Journal of Phonetics (2001) **29,** 231–271
doi:10.006/jpho.2001.0142
Available online at http://www.idealibrary.com on IDEAL®

Phonetic structures of Aleut

Alice Taff and Lorna Rozelle

Department of Linguistics, Box 354340, University of Washington, Seattle, WA 98195, U.S.A.

Taehong Cho and Peter Ladefoged*

Department of Linguistics, UCLA, Los Angeles, CA 90095 1543, U.S.A.

Moses Dirks

Unalaska City Schools, Pouch 260, Unalaska, AK 99685, U.S.A.

Jacob Wegelin

Department of Statistics, 354322, University of Washington, Seattle, WA 98195, U.S.A.

Received 26th January 1999, and accepted 12th April 2001

A detailed analysis of the phonetic structures of Aleut, a moribund language spoken in Alaska, shows how much general phonetic information can be gathered from the investigation of an endangered language. Aleut has an unusual distribution of consonants, with varying functional loads. There are no bilabial stops. Among alveolar, velar and uvular stops, VOT is shorter for alveolar than for velar or uvular stops, but, despite current general phonetic theories, there is no difference in VOT between velar and uvular stops. VOT is also longer before long vowels than before short vowels. Uvular consonants have a significant effect on the formants of following high vowels. There are three vowels and a vowel length contrast. Again, despite previous general phonetic predictions concerning languages with vowel-length contrasts, length plays a role in characterizing stress. Contrastive ratios are maintained between short and long vowels and stressed syllables are longer. Analyses of intonation show that each content word has a peak at its beginning and a trough at its end. Word contours combine with sentence downtrends to form sentence contours of cascading $F0$, each word a step in the cascade. But, unlike the situation in other languages, yes/no questions have the same intonation contours as declarative sentences. Phonetic descriptions and measurements of general theoretical interest include C and V inventories, VOT measurements for coronal, velar and uvular stops, vowel-quality plots, vowel-duration measurements, question and declarative pitch tracks, pitch-track smoothing, and question *vs.* declarative sentence pitch medians.

*E-mail: oldfogey@ucla.edu

0095-4470/01/030231 + 41 $35.00/0

1. Introduction

We know a great deal about the phonetic structures of many of the 10 most widely spoken languages in the world: Mandarin Chinese, Spanish, English, Bengali, Hindi, Portuguese, Russian, Japanese, German and Wu Chinese (Grimes, 1999). Many other languages that have political, commercial or academic importance, such as French, Dutch, Danish and Estonian have also been well described. But by far the majority of the 95 language families listed by Grimes (1999) in the *Ethnologue* have not even a single language that has been well described phonetically. The languages for which we have quantified phonetic observations are a very small sample of the 6809 languages in the *Ethnologue*. If we want adequate phonetic theories we must increase our database.

Several problems arise when trying to extend our knowledge of the sounds of the world's languages. The first is that over 25% of all languages are spoken by less than 1000 people who are comparatively hard to contact. They cannot be paid $10 and asked to come into a phonetics lab to be recorded in various ways. Another problem is that many small languages are becoming endangered in a world in which governments and industries are attracted by the financial benefits of using mass media, forcing people to learn more widely spoken languages, as documented elsewhere, e.g. Fishman (1991), Grenoble & Whaley (1998). Often, endangered languages are spoken only by aging populations, many of whom, being deaf and without their teeth, are not the best subjects for phonetic experiments. But if we are to know the range of possible speech sounds, it is important that we do our best to ensure that our database includes information on as many languages as possible.

The starting point for this paper is the consideration of some theoretically suggestive data from Aleut, an endangered language that has received relatively little attention from phoneticians to date. The fieldwork reported here may be the last to record systematic data from a reasonable sample of Aleut speakers. Consideration of these data is important from the point of view of assessing the overall nature of the sounds of the world's languages. These data have limitations. The rigorous experimentation used in laboratory experiments with speakers of more well-known languages was not possible in the fieldwork conditions of the present investigation, which involved mainly elderly speakers, limited time, and portable equipment that was subject to the vagaries of transportation in a near-arctic climate.

Obtaining data on the sounds of an endangered language is not a process of hypothesis testing as occurs in most papers on phonetic topics. It is a matter of making systematic investigations that can produce limited findings some of which may not be generalizable. But it must also be remembered that data from well-controlled experiments on well-known languages may not be generalizable to a wide range of languages. We need as much information as we can get about all the different ways that humans can use spoken language.

Aleut, a moribund member of the Eskimo–Aleut language family, is an interesting language for the theoretician to study because it provides evidence for the following issues: the differences between velar and uvular consonants, the dispersion of vowels in a three-vowel language, the interaction between vowel length, pitch and stress, and the pitch patterns in sentences with different syntactic structures. Some of the findings reported here for Aleut are commonly found phonetic phenomena, but others are harder to explain. Thus, VOT for Aleut laminals is shorter than for velars and uvulars but VOT for velars and uvulars is the same. VOT is longer before long vowels. The three-vowel

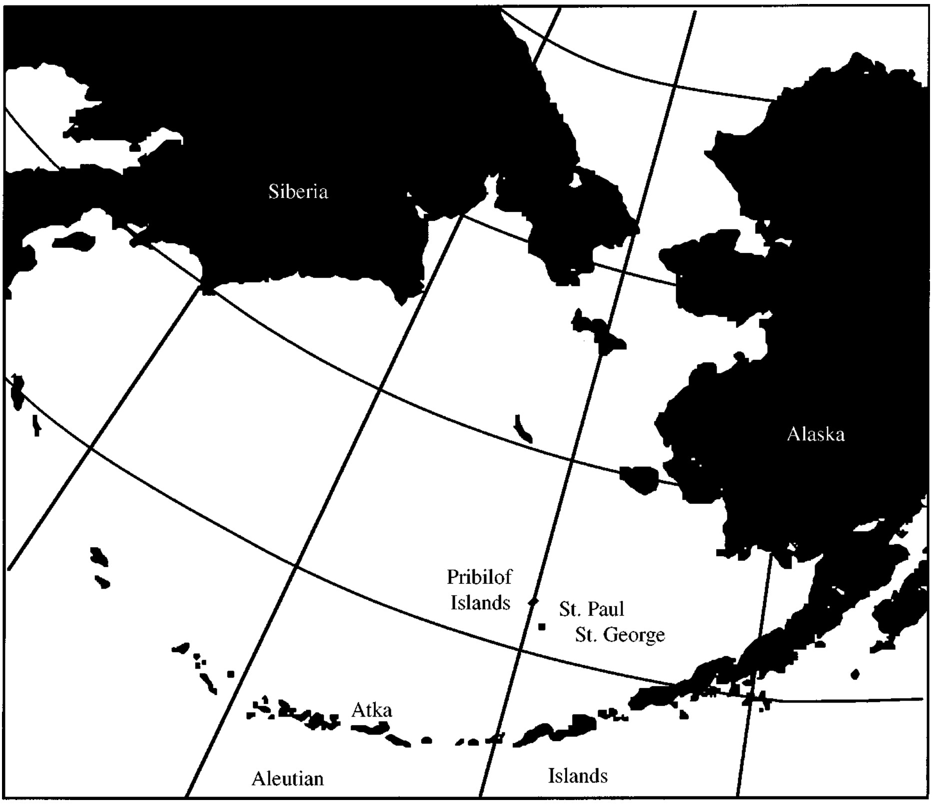

Figure 1. Eastern Aleut, spoken in St. Paul and St. George; Western Aleut, spoken in Atka.

system is dispersed to the corners of the vowel space but **u** is somewhat lower than **i**. (In this paper, phonetic symbols are in bold type). Unlike in many languages, the quality of the longer vowel **ii** is not nearer the outside of the vowel space than **i**. The quality of **u** and **uu** are also alike; but **a** differs from **aa**. Vowel quality is considerably affected by uvulars, but some cases of uvularization suggest lexicalization of these effects. Stress adds length to vowels even though vowel length is contrastive. Word intonation shape is peak–trough. Sentence intonation shape is a cascade of peak–trough words. Y/N questions have the same shape as declaratives. This paper reports a survey of the phonetic structures of Aleut, including an account of all these phenomena.

Aleut is the smallest branch of the Eskimo–Aleut language family. The ancestral territory of the Aleut language community is the Aleutian and Pribilof Islands in the Bering Sea (see Fig. 1). There are also a few speakers in Russia on the Commander Islands and scattered elsewhere in the U.S. The two main dialects are Eastern Aleut (own name Unangan) and Western Aleut (own name Unangas). Eastern Aleut is the indigenous language of the Eastern Aleutian Islands and the Pribilof Islands. There are approximately 200 fluent native speakers, all over the age of 50, and all fluent speakers of English. Our speakers were recorded in St. Paul and St. George, in the Pribilof Islands. Western Aleut is spoken in the western isles of the Aleutian chain. According to Bergsland (1994, p. xvi) there are 60–80 speakers, including some children. Our speakers of this dialect were recorded in Atka. We did not find any children who were fluent speakers in Atka.

1.1. *Phonological inventory*

The first point of general phonetic interest in Aleut is the phonological inventory and the functional load of its various components. Both dialects of Aleut have only 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 Table I. Eastern Aleut has a simpler system in that the contrasts between voiced and voiceless nasals, sibilants and approximants have been largely lost. Table II shows the Aleut orthography for the consonants. This is the practical school orthography developed in 1972 (Bergsland, 1994), in which $\hat{g}$ = a voiced uvular fricative and $\hat{x}$ = a voiceless uvular fricative. In the intonation section of this paper, where the segmental phonology is less important, we will use this orthography.

As may be seen from Table I, Aleut is in the 1% of the world's languages that do not have **p** and **b** (Maddieson, 1984). There are well-established cognate sound correspondences between Aleut and other members of the Eskimo–Aleut family. The lack in Aleut of all labial consonants but **m** is described by Bergsland (1986). Eskimo **p-** and **m-** correspond to Aleut **h-**, **h** being lately lost in Eastern Aleut. Eskimo **-p-**, **-v-**, **-m-** correspond to Aleut and **-m̥-** and **-m-**.

Aleut contrasts velar and uvular stops, initially, medially and finally. There are also voiced and voiceless velar and uvular fricatives but they 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.

As an aid in developing of the word list discussed in the next section, the main entries in a small dictionary of Eastern Aleut designed for use in schools (Bergsland & Dirks,

TABLE I. The consonants of Western Aleut

	Bilabial	Dental	Alveolar	Palato-alveolar	Velar	Uvular	Labial velar
Stop		**t**		**tʃ**	**k**	**q**	
Nasal	**m̥ m**	**n̥ n**			**ŋ̊ ŋ**		
Fricative		**ð̥**	**s z**		**x ɣ**	**χ ʁ**	
Lateral		**l̥ l**					
Approximant				**ç j**			**ʍ w**

TABLE II. The Aleut orthography for consonants

	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

TABLE III. 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

TABLE IV. The relative frequencies of the consonantal segments of Aleut

Segment		Initial	Final	Total	% Initial	% Total
n	n	16		393	2.0	11.3
s	s	112		389	16.0	11.3
ɣ	g	3		346	0.5	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	0.1	4.3
ʁ	ĝ	0		142	0.0	4.1
χ	x̂	5	(488)	615	0.5	3.7
j	y	11		116	1.0	3.3
ŋ	ng	2	(9)	115	0.3	3.1
n̥	hn	0		25	0.0	0.7
m̥	hm	0		18	0.0	0.5
l̥	hl	0		15	0.0	0.4
w	w	7		13	1.0	0.4
Total		703	(1436)	4900	99.4	100.1

1978) were entered into a computer. Only words illustrating segments in the original Aleut phonological inventory (Bergsland, 1994, p. xvii) were included. In all, there were 8744 segments in 1194 words. In addition to providing a 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 are shown in Table III. 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. The relative frequencies of the consonants are shown in Table IV. The total numbers for **x**, **χ** and (to a lesser extent) **ŋ** are inflated by the fact that these are 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 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 given in parentheses in Table IV, making the total number of consonants considered for this purpose 3464 instead of 4900.

It is apparent from the 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" was 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, and **j**, although slightly more common, occurs almost only before **a**. As we have already noted, bilabial stops do not appear in native Aleut words.

Some of these findings can be explained in terms of general phonetic principles. We suspect that there is a universal phonetic tendency among languages with a small number of vowels for **a** and **i** to be much more common than **u**, although we do not have enough quantified data to be able to prove this statistically. The high flow rate in the voiceless nasals and lateral absorbs more respiratory power (the ultimate source of energy for nearly all speech sounds) than their voiced counterparts, but these sounds have a comparatively low intensity and are not auditorily very distinct from one another, both factors making them less preferred. The rarity of **w** may be associated with the lack of labials in general, although why Aleut should be in the 1% of the world's languages that do not have a **p** and **b** (Maddieson, 1984) is not clear. The greater frequency of **j** before **a** may be because it has greater auditory prominence in this sequence.

1.2. *Data collection*

A list of words was developed to illustrate 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** (for which no **u/uu** contrast could be found) 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 nonsibilant fricatives do not contrast in initial position, and accordingly the list includes 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 **u**. These sounds were included in medial position. For the vowels, pairs of words were chosen that showed differences in vowel length, with position in word, stress, and consonantal context being kept constant.

Only the original sounds of Aleut were considered. In the second half of the 18th 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 **tuːtkaχ** *tuutkax̂* "aunt" from Russian *tyótka*, but not **suːpaχ**, *suppax̂* "soup" from English *soup*, because there are no bilabial plosives in the original Aleut inventory.

In addition to the word list, 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. Although the responses were not uniform across speakers, no amendments were suggested so that the responses remained as natural as possible. There were also sentences differing mainly in the location of the clause breaks, and in the degree of emphasis expected on each word.

At the end of each recording, the speaker was asked to talk for a minute 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, six men and six women being speakers of Western Aleut, and four men and eight 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, but did not speak it often. Each speaker was recorded individually, usually using a DAT recorder and a noise-canceling close-talking microphone, giving a signal/noise ratio better than 48 dB. Due to the accidents of fieldwork, five of the speakers of Western Aleut had to be recorded with a backup tape recorder. The frequency response was flat (± 2 dB) 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 immediately know the word required, they were prompted in Aleut. If they were completely unfamilar 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.).

2. Consonants

In this section, our principal phonetic investigations will be limited to voice onset time (VOT) and characteristics of burst spectra in both Eastern and Western dialects of Aleut.

2.1. *Voice onset time*

The VOTs of stop consonants in Eastern and Western Aleut were investigated by reference to the words in Table V. The VOT of stops in word initial position was examined in an effort to investigate how VOT varies according to (1) different dialect (Eastern *vs.* Western), (2) gender differences, (3) context of following vowel (short *vs.* long), and (4) place of articulation, to which special attention was paid. It is well established 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 & Lehiste, 1960; Cho & Ladefoged, 1999). If the variability of VOT is due simply to the distance between the open end of the vocal tract and the source of compression, then, as a working hypothesis, the VOT for a uvular stop will tend to be longer than that for a velar stop.

The data recorded on DAT tapes were down-sampled to 20 000 Hz on the Kay Elemetrics Computerized Speech Lab (CSL). A total of 816 tokens (24 speakers $\times$ 17 words $\times$ 2 repetitions) were digitized, 755 of which were measured. Data from one female Eastern Aleut speaker were excluded for segmental analysis due to sudden background

TABLE V. Words for comparing properties of stop consonants in Eastern and Western Aleut

	Eastern Aleut		Western Aleut	
ti	**tiχlaχ**	bald eagle	**tiʁlaχ**	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

noise and speech errors. Some additional tokens could not be analyzed for similar reasons. 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- or two-factor ANOVAs.

There were differences between the two dialects and between the male and female speakers. The result of a one-way ANOVA reveals that the difference in VOT between these two dialects is highly significant ($F\,[1,\ 753] = 96.934$, $p < 0.0001$), the VOTs in Western Aleut being about 24% longer than those in Eastern Aleut. In the following discussion, the results for the two dialects are presented separately. But, as will be seen, the same significant differences occur in both dialects.

The VOTs for alveolar stops tend to be shorter than the other two stops in both dialects, as shown in Fig. 2. One-factor analyses of variance of each dialect reveal that the effect of place was significant ($F\ [2,\ 359] = 34.358$, $p < 0.0001$ for Eastern Aleut and $F\ [2, 390] = 23.871$, $p < 0.0001$ for Western Aleut). In *post hoc* analyses, the alveolar plosives were distinct from the other two at $p < 0.0001$ in both dialects, but there was no significant VOT difference 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 plosives have a longer mean VOT. Cho & Ladefoged (1999) point out that there is a great deal of variation among VOTs in different languages, and that, although some variations in VOT can be attributed to purely mechanical aerodynamic effects associated with different places of articulation, others cannot be explained in this way and must be ascribed to the particular choices made by individual languages.

One of the interesting acoustic events that 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. Several triple bursts as well as double bursts were found in the spectrograms of both the velar and uvular stops. In

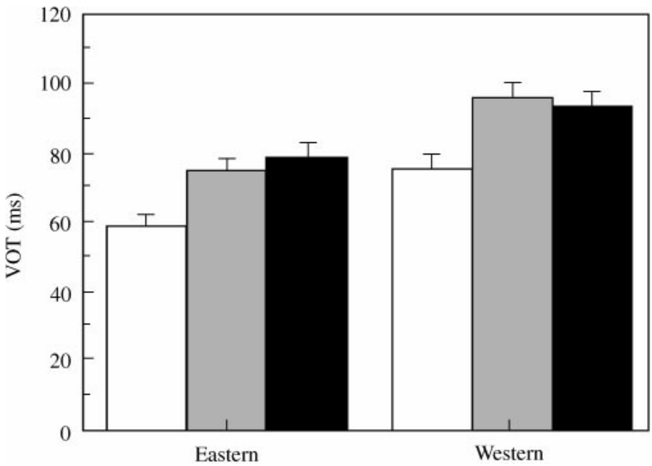

Figure 2. Mean voice onset time (ms) by categories of stop and dialect. (Data averaged across speakers of each dialect of Aleut. Error bars refer to standard errors.): □ t; ▩ k; ■ q.

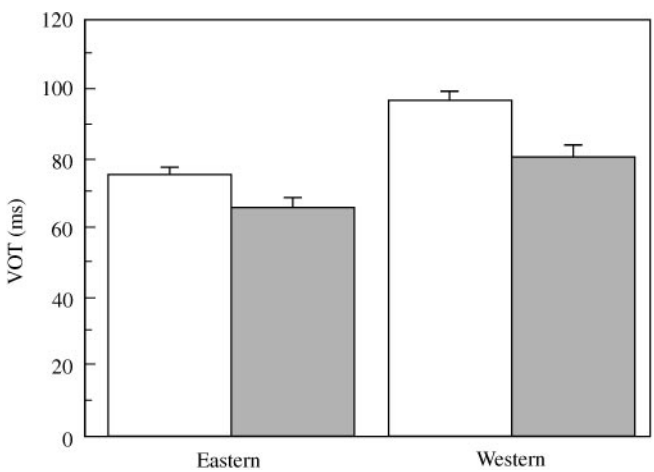

Figure 3. Effect of the following phonemic vowel length on VOT. (Data averaged across speakers of each dialect of Aleut. Error bars refer to standard errors.): □ before long V; ▩ before short V.

addition, the mean VOTs for all three stops are quite long, falling into the category of voiceless aspirated stops (Cho & Ladefoged, 1999). The VOTs for alveolar and velar stops are close to the VOTs reported for aspirated stops in English and Cantonese by Lisker & Abramson (1964).

The mean VOT of the plosives before a long vowel is about 20% longer than that before a short vowel both in Eastern and Western Aleut. A summary is given in Fig. 3. 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 ($p < 0.001$ for Eastern Aleut, $p < 0.0001$ 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 & Rotunno, 1979).

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 ($F\ [2, 356] = 1.131$, $p = 0.3239$ for Eastern Aleut and $F\ [2, 387] = 0.282$, $p = 0.754$ 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.)

In summary, the VOT analyses show that VOT varies between the two dialects, largely due to three less-fluent speakers of Western Aleut. More importantly, it varies according to (1) place of articulation, and (2) the following vowel duration. VOT is shorter for alveolar than for velar or uvular stops, but the difference between velar and uvular was not statistically significant. This is in accord with the view that although some differences in VOT may be determined by aerodynamic factors, others simply reflect the behavior associated with a particular language (Cho & Ladefoged, 1999). The effect of the following vowel duration on VOT is also significant: VOT is longer before long vowels than before short vowels, suggesting that the VOT may be considered as part of the vowel.

2.2. *Burst spectra*

In examining the burst spectra for consonants, there are several suggestions in the literature. Jakobson, Fant & 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, p. 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 (1998), the spectrum of velar stops is less flattening (more compact) compared with that of bilabial or alveolar stops. 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 more slowly, 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 & 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 spectral 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 & 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.

Measurements were taken from the tokens that were used for the VOT measurements for four male and seven female speakers in Eastern Aleut and two male and four female speakers in Western Aleut. The Western Aleut tokens recorded on analog backup tapes were not used, as the high-frequency components were not well captured. Using the Kay CSL 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.

The data were analyzed in terms of the relative frequencies of the primary spectral peak and the compactness of the burst. Figs 4 and 5 show the 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 midfrequencies (somewhere between 500 and 3000 Hz) depending on the following vowel context (cf. Jakobson *et al.*, 1963). In addition, it appears that the spectra for female speakers in both dialects 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.

As discussed earlier, the frequency location of a prominent spectral peak largely depends on the following vowel. Fig. 6 presents data showing that Aleut is in accord with the earlier studies. The frequency location of the spectral peak is higher before **i** than before **a** and **u**. This effect is mainly accounted for by the assumption that the consonant peak is determined by the front-cavity resonance (Keating & 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 the 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 Fig. 6. 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**.

Fig. 6 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 notion that the ratio of the volume of the front cavity to that of the back cavity is higher for the compact

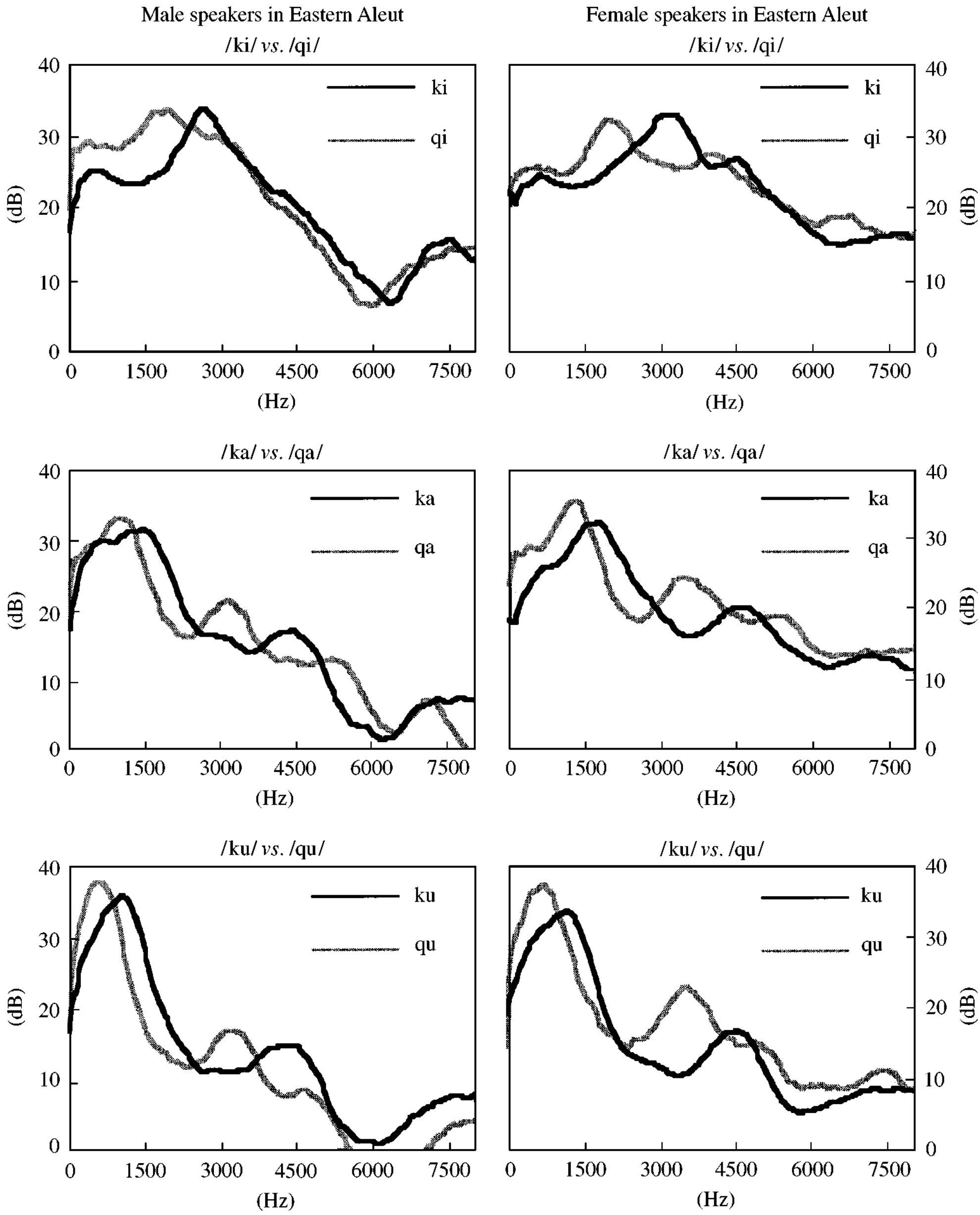

Figure 4. Mean burst spectra of word initial **k** and **q** before **i**, **a** and **u** in Eastern Aleut. Data from four male speakers and seven female speakers.

spectrum as suggested by Jakobson *et al.* (1963). 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.

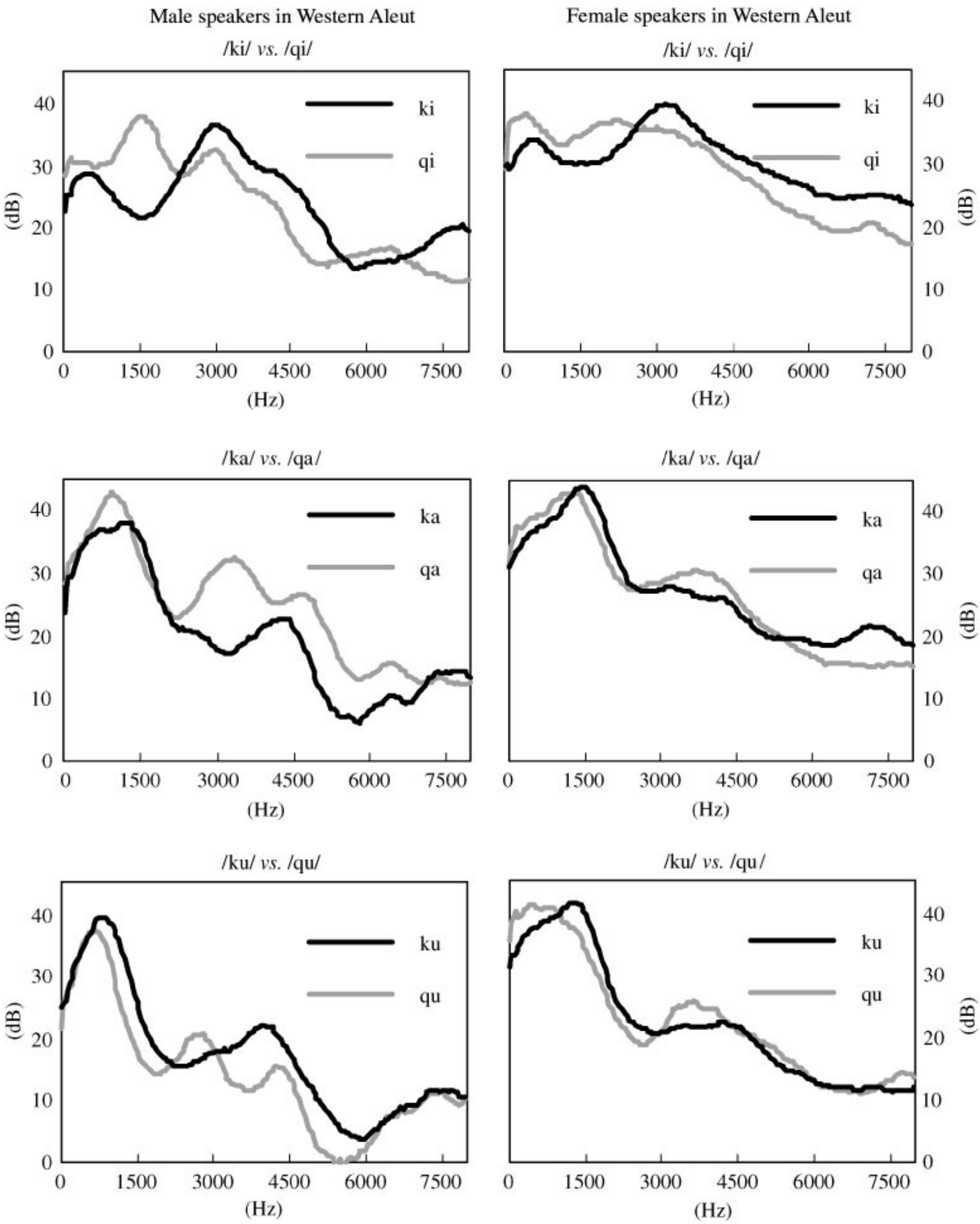

Figure 5. Mean burst spectra of word initial **k** and **q** before, **a**, and **u** in Western Aleut. Data from two male and four female speakers.

Qualitative comparison of the burst and frication noise between the velar and the uvular indicate the possibility that greater overall noise energy of frication and also greater energy in the low-frequency range may serve as significant cues for native speakers of Aleut to distinguish uvular from velar plosives. However, the perceptual status of this acoustic property is clearly a matter for further study. A greater low-range burst intensity for **q** relative to **k** is also reported for Tompson Salish (Mayes, 1979).

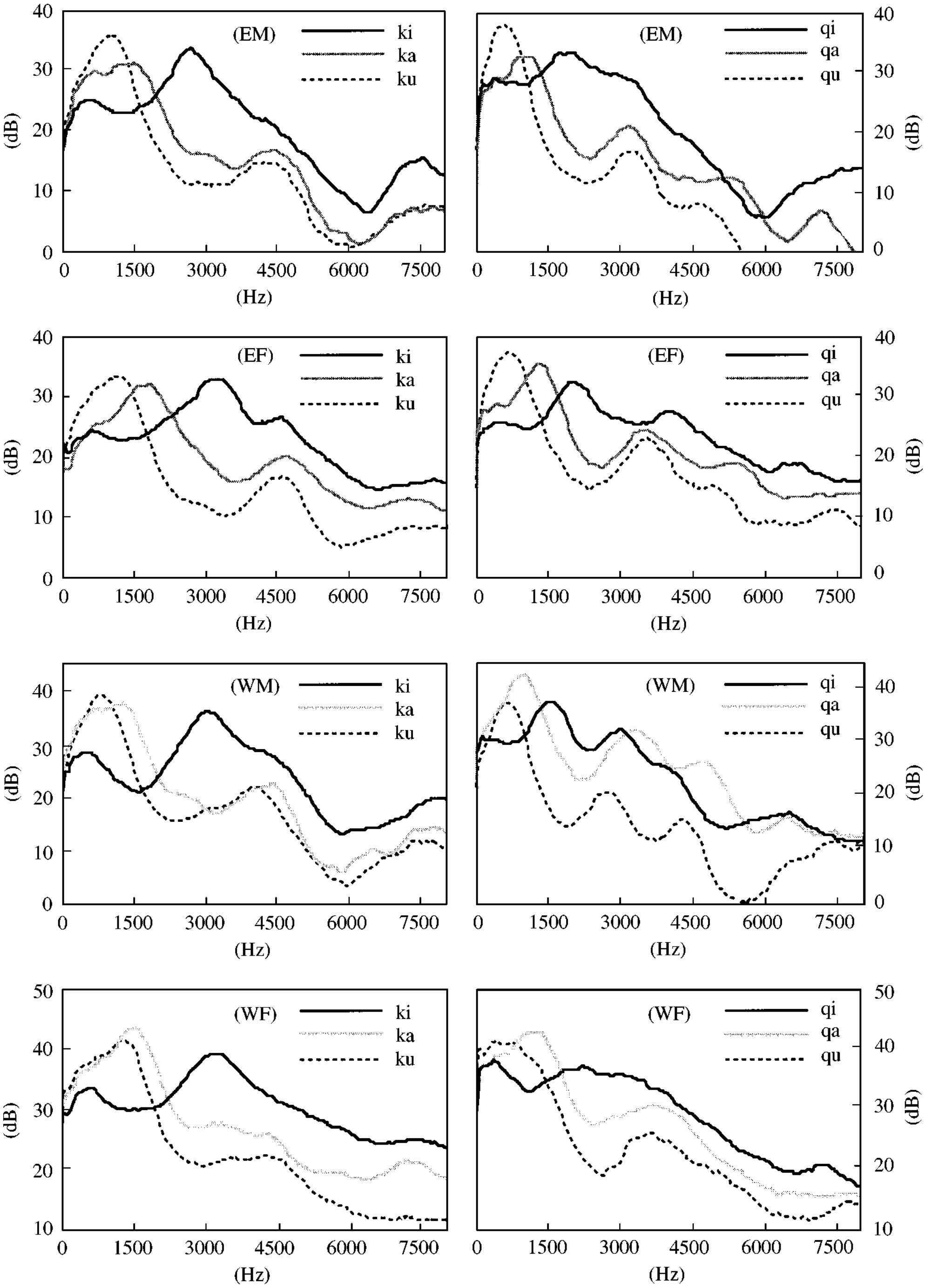

Figure 6. Effect of the following vowel on burst spectra.

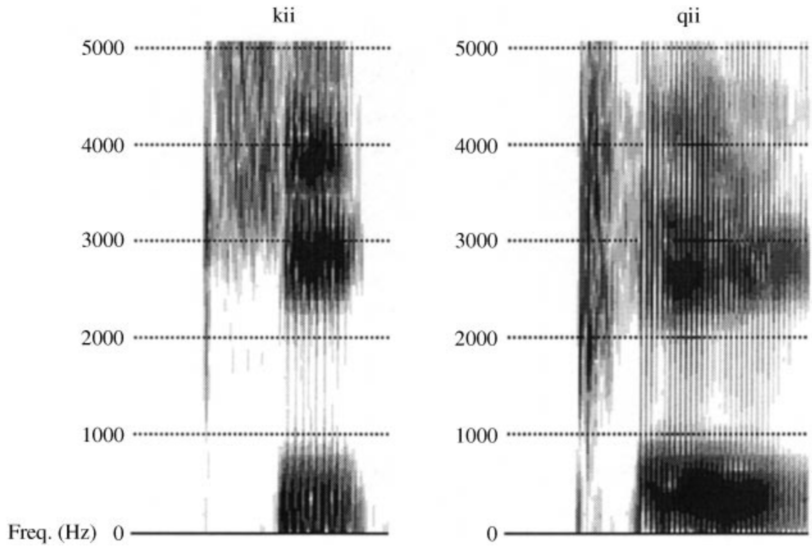

Figure 7. Spectrograms of the first syllables in **kiikaχ**, (cranberry bush) and **qiiɣaχ** (grass) as produced by a female speaker of Eastern Aleut. The uvular **q** is associated not only with more overall noise energy of frication but also with greater low-frequency energy.

The difference between the velar stop **k** and the uvular stop **q** can be mainly characterized by the different frequency location of the spectral peak. The velar stop is associated with a higher-frequency spectral peak as compared with the uvular stop (supporting Hypothesis I). The data also 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 due to the resonance frequency of the cavity in front of the constriction. Illustrative spectrograms for **k** and **q** are shown in Fig. 7. Similar results for a **k/q** effect from the following vowels were obtained for Thompson Salish (Mayes, 1979). Finally, as we will see in the discussion of vowels, the formants during the vowels reflect differences between preceding velar *vs.* uvular stops, providing additional cues distinguishing these two sounds.

2.3. *Other consonantal features*

There are many other interesting properties of Aleut consonants that we hope to discuss in a subsequent paper. We will note here that the consonants listed in the dental column in Table I 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", our observations show that the articulation is more apical and further back. This is in line with the coronals in French, which are typically laminal dental, but with the lateral **l** being more commonly apical and further back (Dart, 1998). This is possibly because laterals may be easier to make if they are apicals, the narrowing of the tongue tending to protrude the tip.

The segment **ð** is usually an interdental fricative, but in initial position, where it occurs mainly in Russian loan words, it is often a stop. The Aleut sibilant **s** varies from an alveolar articulation to a more retracted articulation similar to the palato-alveolar fricative in English. There is no contrast between **s** and **ʃ** in Aleut (some speakers commented that they found this distinction hard in English).

3. Vowels

As we have noted, Aleut has a basic three-vowel system, including a front high vowel **i**, a mid low vowel **a**, and a back high vowel **u**, each with a long counterpart. The Eskimo vowel systems includes **ə** for Yupik as well as **i**, **a**, **u** (Bergsland, 1986; Fortescue, Jacobson & Kaplan, 1994). In Aleut, there are no vowel sequences within a syllable (other than long vowels, which would 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 VI. As we have noted, some of the vowels are much rarer than others, making it difficult to find good, minimally contrasting sets.

The phonetic qualities of the vowels were examined, based upon values for the first three formants, which were measured using the Kay CSL. Two repetitions of each of the words listed in Table VI 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 in each token was identified, and superimposed LPC and FFT spectra were calculated with 30 and 25.6 ms frames, respectively, centered at the midpoint. The formant values were determined from the LPC spectra (with 12 or 14 coefficients), using the FFT spectra (and sometimes formant-history tracking) as supplementary checks.

The most suitable words for comparing the vowels in Table VI are those in the velar column, in which each vowel is preceded by **k** and followed by a velar or coronal consonant. Fig. 8 shows the relative positions of the first two formants of these vowels. These plots were drawn with the UCLA PlotFormants software. The scales are in accordance with the Bark scale, and the ellipses are drawn with radii of two standard deviations along the axes of the first two principal components.

Of general phonetic interest is the distribution of the vowels in Fig. 8. This pattern supports the form of dispersion theory (Liljencrants & Lindblom, 1972) which predicts that, in a given system, contrastive vowels are spaced with a sufficient contrast

TABLE VI. 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			

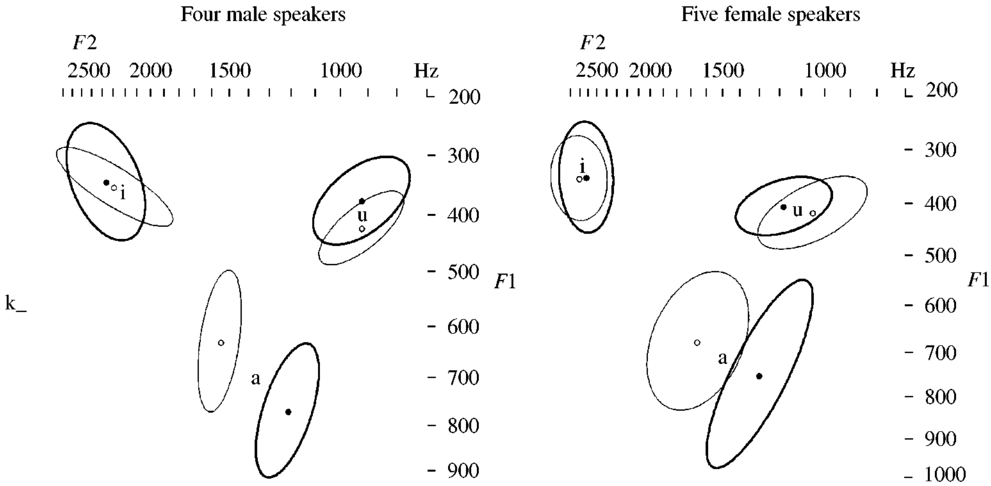

Figure 8. Formant plots of long (heavy ellipses) and short (light ellipses) vowels after **k** in the words in Table VI as produced by four male speakers and five female speakers.

(Lindblom, 1986; Lindblom & Maddieson, 1988). The vowels in Fig. 8 are more than adequately dispersed, being 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 independent variable, showed that $F1$ was significantly lower for **i** than for **u** ($p = 0.0052$ for the male speakers, $p < 0.0001$ for the female speakers). Ladefoged, Ladefoged, & Everett (1997) found a similar effect in Banawá, an Arawakan language spoken in Brazil. Less formal observations of other three-vowel languages that also show this effect lead us to conclude that there may be a general phonetic tendency for languages with a small number of vowels to have a back vowel that is slightly lower than the corresponding front vowel.

An analysis of variance also showed that the only differences in vowel quality corresponding 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 $p < 0.01$). The long and short high vowels were not significantly different in any way. Some of the difference between long and short **a** may 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, but the tendency of long and short **a** to differ in quality is also reported in Taff (1992). Vowels tend to become fronted after laminal dental consonants in which the tongue blade is raised. If this is the case in these vowels, then it seems that length itself does not have a major effect on vowel quality in Aleut. The perceptual distinction between long and short vowels depends on vowel duration without further support from any vowel-quality difference.

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). Fig. 9 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.

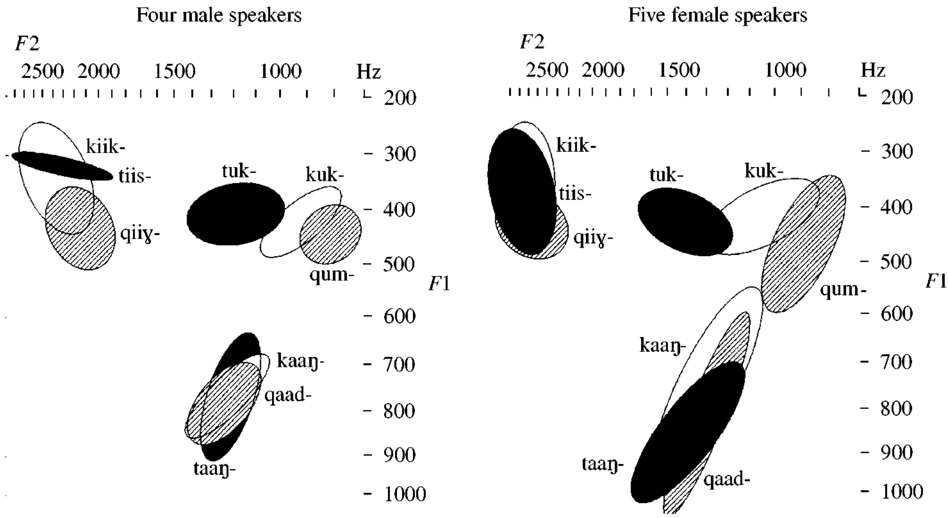

Figure 9. The effect of the preceding consonant on vowel quality. So that the influence of the final consonant may also be considered, the syllables in which the vowels occurred are shown, ellipses velar context, ellipses alveolar context, ellipses uvular context.

Analyses of variance with $F1$ or $F2$ as the dependent variable and place of articulation as the independent variable were made, 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 ($p < 0.0001$). The high back vowel **u** after **q** also tends to be lower ($F1$ is raised) in comparison with when it occurs after **t** ($p = 0.0362$). More strikingly, **u** is more front ($F2$ is raised) when preceded by **t** than when preceded by **k** ($p < 0.001$), and there is a strong tendency for it to be more back when preceded by **q** in comparison with when preceded by **k** ($p = 0.0373$). 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 ($p < 0.01$).

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 VI. As can be seen in Fig. 10, in 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 ($p < 0.01$). In general, uvular consonants have the greatest effect on the quality of adjacent vowels, and the following uvulars probably have a greater effect than the preceding uvulars. We do not have appropriate data to verify this statistically but our research suggests that vowel effects from following uvulars may have become lexicalized.

Taff (1992, 1999) reports that a vowel may retain uvular effects despite final syllable reduction processes (a prosodic marker of sentence end) which delete a word-final uvular

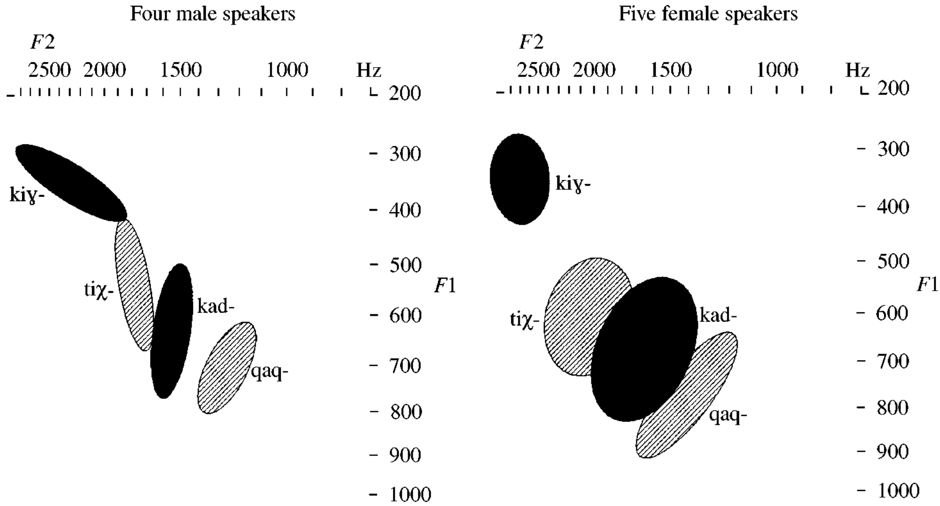

Figure 10. The effect of final uvular consonants (hatched ellipses) compared with final nonuvular consonants (● ellipses) on vowels quality. As these are not minimal pairs, the syllables in which the vowels occurred are shown.

itself. The present tense morpheme **-ku-** becomes [**-ko-**] when followed by the first-person singular absolutive (3abs) **-χ**. For example, /**uluuða-*ku*-χ**/ "it is red" becomes [**uluuðə-*ko***] (Taff, 1999, p. 280). Examples also suggest possible vowel harmony resulting from long-range uvular effects: /**qaχtʃikl*u*-*ku*-χ**/ "it is black" becomes [**qaktʃiklo-*k*-χ**] (Taff, 1999, p. 263). Here the /**u**/ of the present tense morpheme **-ku-** is lost in a devoicing effect of final syllable reduction. But the *previous* /**u**/, part of the stem, is uvularized to [**o**]. Additional analysis along the lines of Bessell's (1998) report on Salish long-distance effects of uvulars is required to establish Aleut lexicalization and vowel-harmony effects related to uvularization. Although frequent in Amerindian languages, uvulars and their effects remain understudied.

4. Vowel length and stress

As we have seen, Aleut has three vowels, **i**, **a**, and **u**, all of which can be both long and short. Aleut also has a quantity sensitive stress system, depending on syllable weight. According to Taff (1992), syllables containing long vowels are heavy, and syllables containing short vowels are light. Stress falls on the last syllable if it is heavy provided that the penultimate syllable is light, as in (1) in Table VII. Otherwise, stress falls on the penultimate syllable, as in (2) and (3). Both long and short vowels can be in the stressed penultimate syllable or in the unstressed final syllable. A further complication arises because final syllables are often deleted in Aleut. We must therefore consider whether stress is assigned before or after the final syllable is deleted.

The situation is still more complex. Bergsland (1994) observed the indeterminacy in Aleut stress assignment and that stress may affect vowel and consonant length. 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 Table VII (4) and (5), and an unstressed long vowel can sound short, as in (6) and (7).

TABLE VII. Variations in stress placement with vowel length in Eastern Aleut

	Orthography	IPA	Gloss
(1)	sichiing	**si'tʃiiŋ**	"nine"
(2)	siching	**'sitʃiŋ**	"four"
(3)	adaadaa	**a'ðaaðaa**	"their father"
(4)	am<u>a</u>xsix	**a'maxsix**	"to spend the night"
(5)	amaagaas<u>a</u>lix	**amaaɣaa'salix**	"to take over there"
(6)	am<u>aag</u>aasix	**amaa'ɣaasix**	"to get there"
(7)	adaad<u>aa</u>	**a'ðaaðaa**	"their father"

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 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 & Lehiste (1960), Umeda (1975), and Crystal & House (1988) have measured vowel duration in English in various phonological environments, including stress. Adams & 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.

Since vowel length is not phonemic in English, none of the studies of English 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 suggested 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, as neither of these languages uses pitch contrastively, Beckman (1986) does not address the question of the neutralization of a contrast by stress. Still, this study is relevant in that it shows that there are cross-linguistic differences in the extent to which one of the phonetic cues to stress can vary.

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 **bɪbɪbɪbɪ**. 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 20 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 languages with phonemic length do not use duration as a correlate of stress.

Many factors influence vowel duration besides distinctive length and stress. Among these factors are vowel quality, postvocalic voicing, postvocalic place of articulation (Peterson & Lehiste, 1960), speaking rate (Crystal & House, 1982), word length (Lehiste, 1972), morphological structure (Bond, 1991), prepausal position (Umeda, 1975; Crystal & House, 1988), and word prominence (the information load the word carries in the message) (Umeda, 1975). According to Peterson & Lehiste (1960) and Umeda (1975) the *pre*vocalic consonant has no consistent effect on vowel length in English. Turk & Shattuck-Hufnagel (2000) summarize the factors affecting vowel duration in words. 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.

4.1. *Experimental procedure for studies of vowel duration and stress*

Data from one speaker, ML, were 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. Although ML, a man in his mid 80s, had been living off of the Pribilof Islands for several years, he spoke Aleut daily with his wife, who is also a native speaker. The other three subjects, BS, LM, and GF, were part of the Eastern Aleut group of speakers who were recorded on DAT equipment as described above. BS, a man in his early 70s, speaks Aleut at home with his wife and with his peers. LM, a woman in her late 60s, speaks Aleut daily with her peers. GF, a man in his early 50s, 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.

All three Aleut vowels were measured in a variety of phonological, morphological and syntactic environments. There were three different sources of data. The first was a connected narrative of approximately 2 min elicited from subject ML who was instructed to tell a short story about something. The second source of data was 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 that seemed unnaturally slow were not used. The third source of data was the set of sentences elicited from each of three subjects, BS, LM and GF, who were part of the Eastern Aleut group. When sentences were repeated, corresponding measurements were averaged.

Every vowel from each source was measured, although not all measurements were included in the analysis. 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

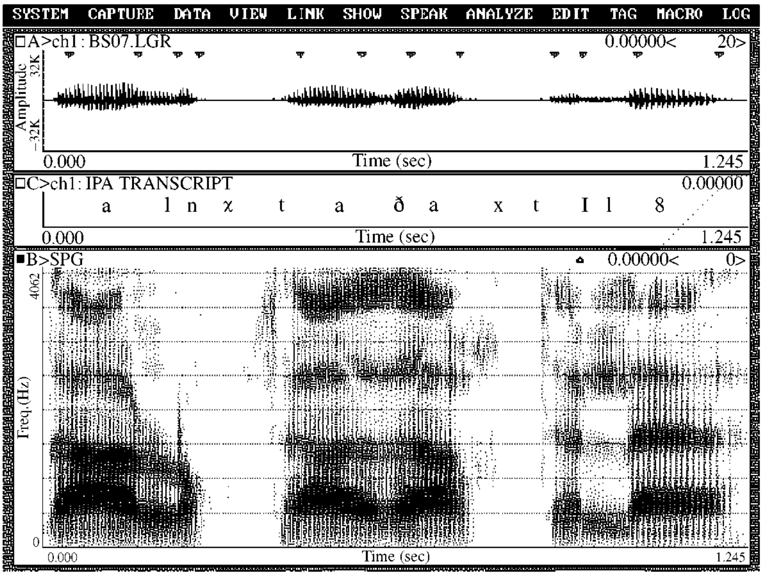

Figure 11. The waveform, transcription and spectrogram of *aalux̂taadax̂til, ii* **aaluχtaaðaχtil ii** "Did (the girls) laugh?" The intervals that were measured are indicated. The transcription beneath the waveform is slightly narrower than the transcription in the text.

deletion has occurred were not included in the primary data set. A total of 537 vowels was measured, 99 short unstressed vowels, 317 short stressed vowels, 83 long unstressed vowels, 38 long stressed vowels.

Vowel duration was measured from the spectral analysis produced on a Kay CSL speech analysis system. In order to mitigate the influence of surrounding consonants on vowel length, the section of the spectrogram where there was formant structure was examined, and the duration of the portion that could be regarded as being simply that of the vowel determined by narrowing the range until the preceding and following consonants were not audible. This proved to be a useful way of eliminating on and off glides and measuring only the vocalic portion of the formant structure. Fig. 11 provides an example.

In order to check the accuracy of the measurements, a random sample of vowels was measured by another phonetician using the same procedures but working independently. Approximately, 7% of the data were checked. These measurements were compared to the corresponding measurements that had been made previously by the author. A paired-sample *t*-test showed that the means of the two sets of measurements were the same at the 0.05 level of significance; therefore the difference between them was not significant.

4.2. *Vowel duration and stress results*

Fig. 12 summarizes the means and standard deviations for short and long vowels in stressed and unstressed positions. 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

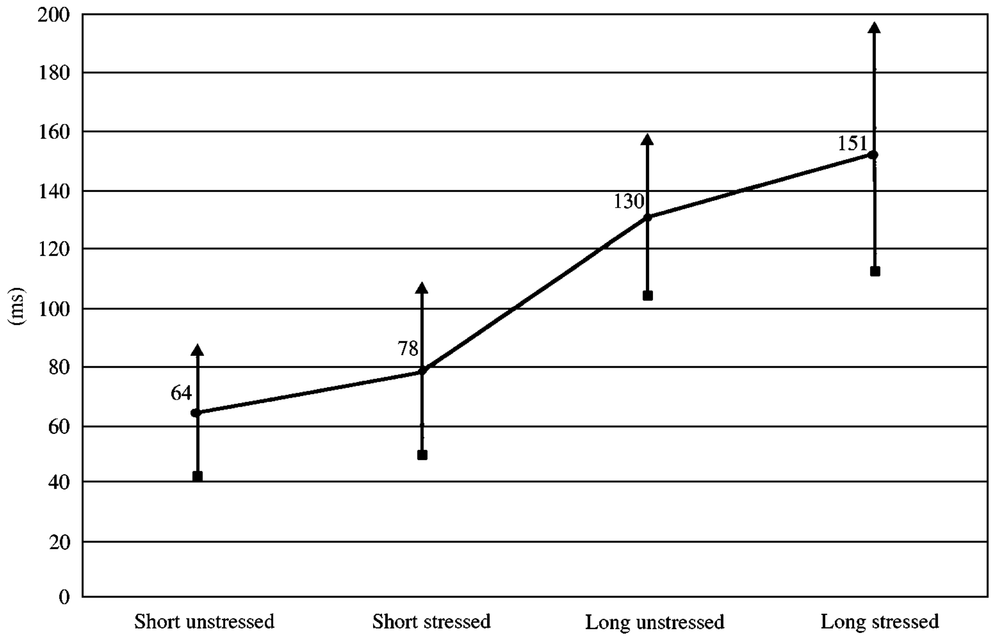

Figure 12. Duration (ms) of stressed and unstressed vowels: ▲ mean + S.D.; ● mean; ■ mean − S.D.

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 nonparametric Mann–Whitney rank sum test was used to compare means (Snedecor & 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 the 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. If there were just random measurement errors then the results of the Mann–Whitney rank sum test would not have been as significant. 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.

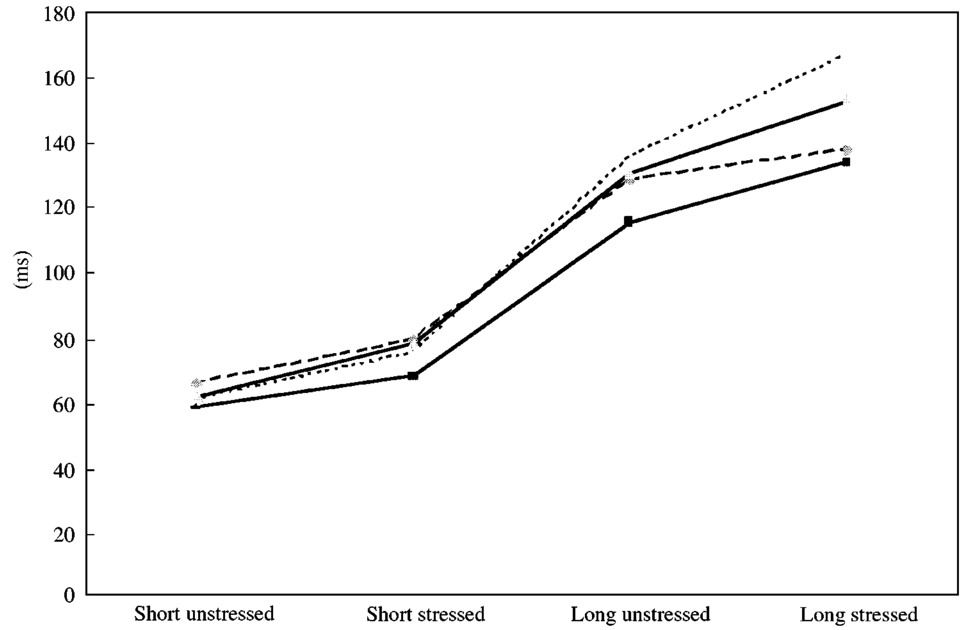

Figure 13. Means of the durations (ms) of four subjects for the stressed and unstressed, long and short vowels: bs; gf; lm; ml.

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. Fig. 13 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.

The durations of short stressed and unstressed vowels for each subject are 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 20% increase in duration between short unstressed and stressed vowels, only 0.5 speaker was required, indicating that the intersubject variability was very small. As is clear from Fig. 13, there is more between-subject variability for long vowels than short 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 20% 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 two subjects are necessary, indicating that there is little between-subject variability amongst the other three subjects. This is in accordance with Kehoe, Stoel-Gammon & Buder (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.

4.3. *Final rhyme deletion*

The relation between duration and stress is affected by rhyme deletion. Final rhymes are frequently deleted in Aleut at the ends of declarative sentences (Taff, 1999). Do such words receive stress before or after the final rhyme is deleted? For some of these words, the question of the rule order is irrelevant. For example, when the final rhyme is deleted, *ayagaadax̂* "girl" is stressed on the underlying penultimate syllable regardless of the order of application of these two rules, as shown in (1) in Table VIII. However, for other words, the two orders yield two different outputs, as shown in (2) in Table VIII. In *igax̂tanax̂* "airplane", if the word receives stress before the final rhyme is deleted, the penultimate syllable is stressed: **iɣaχ'ta**. But, if the final rhyme is deleted before the word receives stress, the antepenultimate syllable is stressed: **i'ɣaχta**.

As 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 rhyme 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 12, four were repetitions whose measurements were averaged, and one was only two syllables. The vowel durations for each of the remaining seven words is shown in Table IX. The rhyme that was deleted is shown in parentheses.

In each of these words, if the word were to receive stress before the final rhyme had been deleted, the outcome would be that penultimate syllable would be stressed. If, on the contrary, the final rhyme were deleted before the word had received stress, then the antepenultimate syllable would be stressed. The measurements show that in each case the vowel with greatest duration is the penultimate, not the antepenultimate. Accordingly, we may conclude that the penultimate syllable is stressed, and words receive stress before the final rhyme is deleted.

TABLE VIII. Some possibilities for stress placement with deletion of the final rhyme

Underlying form	Rule order	Derived forms	Rule order	Derived forms
(1) ayaɣaaðaχ	stress final deletion	**aya'ɣaaðaχ** **aya'ɣaa**	final deletion stress	**ayaɣaa** **aya'ɣaa**
(2) iɣaχtanaχ	stress final deletion	**iɣaχ'tanaχ** **iɣaχ'ta**	final deletion stress	**iɣaχta** **i'ɣaχta**

TABLE IX. Durations (ms) of syllables in words that have undergone final rhyme deletion

	Antepenult	Penult	Final
igaχtan(aχ)	48	67	0
iχtad(a)	64	104	0
kumsiχtad(an)	55	79	0
kumsiχtak(un)	64	71	0
kumsiχtal(ix)	45	57	0
qakchikluk(uχ)	36	49	0
qaxchiklulak(aχ)	46	88	0

Additional support for the conclusion that final rhyme deletion occurs after stress assignment is found in the nature of the deletion process. This process has one of the characteristics of a postlexical rule: it can have variable output (Kiparsky, 1985). Sometimes the final rhyme does not delete, sometimes it deletes entirely, and sometimes it only devoices.

4.4. *Duration as a correlate of stress*

It is apparent that stress in Aleut lengthens vowels, but perhaps 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 would be a weakening of Berinstein's (1979) claim that languages that use length contrastively do not use duration at all as a correlate of stress. So far, her claim is supported only by her own K'ekchi data in which the ratio of stressed short to unstressed short vowels is 1.0. 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 contrast, in a language without length contrasts, like English, the ratio is much greater, 1.6 or 1.7 (Fry, 1995; Crystal & House, 1988, respectively).

In summary, the measurements reported here 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 to which Taff's stress rule assigns main stress are about 20% longer than those not assigned main stress. Furthermore, this acoustic analysis of Aleut shows that duration is indeed a robust correlate of stress, even in a language that uses length phonemically.

5. Intonation

We analyzed the pitch tracks of sentences elicited from eight native speakers of Eastern Aleut, concentrating on the content words. Preliminary analysis indicates that function words may have different intonation characteristics, but this issue requires study beyond the scope of this investigation. Content words make up the bulk of Aleut vocabulary while function words are few. Many sentences have no function words at all. Aleut is an

TABLE X. Declarative *vs.* yes/no question morphology in Aleut*

(a) *Agyuĝum uyungin aduqlakun.*				“Cormorants necks are long”.			
agyuĝu +	m	uyu +	ngin	adu +	qla +	ku +	n
cormorant	RELp	neck	RELp	long	clumsy	prs	ABp
(b) *Agyuĝum uyuungin aduklalix, ii.*				“Are cormorants necks long?”			
agyuĝu +	m	uyu +	ngin	adu +	qla +	lix	ii
cormorant	RELp	neck	RELp	long	clumsy	cnj	Q

* Abbreviations: RELp, relative phrasal; Abp, third person absolutive plural; prs, present; cnj, conjunctive; Q, question word.

agglutinating language with the possibility of affixing several suffixes to a lexical (root) morpheme.

Yes/no questions are marked morphologically by the conjunctive morpheme, -*lix*, **liχ**, 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 word which is *ii*, in the orthography, but which is realized as a different form, **hæ** or **ε** in the Eastern dialect. In Table X, the underlined morphemes show the contrast between declarative and yes/no question morphology. 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 (a) is *qlak* and the final syllable of (b) is *qlal* (+ ii).

5.1. *Experimental procedure for studies of intonation*

In the process of recording the 12 speakers of Eastern Aleut, the researcher gave a stimulus sentence orally in English, and the speaker then provided an oral translation in Eastern Aleut. Each speaker recorded the same list of around 18 sentences. The list included declarative *vs.* yes/no question pairs and simple *vs.* complex sentences. Statistical analysis was limited to eight speakers to get an equal number of men and women.

To avoid gaps in pitch tracks and pitch perturbations caused by voiceless segments, sentences were designed to contain words with all voiced and mostly sonorant segments. However, as 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 Aleut orthography. Translations, glosses, and speaker intentions (e.g., question *vs.* declarative) were later confirmed by a native speaker.

Utterances were sampled at 10 000 Hz on Kay CSL. Pitch tracks were generated using a frame length of 25 ms and a frame advance of 20 ms. Typical pitch tracks are illustrated in Figs 14–18. Figs 14 and 15 are responses from two different speakers to the English

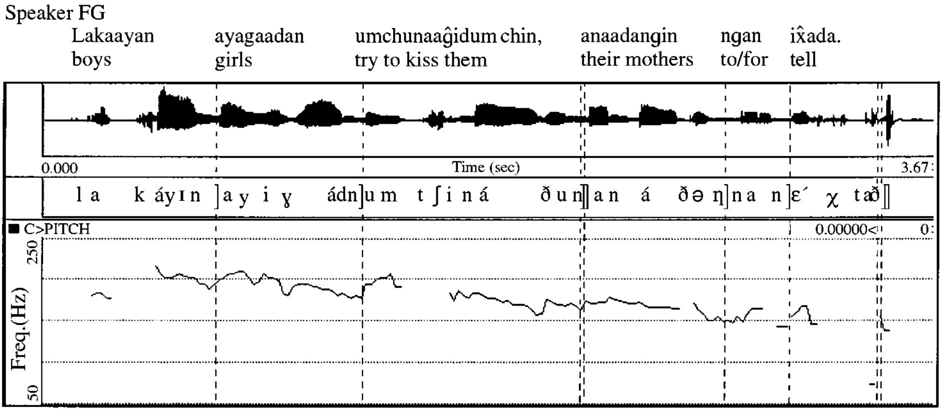

Figure 14. Speaker FG's response to the English stimulus, "When boys try to kiss girls, tell their mothers".

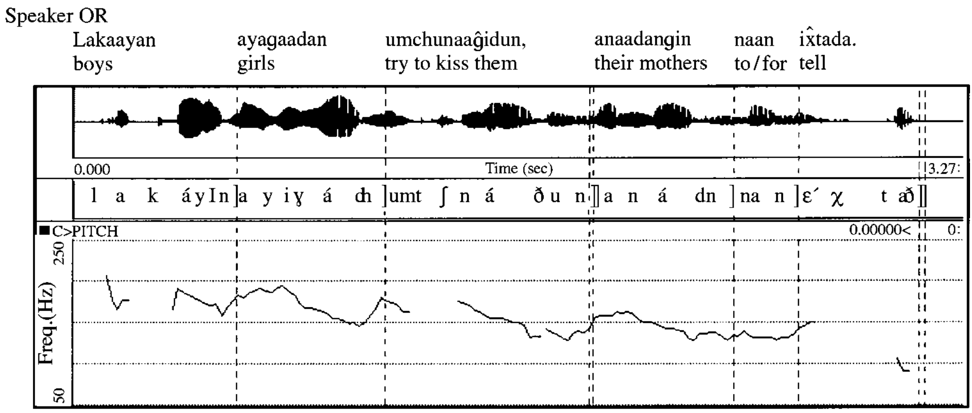

Figure 15. Speaker OR's response to the English stimulus, "When boys try to kiss girls, tell their mother".

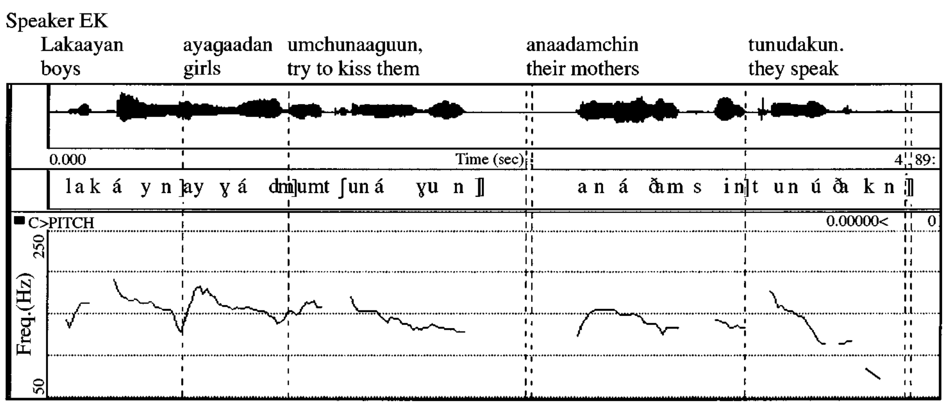

Figure 16. Speaker EK's response to the English stimulus, "When boys try to kiss, girls tell their mothers".

stimulus, "When boys try to kiss girls, tell their mothers". Fig. 16 is the response from a third speaker to the English stimulus, "When boys try to kiss, girls tell their mothers." Fig. 17 is the response to "Cormorants' necks are long". Fig. 18 is the response to "Are cormorants' necks long?" Each example includes an orthographic rendering, gloss, waveform, 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. In Fig. 14, the word *ngan* is a function word as is *naan* in Fig. 15.

Numerical results of the entire pitch track for each sentence were entered into a database for analysis. When the amplitude of the wave was too low for the pitch tracking function of the CSL, pitch was calculated from the waveform. Tokens (sentences) were grouped by speaker and plotted together. To make words comparable, word durations were converted from time units (i.e., seconds) to word units. Thus, in each of the graphs in Figs 19 and 20, the first words start at 0.0 and end at 1.0; the second words

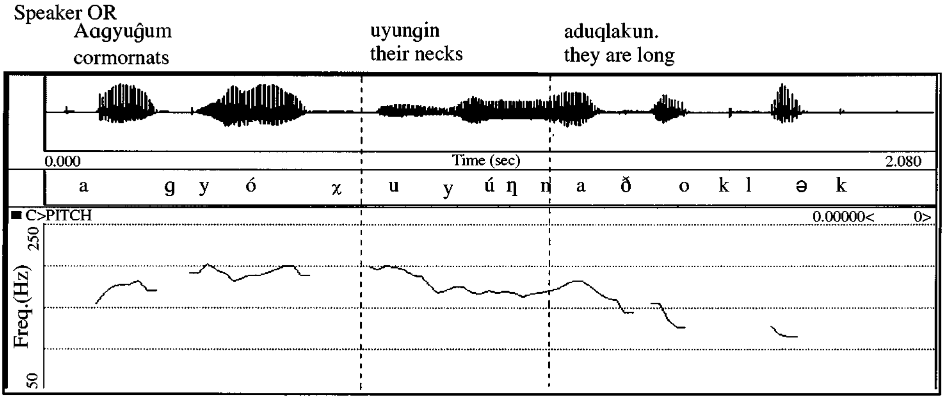

Figure 17. Speaker OR's response to the English stimulus. "Cormorants' necks are long".

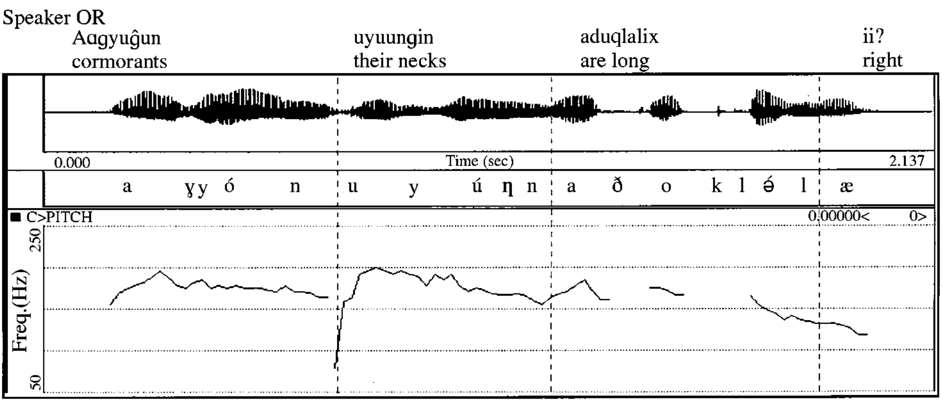

Figure 18. Speaker OR's response to the English stimulus, "Are cormorants' necks long?"

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 (Friedman, 1984).) This summary is visible as the bold curve on each plot. Given data points x_i and y_i ($i = 1, \ldots, n$), the 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 are 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 Figs 19 and 20.

Smoothed pitch tracks of three-word declaratives and yes/no questions are shown in Fig. 19 for the four female speakers, FG, RM, EK, and MM, and in Fig. 20 for the four male speakers GR, AK, BS, and GF. The declaratives appear in the left-hand column and the yes/no questions appear in the right-hand column. Note that in Eastern Aleut, the question word, *ii*, exemplified in Table X(b) and Fig. 18, makes three-word declaratives convert to four-word interrogatives.

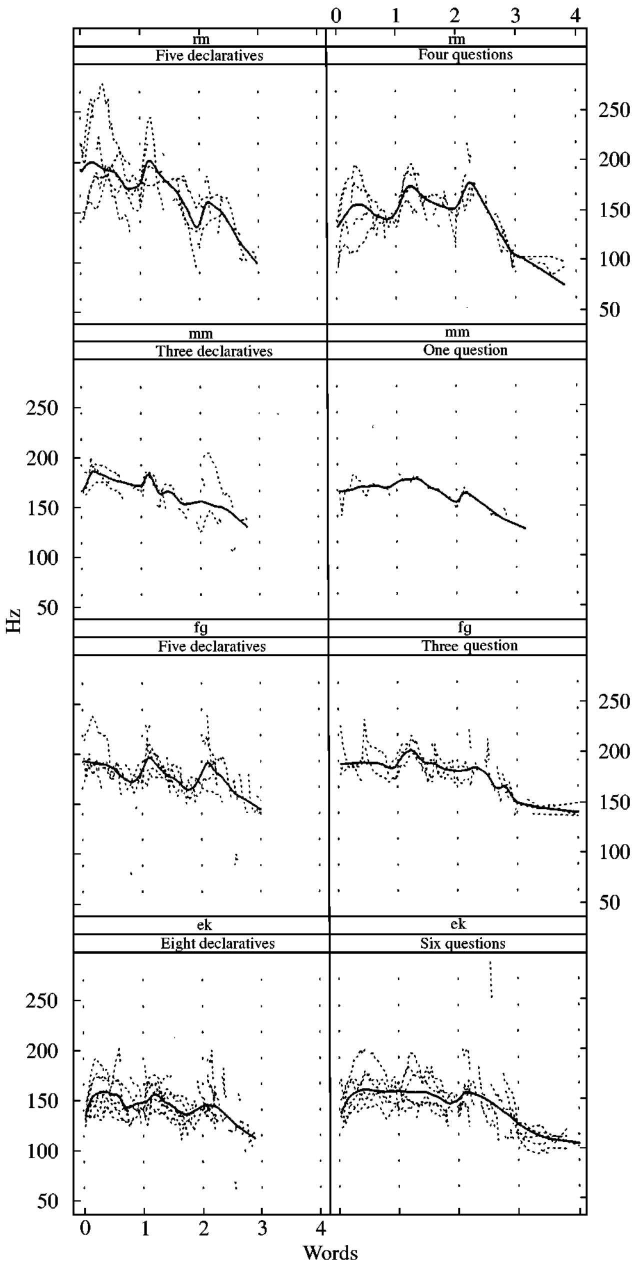

Figure 19. Smoothed pitch tracks for the four female speakers, FG, RM, EK, and MM.

5.2. *Word contours*

Pitch tracks of declarative sentences suggest that each content word is mapped to its own pitch contour. Each content word has a peak near its beginning and a trough near its end. The smoothed results of the declarative sentences of the four women, RN, MM, FG and EK in Fig. 19, and the first man, GR, in Fig. 20, show a clear mapping between word

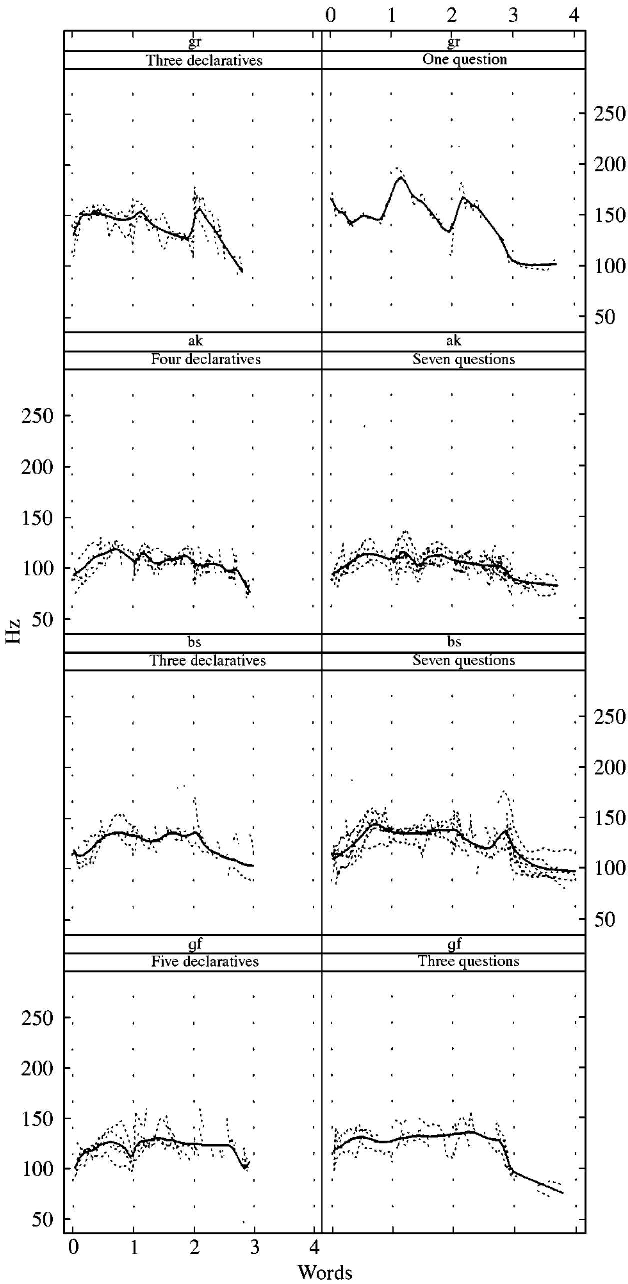

Figure 20. Smoothed pitch tracks for the four male speakers GR, AK, BS, and GF.

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 Eastern Aleut each word is matched to its own intonation contour, with a peak near the beginning and trough near the end.

Results from the declarative sentences of the other three men, AK, BS, and GF, in Fig. 20 are not to easy to analyze. AK shows distinct troughs at word boundaries but also has a small trough in the middle of the second word and a peak preceding the troughs at the end of the first, second, and third words. It is still possible to say, however, that word-length information chunks are identified intonationally by a particular pitch contour. Results from speaker BS also make a case for word-length intonation contours but here the contour is the opposite of that described for the first five speakers. BS shows troughs at word beginnings and peaks near word ends (except for the final word). Results from speaker GF show the least indication of word boundaries being marked by intonation. He has only one trough, at the boundary between the first and second words.

If a word-sized domain does function as an intonational domain it may be related to the morphological structure of Aleut. Perhaps each word is treated intonationally as an individual phrase within the sentence because it must have inflectional suffixes and may have numerous derivational suffixes which provide it the potential for standing alone as a sentence. Of general typological interest, this kind of one-to-one isomorphy between intonation contours and each (content) word in a sentence has not been found in Japanese (Pierrehumbert & Beckman, 1988), English (Pike, 1945 and numerous others), and Bengali (Hayes & Lahiri, 1991). However, in Cup'ik (Central Alaskan Yupik), geographically proximate to and distantly related to Aleut, content words have characteristic intonation contours (Woodbury, 1993) but in Cup'ik such words begin with troughs and end with peaks.

5.3. *Sentence contours*

Pitch tracks in Figs 14–18 and the smoothed sentences in Figs 19 and 20 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 and each successive trough 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.

The issue of an intonation contrast between declaratives and yes/no questions might seem pointless; there is no need for an intonation contrast between these two sentence types in Eastern Aleut 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.

The smoothed intonation contours in Figs 19 and 20 do not show that yes/no interrogatives have intonation contours that clearly pattern differently from declaratives. The appearance of a long, flat final word in questions is misleading (and points out a flaw in this word-stretching and squeezing methodology). The final question word appears as a flat, near-horizontal line in the smoothed examples in Figs 19 and 20 because its duration is stretched. The fourth and final word in each of the yes/no questions is the word, *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. (See the example in Fig. 18, which should be compared with the statement without a final question word in Fig. 19.)

Looking at the smoothed contours of yes/no questions in the right-hand columns of Figs 19 and 20, most speakers show peaks near word beginnings and troughs near word ends for the questions much like those for the declaratives. The first speaker in Fig. 19, RM, has a slightly different pattern in that the downdrift of troughs is not a feature of her yes/no questions. Speaker AK in Fig. 20 has less dip at the second word boundary for questions than for declaratives, otherwise there is not much contrast between his declaratives and 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 word of his questions; otherwise, both his declaratives and questions can be said to be fairly flat.

There may be a 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 word. 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 Fig. 19) 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.

Another possibility is that questions may have a higher overall pitch than declaratives. In order to test this possibility, the median pitch was computed for each sentence and plotted by speaker and sentence type. Results from 172 sentences appear in Fig. 21. The plots in this figure include all sentences in the data, whereas those in Figs 19 and 20 are just the three-word (plus question word) tokens. In Fig. 21, sentence median pitches are plotted as open circles and compared by speaker and sentence type. Note that pitch is plotted on the horizontal axis. Pitch data points are spaced out on the vertical axis, so that the reader can see how many there are. Results from female speakers appear in the top half of the figure, those from male speakers in the bottom half.

Median pitches per sentence plotted in Fig. 21 show that the range of the medians of yes/no questions is within that of declaratives; the centre 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 artefact of the data; there are more declarative sentences than yes/no questions. The speakers with the most questions show the least difference in range between declaratives and questions.

The data analysis presented here is exploratory and incomplete; further work on Aleut intonation appears in Taff (1999). Considering all the data, it seems 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 of questions. Eastern Aleut sentence contours are serial peak–trough word contours. These are affected by downdrift 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, and end with a sharp drop.

6. Conclusion

This detailed account of the phonetic structures, of Aleut provides further data for studies of the range of phonological inventories that can occur and the general phonetic

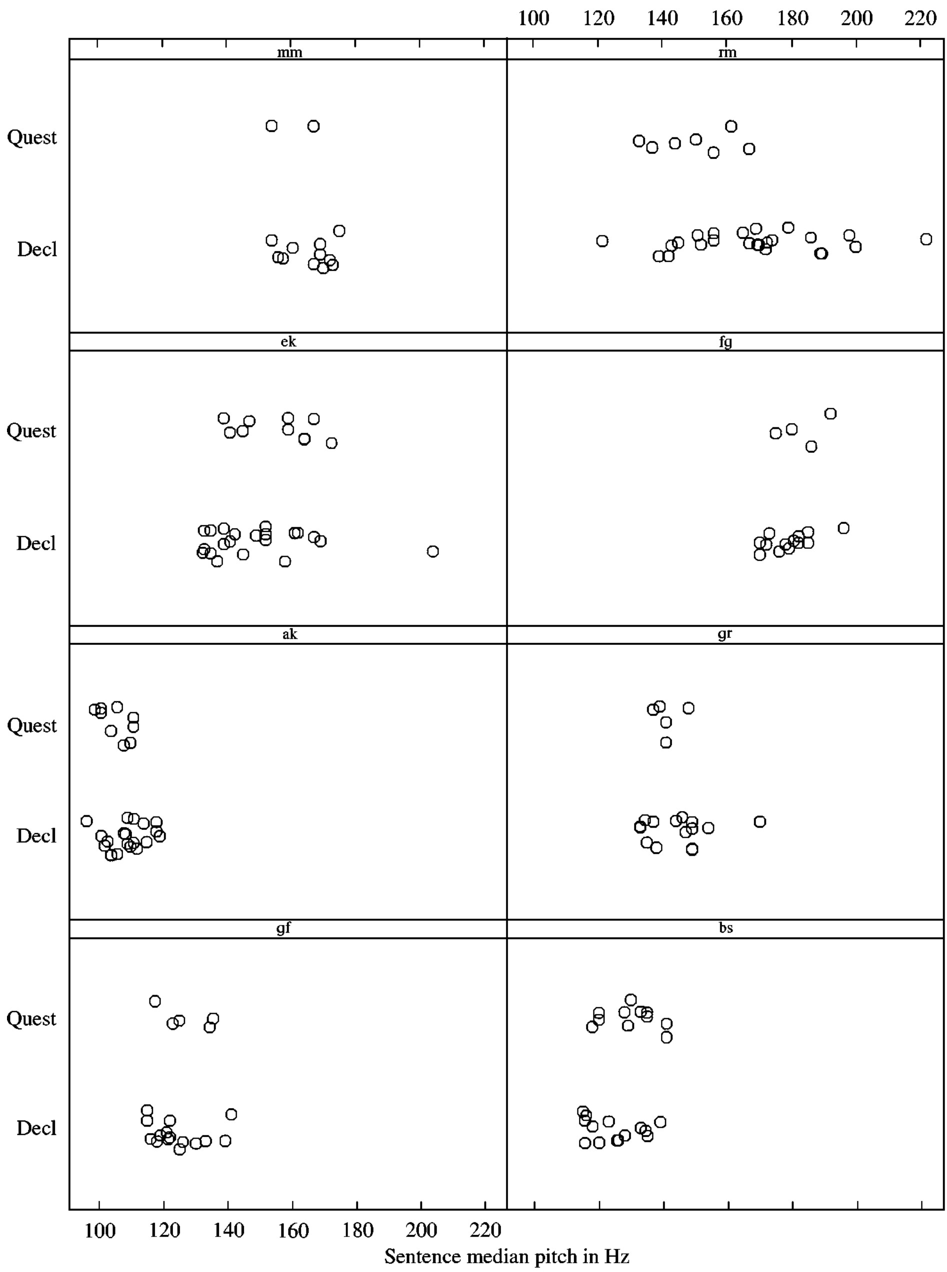

Figure 21. Sentence median pitches by speaker and sentence type.

tendencies that are present in the languages of the world. Aleut has an unusual inventory. It is among the 1% of the worlds' languages that lack both **p** and **b**. It is slightly less unusual in having both velar and uvular stops, a relatively under-documented contrast that occurs in 14% of the world's languages. It is among the 6% of languages that have only three vowels. (All these percentages are from Maddieson, 1984). In addition, Western Aleut contrasts voiced and voiceless nasals and laterals, a contrast no longer present in Eastern Aleut. We speculate that the loss of this contrast is because these sounds have a high airflow rate that requires considerable articulatory effect, but they are nevertheless auditorily indistinct in that they have a low intensity. Both these factors make them less preferred.

The difference between velar and uvular stops in Aleut provides evidence that languages do not always behave in accord with universal phonetic tendencies. It may be simplest to make the VOT of uvular stops longer than that of velar stops, but Aleut does not do this. Languages can choose to implement contrasts such as these in various ways (cf. Cho & Ladefoged, 1999). The major differences between velar and uvular stops in Aleut are in the spectral peaks in the bursts, which are, in a given vowel context, higher for velars than for uvulars. In addition, uvular consonants have a considerable effect on the quality of adjacent vowels.

In connection with the vowels, we noticed a phenomenon that we suspect might be a general phonetic tendency among languages with three vowels: the high back vowel is a little lower than the high front vowel. In making this remark we should point out that we are using the term "high" to designate an auditory property, not an articulatory one designating tongue height. The reason for this tendency, which may also account for the tendency for **a** and **i** to be more common than **u**, is that it is, from an articulatory point of view, more difficult to produce vowels in the high back portion of the auditory vowel space (Ladefoged, forthcoming).

Aleut is an exception to a phonetic universal that has been proposed about duration as an indicator of stress. It has been said that duration is unlikely to be used as a marker of stress in languages that contrast long and short vowels (Berinstein, 1979). There is some individual variation among Aleut speakers in this matter, but the general conclusion seems unassailable: in Aleut, duration is a concomitant of stress, even though it has contrastive vowel length.

Aleut intonation appears to be somewhat unlike intonation in other languages, although here we are on more unsure ground because intonation has not been well documented in many languages. The complex morphological structure of Aleut lends itself to the possibility that each word may be treated intonationally as an individual phrase within the sentence. Sentence contours are concatenations of simple peak–trough word contours. This is true of both statements and questions. Most speakers have peaks near the beginning of each word and troughs near the end.

Aleut has much of interest to phoneticians, but it will not be possible to study it much longer. Children are no longer learning the language at home. This will mean that one more of the comparatively few languages that contrast uvular and velar stops will no longer be available. Three-vowel languages are also not all that common, and we need to see how many of them have vowels distributed in the vowel space as in Aleut, with **u** being lower than **i**. Even more exceptional are languages that have vowel-length contrasts, and yet have vowel duration as a marker of word stress. Finally, intonation has been well studied in only a small percentage of the world's languages, and none of them have the morphological complexity of Aleut. We need more detailed studies of languages such as Aleut.

Our greatest debt is to out Aleut consultants, who were always helpful and eager to see further investigation of their endangered language. The work was supported by a National Science Foundation grant SBR9511118 to

Peter Ladefoged and Ian Maddieson, “Phonetic Structures of Endangered Languages”. Funding for data corroboration was provided by the Jacobs Research Funds, 1997, “Aleut Intonation”. Thanks also go to Richard Wright and two anonymous reviewers, and the editor, whose helpful comments strengthened this paper.

References

Adams, C. & Munro, R. R. (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, M. (1986) *Stress and non-stress accent*, Dordecht: Foris Publications.

Bergsland, K. (1986). Comparative Eskimo-Aleut phonology and lexicon, *Suomalais-Ugrilaisen Seuran Aikakauskira*, **80**, 1–75.

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

Bergsland, K. & Dirks, M. (1978) *Quwalangim tunugon kadug̑ingin.: Eastern Aleut grammar and lexicon.* National Bilingual Materials Development Center, Rural Education Affaris, University of Alaska.

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

Bessell, N. (1998) Local and non-local consonant-vowel interaction in Interior Salish, *Phonology*, **15,** 1–40.

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

Cho, T. & Ladefoged, P. (1999) Variations and universals in VOT: evidence from 18 languages, *Journal of Phonetics*, **27**(2), 207–229.

Crystal, T. & House, A (1982) Segmental durations in connected speech signals: preliminary results, *Journal of the Acoustical Society of America*, **72,** 705–716.

Crystal, T. & House, A. (1988) Segmental durations in connected speech signals: syllabic stress, *Journal of the Acoustical Society of America*, **83,** 1574–1585.

Dart, S. (1998) Comparing French and English coronal consonant articulation, *Journal of Phonetics*, **26,** 71–94.

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.

Fishman, J. A. (1991) *Reversing language shift*. Clevedon: Multilingual Matters.

Friedman, J. (1984) Avariable span smoothers. *Technology report #5*. Laboratory of Computational Statistics, Department of Statistics, Stanford University.

Fortescue, M., Jacobson S. & Kaplan, L. (1994) *Comparative Eskimo Dictionary with Aleut Cognates.* Fairbanks: Alaska Native Language Center.

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.

Grenoble, L. A. & Whaley, J. L. (1998) (editors) *Endangered languages, language loss and community response*, Cambridge: Cambridge University press.

Grimes, B. E. (1999) (editor) *Ethnologue: Languages of the World*. Dallas: Summer Institute of Linguistics.

Hayes, B. & Lahiri, A. (1991) Bengali intonational phonology, *Natural Language and Linguistic Theory*, **9,** 47–96.

Jakobson, R., Fant, G. & Halle, M. (1963) *Preliminaries to speech analysis*, 4th edition. Cambridge, MA: MIT Press.

Keating, P. & Lahiri, A. (1993) Fronted velars, palatalized velars, and palatals, *Phonetica*, **50,** 73–101.

Kehoe, M., Stoel-Gammon, C. & Buder, E. H. (1995) Acoustic correlates of stress in young children's speech, *Journal of Speech and Hearing Research*, **38,** 338–350.

Kiparsky, P. (1985) Some consequences of lexical phonology, *Phonology Yearbook*, **2,** 85–138.

Ladefoged, P. (forthcoming) The relative density of different parts of the auditory vowel space.

Ladefoged, P., Ladefoged, J. & Everett, D. (1997) Phonetic structures of Banawá, an endangered language, *Phonetica*, **54,** 94–111.

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

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

Lindblom, B. (1986) Phonetic universals in vowel systems. In *Experimental phonology* (J. Ohala & J. J. Jaeger, editors), pp. 13–44. New York: Academic Press.

Lindblom, B. & Maddieson, I. (1988) Phonetic universals in consonant systems. In *Language, speech and mind* (L. Hyman & C. Li, editors), pp. 62–78. London: Routledge.

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

Maddieson, I. (1984) *Patterns of sounds*. Cambridge: Cambridge University Press.
Mayes, S. (1979) An acoustic analysis of Thompson velar/k/ and uvular/q/, *University of Hawaii Working Papers in Linguistics*, **11,** 11–22.
McCarthy, J. (1979) *Formal problems in semitic phonology and morphology*. PhD dissertation, MIT.
Peterson, G. E. & Lehiste, I. (1960) Duration of syllabic nuclei in English, *Journal of the Acoustical Society of America*, **32,** 693–703.
Pierrehumbert, J. & Beckman, M. (1988) *Japanese tone structure*. Cambridge, MA: MIT Press.
Pike, K. (1945) *The intonation of American English*. Ann Arbor: University of Michigan Press.
Port, R. F. & Rotunno, R. (1979) Relation between voice-onset time and vowel duration, *Journal of the Acoustical Society of America*, **66,** 3, 654–662.
Snedecor, G. W. & Cochran, W. G. (1980) *Statistical methods*. Iowa: The Iowa State University Press.
Stevens, K. (1998) *Acoustic phonetics*. Cambridge, MA: MIT Press.
Taff, A. (1992) *Uvularization of vowels in Aleut*. MA thesis, University of Washington.
Taff, A. (1999) *Phonetics and phonology of Unangan* (*Eastern Aleut*) *intonation*. PhD dissertation, University of Washington.
Turk, A. & Shattuck-Hufnagel, S. (2000) Word-boundary related duration patterns in English. *Journal of Phonetics*, **28**(4), 397–440.
Umeda, N. (1975) Vowel duration in American English, *Journal of the Acoustical Society of America*, **58,** 434–445.
Woodbury, A. (1993) Against intonational phrases in Central Alaskan Yupik Eskimo. *Presentation at Linguistic Society of America Meeting*, Los Angeles, January.

Appendix A: Aleut word list used as a basis for making recordings

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

A.1. Western Aleut List

A.1.1. Stops in initial position

	Aleut	English	Transcription and notes
ti	tiĝlax̂	bald eagle	**ˈtiʁlaχ**
tii	tiistax̂	dough, piecrust	**ˈtiistaχ**
	tiihnax̂	mound	**ˈtiin̥aχ**
ta	tanix̂	forehead	**ˈtaniχ**
taa	taangax̂	water	**ˈtaaŋaχ**
tu	tukux̂	chief	**ˈtukuχ**
	tulax̂	forearm	**ˈtulaχ**
tuu	tuutx̂ax̂	stalk of cow parsnip	**ˈtuutχaχ**
chi	chis	to stick, adhere, cleave	**ˈtʃis**
	chisil	scatter	**ˈtʃisil**
	chigdax̂	Aleut raincoat	**ˈtʃiʁdaχ** **d** or **ð**
chii	chiidax̂	baby animal, e.g., seal	**ˈtʃiiðaχ**
cha	chagix̂	halibut	**ˈtʃaɣiχ**
chaa	chaasxix̂	cup	**ˈtʃaasxiχ**
chu	chugux̂	sand	**ˈtʃuɣuχ**
chuu	chuulkix	socks	**ˈtʃuulkix**
ki	kigil	to bite	**ˈkiɣil**

kii	kiin	who	**ˈkiin**
ka	kadan	in front of the speaker	**ˈkaðan**
kaa	kaangux̂	healthy	**ˈkaaŋuχ**
ku	kukax̂	grandmother	**ˈkukaχ**
kuu	kuusxix̂	cat	**ˈkuusxiχ**
qi	qiqix̂	slime	**ˈqiqiχ**
qii	qiigax̂	grass	**ˈqiiɣaχ**
qa	qalgadax̂	food	**qalˈɣaðaχ**
qaa	qaadas	dolly varden (fish)	**ˈqaaðas**
qu	quhmax̂	white	**ˈqum̥aχ**

A.1.2. Sibilants

si	sisax̂	hundred	**ˈsisaχ**
sii	siintax̂	cent	**ˈsiintaχ**
sa	sas	birds	**ˈsas**
saa	saalax̂	lard, reindeer fat	**ˈsaaluχ**
	saakux̂	king eider	**ˈsaakuχ**
su	susux̂	pus	**ˈsusuχ**
suu	suuskax̂	baby's bottle	**ˈsuuskaχ**
zi			
zii	ziilitax̂	vest	**ˈziilitaχ**
za	azax̂	to be (no time marker)	**ˈazaχ**
zaa	zaavtrikax̂	breakfast	**ˈzaavtrikaχ**
zu	huzus	to make plenty	**ˈhuzus**
zuu	zuulutax̂	gold	**ˈzuulutaχ**

A.1.3. Voiced fricatives

Orthographic **d**, **g** are normally fricatives, but in initial position (where they are comparatively rare) they may be stops, particularly in Russian and English 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	**ˈdiχ**
dii	diikal	to be mischievous	**ˈdikal**
da	dax̂	eye	**ˈdaχ**
daa	daax̂tux̂	kidney	**ˈdaaχtuχ**
du	duskax̂	washboard, plank	**ˈduskaχ**
gi	gil	he is envious	**ˈɣil**
gu	agul	to make, build	**ˈaɣul**
ĝaa	x̂aayax̂	steam bath	**ˈχajaχ**
x̂a	x̂ax̂	(halibut) stomach	**ˈχaχ**
gaa			
gu	gul	to go through	**ˈɣul**
hi	hinus	piece to sod	**ˈhinus**
ha	halal	to turn the head	**ˈhalal**
hu	hudax̂	dried fish	**ˈhudaχ**
ĝ	aĝal	to open, ebb tide	**ˈaʁal**

A.1.4. 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.)

uχ	asxinux̂	(one) girl	**asˈxinuχ**
ux	alax asxinux	two girls	**ˈalax asˈxinux**
iχ	(dix̂)	soot	**ðiχ**
ix	(chuulkix)	socks	**ˈtʃuulkix**
aχ	(taangax̂)	water	**ˈtaaŋaχ**
ax	hizax	almost	**ˈhizax**

A.1.5. Voiced and voiceless nasals and laterals

mi	miilax̂	soap	**ˈmiilaχ**
ma	mal	to do	**ˈmal**
mu	mukax̂	flour	**ˈmukaχ**
ni	Niiĝuĝis	people of Atka	**ˈniiʁuʁis**
na	naga	it's inside	**ˈnaɣa**
na	anax̂	mother	**ˈanaχ**
nu	qanul	smells of fish	**ˈqanul**
ngi	qungix̂	hump (back)	**ˈquŋiχ**
nga	hangal	to go up	**ˈhaŋal**
ngu	qangul	to go in	**ˈqaŋul**
hmi	hmiichix̂	ball	**ˈm̥iitʃiχ**
hma	hmachil	to get stuck	**ˈm̥atʃil**
hmu	hmuqatix̂	fish's gill cover	**m̥uˈqatiχ**
hni	chuhnil	to poke	**ˈtʃun̥il**
hna	ahnatix̂	marker	**aˈn̥atiχ**
hnu	hnul	to reach	**ˈn̥ul**
hngi	kahngil	to bend	**ˈkaŋ̥il**
hnga	ahngal	to acknowledge	**ˈaŋ̥al**
hngu	qahngus	seaweed	**ˈqaŋ̥us**

A.1.6. Approximants

wi			
wa	waya	right here, now	**ˈwaja**
wu			
li	lil	to look like, appear	**ˈlil**
la	lal	to pick up, gather	**ˈlal**
lu	lul	to believe	**ˈlul**
yi			
ya	yas	reef	**ˈjas**
yu	ayul	to fall	**ˈajul**
hwi			
hwa	hwaĝix̂	smoke	**ˈʍaʁiχ**

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**

A.1.7. 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**

A.1.8. 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	**'l̥am ta'nii**

A.1.9. Velar vs. uvular influence on the vowels

kak	kakix̂tal	looks up	**ka'kiχtal**
qaq	qaqaĝix̂	arctic loon	**qa'qaʁiχ**
kix	kixs	to bite	**'kixs**
qiq	(qiqix̂)	slime	**qiqiχ**
kuk	(kukax)	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	**'ðaaχtuχ**
ak	akalux̂	way, path	**a'kaluχ**
aq	aqal	stretch	**a'qal**

A.1.10. *Illustrative clusters*

chng	chngax̂	fur, pelt, yarn	**ˈtʃŋnaχ**
mg	humgix̂	lung	**ˈhumɣiχ**
mx	amxix̂	flesh of fish	**ˈamxiχ**
mv			
mχ			
mɣ	umĝix̂	bait	**ˈumʁiχ**
kd	kdax̂	ice	**ˈkðaχ** **ð** may be **θ**
sd	sdax̂	star	**ˈsðaχ** **ð** may be **θ**

A.2. *Intonation* (*Eastern Aleut*)

Tayuux̂ slutuu angqaxtal, qawax̂ ilgakux̂.	*Old men hunted sea mammals.*
Tayuux̂ slutuu angqaxtal, qawanaaĝil, ii.	*Did old men hunt sea mammals?*
Anaadax̂ ayagaadan kumsix̂takux̂.	*The mothers carried the girls.*
Anaadax̂ ayagaad kumsix̂tal, ii?	*Did the mothers carry the girls?*
Ayagaadan aalux̂taadakun.	*The girls laughed.*
Ayagaadan aalux̂taadaax̂tul, ii?	*Did the girls laugh?*
Ayagaadam angunangin aalux̂taadakun.	*The big girls laughed.*
Ayagaadam angunangin aalux̂taadax̂tul, ii.	*Did the big girls laugh?*
Aagyuĝum uyungin aduqlaadakun.	*Cormorants' necks are long.*
Aagyuĝum uyungin aduqlaadalix, ii.	*Are cormorants' necks long?*
Ayagaadan lakaayam umchinaaĝigumchin, anaadamchinaan ix̂taachin.	*If boys try to kiss girls, — tell their mothers.*
Lakaayan ayagaadan umchunaaguun, anaadamchin tunudakun.	*If boys try to kiss, — girls tell their mothers*
Adaadan nuĝin kumsix̂takun.	*Their fathers carried rock.*
Adaadan nuĝin kumsix̂takuniin, ayagaadan aalux̂taadakun.	*When their fathers carried rocks, — the girls laughed.*
Amaatxan akun.	*We were out there* (*camping*)
Igax̂tadax̂ waaĝakux̂.	*The airplane came.*
Amaadan akuniin, igax̂tadax̂ waaĝaax̂tukux̂.	*When we were out there, — the airplane came.*
Ulax̂ uluudalakax̂; qax̂chiklukux̂ agach.	*Their house is black, it isn't red.*