THE ACOUSTICS OF UVULARS IN TLINGIT

by RYAN E DENZER-KING

A thesis submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey
In partial fulfillment of the requirements
For the degree of
Master of Arts
Graduate Program in Linguistics
Written under the direction of
Shigeto Kawahara
And approved by

Shigeto Kawahara (chair)

Akin Akinlabi

Bruce Tesar

New Brunswick, New Jersey
January 2013

ABSTRACT OF THE THESIS

The Acoustics of Uvulars in Tlingit

By RYAN DENZER-KING

Thesis Director:
Shigeto Kawahara

This paper looks at the acoustics of uvulars in Tlingit, an Athabaskan language spoken in Alaska and Canada. Data from five native speakers was used for acoustic analysis for tokens from five phoneme groups (alveolars, plain velars, labialized velars, plain uvulars, and labialized uvulars). The tokens were analyzed by computing spectral moments of plosive bursts and fricatives, and F_2 and F_3 values for post-consonantal vowels, which were used to calculate locus equations, a descriptive measure of the relationship between F_2 at vowel onset and midpoint. Several trends were observed, including a greater difference between F_2 and F_3 after uvulars than after velars, as well as a higher center of gravity (COG) and lower skew and kurtosis for uvulars than for velars. The comparison of plain versus labialized consonants supports the finding of Suh (2008) that labialization lowers mean burst energy, or COG, and additionally found labialization to raise skew and kurtosis.

Acknowledgements

There are numerous people without whom this work would not have been possible. I mention some of them here, but there are many more whom I relied upon for personal and academic support during its writing. First and foremost I must thank my chair Shigeto Kawahara, who suggested that this might be an interesting project worth pursuing, and who taught me a great deal about phonetics, statistics, and the useful relationship that exists between the two. Akin Akinlabi provided invaluable encouragement, insight, and criticism throughout the course of this research. Bruce Tesar, in addition to helpful comments throughout, was a major force in the focusing and presentation of the final result of this research, and his ability to bring clarity to all he touches is hopefully in evidence here. Aaron Braver answered numerous questions as I struggled through learning R and Praat scripting, as well as contributing mightily to my knowledge of LaTeX so that I could provide a clearer format for my research. Will Bennett, Jeremy Perkins, and Alan Prince contributed valuable comments at the early stages of this research. I must also thank James Crippen, Seth Cable, audiences at the 2010 Athapaskan/Dene Languages Conference and RUMMIT 2010, and students in Shigeto Kawahara's Phonology III seminar for useful comments on a related project which inspired the present work. Last and certainly not least I must thank my wife Amanda, who helped me emotionally and editorially at all stages of this project.

Contents

List of Tables

List of Figures

1 Introduction

This paper investigates the acoustic properties of uvulars in Tlingit by examining the correlates of place of articulation and labialization in Tlingit obstruents at the alveolar, velar, and uvular places of articulation, including labialized velars and uvulars. Measurements were made to determine the properties of plosive burst and fricative spectra and formant values for post-consonantal vowels, which were used to calculate locus equations for alveolar, velar, and uvular obstruents. Identifying the acoustic correlates of a phoneme group is a necessary first step in the study of how the acoustic signal is mapped onto phonemes that enter into phonological and other linguistic processes, and in fact the dispersion theory of contrast (Flemming 2002) argues that such acoustic properties are directly involved in phonological perception and production. Taking into account these acoustic facts allows us to predict a number of empirical phonological facts (e.g., preferentially rounding back vowels) without requiring stipulation (e.g., a redundancy rule stipulating [+back] → [+round]).

Previous research indicates that formant transitions relative to a vowel, the first four spectral moments for plosive bursts and fricatives, and locus equations are likely starting places for acoustically distinguishing the various places of articulation (Stevens & Blumstein 1978, Walley & Carrell 1983, Werker & Tees 1984, Stevens et al. 1986, Forrest et al. 1988, Alwan 1989, Sussman et al. 1993, Jongman et al. 2000, Gordon et al. 2002, Suh 2008, Żygis & Padgett 2010). The interaction of these acoustic cues allow speakers to accurately assign each segment of the acoustic signal to a phoneme group. F_1 drops when any constriction is made in the oral cavity, which means that F_1 will lower into any consonant. The movement of F_2 and F_3 is variable according to place of articulation (Jakobson et al. 1952, Stevens & Blumstein 1978).

When an oral constriction is released, a noise burst is released from the site of the constriction. The distribution of spectral energy in these bursts correlates with place of articulation. The distribution of formant frequencies in the portion of a vowel adjacent to a consonant also correlates with place of articulation. A neutral vocal tract configuration with no constriction has peaks for second through fifth formants around 1500, 2500, 3500, and 4500Hz (Stevens & Blumstein 1978). Labial consonants lower the amplitude of these higher formants, and result in lower formant frequencies. Alveolar consonants raise the amplitude of higher formants, and shift formant frequencies upwards. Velar consonants have bursts with high and close F_2 and F_3 amplitudes. Uvulars have received less attention in the phonological and phonetic literature than labials, alveolars, and velars (though see Werker & Tees 1984, Alwan 1989). Werker & Tees (1984) find "little information as to the acoustic differences between velar and uvular sounds."

Another known acoustic correlate of place of articulation is the slope and y-intercept of the locus equation for a phoneme group. Locus equations (Sussman et al. 1993) are a descriptive measure of formant transitions obtained by plotting F_2 at the onset of a vowel following a consonant (F_2 onset) against F_2 at the midpoint for the vowel (a point at which coarticulation is minimal and the idealized target frequency for that vowel has been reach, hence called F_2 target) and calculating a regression line, as described in Sussman et al. (1993). The equation thus calculated represents the extent of coarticulation for the phoneme group. The relationship between F_2 onset and F_2 target represents an important acoustic cue in determining place of articulation (Fruchter & Sussman 1997, Waibel et al. 1989). While F_2 varies with both the preceding consonant and the vowel target, the slope and y-intercepts of locus equations for consonants is relatively stable. Labials have steep slopes and low y-intercepts, alveolars have shallow slopes and high y-intercepts, and velars have steeper slopes than alveolars and mid-range y-intercepts. Locus equations of uvulars are understudied (though see Shosted 2011).

Tlingit[1] is a Na-Dene language spoken in Alaska, British Columbia, and the Yukon. The Modern Tlingit phonological inventory is shown in Table 1. Currently the language is spoken by between 150 and 300 speakers (James Crippen, p.c.), with most speakers elderly. Because of these factors and others, this study used existing historical recordings, described in section 2.2. This paper first describes the method used in this paper, including background information about the recordings used and the speakers, then gives the results of the acoustic analysis. The paper concludes with discussion of the acoustic data and its implications for phonological theory.

---

[1] The preferred English pronunciation among members of the Tlingit speech communities is [ˈklɪŋ.kɪt].

	alveolar	*postalveolar*	*palatal*	*lateral*	*velar*	*lab. velar*	*uvular*	*lab. uvular*	*glottal*
unaspirated plosive	t				k	k^{w}	q	q^{w}	ʔ
aspirated plosive	t^{h}				k^{h}	k^{hw}	q^{h}	q^{hw}	
ejective plosive	t’				k’	k’^{w}	q’	q’^{w}	
unaspirated affricate	ts	tʃ		tɬ					
aspirated affricate	ts^{h}	tʃ^{h}		tɬ^{h}					
ejective affricate	ts’	tʃ’		tɬ’					
fricative	s	ʃ		ɬ	x	x^{w}	χ	χ^{w}	h
ejective fricative	s’			ɬ’	x’	x’^{w}	χ’	χ’^{w}	
nasal	n								
approximant			j			w			

Table 1: Tlingit phonological inventory

2 Method

The core method of this paper seeks to replicate aspects of three previous studies of English and other languages, looking at spectral moments of plosive bursts (Forrest et al. 1988) and fricative intervals (Jongman et al. 2000), as well as F$_2$ and F$_3$ values to determine formant transitions and locus equations for all tokens (Sussman et al. 1993). Extracting spectral moments from an interval relies on treating the phonetic spectrum as a probability distribution and calculating the first four moments around the mean: mean (generally referred to as center of gravity or COG), standard deviation (not discussed here as it typically does not play as large a role in phoneme discriminability as do the other moments), skewness or skew (the location of the peak relative to the distribution), and kurtosis (how peaked the distribution is). Comparisons of spectral moments have been shown to accurately distinguish between groups of phonemes in a number of studies (Stevens & Blumstein 1978, Behrens & Blumstein 1988, Forrest et al. 1988, Sussman et al. 1993, Jongman et al. 2000, Sundara 2005, Suh 2008, among others). For formant transitions and locus equations, F$_2$ and F$_3$ were measured at vowel onset and vowel midpoint (representing the articulatory or acoustic target of the vowel).

The data consulted for this paper resides in two primary collections: one of recordings made by Richard and Nora Dauenhauer between 1952 and 1997, and one of recordings made by Margaret Gross-Hope and Margaret Sturtevant (date unknown, but thought to be late 20th century).

2.1 Speakers

The Dauenhauer recordings comprise some forty speakers. I excluded recordings with excessive static or noise in the 0–10 kHz range. Four recordings were selected that were judged to be of suitable length (greater than 60 minutes) and suitable quality (the majority of words in the recordings were enunciated clearly and accurately picked up by the microphone). These recordings consist of speakers AJ (male), JM (female), and WM (male), who speak the Northern dialect of Tlingit, and speaker HD (male), who speaks the Sanya dialect (a subgrouping of Southern Tlingit). Dialectal differences are minor; the most relevant differences for this study are slight differences in vowel realization, which do not affect the acoustic characteristics studied here (see Leer et al. 2001 for discussion of the differences among interior Tlingit dialects). Recordings range from 61 to 186 minutes of audio per speaker. The Gross-Hope/Sturtevant recordings comprise three speakers, with one recording of acceptable quality: speaker FD (male), who is a speaker of the Northern Tlingit dialect. The recording comprises 62 minutes of audio.

2.2 Data

All recordings have been digitized in stereo (except for speaker HD, which recording is mono) at a 44.1kHz sample rate from reel-to-reel and audio cassette tapes. The majority of the recordings used in this study were not annotated, and thus most tokens were selected by listening to the recordings and extracting CV tokens whenever a clear token of an alveolar, velar, or uvular obstruent was heard. One goal was to ensure that the tokens extracted from the recordings were clear and easily identifiable, so as to collect prototypical tokens. On the other hand, it is desirable to obtain a sufficient number of tokens for statistical analysis. In an attempt to balance these two goals, two passes over the recordings were made. The first pass covered roughly the first 75% of each recording, and was done extremely conservatively, extracting only easily perceivable tokens in areas of the recordings with minimal background noise. Extracting only these easily perceivable tokens biases the data towards representative tokens rather than ones that are marginally perceptible. Tokens from the first pass represent 20–30% of the total number of tokens for each speaker. The second pass covered roughly the last 25% of each recording, and was done very liberally, only excluding tokens where the place of articulation could not be determined, increasing the number of tokens extracted so that the sample would be large enough to be statistically meaningful. Table 2 shows the number of tokens per speaker by place of articulation. Table 3 shows the number of tokens by manner of articulation.

Speaker	**alveolars**	**velars**	**lab. velars**	**uvulars**	**lab. uvulars**	**total**
AJ	330	148	47	77	63	665
FD	194	79	57	46	33	409
HD	350	126	53	100	48	677
JM	168	41	37	22	14	282
WM	66	24	22	24	10	146
total	1108	418	216	269	168	2179

Table 2: Number of tokens per speaker by place of articulation

2.3 Procedure

The tokens were annotated using Praat (Boersma & Weenick 2011). For unaspirated plosive tokens, the plosive burst onset was marked at the beginning of aperiodic noise, while the burst offset/vowel onset was marked at the beginning of voicing, as shown in Figure 1. A similar procedure was followed for aspirated plosive tokens and fricative tokens, as shown in Figures 2 & 3. For affricate tokens, no attempt was made to distinguish between the stop and fricative portion of the consonant (as shown in Figure 4), since only prevocalic word-initial and intervocalic word-medial tokens were used, to ensure clarity, and neither the stop portion nor the fricative portion fit either of these contexts. Affricate tokens were

Manner	**alveolars**	**velars**	**lab. velars**	**uvulars**	**lab. uvulars**	**total**
C	594	129	123	64	19	929
C^h	213	206	41	42	57	559
C’	6	31	7	6	4	54
C^F	84	0	0	0	0	84
F	206	17	27	132	69	451
F’	5	35	18	25	19	102
total	1108	418	216	269	168	2179

Table 3: Number of tokens by manner and place of articulation

C = plosive, F = fricative, C^h = aspirated plosive, C’ = ejective plosive, C^F = affricate, F’ = ejective fricative

only used for calculating locus equations, and thus the vowel formant structure was the only portion of the token on which analysis was conducted. For ejective plosive tokens, the plosive burst onset was marked at the beginning of aperiodic noise, while the burst offset was marked at the end of aperiodic noise, as shown in Figure 5. Since the acoustic properties of ejective fricatives are understudied, these tokens were left out of fricative calculations to avoid any unforeseen changes they could introduce. Because these tokens were only used for calculating locus equations, only the vowel onset was marked (as shown in Figure 6). Only labialized dorsal obstruents (i.e., velar and uvular obstruents) occur before (or after) round vowels, and thus there are no instances of plain /ku/ or /qu/ in the corpus. To verify that these sequences should be counted as phonetically labialized, these sequences were initially kept as a separate group. However, initial investigation into the spectral moments of these bursts as compared to the bursts for phonemically labialized plosives found no significant difference between the two groups, so for all the data below these are pooled with the phonemically labialized plosives.

For unaspirated stops, the burst interval was queried for the first four spectral moments (center of gravity, standard deviation, skewness, kurtosis). Center of gravity (COG), or mean, indicates the frequency at which energy is most concentrated. Since strident fricatives

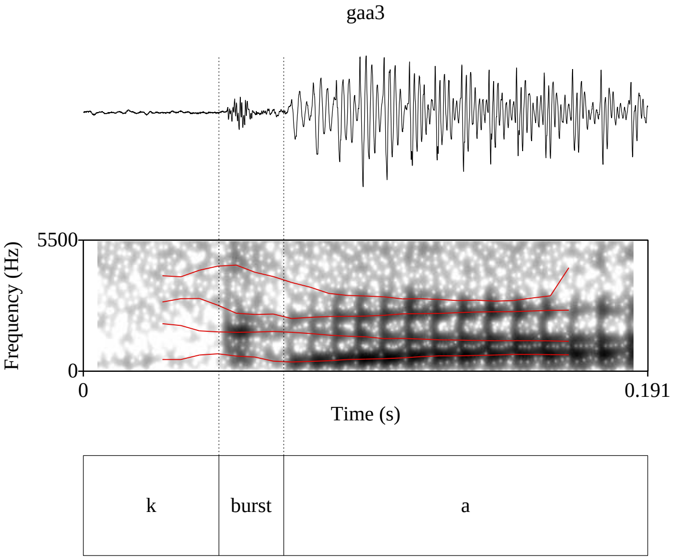


Figure 1: Spectrogram for syllable [*kaː*] from speaker HD

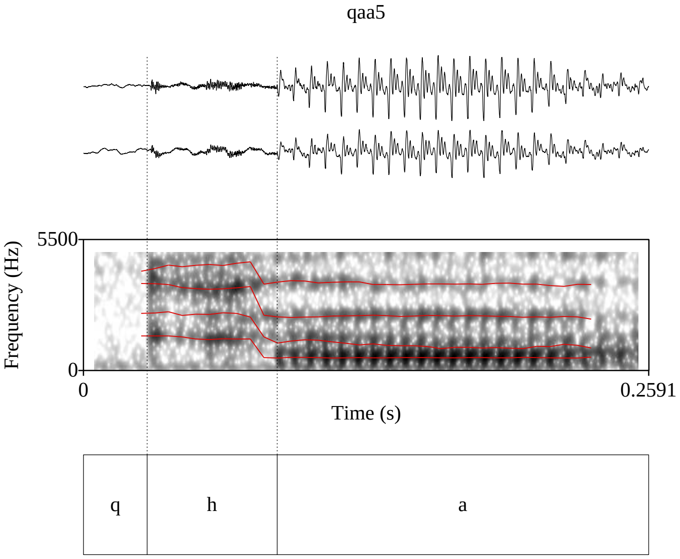


Figure 2: Spectrogram for syllable [q^h*aː*] from speaker FD

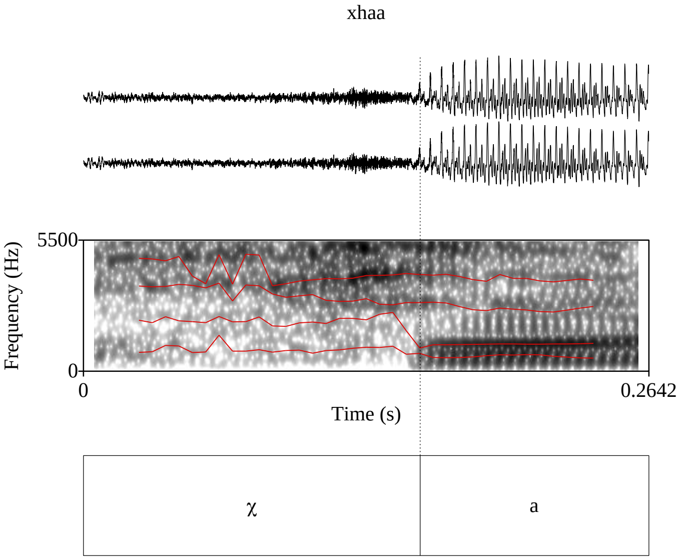


Figure 3: Spectrogram for syllable [*χaː*] from speaker AJ

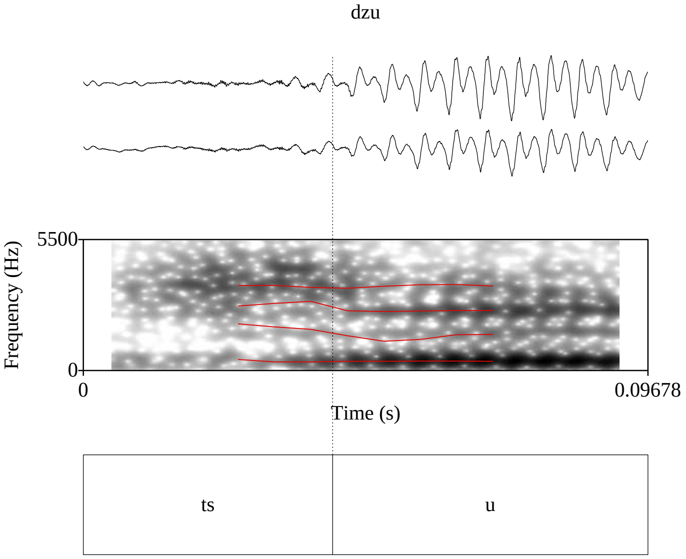


Figure 4: Spectrogram for syllable [*tsʊ*] from speaker FD

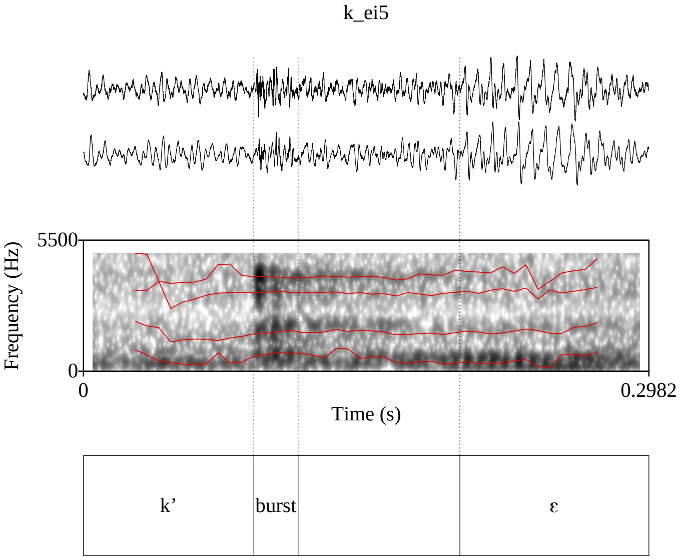


Figure 5: Spectrogram for syllable [*k'ɛ*] from speaker JM

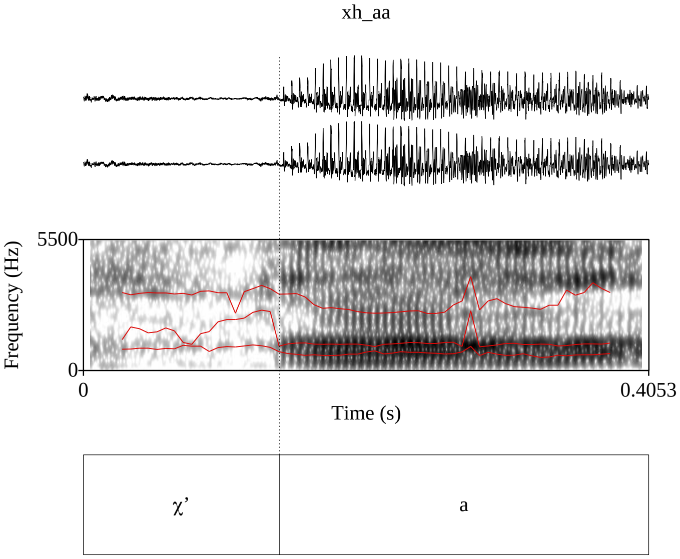


Figure 6: Spectrogram for syllable [*χ'aː*] from speaker AJ

feature high-frequency noise, these consonants have among the highest COGs. COG is computed for plosive bursts and fricatives, because these intervals consist of aperiodic noise. The third moment, skewness or skew, describes the skew or tilt of the spectrum, and is also called spectral tilt. A high positive skew indicates a much longer tail on the right side of the distribution than the left side, i.e., acoustic energy more sharply concentrated in the lower frequencies of the distribution than in the higher ones. The fourth moment, kurtosis, indicates how peaked the distribution is. A high kurtosis value indicates a relatively sharply pointed distribution, while a low value indicates a more shallow curve, i.e., a spectrum with a high kurtosis value will be very peaked, showing a high concentration of acoustic energy in a narrow frequency range. Following Forrest et al. (1988), only center of gravity, skewness, and kurtosis are compared in this paper. All intervals were resampled at 22kHz, and all extracted windows were Gaussian. For aspirated stops, the first 20ms of the aspiration interval was extracted and queried for spectral moments.

The procedure followed for fricatives duplicated that of Jongman et al. (2000). Fricative intervals were low-pass filtered at 11kHz, and four windows were extracted from the token: (1) the first 40ms after the fricative onset, (2) a 40ms window centered on the fricative midpoint, (3) the last 40ms before the fricative offset, and (4) a 40ms window centered on the fricative offset. These four windows were then queried for spectral moments. Locus equations were calculated following Jongman et al. (2000) and Sussman et al. (1993). Two windows were extracted from vowel intervals: (1) the first 23.3ms of the vowel interval, and (2) a 23.3ms window centered on the vowel midpoint. For each window, mean F_2 and F_3 were queried. Locus equations were calculated by creating scatter plots of F_2 onset (window 1) versus F_2 target (window 2) and performing linear regression analysis in R (R Development Core Team 2011). For speakers with at least three tokens per condition, Chauvenets criterion was used to exclude outliers. The decision to exclude outliers was made based on the fact that the recordings used contain a significant amount of background noise, and thus tokens that did not meet the criterion were judged likely to be errors in the computation of the spectral moments rather than genuine measurements far away from the mean. The cumulative probability of each token in each condition was calculated using the normal distribution function, then was multiplied by the number of tokens for the condition. Any tokens with a resulting statistic value of less than 0.5 or greater than n−0.5 were excluded from analysis; *p*-values were calculated using a Welch two-sample t-test in R.

3 Results

This section gives the results for spectral analysis of plosive bursts and fricatives, as well as those for formant transitions and locus equation analysis for each phoneme group.

3.1 Plosive bursts

Tables 4–6 show the results of plosive burst analysis for all speakers. Because of the high degree of variability at many different levels of the recording process, results for aperiodic noise are given speaker by speaker rather than pooled. Two speakers (AJ, JM) had lower COG for uvulars, while three had higher, one of them (FD) significantly ($p < .05$). Two speakers (HD, JM) had higher skew for uvulars, one (HD) significantly ($p < .01$) and three had lower skew for uvulars, one (FD) significantly ($p < .05$). Four speakers (AJ, FD, JM, WM) had lower kurtosis for uvulars than velars, one (FD) significantly ($p < .001$). Controlling for vowel quality does not seem to change these numbers, as running a similar analysis limited to stops before /a/ (the most common vowel in the corpus, and the only one present for all speakers for all phonemes) showed no significant difference.

Following Suh (2008), the expectation was that labialization would lower COG for plosive bursts, which is supported by the data in Table 4. For /k/ vs. /kw/, all speakers had a lower COG for /kw/; all but one of these differences (FD) were significant (at the .01 level for JM and WM, and at the .001 level for AJ and HD). For /q/ vs. /qw/, all but one speaker (AJ) had a lower COG for /qw/. The difference for the one speaker with higher COG for /qw/ was not significant. Of the four speakers with lower COG for /qw/, three showed a significant difference (at the .05 level for HD, at the .001 level for FD and WM).

Skew also seems to provide a clear distinction between plain and labialized dorsal plosives in Tlingit. All speakers showed a significantly higher skew for labialized velars than for plain velars (WM $p < .05$, FD $p < .01$, AJ, HD, JM $p < .001$), and all but one speaker (AJ) showed a higher skew for labialized uvulars than for plain uvulars, all but one significantly (HD $p < .01$, FD, WM $p < .001$).

Higher kurtosis, too, seems to be a reliable feature of labialization. Every speaker had higher kurtosis for labialized velars and uvulars than for plain velars and uvulars. The

Speaker	k	kw	q	qw	p($k \neq q$)	p($k \neq kw$)	p($q \neq qw$)
AJ	1772 (102)	1140 (230)	1675 (457)	2155 (1010)[a]	.65	$< .001$	.33
FD	498 (66)	427 (54)	729 (163)	409 (75)	$< .05$	.10	$< .001$
HD	1169 (53)	927 (54)	1251 (73)	1133 (70)	.07	$< .001$	$< .05$
JM	646 (70)	516 (48)	589 (89)	527 (129)	.29	$< .01$	.38
WM	1985 (294)	1296 (312)	2465 (505)	1190 (99)	.10	$< .01$	$< .001$

Table 4: COG (95% confidence interval)

---

[a] The wide range of the 95% CI here (and throughout) is caused by a small number of tokens rather than a high degree of variance; this particular value for AJ was calculated from only 7 tokens of aspirated and unaspirated plosives.

Speaker	k	kw	q	qw	p($k \neq q$)	p($k \neq kw$)	p($q \neq qw$)
AJ	2.9 (0.2)	6.2 (1.8)	2.4 (0.8)	2.3 (1.7)	.22	$< .001$	.90
FD	3.6 (0.6)	5.5 (1.1)	2.8 (0.6)	5.6 (1.2)	$< .05$	$< .01$	$< .001$
HD	2.9 (0.2)	5.0 (0.4)	3.4 (0.3)	4.1 (0.4)	$< .01$	$< .001$	$< .01$
JM	5.1 (0.6)	7.8 (0.5)	5.1 (1.0)	6.7 (1.8)	.94	$< .001$	.10
WM	2.0 (0.4)	3.0 (0.7)	1.3 (0.6)	3.0 (0.5)	.06	$< .05$	$< .001$

Table 5: Skew (95% confidence interval)

Speaker	k	kw	q	qw	p($k \neq q$)	p($k \neq kw$)	p($q \neq qw$)
AJ	11 (2)	85 (51)	7 (5)	9 (11)	.13	$< .01$	.82
FD	37 (9)	122 (37)	17 (6)	95 (33)	$< .001$	$< .001$	$< .001$
HD	19 (2)	45 (7)	21 (4)	28 (6)	.30	$< .001$	$< .05$
JM	42 (9)	87 (12)	35 (16)	74 (44)	.43	$< .001$	.08
WM	5 (2)	12 (6)	3 (3)	10 (3)	.21	.05	$< .01$

Table 6: Kurtosis (95% confidence interval)

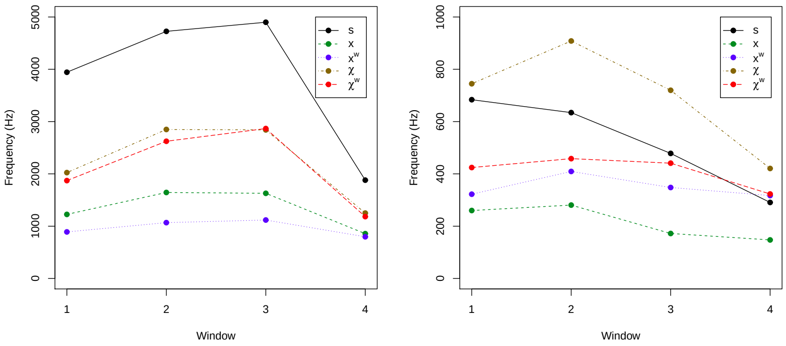
difference in kurtosis was significant for all but one speaker (WM) for velars (AJ $p < .01$, FD, HD, JM $p < .001$), and was significant for all but two speakers (AJ, JM) for uvulars (HD $p < .05$, WM $p < .01$, FD $p < .001$).

3.2 Fricatives

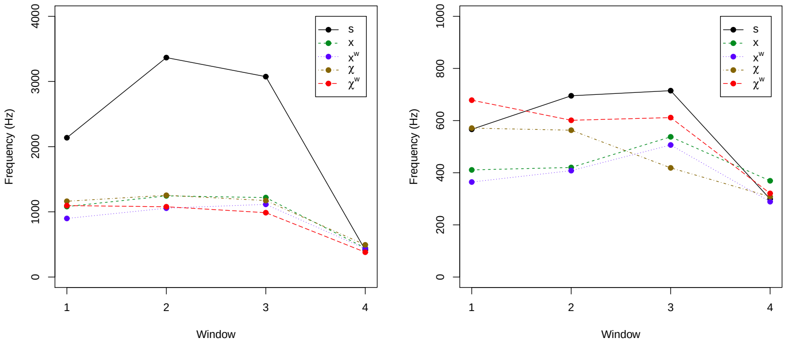
Figure 7 shows COG in the four windows described in section 2.3. Each subfigure represents the spectral moment for a single speaker, with each line indicating the values for a single phoneme group. In window one all speakers but WM had higher COG for uvulars, two of them significantly. Window two showed similar results, with all but speaker WM having higher COG for uvulars, three significantly. Window three displayed more variability, with three speakers showing higher COG for uvulars (two significantly), and two for velars (neither significantly). Window four showed similarity to the first two windows, with all but one speaker having a higher COG for uvulars, two significantly so. Overall, three speakers showed a significant difference in COG in at least one of the windows, two of these consistently in every window.

Results for plain versus labialized fricatives support the finding of the previous section that labialization lowers COG, meaning that labialized consonants have more acoustic energy concentrated in lower frequencies. Three speakers showed consistently lower COG for /kw/ than /k/, one showed consistently higher COG for /kw/, and one showed lower COG for /kw/ in the first three windows and higher in the fourth. The results were similar but not identical for /q/ versus /qw/. Three speakers showed a consistently equal or lower COG for /qw/, one showed lower COG in all windows except window three, and one showed consistently higher COG for /qw/. Overall, window three showed the smallest p-values for distinguishing plain from labialized fricatives. FD and JM showed unexpectedly low values for COG. Most likely this was due to an asymmetry in the frequency response for the microphones used to record these speakers, resulting in these recordings being artificially skewed towards lower frequencies. Figure 7 shows different y-axes for each speaker to adjust for the high between-speaker variability.

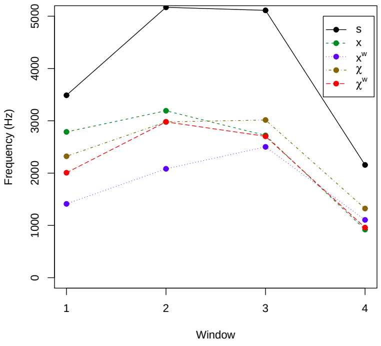
Skew for fricatives is shown in Figure 8. Three speakers had higher skew for velars in all windows (two significantly in all windows), one speaker had higher skew for velars in



(a) Speaker AJ (b) Speaker FD

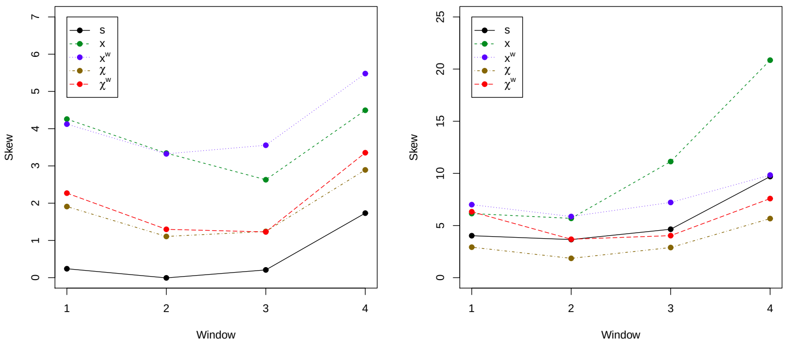


(c) Speaker HD (d) Speaker JM

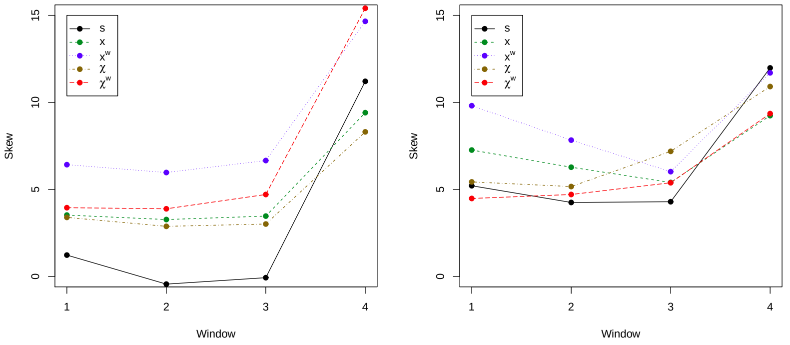


(e) Speaker WM

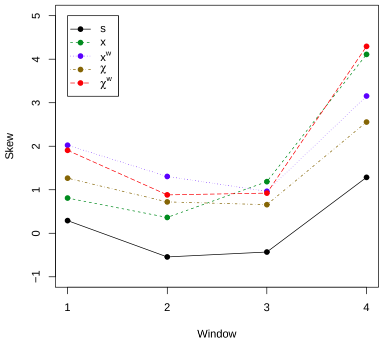
Figure 7: Fricative COG



(a) Speaker AJ (b) Speaker FD

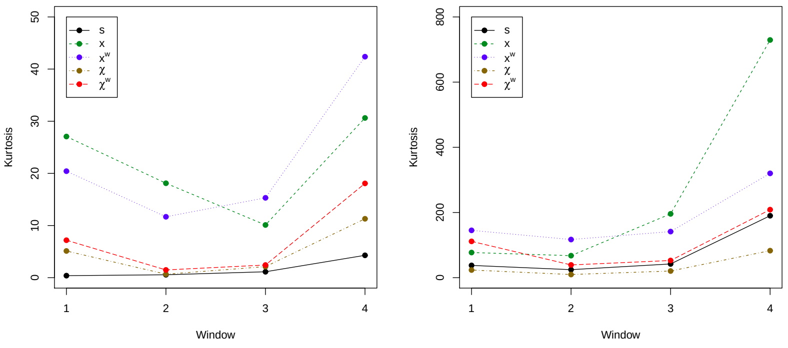


(c) Speaker HD (d) Speaker JM



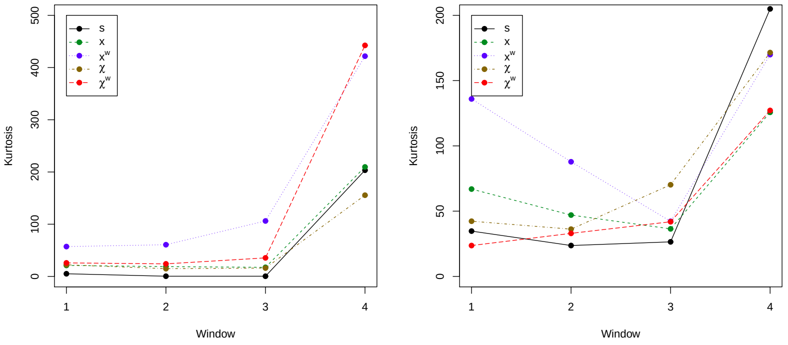
(e) Speaker WM

Figure 8: Fricative Skew



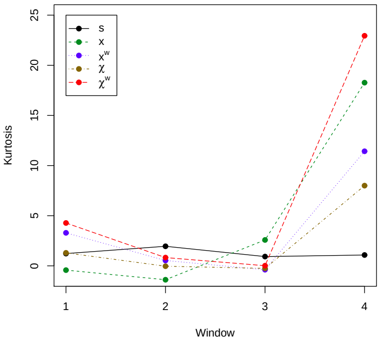
(a) Speaker AJ

(b) Speaker FD



(c) Speaker HD

(d) Speaker JM



(e) Speaker WM

Figure 9: Fricative Kurtosis

the first two windows and higher skew for uvulars in the last two, and one speaker had higher skew for uvulars in the first two windows, and higher skew for velars in the last two. Higher skew for velars means that velars had acoustic energy concentrated in lower frequencies than uvulars. Plain versus labialized fricatives tended to show higher skew on labialized phonemes, which, as in the COG results, shows that labialized phonemes have energy concentrated more in lower frequencies. As in Figure 7, y-axes vary by speaker because of the differences in the original recordings.

As for plosive bursts, kurtosis for fricatives seems to be lower for uvulars than for velars, as can be seen in Figure 9 (y-axes are different for each speaker). Speakers AJ and FD had lower kurtosis for uvulars in all windows, AJ significantly in all but window one, and FD significantly in all windows. HD had lower kurtosis for uvulars in all but window one. JM and WM had lower kurtosis in two windows (one and two for JM, three and four for WM). Labialized phonemes tended to have higher kurtosis, as in plosive bursts. If the amount of acoustic energy is the same for plain and labialized phonemes, higher kurtosis for labialized phonemes is expected, since the same amount of energy will be concentrated in a smaller frequency range. Speaker HD consistently showed higher kurtosis for labialized phonemes, with /x^{w}/ higher than /x/ and /χ^{w}/ higher than /χ/ in all windows. Speakers FD and WM consistently showed higher kurtosis for labialized uvulars. Speaker JM showed consistently higher kurtosis for labialized velars, but consistently lower kurtosis for labialized uvulars. Speaker AJ had higher kurtosis for labialized uvulars but no discernible relationship for velars.

3.3 Formant Transitions & Locus equations

Formant transitions were investigated by measuring F_2 and F_3 at two points: (1) the onset of the vowel (the first 23.3ms of the vowel interval), and (2) the target quality of the vowel (a 23.3ms window centered on the midpoint of the vowel interval). Locus equations were derived from these measurements by plotting F_2 onset against F_2 target and performing linear regression analysis. As discussed in Sussman et al. (1993), locus equations differ between places of articulation but remain for the most part constant across speakers and vowel qualities within phoneme groups. Locus equations differ in both the slope of the

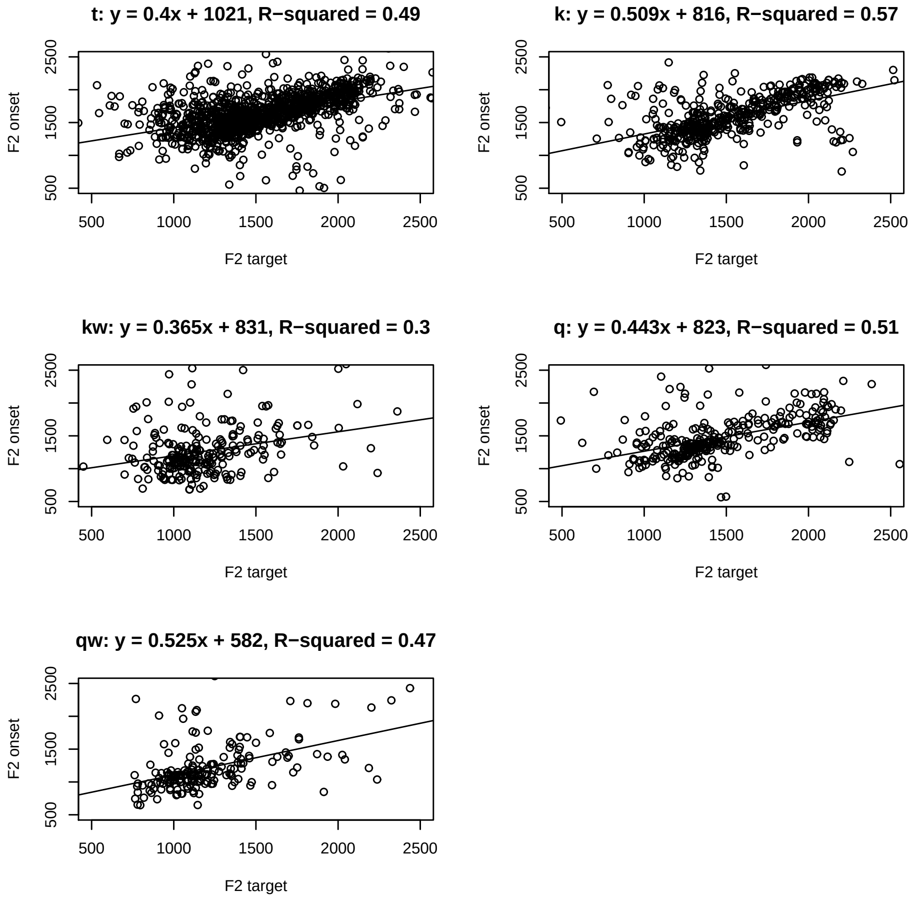


Figure 10: Locus Equations

regression line and the y-intercept of the line, and represent the formant transitions from the onset of the vowel to the midpoint or target of the vowel. Table 7 and Figure 10 show the locus equations for each phoneme, with tokens pooled across speakers. The decision to pool speakers was made because locus equations represent regression lines calculated from scatter plots of related measurements: F_2 at vowel onset (F_2 onset) and the steady-state target of the vowel (F_2 target). While frequency ranges in the recordings vary from speaker to speaker, the relationship between F_2 onset and F_2 target may still remain constant.

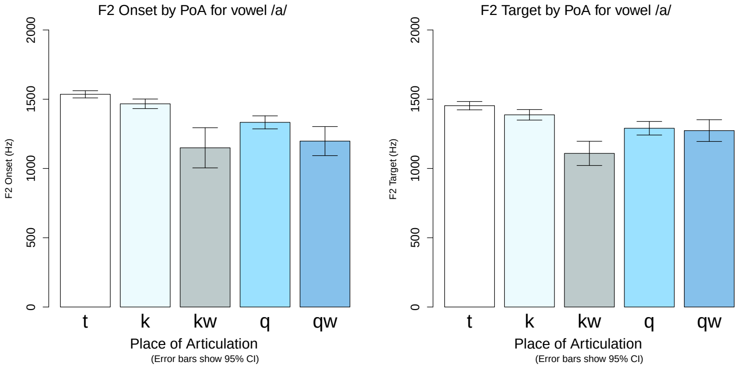
Figures 11–12 show F_2 and F_3 values at vowel onset and vowel target for vowels /a/ and /aː/. For /a/ and /aː/, all phoneme groups showed significantly different F_2 onset values (for /a/, alveolars = 1536Hz, velars = 1467Hz, uvulars = 1333Hz; for /aː/, alveolars = 1511Hz, velars = 1436Hz, uvulars = 1356Hz). Velars had significantly higher ($p < .05$) F_2

Place	Slope	Y-Intercept	R^2
t	0.400	1021	0.49
k	0.509	816	0.57
kw	0.365	831	0.30
q	0.443	823	0.51
qw	0.525	582	0.47

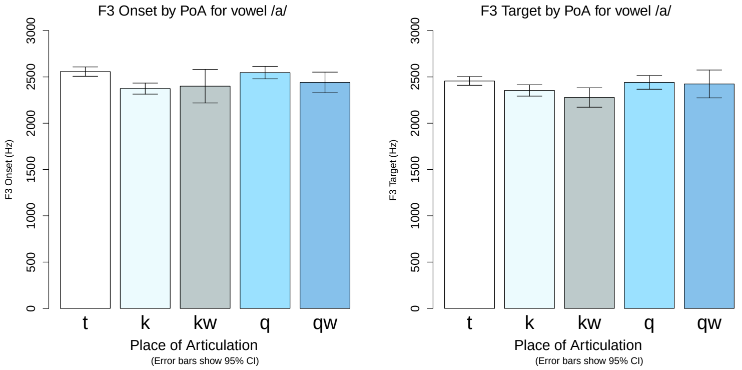
Table 7: Slope, y-intercept, and R^2 for locus equations

onset values than uvulars. Velars and uvulars also showed a significant difference between F_3 onset values, with uvulars having significantly higher ($p < .01$) F_3 onset values than velars (for /a/, velars = 2374Hz, uvulars = 2546Hz; for /aː/, velars = 2462Hz, uvulars = 2632Hz). The fact that F_2 and F_3 are separated by a significantly wider margin in uvulars than velars is likely an important acoustic cue differentiating the two places of articulation. Coarticulation seems to extend somewhat into the steady-state vowel, though for the most part the values of F_2 target are in line with the findings for F_2 values for /a/ in Maddieson et al. (2001).

Alveolars had the highest y-intercept of all phoneme groups, indicating the least coarticulation. The slope for alveolars was the lowest of the basic places of articulation (excluding the labialized sets as separate groups), indicating a lower degree of coarticulation than for other phonemes. Velars have the steepest slope of the plain phoneme groups, indicating the highest degree of coarticulation, and a lower y-intercept than alveolars. Labialized velars have a less steep slope than plain velars, indicating less coarticulation at F_2 onset or more at F_2 target, and a similar (though slightly higher) y-intercept. Plain uvulars show a slope intermediate between alveolars and velars, indicating an intermediate degree of coarticulation (more than alveolars, less than velars), and a y-intercept similar to velars and significantly lower than alveolars. Labialized uvulars show the opposite trend regarding plain uvulars than labialized velars do to plain velars. Labialized uvulars have a steeper slope and a significantly lower y-intercept than plain uvulars. Previous studies involving locus equations have not looked at labialized phonemes, and so it is unclear if labialization exerts a consistent effect on the slope and y-intercept of locus equations, or if so, what that effect is.

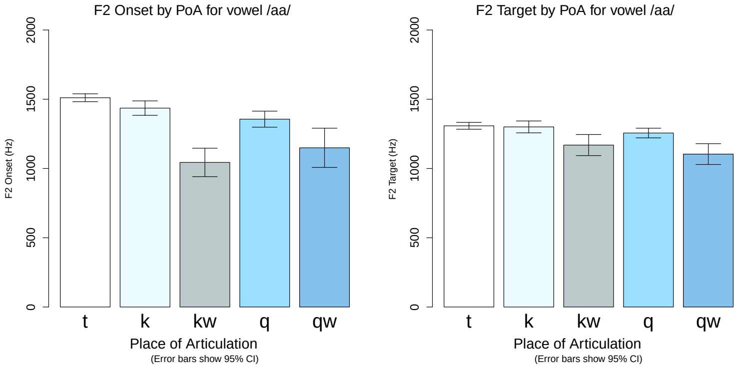


(a) F_2 onset (b) F_2 target

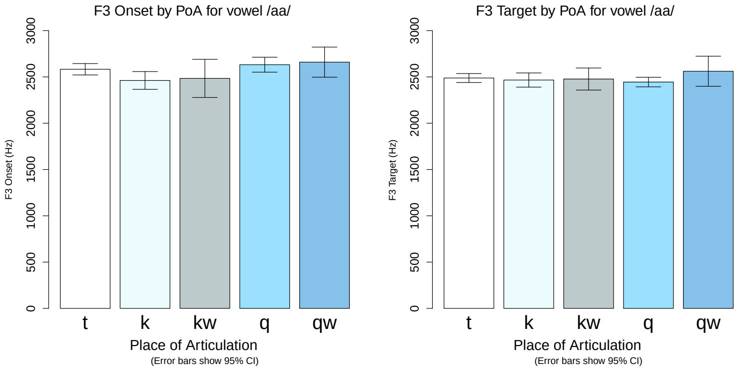


(c) F_3 onset (d) F_3 target

Figure 11: Formant onset and target results for vowel /a/



(a) F_2 onset (b) F_2 target



(c) F_3 onset (d) F_3 target

Figure 12: Formant onset and target results for vowel /aː/

4 Discussion

This study found that one clear acoustic cue distinguishing velars and uvulars is the separation of F_2 and F_3: uvulars have a significantly lower F_2 and significantly higher F_3 in following vowels. One of the more consistent results in looking at aperiodic noise was that labialization lowers COG of plosive burst and fricative spectra. Across speakers, labialized velars and labialized uvulars had a lower COG than their plain counterparts, especially in plosive bursts but also consistently in fricatives. The consistent lowering of COG for labialized phonemes supports the finding of Suh (2008) that labialization lowers mean burst energy for plosives. Uvulars generally had a higher COG and lower skew and kurtosis than velars, meaning that the energy distribution in aperiodic noise associated with uvulars (plosive bursts and fricative intervals) is less concentrated than velars, and centered in higher frequency ranges. The locus equations for uvulars were similar to those for velars, with velars having a steeper slope. Significant differences can be seen in F_2 onset, both among the different places of articulation and between plain and labialized consonants. F_2 onset decreases as the place of articulation moves further back in the oral cavity, i.e., as the resonance chamber increases in length. F_2 onset also decreases with labialization, which also increases the length of the oral cavity by moving the lips forward.

Higher COG and lower skew for uvulars indicates that uvulars have acoustic energy concentrated in higher frequency ranges than velars in Tlingit. Frequency decreases with the length of the resonance chamber, and thus the finding that uvulars seem to have a higher COG than velars is unexpected. Since the portion of the oral cavity in front of the constriction functions as a resonator for the produced sound, the backer the point of constriction, the lower the resonant frequencies should be. The lowering of F_2 with backer points of constriction can be seen in the lowering of F_2 in vowels as the tongue body moves toward the back of the mouth. It may be that the so-called “gutturals” have unexpected acoustic properties because their points of articulation are so far back in the vocal tract and do not involve constriction near the hard palate; the acoustics of uvulars and other gutturals is still understudied (though see Alwan 1989, McCarthy 1994). The finding of lower kurtosis for uvulars means that acoustic energy is spread out over a wider range of frequencies than in velars, or that uvulars have a more diffuse spectral peak compared to velars.

Lower COG and higher skew for labialized velars indicates that labialized velars have energy concentrated in the lower frequencies, with a long sloping tail in higher frequencies. Lower mean frequency for labialized phonemes is expected, since frequency decreases as the length of the resonating chamber increases, and pursing the lips in labialization extends the length of the oral cavity. FD and JM, who did not show significantly lower COGs for velars and uvulars, respectively, had the lowest COGs of all the speakers, in fact lower than expected. These unexpectedly low COG measurements may mean that the microphones used to record these speakers did not have sufficient sensitivity in higher frequencies, and that the uneven frequency response of the microphones resulted in artificially low COGs for plain velars and uvulars, which have higher COGs, and relatively accurate COGs for labialized velars and uvulars, which have lower COGs, thus diminishing the difference between these groups such that it was not always possible to see a significant difference between the two. Higher values for kurtosis indicate a more peaked distribution, i.e., more concentrated acoustic energy, for labialized consonants as compared to plain consonants.

Formants transitions are generally what would be expected, with the exception of the relatively flat F_3 trajectory for velars. The slope of the regression line of the locus equation for each phoneme group indicates the amount of coarticulation for that group, if target values for vowel formant frequencies are assumed to be relatively invariant. Coarticulation is assumed to arise from gestural overlap (Hardcastle & Hewlett 2006). This movement away from the ideal or prototypical target frequency should theoretically be restricted to the portion of the vowel adjacent to the preceding segment, represented in this case by F_2 onset. However, as detailed in Flemming (1997), gestural undershoot often results in effective coarticulation through the target of the vowel, and thus in many languages coarticulation effects are seen even at the vowel midpoint. Sussman et al. (1993) shows that slopes and y-intercepts of locus equations (and thus the degree of coarticulation at a given point of articulation), like absolute formant frequencies of various vowels, vary across languages, and thus learning how these measures correspond to phoneme categories is a part of learning the phonetics of a given language. Such variation suggests that while there are universal trends (e.g., all languages studied in Sussman et al. 1993 had higher slopes, and thus greater coarticulation, for velars than for alveolars), to some extent phonetics is language-specific and learned.

The differentiation of F_2 and F_3 among different places of articulation is expected, since these differences across phoneme groups is an acoustic effect of coarticulation and serve to differentiate the phonemes. Lower F_2 for labialized phonemes is also expected, since F_2 decreases as the resonance chamber lengthens, and rounding and pursing the lips in labialization increases the length of the oral cavity. If coarticulation extends significantly into the steady-state vowel, we should expect a similar slope for the locus equations of labialized phonemes as for their plain counterparts, and if coarticulation is restricted to the vowel onset, we should expect a steeper slope for the locus equations of labialized phonemes relative to their plain counterparts. A differing extent of CV coarticulation in velars versus uvulars might explain the difference in labialization's effect on locus equation slope in velars versus uvulars. For instance, the shallower slope for labialized velars could result from either less coarticulation at F_2 onset or greater coarticulation at F_2 target.

Extraneous background noise in the recordings used herein likely contributed to the inconsistent results for spectral moment analysis, since this type of analysis looks at aperiodic noise, which in principle cannot accurately be distinguished acoustically from background noise. It could be that a future study using discriminant analysis may find more consistent differences among phoneme groups considering COG, skew, and kurtosis simultaneously (as was done in Forrest et al. 1988), though individual differences in these measures were not always significant and consistent from speaker to speaker. Since a periodic signal is much less affected by such background noise, formant analysis yielded clear results, showing that uvulars have a significantly wider gap in F_2 and F_3 of following vowels than do velars.

The acoustic properties of uvulars have ramifications for both phonological theory in the abstract and phonological processes active in the Tlingit language. The dispersion theory of contrast (Flemming 2002) argues that markedness is a property of contrasts rather than of segments, and uses this theory to account for a number of phonological universals, such as the correlation of [+back] and [+round] in vowels, by referencing acoustic properties and perception. Dispersion theory predicts that these features co-occur because backing and rounding both lower F_2, and thus applying both of these features maximizes contrast with non-back and unround segments. Since backing likewise lowers F_2 in post-consonantal vowels, as in Figures 11 & 12, the incorporation of acoustic features in phonology may also help explain the fact that dorsals such as velars and uvulars are more often contrastively labialized. The acoustic properties of uvulars may also be relevant for loanword adaptation in Tlingit. Labials in foreign loanwords are mapped to labialized velars, since Tlingit lacks labials. Since labialized velars and labialized uvulars represent equally faithful adaptations in terms of preserving the feature [+labial], it may be that the acoustic properties of the phoneme groups play a role in determine how foreign borrowings are mapped to native phonemes.

5 Conclusion

This paper has looked at the acoustics of uvulars in Tlingit in order to determine what acoustic properties are associated with the uvular place of articulation and how uvulars are distinguished from other places of articulation. Spectral moments were analyzed for plosive bursts and fricative intervals, and formant transitions and locus equations were analyzed for post-consonantal vowels, using 2326 CV tokens from five Tlingit speakers. Because the recordings used in this study were primarily digitizations of old reel-to-reel recordings not intended for phonetic analysis, the results obtained from analysis of aperiodic noise are necessarily preliminary, but clear results were obtained for formant transitions. This study found that uvulars consistently showed a significantly greater separation of F_2 and F_3 than velars. Results for plosive burst analysis showed that, supporting Suh (2008), labialization lowers COG for plosive bursts. It also appears that plosive labialization is associated with higher skew and also higher kurtosis (indicating a more peaked distribution of frequencies). Fricative results also generally showed a lower COG for labialized phonemes as well as higher skew and kurtosis. Uvulars seemed to show a higher COG and lower skew and kurtosis than velars. As expected from the work reported in Sussman et al. (1993), locus equations, which incorporate the relationship between F_2 at the onset and midpoint a post-consonantal vowel, provide a stable cue for place of articulation. As expected, alveolars showed a relatively shallow slope (indicating no great degree of coarticulation) and velars a relatively steep slope (indicating a high degree of coarticulation). Uvulars turned out to be intermediate between these two, with a slope and y-intercept between alveolars and velars.

Bibliography

Alwan, Abeer. 1989. Perceptual cues for place of articulation for the voiced uvular and pharyngeal consonants. *Journal of the Acoustical Society of America* 86. 549–556.

Behrens, Susan J. & Sheila E. Blumstein. 1988. Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *Journal of Phonetics* 16. 295–298.

Boersma, Paul & David Weenick. 2011. Praat: Doing phonetics by computer. `http://www.praat.org/`. [Computer program].

Flemming, Edward S. 1997. Phonetic optimization: Compromise in speech production. In *University of Maryland Working Papers in Linguistics 5: Selected phonology papers from H-OT-97*, 72–91.

Flemming, Edward S. 2002. *Auditory representations in phonology.* New York: Routledge.

Forrest, Karen, Gary Weismer, Paul Milenkovic & Ronald N. Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America* 84. 115–123.

Fruchter, David & Harvey M. Sussman. 1997. The perceptual relevance of locus equations. *Journal of the Acoustical Society of America* 102. 2997–3008.

Gordon, Matthew, Paul Barthmaier & Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32. 141–174.

Hardcastle, William J. & Nigel Hewlett. 2006. *Coarticulation: Theory, data and techniques.* Cambridge: Cambridge University Press.

Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to speech analysis.* Cambridge, MA: MIT Press.

Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108. 1252–1263.

Leer, Jeff, Doug Hitch & John Ritter. 2001. *Interior Tlingit noun dictionary.* Yukon Native Language Centre.

Maddieson, Ian, Caroline L. Smith & Nicola Bessell. 2001. Aspects of the phonetics of Tlingit. *Anthropological Linguistics* 43. 135–176.

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In Patricia A. Keating (ed.), *Papers in Laboratory Phonology III: Phonological structure and phonetic form*, 191–233. Cambridge: Cambridge University Press.

R Development Core Team. 2011. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing Vienna, Austria. `http://www.R-project.org/`. ISBN 3-900051-07-0.

Shosted, Ryan. 2011. Towards a glottalic theory of Mayan. In Heriberto Avelino (ed.), *New perspectives in Mayan linguistics*, Newcastle upon Tyne: Cambridge Scholar's Publishing.

Stevens, Kenneth N. & Sheila E. Blumstein. 1978. Invariant cues of place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64. 1358–1368.

Stevens, Kenneth N., Samuel Jay Keyser & Haruko Kawasaki. 1986. Toward a phonetic and phonological theory of redundant features. In Joseph S. Perkell & Dennis H. Klatt (eds.), *Invariance and variability in speech processes*, 426–449. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Suh, Yunju. 2008. The effect of secondary labialization on stop burst spectra. *Journal of the Acoustical Society of America* 124. 2556. Abstract.

Sundara, Megha. 2005. Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *Journal of the Acoustical Society of America* 118. 1026–1037.

Sussman, Harvey M., Kathryn A. Hoemeke & Farhan S. Ahmed. 1993. A cross-linguistic description of locus equations as a phonetic descriptor for place of articulation. *Journal of the Acoustical Society of America* 94. 1256–1268.

Waibel, Alexander, Toshiyuki Hanazawa, Geoffrey Hinton, Kiyohiro Shikano & Kevin J. Lang. 1989. Phoneme recognition using time-delay neural networks. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37. 328–339.

Walley, Amanda C. & Thomas D. Carrell. 1983. Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants. *Journal of the Acoustical Society of America* 73. 1011–1022.

Werker, Janet F. & Richard C. Tees. 1984. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75. 1866–1878.

Żygis, Marzena & Jaye Padgett. 2010. A perceptual study of Polish fricatives, and its implications for historical sound change. *Journal of Phonetics* 38. 207–226.