which has no roundedness contrasts, is affecting the Gaelic one. This may be especially true for speakers who have been bilingual from a very young age, perhaps since birth. Speakers may be adopting English vowel qualities and contrasts.

Finally in this discussion of the short vowels, we must again note the very real possibility that the word set we used may have simply been inadequate. There may be several sets of words showing a three way contrast between **u**, **ɯ** and **ɤ**, but we just didn't use them.

3.3. Long Vowels

Table 3.4 shows the elicitation list for the words containing long vowels. As for the short vowels, we tried as far as possible to elicit each vowel in two environments: (1) following an initial alveolar or palatalized consonant, and (2) following an initial labial consonant. This proved difficult for **ɛː**, perhaps not surprisingly, as Borgstrom describes **ɛː** as being fairly rare. We also did not expect to find a common post-alveolar instance of **ɤː**, although it turned out that all our speakers pronounced *daoine* with **ɤː** instead of the expected **ɯː**.

In general, the environment following the test vowel was not controlled. For **uː**, we included contexts which were expected to elicit the default (more front) articulation as well as the more back allophone which occurs in front of retroflex allophones such as **ʈ**, **ɳ**. The more front (default) allophone is transcribed as **uː**, and the back allophone as **u̠ː** . Some of the words were elicited only for a subset of speakers; these are listed in Table 3.5. These words were not part of the original set, but were thought of during the course of the fieldwork, and were added in an attempt to have more examples of the back, non-low vowels.

Table 3.4. Word-list used for investigating long oral vowels.

iː	dìobhairt	**tʲiːvəʈ**	vomit
eː	eud	**eːt**	zeal, jealousy
iː	dìnnear	**tʲiːnʲərʲ**	dinner
iː	bìodach	**piːtəx**	tiny
eː	dèineas	**tʲeːnʲəs**	vehemence
eː	beudach	**peːtəx**	harmful
ɛː	nèamh	**nʲɛ̃ːv**	heaven
ɛː	cha b’e	**xa bɛː**	it wasn’t
ɛː	Gaidheal	**kɛː-əlˠ**	Gael
aː	deàrrsadh	**tʲaːrsəɣ**	shining
aː	bàidse	**paːtʲə**	musician’s fee
ɔː	Deòrsa	**tʲɔːʂ̥ə**	George
ɔː	bòid	**pɔːtʲ**	oath
oː	deònach	**tʲoːnəx**	willing
oː	bò	**poː**	cow
ɤː	foghlam	**fɤːlˠəm**	education
ɤː	adhbhar	**ɤːvər**	reason
uː	diùrrais	**tʲuːriʃ**	secret
uː	bùth	**puː**	shop
u̠ː	dùirn	**tu̠ːɳ**	fists
u̠ː	bùird	**pu̠ːʈ**	tables
ɯː	daoine	**tɤːnʲə**	people
ɯː	baothair	**pɯːhəðʲ**	idiot

Table 3.5. Supplementary word-list used for investigating long oral vowels.

uː	ùradh	**uːrəɣ**	making new
ɯ	saoradh	**sɯːrəɣ**	freeing
ɯː	faoileag	**fɯːlak**	seagull
ɯː	faobhar	**fɯːvər**	edge of blade
ɤː	adhradh	**ɤːrəɣ**	worshiping
ɯː	faolainn	**fɯːlˠinʲ**	exposed beach
ɯː	faoilleach	**fɯːlʲəx**	January

Measurements and criteria for including tokens in the analyses were made as described for the short vowels. There were only a few elicitation problems. Speaker 2 produced the first vowel of *deineas* as **iː** instead of **eː** as expected, and all speakers produced the first vowel in*daoine* as **ɤː** instead of **ɯː** (judged auditorily and confirmed by looking at the plots). The categorization of these vowels was changed accordingly. We failed to elicit any tokens of *baothair* for speaker 2.

The vowel space representing the long vowels in Table 3.4 is shown in Figure 3.5. As in the case of the short vowels, the pattern for each speaker is generally well represented by the pattern for the means, indicated by the large vowel symbols, with the overlap being largely due to properties of the individual vocal tracts. The overall pattern is much as one might expect, with the possible exception of the placement of the back vowels **uː**, **ɯː** and **ɤː**. It must be remembered, however, that this is a plot of the acoustic properties of vowels, which is similar to the traditional vowel chart only for front unrounded and back rounded vowels.

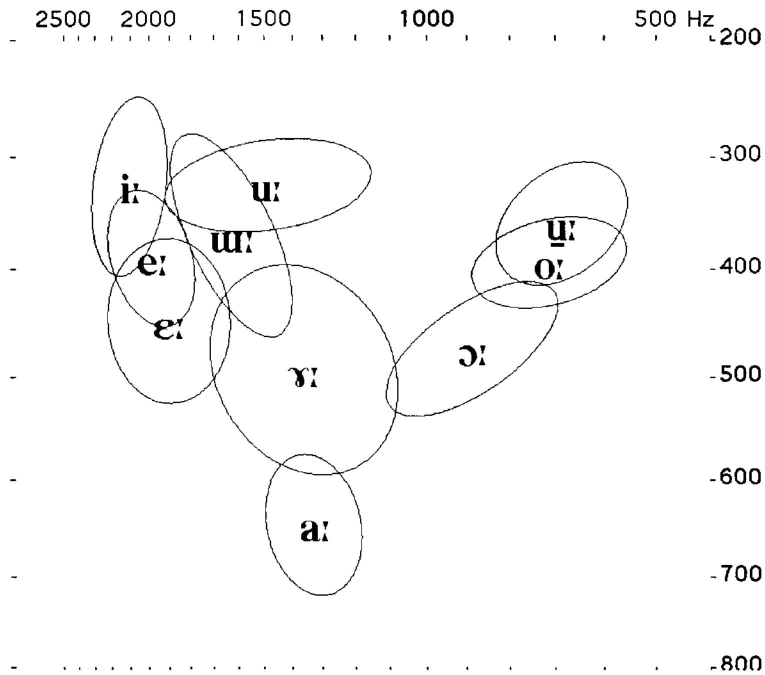


Figure 3.5. The long vowels in the words in Table 3,4 as produced by 7 male speakers of Gaelic.

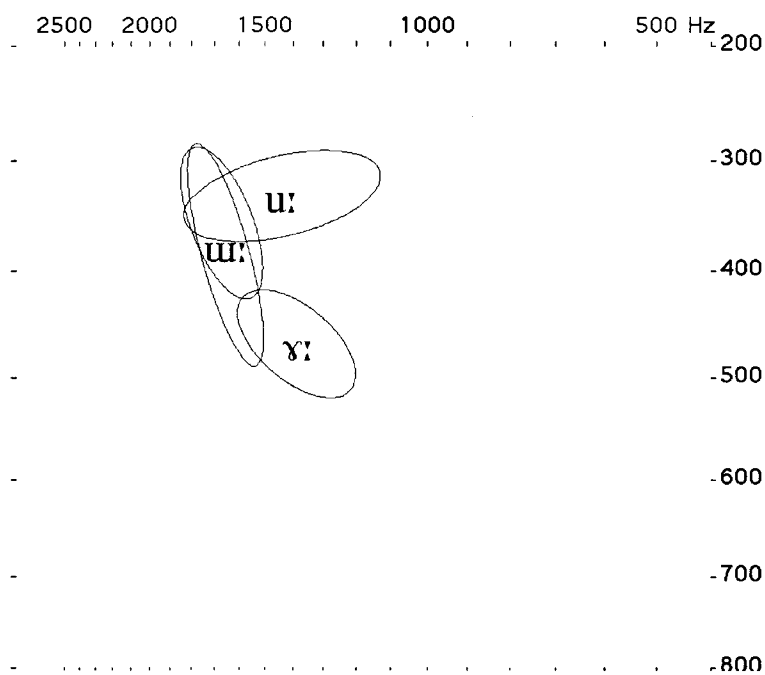


Figure 3.6. The vowels **uː** and **ɯː** in *uradh* **uːrəɣ** and *saoradh* **sɯːrəɣ** as produced by speakers 1-6 (upper pair of ellipses), and **ɯː** and **ɤː** in various words as noted in the text (lower pair of ellipses.

For most of the speakers, the non-low non-front vowels **ɯː** and **ɤː** are clearly distinct from each other as well as from the front and back allophones of **uː**. However, some of the speakers pronounced words with the spelling 'ao' (used to elicit the **ɯː** vowel) with a high F1, making them more similar to **ɤː**. This was true for speaker 2's *faobhar* and speaker 3's *faolainn*. A striking difference between the plots for the short vowels vs. long vowels is that long **ɛː** is much higher than its short counterpart, and overlaps in many cases with **eː** . We also noted that speakers 4, 6, and 7 produced the first vowel in the word *deònach* (classified as **oː**) with a much higher F1 than in their renditions of the **oː** in *bò*, possibly due to the following nasal consonant. This vowel has not been included in the plot in Figure 3.5.

Statistical analyses of subsets of the non-front vowels were conducted to confirm that **ɯː**, **ɤː** and **uː** are indeed separate categories for speakers of the Bernera dialect. First, we compared the near-minimal pair*uradh* **uːrəɣ** and *saoradh* **sɯːrəɣ** for the six speakers who had instances of both words (speakers 1-6). The vowels in these two words are represented by the upper two ellipses in Figure 3,6.

The ellipses for **uː** and **ɯː** show some degree of overlap, and indeed, the formant values for Speaker 1's **ɯː** and **uː** are almost identical. However, even when including this speaker in an analysis across all six speakers, the vowels are distinct ($p < 0.05$). Repeated measures ANOVAs with Vowel (**ɯː** vs. **uː**) as a within-speakers factor are shown in Table 3.5

Table 3.5. Means in Hz and standard deviations (in parentheses) from a repeated measures ANOVA, for **ɯː** vs. **uː** with each of the formants as dependent variable.

	uː MEAN (S.D.)	**ɯː** MEAN (S.D.)	F(1,5)	p
F1	332 (21)	356 (36)	6.61	< .05
F2	1450 (178)	1673 (89)	8.18	< .05
F3	2063 (74)	2183 (94)	22.99	< .05

The **ɯː** vowel thus appears to be slightly lower and less rounded (and/or more front) than **uː**. Examination of the plots for the individual speakers (not included here) indicates that the two vowels are more distinct for some speakers than for others.

Table 3.6. Means in Hz and standard deviations (in parentheses) from a repeated measures ANOVA, for **ɯː** vs. **ɤː** with each of the formants as dependent variable.

	ɯː	**ɤː**		p
	saoradh **sɯːrəɣ**	adhradh **ɤːrəɣ**	F (1,2)	
F1	356 (44)	481 (31)	39.83	< .05
F2	1700 (103)	1433 (43)	17.36	n.s.
F3	2195 (91)	2251 (34)	1.0	n.s.
	faobhar **fɯːvər**	adhbhar **ɤːvər**:	F (1,5)	
F1	386 (53)	468 (27)	12.83	< .05
F2	1657 (84)	1393 (107)	41.88	< .001
F3	2221 (226)	2324 (91)	2.52	n.s.
	faolainn **fɯːlˠinʲ**	foghlam **fɤːlˠəm**	F (1,1)	
F1	393 (79)	517 (47)	12.25	n.s.
F2	1453 (264)	1171 (41)	1.71	n.s.
F3	2162 (115)	2354 (67)	3.76	n.s.

There were three near-minimal pairs illustrating the difference between **ɯː** and **ɤː** : *saoradh* **sɯːrəɣ** and *adhradh* **ɤːrəɣ** (for speakers 3,4, and 6);*faobhar* **fɯːvər** and *adhbhar* **ɤːvər** , for speakers 1-6, and*faolainn* **fɯːlˠinʲ** and*foghlam* **fɤːlˠəm** for speakers 3 and 6. Results from repeated measures

ANOVAs with Vowel (**ɯː** vs. **ɤː**) as a within-speakers factor are shown in Table 3.6. These means and the data in Figure 3.7 show that the **ɯː** vowels are higher (lower F1) and more front (higher F2) than **ɤː**. The statistical tests for *faolainn* vs. *foghlam* did not show a significant difference, probably because of the small number of tokens (4 tokens of *faolainn* vs. 6 tokens of *foghlam*) that went in to the analysis, and also because speaker 3's *faolainn* showed a relatively high F1.

Taken all together, the plots and statistical tests suggest that for some speakers at least, there are indeed 9 distinct long vowels, as noted by Borgstrom (1940). These vowels include two non-front non-low back unrounded vowels, **ɯː** and **ɤː**.

4. Gaelic syllable structure

One of the most interesting aspects of Gaelic phonology is the syllable structure. It is often hard to say how many syllables there are in a word. Epenthesis occurs as a fairly general process, inserting a vowel following the initial stressed syllable whenever certain consonantal sequences follow, and perhaps adding another syllable. For example, the word *carbad* (coach) is pronounced as **karabəd**, with three distinct vowels. But native speakers think of this word as two syllables rather than three. A number of words that were historically one syllable now seem like two, because of the insertion of a vowel between two consonants. Thus the second vowel in *balg* **palˠak** (belly) was inserted to break up the final consonant cluster which is still reflected in the spelling. The question at issue is how these words differ from other words such as **palˠ.ak** (skull) which is spelled 'ballag' and is historically two syllables. In transcribing these pairs of words we have marked a syllable break in words such as *ballag* **palˠ.ak** (skull) by using the IPA symbol for a syllable division, a period. We have not marked a syllable break in the words with epenthetic vowels, such as *balg* **palˠak** (belly), which native speakers of Gaelic regard as monosyllables. In linguistic terms, we might consider the words with an epenthetic vowel to have something akin to a consonantal interlude (Hockett 1955) rather than a syllable break. As has been noted elsewhere (Ladefoged 1971, 1995), the concept of a syllable cannot be easily defined in phonetic terms. Syllabicity is a phonological property, an organizing principle, and may not be a phonetically observable feature.

The distinction between words with epenthetic vowels and similar words with two underlying vowels has a number of implications for phonological theory. Borgstrøm (1940) takes a different view from that suggested above, in that he regards both *balg* **palˠak** (belly) and *ballag* **palˠ.ak** (skull) as having two syllables. He suggests that the contrast between them arises because the consonant is the onset of the second syllable in the words with an epenthetic vowel, whereas in the regular two syllable words there is backwards syllabification, in that the intervocalic consonant forms the coda of the first syllable. Using IPA symbols, he would presumably transcribe the examples we have been discussing as *balg* **pa.lˠak** (belly) and *ballag* **palˠ.ak** (skull).

This is a complex topic which is currently being investigated by Mark Ellison and Kevin Hind. As they point out "Borgstrøm's claim has been accepted widely in the phonological literature dealing with syllable structure (Kenstowicz and Kisseberth 1979, Clements 1986, Bosch 1988, Goldsmith 1990, Bosch 1995). In fact, articles which reference backwards syllabification in Gaelic exclusively cite Borgstrøm and ... the claim of backwards syllabification in Gaelic therefore depends solely on the fieldwork of Borgstrøm." (Ellison and Hind, p.c.; we should note that Bosch has considerable fieldwork data of her own.)

Although our field recordings contain examples of epenthetic vowels corresponding to 8 of the 9 short vowels, we will report here only preliminary observations of the contrast between *balg* **palˠak** (belly) and *ballag* **palˠ.ak** (skull). We must await the results of Ellison and Hind's more extensive work on this topic before we can fully evaluate Borgstrøm's proposal that the consonant in *ballag* **pal$^{\gamma}$.ak** (skull) is part of the first syllable.

The most noticeable difference between these two words is in the pitch pattern. Figure 4,1, shows the pitch of theses two words as produced by speaker 2, the speaker who made the clearest distinction. As may be seen, the pitch in the second part of the word *balg* **pal$^{\gamma}$ak** (belly) is higher than in *ballag* **pal$^{\gamma}$.ak** (skull). In the first word the pitch rises 15 Hz during the last vowel; in the second word it falls 50 Hz during the corresponding period.

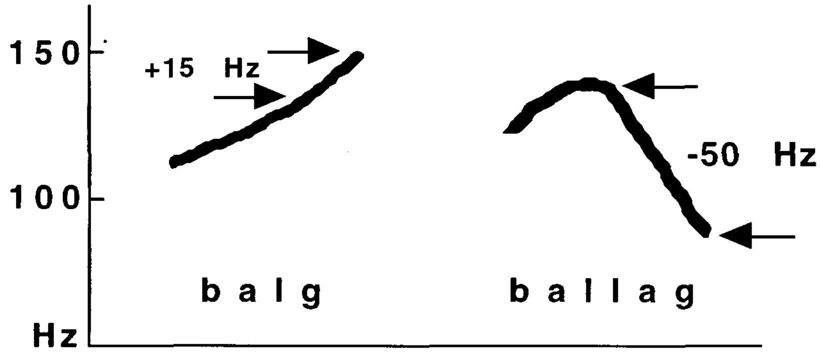


Figure 4.1 Pitch curves illustrating the differences in pitch in the words *balg* **pal$^{\gamma}$ak** (belly) and *ballag* **pal$^{\gamma}$.ak** (skull) as spoken by speaker 2.

The pitch does not always rise for the one word and fall for the other, as it does for the examples shown in Figure 4.1. The words in our data set were always elicited within a word list, and were colored by whatever list intonation pattern the speaker used. All of our speakers had a rising pitch in *balg* **pal$^{\gamma}$ak** (belly) and one of them had a rise, a smaller one, in *ballag* **pal$^{\gamma}$.ak** (skull) . The remainder had a fall in this word. Every speaker used pitch to distinguish between these two words, although sometimes, as in the case of speaker 4, the difference was small. Figure 4.2 compares the pitch changes in the second part of each of these two words as produced by the six male speakers for whom we have valid data. The height of the bar represents the difference in pitch as measured at the onset of the final vowel, immediately after the lateral, and the offset of this vowel, just before the final stop.

An extensive investigation of the phonetic facts underlying contrasts involving epenthetic vowels produced by one speaker has been made by Bosch and de Jong (1997). They analyzed recordings made during the course of fieldwork, and noted not only pitch differences similar to those we have been describing above, but also differences in the relative lengths of epenthetic and other vowels. They suggest that "the epenthetic vowel is the head of an exceptionally stressed syllable in an otherwise fixed-stress quantity-insensitive system." On the basis of their work it certainly appears that epenthetic vowels might add an extra syllable to a word, and this syllable could be said to carry a pitch accent. But it is also possible to interpret their findings within an analysis that regards an epenthetic vowel as an additional vowel within the same syllable. Moreover there is further evidence for thinking that the epenthetic vowels do not constitute separate syllables.

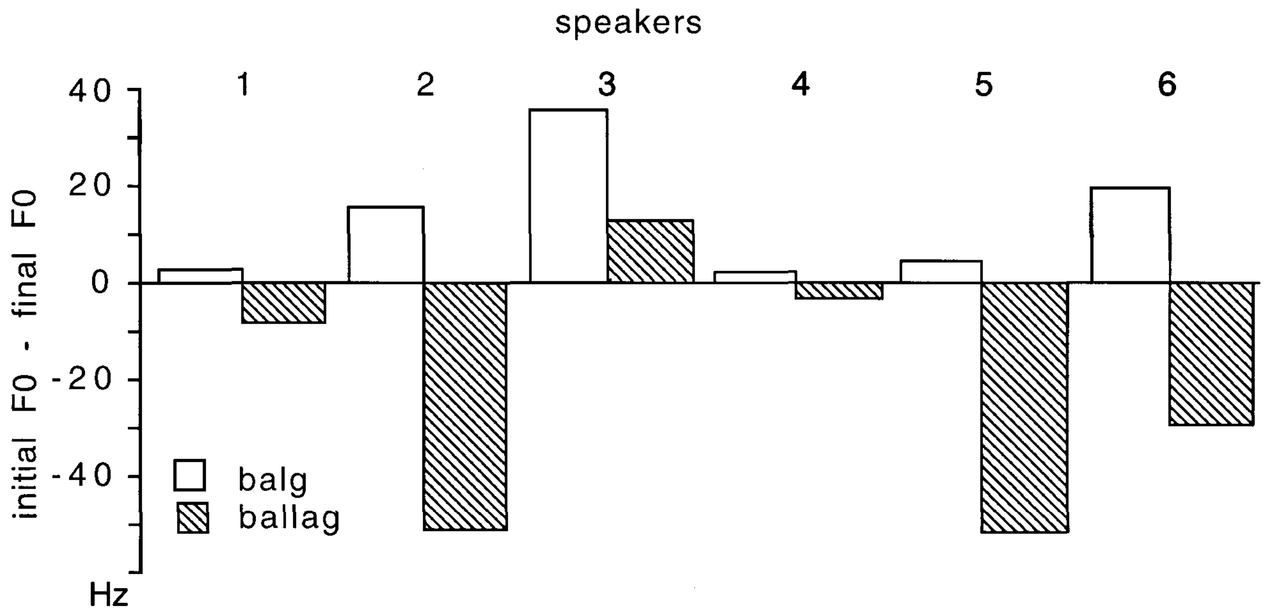


Figure 4.2. The extent of the rise or fall in the second vowels in the word*balg* **pal$^{\gamma}$ak** (belly) and *ballag* **pal$^{\gamma}$.ak** (skull).

A further feature of Gaelic is that there is another set of words in which there are questions concerning syllable structure. These words illustrate what might be considered the reverse of the previously discussed situation, the loss of a consonant between two vowels, as in *dubhan* **tu.an** (hook), which contrasts with *duan* **tuan** (song). We have transcribed the first of these two words as if it were two syllables, and the second as one, which is the intuition that native speakers have about them. Interestingly, these two words have very similar pitch differences to those between *ballag* **pal$^{\gamma}$.ak** (skull) and *balg* **pal$^{\gamma}$ak** (belly). Figure 4.3, shows the pitch curves for speaker 2. It is difficult to say where the second vowel begins in each of these words, but in so far as we can make a judgment it appears that the pitch rises in the second part of the monosyllabic word *duan* **tuan** (song), and falls during the corresponding part of *dubhan* **tu.an** (hook), which speakers regard as disyllabic. Figure 4.4 shows the results for the six speakers saying these words. As in the case of *balg* **pal$^{\gamma}$ak** (belly) and *ballag* **pal$^{\gamma}$.ak** (skull), each of our speakers had a greater fall in pitch on the second part of the word that was clearly two syllables, *dubhan* **tu.an** (hook), than on the monosyllabic form *duan* **tuan** (song).

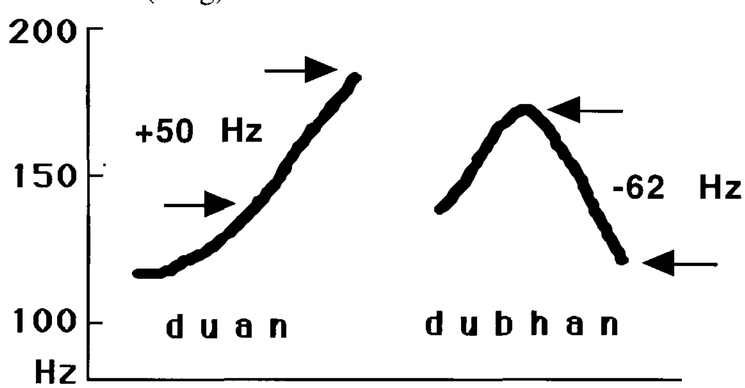


Figure 4.3. Pitch curves illustrating the differences in pitch in the words *duan* **tuan** (song) and *dubhan* **tu.an** (hook) as spoken by speaker 2.

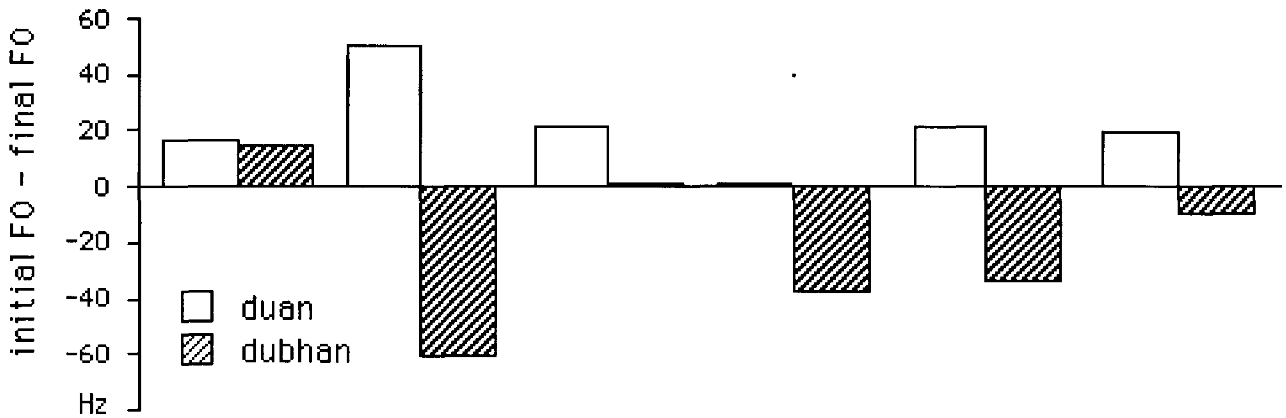


Figure 4.4. The extent of the rise or fall in the second vowels in the words *duan* **tuan** (song) and *dubhan* **tu.an** (hook)

Finally, we will consider briefly a pair of words in which the first is plainly a single syllable with a long vowel, and the second has two adjacent vowels that form two syllables: *bò* **poː** (cow) and *bogha* **po-ə** (underwater rock). Here, again, the one syllable word has less of a fall than the two syllable word, when said in this particular elicitation context. For the speaker we have been illustrating, speaker 2, there is a pronounced rise on the monosyllabic word and a rise fall on the two syllable word, as shown in Figure 4.5.

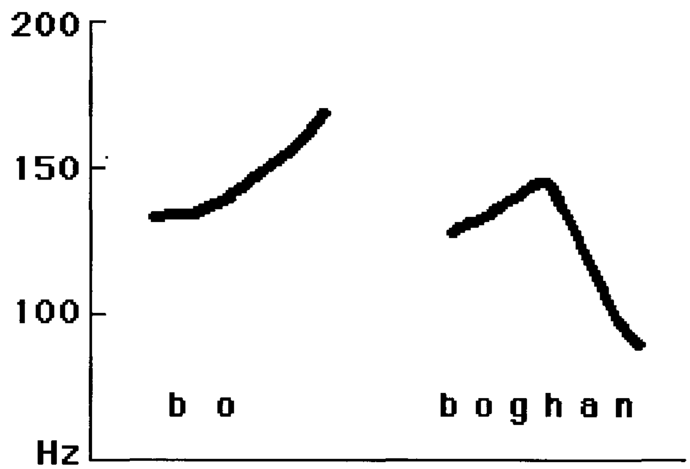


Figure 4.5. Pitch curves illustrating the differences in pitch in the words *bò* **poː** (cow) and *bogha* **po-ə** (underwater rock) as spoken by speaker 2.

The analysis of both epenthetic and elided vowels outlined above is very much in accordance with that suggested by Clements (1986), who regards words such as *ballag* **palˠ.ak** (skull) and *dubhan* **tu.an** (hook) as having underlying syllabification, and words such as *balg* **palˠak** (belly) and *duan* **tuan** (song) as having surface syllabification. We are not sure that there is any need to make the distinction in this way. We would rather just say that *ballag* **palˠ.ak** (skull) and *dubhan* **tu.an** (hook) are two syllables, and *balg* **palˠak** (belly) and *duan* **tuan** (song) are monosyllables. But however this is phrased, it is clear that this is just one of the many aspects of the phonetic structure of Gaelic that makes it a fascinating language, with many properties of interest to both phoneticians and phonologists.

Acknowledgments

We are very grateful for all the cooperation and hospitality we enjoyed during our stay in Bornera. Particular thanks are due to the speakers we recorded: Donald MacAuley, Neil John Macauley, Reverend Donald Macaulay, Callum Murdo McLean, Donald Zachariah Macaulay, Annie Macaulay, Alasdair Maclennan, Floraidh Macdonald, Mary Macdonald, (another) Mary Macdonald and John Grant Maciver. May they and their language continue to flourish.

References

Blankenship, B., Ladefoged, P.,Bhaskararao, P., & Chase, N. (1993). Phonetic structures of Khonoma Angami. *Linguistics of the Tibeto-Burman area*, 16(2), 69 - 88.

Borgstrøm, C. H. (1940). A linguistic survey of the Gaelic dialects of Scotland: Vol 1 - The dialects of the Outer Hebrides. *Norsk Tidsskrift for Sprogvidenskap,* 1 (supp.).

Bosch, A. (1988). VC syllable structure in Scottish Gaelic: some implications for syllable theory. Unpublished manuscript. University of Chicago.

Bosch, A. (1995). A gestural analysis of epenthesis in Scottish Gaelic. Paper presented at Celtic Linguistics Conference, Dublin, June 22-23, 1995.

Bosch, A, & de Jong, K. (1997). The prosody of Barra Gaelic epenthetic vowels. *Linguistic Sciences* (in press).

Clements, G. N. (1986). Syllabification and epenthesis in the Barra dialect of Gaelic. In K. Bogers,H. van der hulst, & M. Mous (Eds.), *The phonological representation of suprasegmentals* (pp. 317-336). Dordrecht, Holland: Foris.

Dorian, N.C. *East Sutherland Gaelic: the dialect of the Brora, Golspie, and Embo fishing communities.* 1978

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology: An introduction.* Oxford: Basil Blackwell.

Hockett, C.F. (1955). *A manual of phonology.* Indiana University Publications in Linguistics and Anthropology, Memoir II. Baltimore,

Kenstowicz, M., & Kisseberth, C. (1979). *Generative phonology: Description and theory.* New York: Academic Press.

Ladefoged, P. (1971) *Preliminaries to linguistic phonetics.* Chicago: University of Chicago Press.

Ladefoged, P.(1995). *A course in phonetics.* Orlando:Harcourt Brace.

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

Ní Chasaide, A. (1979). Laterals in Gaoth-Dobhair Irish and Hiberno English. In D. P. O. Baoill (Ed.), *Papers in Celtic phonology* (pp. 54-78). Coleraine, N. Ireland: Linguistics Department, New University of Ulster.

Ternes, E. (1973). *The phonemic analysis of Scottish Gaelic. Based on the dialect of Applecross, Ross-shire.* Hamburg: Helmut Buske Verlag.

Gaelic word list

This word list, together with some notes made on our fieldwork computer, is provided for the use of other researchers who wish to use our recordings. We attempted to find words illustrating consonants preceding **a** in initial and medial position, and vowels after labials and coronals. Examples of some salient allophones and certain prosodic features were also sought. The sound being exemplified is listed in the first column. Gaps have been left where we were unable to find a suitable word.

Because the list evolved during the course of the fieldwork, the order in which words appear is not entirely logical, in that words were added at the top or bottom of various sections as convenient. This order has been maintained here so as to make it easier to find words on the recordings. Some of the speakers did not know some of the words, and the list should not be regarded as an accurate transcription of any particular recording.

Consonants

m^{j}	meàirleach	**m^{j}a:ləx**	thief
m^{j}	mealladh	**m^{j}al^{ɣ}əɣ**	deceiving
p^{h}	pana	**p^{h}anə**	pan
p^{h}	Pabach	**p^{h}apəx**	person from Pabaigh
p^{h}	apag	**a^{h}pak**	little ape
t^{h}	tana	**t^{h}ana**	thin
t^{h}	tagairt	**t^{h}akəf**	claim
t^{h}	atadh	**a^{h}təɣ**	swelling
k^{h}	cana	**k^{h}anə**	can
k^{h}	cagair	**k^{h}akəð^{j}**	whisper
k^{h}	aca	**a^{h}kə**	at them
p	bad	**pat**	sod
p	baga	**pakə**	bag
p	abaich	**apiç**	ripe
t	dad	**tat**	anything
t	daga	**takə**	pistol
t	adag	**atak**	stook
k	gad	**kat**	with
k	gagach	**kakəx**	stammering
k	agairt	**akəf**	claiming
p^{hj}	peallag	**p^{hj}al^{ɣ}ak**	rag
p^{hj}	peann	**p^{hj}aun^{ɣ}**	pen
p^{hj}			no medial **p^{hj}** (= **^{h}p^{j}**)
t^{hj}	teannadh	**t^{hj}an^{ɣ}əɣ**	tightening
t^{hj}	teallach	**t^{hj}al^{ɣ}əx**	fireplace
t^{hj}	aiteal	**a^{h}t^{hj}əl^{ɣ}**	glimpse
k^{hj}	ceannach	**k^{hj}an^{ɣ}əx**	buying
k^{hj}	cearrag	**k^{hj}ar^{ɣ}ak**	left-hander

kʰʲ	se taice a thuge dha	**... tʰaʰkʲə ...**	it’s support that he gave him
pʲ	beannachd	**pʲanˠəxk**	blessing
pʲ	bealach	**pʲalˠəx**	pass
pʲ			no medial **pʲ**
tʲ	dealan	**tʲalˠan**	electricity
tʲ	deannan	**tʲanˠan**	rush
tʲ	aideachadh	**atʲiç**	admit
kʲ	gealag	**kʲalˠak**	sea trout
kʲ	gearradh	**kʲarˠəɣ**	cutting
kʲ	aigeann	**akʲənˠ**	abyss
f	fanaidh	**fani**	stay (future tense)
f	fada	**fatə**	long
f	tafann	**tafənˠ**	offer
s	sadadh	**satəɣ**	throwing
s	sabaid	**sapɛtʲ**	fight
s	asal	**asəlˠ**	ass
x	chaidh	**xaj**	went
x	chaidil	**xatʲil**	slept
x	achadh	**axəɣ**	field
v	bha	**va:**	was
v	a bhaga	**əvakə**	his bag
v	abhag	**avak**	terrier
ɣ	dh’fhalbh	**ɣ[alˠa]v**	went
ɣ	dhachaigh	**ɣaxi**	homewards
ɣ	langa	**laɣa**	ling (fish)
fʲ	feannadh	**fʲanˠəɣ**	flaying
fʲ	fealladh	**fʲalˠəɣ**	deceit
fʲ			(no medial)
ʃ	seachad	**ʃaxət**	past
ʃ	sealladh	**ʃaləɣ**	sight
ʃ	aiseal	**aʃikʲ**	ferry (noun)
ç	cheannaich	**çanˠiç**	bought
ç	cheartaich	**çaşfiç**	corrected
ç	faiche	**façə**	meadow
ɣʲ	gheàrr	**ɣʲa:rˠ**	cut
ɣʲ	gheall	**ɣʲaulˠ**	promised
ɣʲ	laigh	**lˠaɣʲ**	lie
vʲ	a’ bheannachd	**əvʲanˠəxk**	the blessing
vʲ	a bhealach	**əvʲalˠəx**	his pass
j	e`arlas	**ja:lˠəs**	token
j	earrach	**jarˠəx**	spring
j	laighe	**lˠajə**	lying
m	manach	**manəx**	monk

m	magadh	**makəɣ**	mocking
m	amas	**aməs**	aim
n^{ɣ}	nasgadh	**n^{ɣ}askəɣ**	binding
n^{ɣ}	nathair	**n^{ɣ}ahəð^{j}**	snake
n^{ɣ}	Annag	**an^{ɣ}ak**	Anne
n^{j}	neart	**n^{j}aʂt**	strength
n^{j}	an e`arlas-san	**ən^{j}aːl^{ɣ}əs sən**	their token (emphatic)
n^{j}	ainneamh	**an^{j}əv**	seldom
n	nach	**nax**	not
n	a nathair-sa	**a nahəð^{j}-sə**	his snake
n	fanadh	**fanəɣ**	stay (3rd. sing. imp.)
ŋ	banca	**baŋkə**	bank
ŋ^{j}	taingeil	**taŋ^{j}k^{j}al**	thankful
l^{ɣ}	lagan	**l^{ɣ}akan**	bay
l^{ɣ}	latha	**l^{ɣ}a-ə**	day
l^{ɣ}	balla	**pal^{ɣ}ə**	wall
l^{j}	leannan	**l^{j}an^{ɣ}an**	sweatheart
l^{j}	leantainn	**l^{j}an^{ɣ}tin^{j}**	following
l^{j}	caillidh	**k^{h}al^{j}i**	lose (future tense)
l	mo leannan	**mə lan^{ɣ}an**	my sweetheart
l	leat-sa	**la^{h}tsə**	with you (emphatic)
l	baile	**palə**	town
r^{ɣ}	rabaid	**r^{ɣ}apat^{j}**	rabbit
r^{ɣ}	rathad	**r^{ɣ}a-ət**	road
r^{ɣ}	bearradh	**p^{j}ar^{ɣ}əɣ**	shearing
ð^{j}	riamh	**ð^{j}iəv**	ever
ð^{j}	ris	**ð^{j}iʃ**	to him
ð^{j}	aire	**að^{j}ə**	attention
r	rinn	**rain^{j}**	did
r	rannsaich	**r^{ɣ}ãũsiç**	search
r	aran	**aran**	bread
h	tha	**haː**	is
h	fhathast	**ha-əst**	still
h	athair	**ahəð^{j}**	father
m^{h}	am pana	**əm^{h}anə**	the pan
m^{h}	Am Pabach	**əm^{h}apəx**	the person from Pabaigh
n^{ɣh}	an talla	**ən^{ɣh}al^{ɣ}ə**	the hall
n^{ɣh}	an tagairt	**ən^{ɣh}akəʃ**	the claim
ŋ^{h}	an cana	**əŋ^{h}anə**	the can
ŋ^{h}	an cagair	**əŋ^{h}akəð^{j}**	the whisper
m^{jh}	am peallag-san	**əm^{jh}al^{ɣ}aksən**	their rag
m^{jh}	am peann	**əm^{jh}aun^{ɣ}**	the pen
n^{jh}	an teallach	**ən^{jh}al^{ɣ}əx**	the fireplace

n^{jh}	an teannadh	**ən^{jh}an^{ɣ}əɣ**	the tightening
ŋ^{jh}	an ceannach	**əŋ^{jh}an^{ɣ}əx**	the purchase
ʈ	freagairt	**frɤkəʈ**	answer
ɖ	bàrdachd	**pafɖəxk**	poetry
ʂ	arsa	**aʂə**	said
ɳ	beàrn	**p^{j}aːɳ**	gap
l̥^{ɣ}	altaich	**al̥^{ɣ}tiç**	move joints
l̥^{j}	oillteil	**ɤil̥^{j}t^{j}al**	terrifying
l̥	bailtean	**pal̥t^{j}ən**	towns
r̥^{j}	pàircean	**p^{h}aːr̥^{j}k^{j}ən**	parks
r̥	bàrcadh	**paːr̥kəɣ**	rushing in a torrent

Short oral vowels

ɔ	toddy	**t^{h}ɔdi**	toddy
i	chatig	**t^{j}ik^{j}**	won't come
i	binneas	**piɲəs**	harmony
i	biodag	**pitak**	dagger
e	deireadh	**t^{j}eð^{j}əɣ**	end (noun)
e	beinne	**peɲə**	mountain
e	beithe	**pehə**	birch tree
ɛ	dearcadh	**at^{j}ɛr̥kəɣ**	looking
ɛ	beatha	**pɛhə**	life
ɛ	bean	**pɛn**	woman
ɛ	beadaidh	**pɛti**	ill-mannered
a	tagsaidh	**t^{h}aksi**	taxi
a	bannag	**pan^{ɣ}ak**	new year's cake
a	badan	**patan**	little sod
ɔ	togsaid	**t^{h}ɔksat^{j}**	barrel
ɔ	bonnach	**pɔn^{ɣ}əx**	scone
ɔ	bodach	**pɔtəx**	old man
o	togail	**t^{h}okal**	lifting
o	bogadh	**pokəɣ**	soaking
ɤ	tagh	**t^{h}ɤɣ**	choose
ɤ	beagan	**pɤkan**	a little
u	tugainn	**t^{h}ukiɲ**	com on!
u	bun	**pun**	foot (of hill)
u	buideal	**put^{j}əl^{ɣ}**	bottle
ɯ	turadh	**t^{h}ɯrəɣ**	dry weather after wet
ɯ	uinneag	**ɯɲak**	window

Long oral vowels

iː	diòbhairt	**t^{j}iːf**	vomit
eː	eud	**eːt**	zeal

iː	dìnnear	**tʲiːɲər**	looking
iː	biòdach	**piːtəx**	tiny
eː	dèineas	**tʲeːɲəs**	vehemence
eː	beudach	**peːtəx**	harmful
ɛː	nèamh	**ɲɛːv**	heaven
ɛː	cha b'e	**xa bɛː**	it wasn't
ɛː	Gaidheal	**kɛː-əlˠ**	Gael
aː	deàrrsadh	**tʲaːrsəɣ**	shining
aː	bàidse	**paːtʲə**	musician's fee
ɔː	Deòrsa	**tʲɔːrsə**	George
ɔː	bòid	**pɔːtʲ**	oath
oː	deònach	**tʲoːnəx**	willing
oː	bò	**poː**	cow
ɤː	foghlam	**fɤːləm**	education
ɤː	adhbhar	**ɤːvər**	reason
uː	diùrrais	**tʲuːriʃ**	secret
uː	bùth	**puː**	shop
Uː	dùirn	**tUːɳ**	fists
Uː	bùird	**pUːɖ**	tables
ɯː	daoine	**tɯːɲə**	people
ɯː	baothair	**pɯːhəðʲ**	idiot

Short nasal vowels

ĩ	nimheil	**nʲĩval**	poisonous
i	mhilis	**vili∫**	sweet
ɛ̃	teanga	**tʲʰɛ̃ɣə**	tongue
ɛ̃	faicinn	**fɛ̃ʲʰkiɲ**	seeing
ã	damh	**tãv**	ox
ã	famh (mhallachd)	**fãv (vãləxk)**	mole
ɔ̃	tomhais	**tõ-iʃ**	measure
ɔ̃	mholach	**vɔ̃ləx**	hairy
ũ	mhullach	**vũləx**	curse (lenited)
ũ	coimhead	**kʰũ-at**	look

Long nasal vowels

ĩː	cha mhì	**xavĩː**	it's not me
ĩː	innse	**ĩːʃə**	telling
ɛ̃ː	ràimh	**rˠɛ̃ːv**	oars
ɛ̃ː	gnè	**kðʲɛ̃ː**	nature (of something)
ãː	tàmh	**tãːv**	quiet
ãː	a Mhàiri	**əvãːðʲi**	Mary (vocative)
ãː	àmhghar	**ãːɣər**	distress
ɔ̃ː	tòimhseachan	**tɔ̃ːʃəxan**	puzzle

ɔ̃:	a Mhòd	**əvɔ̃:d**	his mod
ũ:	dùmhail	**tũ:-al**	dense
ũ:	unnsa	**ũ:sə**	ounce
ɯ̃:	caomh	**kʰɯ̃:v**	kind

Epenthetic Vowels

i	tilg	**tʰʲ[ili]kʲ**	throw
e	deirge	**dʲ[eðʲe]kʲ**	redness
e	meirg	**m[eðʲe]kʲ**	rist
ɛ	tairbh	**tʰ[ɛðʲɛ]v**	bulls
ɛ	meirbh	**m[ɛðʲɛ]v**	digest
a	Tarmod	**tʰ[ara]mɔt**	Norman
a	balg	**p[alˠa]k**	belly
a	ballag	**palˠak**	skull
ɔ	dorcha	**t[ɔrɔ]x**	dark (adj)
ɔ	borb	**p[ɔrɔ]p**	savage
ɤ	mairbh	**m[ɤðʲɤ]v**	the dead
ɤ	failm	**f[ɤlɤ]m**	helm
u	duirche	**t[uðʲu]çə**	dark (gen)
u	builg	**p[ulu]kʲ**	bellies
ɯ	aonghas	**[ɯnɯ]-əs**	Angus

Oral dipthongs

ai	daingeann	**tai-ən**	firm
ai	àillse	**ailʲʃə**	cancer
ɛi	teinntean	**tʲʰɛiɲʲtʲʰan**	hearth
ei	teinn	**tʲʰeinʲ**	distress
ei	beinn	**peinʲ**	mountain
ui	laoigh	**lˠui**	calves
ui	muillear	**muilʲɛðʲ**	miller
ɤi	doill	**tɤilʲ**	blind
ɤi	boill	**pɤilʲ**	members
au	dall	**taulˠ**	blind
au	ball	**paulˠ**	member
ɔu	toll	**tʰɔulˠ**	hole
ɔu	bonn	**pɔunˠ**	sole of foot
iə	dia	**tʲiə**	god
iə	biadh	**piəɣ**	food
ia	deur	**tʲiar**	tear
ia	beul	**pialˠ**	mouth
iɛ	Iain	**iɛnʲ**	Ian
uə	duais	**tuəʃ**	prize
uə	buaidh	**puəʲ**	victory
ua	duan	**tuan**	song

ua	uan	**uan**	lamb
uɛ	uain	**uɛnʲ**	lambs
ii	dinn	**tiinʲ**	squeeze
ii	binn	**piinʲ**	judgement
Uu	diùlt	**tʲUul̥ˠt**	refuse
Uu	punnd	**pʰUunˠt**	pound

Nasal diphthongs

ɛ̃ũ	geamhradh	**kʲɛ̃ũrəɣ**	winter
ɛ̃ĩ	sreang	**strɛ̃ĩɣ**	string
ãũ	samhradh	**sãũrəɣ**	summer
ãũ	ban-rìgh	**pãũri**	queen
ãũ	Amhlaigh	**ãũlˠaj**	man's name
ũã	uamh	**ũãv**	cave
ĩə̃	dìomhain	**tʲĩə̃vinʲ**	idle
ĩə̃	a' mhìos	**əvĩə̃s**	the month

Hiatus

uː	dùn	**tuːn**	fort
ua	duan	**tuan**	song
u-a	dubhan	**tu-an**	hook
iə	fiach	**fiəx**	worth
i-ə	fitheach	**fi-əx**	raven
oː	bò	**poː**	cow
o-ə	bogha	**po-ə**	underwater rock
aː	ràdh	**raː**	saying
a-ə	rathad	**ra-ət**	road
ai	brainn	**brainʲ**	womb (dat. sing.)
a-i	brathainn	**bra-inʲ**	quern (gen. sing.)
iː	dìs	**tʲiːʃ**	dice
i-i	dithis	**tʲi-iʃ**	two (people)
eː	dèineas	**tʲeːnʲəs**	vehemence
ei	teinn	**tʲʰeinʲ**	distress
e-i	as a dèidh	**as ə tʲe-i**	after her