# Acoustic characteristics of Deg Xinag fricatives

Sharon Hargus, Gina-Anne Levow, Richard Wright
Department of Linguistics, University of Washington
Box 352425, Seattle WA 98195-2425
sharon@uw.edu, levow@uw.edu, rawright@uw.eduw

## Abstract

This article presents an acoustic study of the contrast between seven fricatives in two positions (before [a], after [a]) using data from 8 speakers of Deg Xinag, an Athabaskan language of Alaska. An initial set of 52 measures of various properties of the fricative and adjacent [a] was narrowed to 13 non-correlated measures. Although no single measure distinguished all pairs of fricatives in either position, fricatives before [a] were generally differentiated by energy profile whereas differentiation of fricatives after [a] relied more heavily on formant transitions. In a subsequent experiment, statistical analyses were performed to help understand why **ɬ* and **θ*, which are distinct in Deg Xinag, have merged as /ɬ/ in the closely related language Koyukon. An additional experiment determined that one contentious Deg Xinag fricative, which has been variously transcribed as [χ] or [h], has more of the characteristics of /χ/ despite considerable inter-speaker variation.

## Keywords

fricative, acoustic, spectrum, formant transition, MFCC, Deg Xinag, Athabaskan

## 1 Acoustic differences between fricatives

The primary goal of this article is to describe the acoustic characteristics of the fricatives of Deg Xinag (DX) (ISO 639-3 ing), an Athabaskan (a.k.a. Dene, Hargus 2013) language with seven contrastive fricative places of articulation. A second, related goal of this article is to determine which acoustic measures distinguish the DX fricatives from each other, testing a battery of measures within a single language.

In deciding on our set of acoustic measures, we found ourselves in agreement with Spinu & Lilley 2016 that there is little consensus on the measures (and statistical tests) for investigating differences between fricatives. There are at least four different types of measures. Three of these measure properties of the fricative itself (spectral, intensity, duration), and another type of measure assesses effects on the quality of an adjacent vowel (formant transitions). A brief survey of the literature on the acoustics of fricatives reveals selective rather than comprehensive use of these four kinds of measures within a single study. For example, Strevens 1960 reports only intensity, Jesus & Shadle 2002 various amplitude measures and Elmazouzi et al. 2014 duration and formant transitions for sets of fricatives. Other studies perform an LPC analysis (Tronnier & Dantsuji 1993, Johnson 1997) or FFT analysis (Nartey 1982), sometimes in conjunction with other measures (duration, amplitude, LPC, Behrens & Blumstein 1988; FFT, spectral slope, Bladon et al. 1987). Many acoustic studies of fricatives measure some or all of the first four spectral moments---center of gravity, standard deviation, skewness (asymmetry in frequency distribution), and kurtosis (peakedness) (Forrest et al. 1988, Johnson 1993, Munson 2001). Some studies employ spectral moments in conjunction with other measures: spectral moments, intensity (Svantesson 1986); spectral moments, formants, intensity, f0 (Norlin 1983); spectral moments, duration, amplitude, frequency measures (van Dommelen 2003); spectral moments, formant frequency and amplitude measures, spectral peak frequency (Jassem 1968); spectral moments, spectral peak location, locus equations, F2 onset, amplitude measures, duration

(Jongman et al. 2000); spectral moments, duration, formant transitions, spectral slopes, f0 of following vowel, amplitude measures, harmonics-to-noise ratio, energy below 500 Hz (Maniwa et al. 2009); and spectral moments and cepstral coefficients (Spinu & Lilley 2016). Often fricatives are sampled at more than one location, not just midpoint. For example, Reidy 2016 measured a psychoacoustic version of peak frequency at 17 evenly spaced points across some of the fricatives of English and Japanese.

## 2 Acoustic differences between Athabaskan fricatives

Since Deg Xinag is an Athabaskan language, next we survey acoustic studies of fricatives in related languages. Not surprisingly, we also find a variety of measures in this subfield. Table 1 provides the fricative inventories of each language along with study measures. Two of these languages, Navajo and Western Apache, are relatively closely related, both Apachean.

Table 1. Quantitative acoustic studies of other Athabaskan languages

| | | | |
|---|---|---|---|
| Hupa | Gordon 1995, Gordon et al. 2002 | /s ʃ ɬ x xʷ x̫ʷ/[1] | duration, center of gravity, formant transitions |
| Western Apache | Gordon et al. 2001, Gordon et al. 2002 | /s ʃ ɬ x/ | duration, center of gravity |
| Navajo[2] | Nartey 1982 | /s z ɬ ʃ ʒ x ɣ/ | FFT generated (22 bands) |
| Navajo | McDonough 2003 | /s z ɬ l ʃ ʒ x ɣ/ | center of gravity, duration, points (uncorrected), standard deviation[3], skew[4] |

Some studies found that the fricatives were distinguishable in one or more spectral moments. /s/ was found to have the highest center of gravity (centroid) for all studies which investigated spectral moments. In Hupa, Gordon 1995 noted that /s/ had the highest centroid, /ɬ ʃ/ intermediate, and /x xʷ x̫ʷ/ the lowest (only descriptive statistics were presented). In Western Apache, Gordon et al. 2001 found significant differences in centroid among the voiceless fricatives, with /s/ > /ʃ/ > /ɬ/ > /x/. In Navajo, McDonough 2003 found /s/ had a significantly higher centroid than /ʃ/, which was in turn higher than /ɬ/. For the other spectral moments, McDonough 2003 found that /ʃ/ and /ɬ/ had significantly greater skewness than /s/ and /ɬ/ had significantly greater standard deviation than /ʃ/. There were no differences in kurtosis for these fricatives.

The one study (Hupa, Gordon 1995) that investigated formant transitions found that formant transitions for the vowel before /x xʷ x̫ʷ/ did not distinguish these fricatives from each other. Lowering of both F1 and F2 before all the velar fricatives was observed.

Duration was investigated for Western Apache and Hupa and did not appear to play much role in distinguishing the fricatives of those languages. For Western Apache, Gordon et al. 2001 noted that there were no significant differences in duration among the fricatives, except that [ɬ] was longer than the other fricatives for the three female speakers in their study. For Hupa, Gordon et al. 2002 noted that "/s/ was the longest fricative for the female speaker but is similar in length to /ʃ/ for the male speaker" (descriptive statistics only). For the Hupa male speaker, /ʃ/ was the longest fricative.

---

[1] /x̫ʷ/ = "/x̹ʷ/ more rounded" (< *ʃ/); cf. /xʷ/ = slightly rounded

[2] Iskarous et al. 2012 is an ultrasound study of Navajo [x] with qualitative spectrographic observations.

[3] for /s ʃ ɬ x ɣ w/

[4] for /s ʃ ɬ x ɣ w/

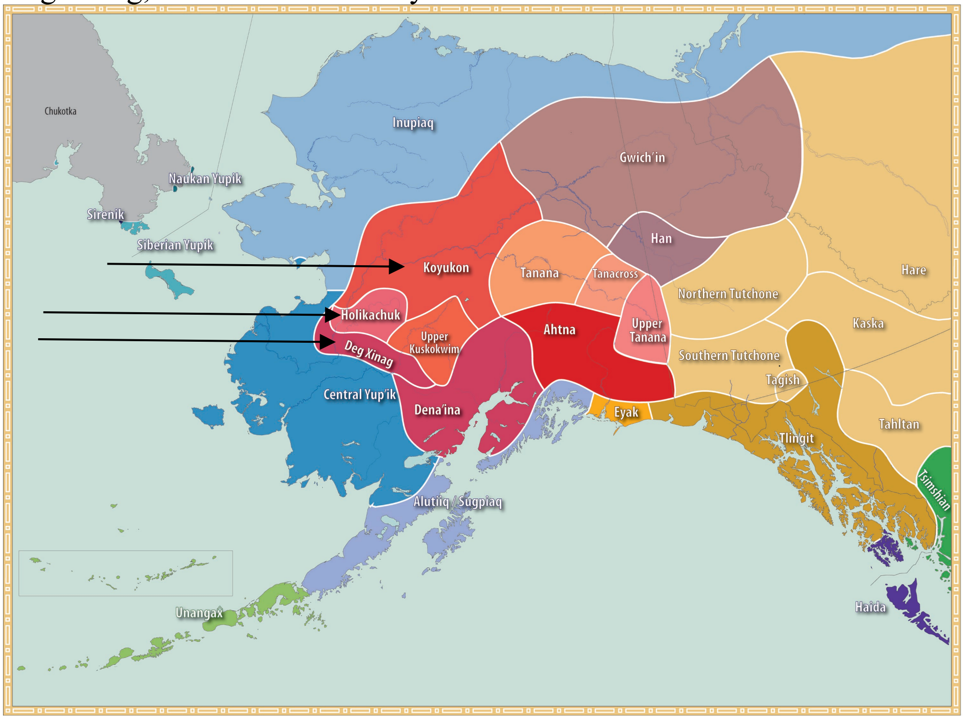
Intensity results are available for Hupa and Navajo, but only as qualitative information. Gordon 1995 noted that Hupa /ɬ/ had less intensity than /ʃ/ for one of the speakers in that study. Similar observations were made about /ʃ/ vs. /ɬ/ in Navajo by both McDonough 2003 and Nartey 1982: "the laterals seemed to have less amplitude and a broader energy spectrum" (McDonough); /ʃ/ was "the most intense of all the Navajo fricatives" (Nartey).

As with the studies discussed in §1, it has been difficult to determine genuine cross-linguistic acoustic differences, if any, between the fricatives of Athabaskan languages due to the variety of methods employed by different studies.

## 3 Deg Xinag

Deg Xinag is an Athabaskan language spoken in the western interior of Alaska along the lower Yukon River. The language has also been known as Ingalik (Osgood 1940) and Deg Hit'an (Krauss 1974). In one approach to Athabaskan-internal classification (Goddard 1996), Deg Xinag, along with upriver Koyukon and Holikachuk (see Figure 1), form a subgrouping, Koyukon-Ingalik, of the Athabaskan languages.

Figure 1. Indigenous peoples and languages of Alaska (Krauss et al. 2011). Arrows point to Deg Xinag, Holikachuk and Koyukon.



Two dialects of Deg Xinag are recognized: Yukon and Kuskokwim (Krauss 1962, Kari 1974). Data from the Yukon dialect, which is the traditional language of the villages of Anvik and Shageluk, are discussed in this article. The language is nearly extinct.

The inventory of Deg Xinag consonant phonemes is given in Table 2 (syllable-initial) and Table 3 (syllable-final, i.e. before a consonant or word-finally). Parenthesized segments are marginal, found only in loan words. In Table 2 note the three-way syllable-initial contrast between voiceless unaspirated, voiceless aspirated and ejective stops and affricates at most places of articulation.[5] Syllable-finally, Deg Xinag, like other Alaskan Athabaskan languages, contrasts voiced and voiceless stops and affricates, as seen in Table 3.[6]

---

[5] The basic inventory of consonantal contrasts was established by Krauss 1962. We have not performed an articulatory study of Deg Xinag. The symbols in Table 2 and Table 3 are simply modern versions of Krauss's symbols, clarified by some acoustic studies (Hargus 2008, Hargus 2010, Hargus 2011, Hargus 2016).

[6] Voicing contrasts developed in syllable-final position when word-final vowels were lost (Krauss 1962).

Table 2. Syllable-initial consonants

| | bilabial | labio-dental | inter-dental | alveo-lar | retro-flex | palato-alveolar | velar | uvular | glottal |
|---|---|---|---|---|---|---|---|---|---|
| plosive | (p pʰ) | | | t tʰ | | | k kʰ | q qʰ | ʔ |
| ejective | | | | t’ | | | k’ | q’ | |
| affricate | | | tθ tθʰ | ʦ ʦʰ | ʈʂ ʈʂʰ | ʧ ʧʰ | | | |
| affricate, lateral release | | | | tɬ tɬʰ | | | | | |
| ejective affricate | | | tθ’ | ʦ’ | ʈʂ’ | ʧ’ | | | |
| ejective affricate, lateral release | | | | tɬ’ | | | | | |
| fricative | | v | θ ð | s z | ʂ ʐ | (ʃ) | | χ ʁ | h |
| lateral fricative | | | | ɬ | | | | | |
| nasal | m | | | n | | | | | |
| approximant | | | | | | j | ŋ | | |
| lateral approximant | | | | l | | | | | |

Table 3. Syllable-final consonants

| | bilabial | labio-dental | inter-dental | alveolar | retroflex | palato-alveolar | velar | uvular | glottal |
|---|---|---|---|---|---|---|---|---|---|
| plosive | (p) | | | t d | | | k g | q ɢ | ʔ |
| affricate | | | tθ dð | ʦ ʣ | ʈʂ ɖʐ | ʧ | | | |
| affricate, lateral release | | | | tɬ dl | | | | | |
| fricative | | v | θ ð | s z | ʂ ʐ | | | χ ʁ | |
| lateral fricative | | | | ɬ | | | | | |
| nasal | m̥ m m̰ | | | n̥ n n̰ | | | ŋ̊ ŋ ŋ̃ | | |
| approximant | | | | | | j̊ j j̃ | | | |
| lateral approximant | | | | l | | | | | |

The laryngealized nasals and semi-vowels are restricted to word-final position. The voiceless nasals and semi-vowels are also found only word-finally, with the exception of [n̥], which also occurs word-internally (syllable-finally). /j̊/, the voiceless palatal approximant, can be more narrowly transcribed [ç].

The Deg Xinag consonant inventory is relatively large for an Athabaskan language. As noted by Krauss 1962, one reason for this is that Deg Xinag has not undergone some of the mergers that have affected other languages of the family. For example, Koyukon-Ingalik interdental fricatives and affricates (_*tθ *tθ$^h$ *tθ’ *θ *ð_[7]) and lateral fricatives and affricates (_*tɬ *tɬ$^h$ *tɬ’ *ɬ *l_) have merged as laterals in Koyukon (Krauss & Golla 1981, Krauss 2000). Also, most other Athabaskan languages have only one or two sibilant places for fricatives and affricates, merging the historic retroflexes with other sibilants. For example, Witsuwit’en (Hargus 2007) has a single set of sibilant obstruents, /ts ts$^h$ ts’ s z/, merged reflexes of Proto-Athabaskan alveolar (_*ts *ts$^h$ *ts’ *s *z_), palato-alveolar (_*ʧ *ʧ$^h$ *ʧ’ *ʃ *ʒ_) and retroflex (_*ʈʂ *ʈʂ$^h$ *ʈʂ’_) consonants.[8]

Syllable-initially in Deg Xinag, there are seven voiceless fricatives: /θ s ʂ ɬ ʃ χ h/. Syllable-finally, there are also seven voiceless fricatives /θ s ʂ ɬ ʃ χ/ and [ç] as allophone of /j̊/. With this large inventory of fricatives, Deg Xinag is thus potentially a good test case for identifying a set of measures for distinguishing fricatives.

## 4 Research questions

This study is motivated by four basic research questions. (1) Which acoustic measures best distinguish the seven voiceless Deg Xinag fricatives from each other? (2) How do Deg Xinag fricatives compare with fricatives from other languages, particularly other Athabaskan languages? Research questions 1-2 are addressed in experiment 1 (§5).

(3) Can analyses from our study shed light on the _*θ, *ɬ_ > _/ɬ/_ merger found in closely related Koyukon? Research question 3 is addressed in experiment 2 (§6).

(4) Deg Xinag field linguists have had difficulty differentiating the uvular fricative [χ] from the glottal fricative [h] in unstressed contexts. Provisionally transcribing this fricative as [“x”], can results from experiment 1 for identifying [χ] and [h] be used to resolve the place of articulation of “x”? Research question 4 is addressed in experiment 3 (§7).

## 5 Experiment 1: Effect of place on fricative characteristics

### *5.1 Materials and methods*

#### *5.1.1 Recordings and word list*

The recordings used in this study were drawn from a larger corpus. Recordings were made in the field from the speech of eight Deg Xinag native speakers (5 female, 3 male) using a professional CD field recorder (Marantz CDR300) and head mounted microphone (Shure SM-10). Participants read from a word list in a fixed random order. Since the distributions of DX fricatives /h/ and [ç] are both contextually restricted, depending on whether they follow or precede a vowel, two word lists were devised to illustrate both sets of fricatives, as shown in Table 4. In all cases the fricative of interest

---

[7] Interdental fricatives in Koyukon-Ingalik (and elsewhere) are actually from Proto-Athabaskan _*ts *ts$^h$ *ts’ *s *z_, respectively. Leer 1996 terms this sound change (part of) “The Great Northern Series Shift.” Leer 2005 provides a current reconstruction of the Proto-Athabaskan sound inventory.

[8] Krauss 1964 first identified the correspondences underlying the sounds which are now reconstructed as retroflex.

was adjacent to the vowel [a], since each of the fricatives occurs in this vocalic context.[9] Due to differences in type frequency of the Deg Xinag fricatives, it was not possible to control for the lefthand environment for pre-[a] fricatives nor the righthand environment for post-[a] fricatives.[10] This limitation of our word list meant that formant transitions could only be studied on one side of each fricative, the one adjacent to [a].

Table 4. Word list for Experiment 1

| | __a (prevocalic fricative) | | a___ (postvocalic fricative) |
|---|---|---|---|
| /θ/ | [“x”əˈθag ts’ən’] ‘(3SG) isn’t talking’ | /θ/ | [ŋəɬˈt’aθ] ‘fry it’ |
| /s/ | [kag nəˈsa] ‘I’m picking berries’ | /s/ | [kənasˈton’]~[kənasˈtoŋ’] ‘I’m full’ |
| /ɬ/ | [ɬats] ‘soil, mud, dirt’ | /ɬ/ | [ətəˈʧaɬ] ‘(3SG) is busy’ |
| /ʃ/ | [ˈʃaχəl] ‘sugar’ | /ʃ/ | [kaˈʃəʁ] ‘white person’ |
| /ʂ/ | [ʂaɬˈt’aŋ̥] ‘woman’ | /ʂ/ | [t’aʂ] ‘charcoal’ |
| /χ/ | [χaɬ] ‘pack’ | /χ/ | [tsaχ] ‘pitch’ |
| /h/ | [kag nəˈha] ‘(3SG) is picking berries’ | [ç] | [“x”əˈnaç] ‘(3SG) is talking’ |

Most of the fricatives on the word list occur in stressed syllables, except for pre-vocalic /ʂ/ and post-vocalic /s/. Also, all of the prevocalic fricatives are syllable initial and most of the postvocalic fricatives are syllable final, except for post-vocalic /ʃ/. These limitations were due to vocabulary restrictions.

If a speaker did not use a word on the list, another word with the segments of interest in the same position in the word was substituted on the word list recording (e.g. [ʂaˈqʰaj] ‘children’ substituted for [ʂaɬˈt’aŋ̥] ‘woman’ for some speakers as an example of the retroflex fricative). The data was recorded at a sampling rate of 44100 Hz (with a 16 bit quantization rate) and later downsampled to 22050 Hz for analysis. Four or five repetitions of the fricatives of interest were recorded from each speaker. Fricatives from words on the target wordlist which were shorter than 70 ms were discarded and suitable substitutions were made from other words recorded at the same time. A total of 478 words were analyzed in Experiment 1.

### *5.1.2 Measures*

First, we performed a visual, qualitative analysis of each fricative. A set of 512-point FFTs (= .02322 sec window length) were generated for each speaker for each repetition of the fricatives on the word list in Table 4. The FFTs were created with Multi-Speech (Kay Elemetrics Corp. 2004) (v. 3.1.2 and previous) from sound files sampled at 22050 Hz with no preemphasis and low smoothing. We inspected graphs of these FFTs to gain a view of the overall shape of each fricative and its relation to place of articulation. Additionally, we used the FFTs to gain some understanding of between-speaker variation. Each speaker’s average FFT for each fricative is included in §10.

Next, using Praat (Boersma & Weenink 2013) (v. 5.3.39 and previous), the recordings for each speaker were 300 Hz-high pass filtered to minimize effects on the intensity of frequency components of the spectrum from voicing which might extend from the vowel into the fricative. Time aligned transcriptions were created in which the

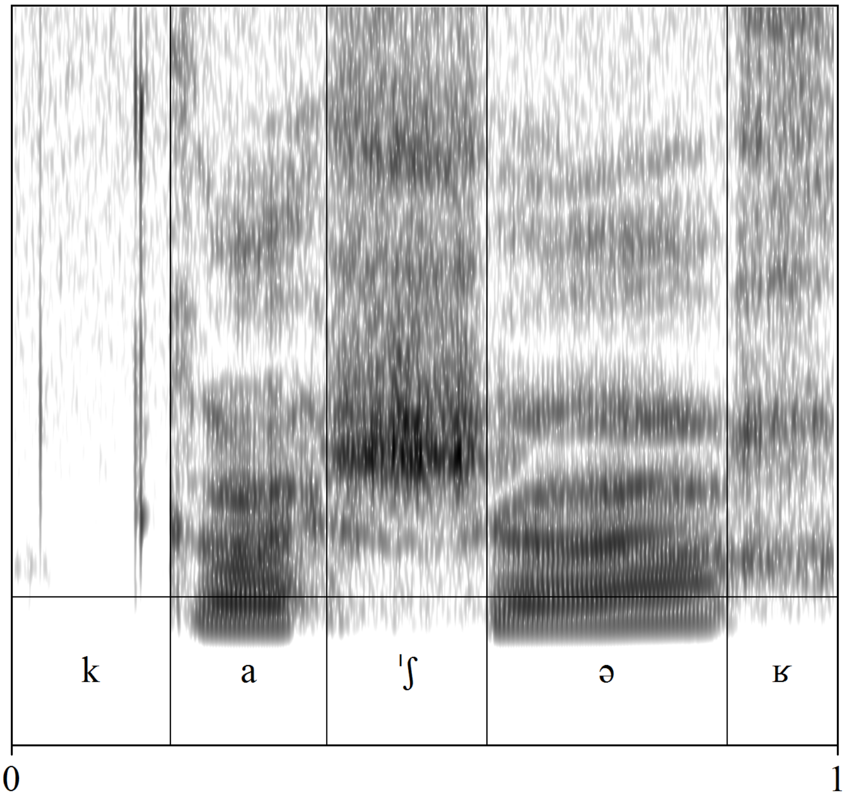
---

[9] The DX vowel inventory is /ə ʊ e o a/. See Hargus 2010 for information on spectral and durational differences between these vowels.

[10] For example, [ʃ] is found in only a handful of loan words, not all of which were used by our speakers, and occasionally as a fast speech variant of [s] before [ʧ]. [s] also occurs in relatively few morphemes in Deg Xinag, compared to [ʂ].

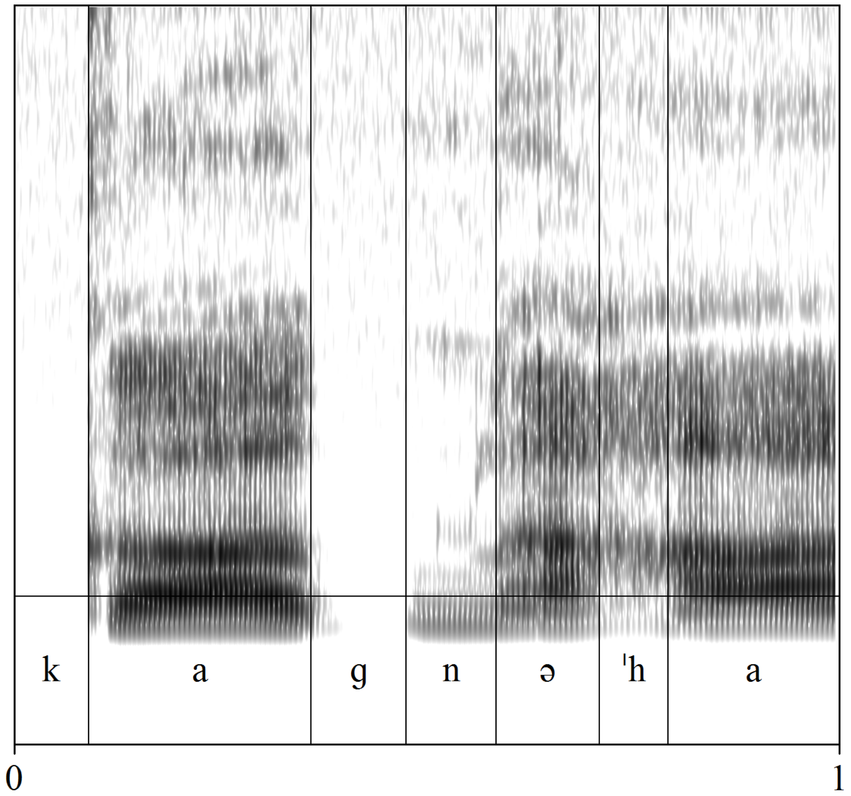
onset and offset of the fricative and adjacent vowel were marked. Vowel onset was considered the release burst following stops, or the offset of frication coincident with overall spectral change, following fricatives. Vowel onset was measured from the rightmost of multiple bursts if separated by 30 or more ms., otherwise the leftmost. Figure 2 shows an example of /aʃ/ following a release burst with multiple spikes, along with the typical segmentation of vowel onset following a fricative.

Figure 2. Spectrogram and segmentation of [kaˈʃəʁ] ‘white person’ (speaker KH). The vertical axis of the spectrogram displays 300-10000 Hz.



The glottal fricative [h] presented some unique challenges.  For some individuals [h] was frequently breathy voiced, as was the adjacent vowel, in which case [h] onset and offset were identified at dips in intensity.  An example is shown in Figure 3:

Figure 3. Spectrogram and segmentation of [kag nəˈha] ‘he’s picking berries’ (speaker LH). The vertical axis of the spectrogram displays 300-10000 Hz.



For vowel-fricative sequences, the onset of turbulence coincident with spectral change was considered the onset of the fricative, even in tokens where turbulence and periodicity overlapped. Word-final vowels often terminated in a period of breathy voice, which was excluded from the vowel segmentation at the point where intensity dropped abruptly. For vowels which were breathy throughout,[11] the offset of the vowel was considered to be the point where there was a marked change in amplitude, similar to decisions involving breathy [h]. For word-final fricatives, the offset of the fricative was marked where there was an abrupt drop in intensity.

Recall from §1 that there is a lack of consensus on the best measures to use to distinguish fricatives. We therefore employed a large set of measures initially and then eliminated correlated measures (5.2.1), roughly similar to the approach used by Garrelek 2020 in investigating voice quality in !Xóõ, another area where a variety of measures have been used in the literature.[12] Using a large number of measures allows us to approximate observations in some of the detailed qualitative descriptions of fricatives which have been around since the advent of the spectrograph (e.g. Harris 1958).

We used four types of measures: duration (both of the fricative and adjacent [a]), intensity (of the fricative), formant transitions (of [a]), and a series of spectral measures (of the fricative), all of which have been used in acoustic studies of fricatives

---

[11] Vowels before fricatives were often breathy voiced and turbulent. Some care was taken not to confuse vowel-final breathy turbulence with fricative turbulence. An example of this kind of segmentation can be seen above in Figure 2.

[12] Garrelek began with 18 measures and then performed a linear discriminant analysis to determine which measures performed the most work to distinguish the six phonation types.

(see §1). Duration, intensity and vowel formants were measured with Praat via custom scripts written by the second author. All fricative spectral measures were extracted using the openSMILE package (Eyben et al. 2010) with a custom configuration file written by the second author.

Fricative duration (*FricDur*) was measured following the segmentation procedures described in 5.1.1.

Fricative intensity (*Intens*) was measured at three points, 25%, 50% and 75% of the fricative, as is often done in studies of fricatives (e.g. Jongman et al. 2000). The intensity of adjacent /a/ was also measured at midpoint. Normalized intensity (*NrmInt*) was then calculated as a ratio of fricative intensity (at each time point)/[a] midpoint intensity. (6 measures: 3 each raw and normalized)

Formant transitions (F2 and F3) of the /a/ adjacent to each fricative were measured at 50%, 70%, and 90% (for [a]-fricative tokens) and 10%, 30%, 50% (for fricative-[a] tokens). To normalize, we then calculated the frequency difference between the midpoint of the vowel (50%) and the point of the vowel nearest the fricative boundary (90% or 10%) (*F2LongΔ*, *F3LongΔ*), and between roughly one-third into the vowel and the fricative edge (90%-70%, 10%-30%) (*F2ShrtΔ*, *F3ShrtΔ*). A slope was calculated between the same pairs of points, 90%-70% and 10%-30% (*F2ShrtSlp*, *F3ShrtSlp*), and 10%-50% and 90%-50% (*F2LongSlp*, *F3LongSlp*). (8 measures)

Fricative spectra were measured at 25%, 50% and 75% of the fricative via the following sets of measures:

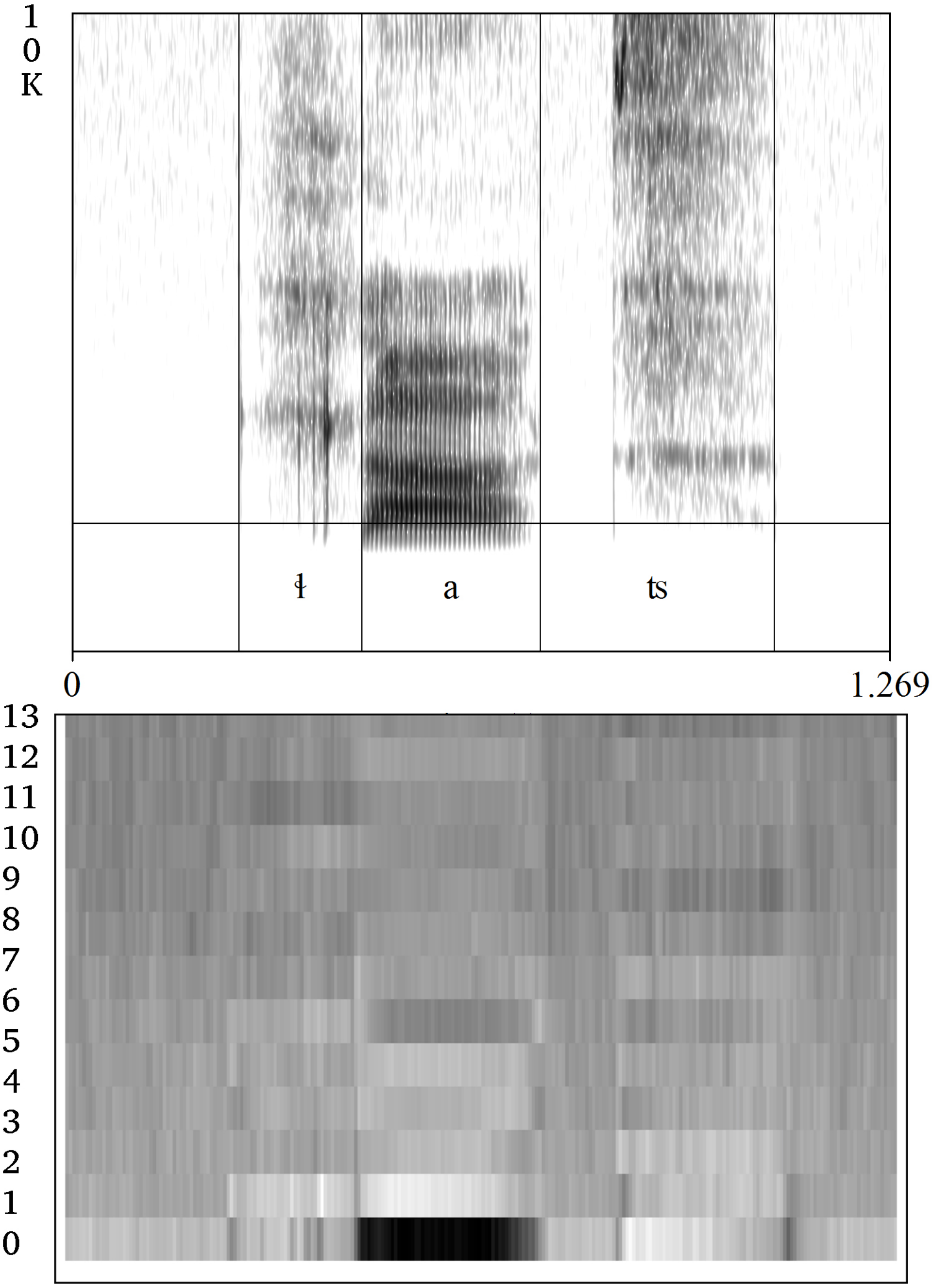
- spectral moments (Centroid (i.e. center of gravity), Variance, Skewness (*Skew*), Kurtosis) (12 measures)
- maximum position (*MaxPos*), the frequency bin with the highest energy (3 measures)
- spectral roll-off, the frequency below which a certain percentage of the total energy of the fricative falls. This was calculated for 25%, 50%, 75%, and 90% energy (*Rlloff25*, *Rlloff50*, *Rlloff75*, *Rlloff90* respectively). (12 measures)
- MFCC (mel frequency cepstral coefficient), first 4 coefficients: These measures (*MFCC0*, *MFCC1*, *MFCC2*, *MFCC3*) are commonly used in speech recognition to create Mel scaled representations of the signal and to decorrelate source and filter (Davis & Mermelstein 1980). In addition, cepstral coefficients have been employed in classification of Romanian fricatives (Spinu & Lilley 2016) and were shown to reliably outperform spectral moments in that task. These measures aim to capture the features of the shape of the spectrum while separating the glottal source from the articulatory configuration of the upper vocal tract. The calculation of MFCCs involves computing the inverse Fourier Transform of the log magnitude of the Fourier Transform.[13] As O'Shaughnessy 2008 states, "the first MFCC [MFCC0] is simply a version of energy," and MFCC1 can be interpreted as "the global energy balance between low and high frequencies". Other MFCCs unfortunately "are difficult to relate to any clear aspects of speech production or perception" although they "contain finer spectral detail." Figure 4 shows a representation of the traditional 13 MFCCs in one Deg

---

[13] As described by O'Shaughnessy 2008: 2973, "an FFT or LPC spectrum is obtained for each speech frame, for which the logarithm is then taken of the spectral amplitude (converting to decibels, and discarding the spectral phase), a set of triangular filters spaced according to the perceptual mel scale weights this result, and finally an inverse FFT is done. The low-order coefficients (e.g., 10–16 in number) of this last step provide the spectral vector for evaluation."

Xinag word, with a labelled spectrogram for comparison. The vertical axis in the MFCC figure shows cepstral coefficient values. MFCCs were calculated with preemphasis of .97 using Hamming windows of 25 ms with 25 ms frame, sampled with a time step of 10 ms. (12 measures)

Figure 4. Wide-band spectrogram (top) and lowest 13 MFCCs (bottom) in [ɬaʦ] ‘dirt’ (ED)



### *5.2 Results*

#### *5.2.1 Correlations among measures*

Given the large number (52) of measures in our study, first we investigated correlations among the measures in order to eliminate as much correlation bias (Toloşi & Lengauer 2011) as possible. In Table 5 we provide a fixed effect correlation table. For measures which were made at three time points, only the midpoint measure was used to investigate correlation.  (Table 5 is broken into three subtables because of the long x axis.)  As there is no invariant number for considering a pair of measures correlated, we selected .5 (or -.5) as our threshold for considering a pair of measures correlated. Correlated measures are italicized and placed in a box.

Table 5. Correlation matrix for measures

| | (Intr) | FricDur | Intens2 | NrmInt2 | F2LongSlp | F2LongΔ | F2ShrtSlp | F2ShrtΔ |
|---|---|---|---|---|---|---|---|---|
| FricDur | 0.043 | | | | | | | |
| Intens2 | 0.068 | 0.024 | | | | | | |
| NormInt2 | -0.067 | -0.054 | *-0.725* | | | | | |
| F2LongSlp | 0.284 | -0.034 | -0.009 | -0.05 | | | | |
| F2LongΔ | 0.265 | -0.068 | 0.017 | -0.048 | *0.958* | | | |
| F2ShrtSlp | 0.294 | -0.002 | 0.018 | 0.012 | *-0.78* | *-0.76* | | |
| F2ShrtΔ | 0.277 | 0.009 | 0.001 | 0.02 | *-0.765* | *-0.799* | *0.961* | |
| F3LongSlp | 0.243 | 0.049 | 0.098 | -0.041 | -0.148 | -0.148 | 0.173 | 0.189 |
| F3LongΔ | 0.253 | 0.045 | 0.065 | -0.031 | -0.171 | -0.187 | 0.195 | 0.221 |
| F3ShrtSlp | -0.295 | 0.072 | -0.104 | 0.079 | 0.067 | 0.07 | -0.26 | -0.259 |
| F3ShrtΔ | -0.282 | 0.097 | -0.086 | 0.078 | 0.093 | 0.112 | -0.276 | -0.296 |
| Rlloff25_2 | 0.079 | 0.198 | -0.163 | -0.052 | 0.04 | 0.004 | 0.004 | 0.043 |
| Rlloff50_2 | -0.041 | -0.06 | 0.068 | -0.03 | 0.011 | 0.037 | -0.017 | -0.071 |
| Rlloff75_2 | -0.002 | 0.01 | -0.024 | 0.039 | -0.046 | -0.06 | 0.011 | 0.053 |
| Rlloff90_2 | -0.001 | -0.026 | 0.021 | -0.058 | 0.003 | 0.01 | 0.017 | -0.016 |
| MaxPos2 | -0.017 | -0.096 | -0.063 | 0.087 | -0.01 | 0.005 | 0.023 | 0.023 |
| 2 | 0.047 | 0.059 | 0.049 | -0.057 | 0.151 | 0.13 | -0.105 | -0.061 |
| Variance2 | -0.054 | -0.022 | -0.061 | 0.127 | -0.202 | -0.183 | 0.154 | 0.126 |
| Skewness2 | 0.028 | 0 | 0.056 | -0.063 | 0.012 | -0.018 | 0.025 | 0.065 |
| Kurtosis2 | -0.033 | -0.046 | -0.199 | 0.154 | 0.017 | 0.044 | -0.025 | -0.069 |
| MFCC0_2 | -0.075 | -0.004 | *-0.784* | 0.25 | 0.044 | 0.015 | -0.046 | -0.028 |
| MFCC1_2 | 0.037 | 0.181 | 0.057 | -0.136 | 0.09 | 0.034 | -0.059 | -0.001 |
| MFCC2_2 | -0.048 | 0.054 | 0.03 | -0.029 | 0.055 | 0.041 | -0.076 | -0.115 |
| MFCC3_2 | 0.022 | -0.001 | -0.148 | 0.194 | -0.037 | -0.087 | 0.086 | 0.074 |

|  | F3LongSlp | F3LongΔ | F3ShrtSlp | F3ShrtΔ | Rlloff25_2 | Rlloff50_2 | Rlloff75_2 | Rlloff90_2 | Maxpos2 |
|---|---|---|---|---|---|---|---|---|---|
| F3LongSlp |  |  |  |  |  |  |  |  |  |
| F3LongΔ | *0.963* |  |  |  |  |  |  |  |  |
| F3ShrtSlp | *0.836* | *-0.793* |  |  |  |  |  |  |  |
| F3ShrtΔ | *-0.798* | *-0.815* | *0.956* |  |  |  |  |  |  |
| Rolloff25_2 | 0.001 | 0.024 | 0.022 | 0.017 |  |  |  |  |  |
| Rolloff50_2 | 0.01 | -0.008 | -0.044 | -0.03 | *-0.602* |  |  |  |  |
| Rolloff75_2 | -0.017 | 0.007 | 0.063 | 0.04 | 0.206 | *-0.724* |  |  |  |
| Rolloff90_2 | 0.033 | 0.024 | -0.061 | -0.041 | -0.102 | *0.536* | *-0.913* |  |  |
| MaxPos2 | -0.056 | -0.065 | 0 | -0.009 | 0.098 | 0.093 | 0.002 | -0.052 |  |
| Centroid2 | -0.034 | -0.069 | 0.037 | 0.041 | -0.091 | -0.215 | 0.139 | -0.319 | -0.434 |
| Variance2 | 0.051 | 0.084 | -0.043 | -0.032 | 0.18 | -0.085 | 0.011 | 0.168 | 0.298 |
| Skew2 | 0.015 | -0.017 | -0.039 | -0.021 | -0.116 | -0.174 | 0.078 | -0.119 | -0.112 |
| Kurtosis2 | -0.052 | -0.011 | 0.025 | 0.009 | 0.028 | 0.147 | -0.066 | 0.073 | 0.026 |
| MFCC0_2 | -0.054 | -0.025 | 0.069 | 0.051 | 0.208 | -0.04 | -0.034 | 0.053 | -0.014 |
| MFCC1_2 | -0.018 | -0.031 | 0.062 | 0.082 | 0.31 | -0.155 | -0.095 | -0.091 | -0.105 |
| MFCC2_2 | -0.004 | -0.020 | -0.02 | 0.002 | -0.073 | 0.093 | 0.083 | 0.048 | 0.039 |
| MFCC3_2 | -0.158 | -0.111 | 0.089 | 0.073 | -0.033 | -0.020 | 0.059 | -0.070 | 0.052 |

|  | Centroid2 | Variance2 | Skew2 | Kurtosis2 | MFCC0_2 | MFCC1_2 | MFCC2_2 | MFCC3_2 |
|---|---|---|---|---|---|---|---|---|
| Centroid2 |  |  |  |  |  |  |  |  |
| Variance2 | *-0.782* |  |  |  |  |  |  |  |
| Skew2 | *0.61* | -0.430 |  |  |  |  |  |  |
| Kurtosis2 | -0.444 | 0.361 | *-0.847* |  |  |  |  |  |
| MFCC0_2 | 0.027 | -0.020 | 0.014 | 0.132 |  |  |  |  |
| MFCC1_2 | 0.390 | -0.084 | 0.307 | -0.419 | 0.097 |  |  |  |
| MFCC2_2 | -0.225 | -0.099 | 0.001 | 0.108 | -0.107 | -0.287 |  |  |
| MFCC3_2 | -0.071 | 0.238 | -0.096 | 0.081 | 0.030 | 0.169 | -0.101 |  |

As might be expected, a number of the measures in our study were correlated with each other:

- MFCC0 and Intens, NormInt2 and Intens
- three of the spectral moments: Skew and Kurtosis, Centroid and Variance, Centroid and Skew
- four of the spectral roll-off measures: Rlloff25 and Rlloff50, Rlloff50 and Rlloff75, Rlloff50 and Rlloff90, Rlloff75 and Rlloff90
- various F2 and F3 transition measures
    - ❖ F2ShrtΔ and F2ShrtSlp, F2LongSlp and F2ShrtSlp, F2LongΔ and F2ShrtSlp, F2LongSlp and F2ShrtΔ, F2LongΔ and F2ShrtΔ, F2LongΔ and F2LongSlp
    - ❖ F3ShrtΔ and F3ShrtSlp, F3LongSlp and F3ShrtSlp, F3LongSlp and F3ShrtΔ, F3LongΔ and F3ShrtSlp, F3LongΔ and F3ShrtΔ, F3LongΔ and F3LongSlp

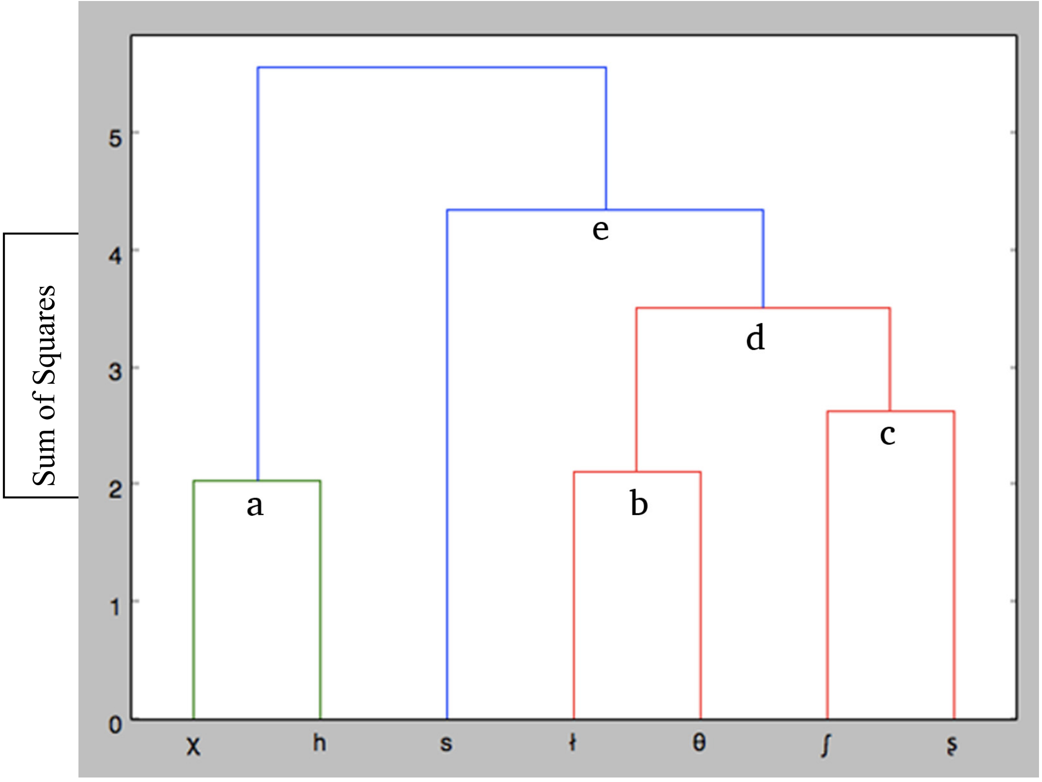
Of the correlated measures, we eliminated Intensity but kept Normalized Intensity (since it is a normalized measure). We also eliminated Skew and Variance since both measures were correlated with Centroid, keeping the latter (as well as Kurtosis) because center of gravity has been more widely used in previous studies of fricatives. Of the correlated roll-off measures, we eliminated Rolloff50 since it was correlated with Rolloff25 and Rolloff75 but kept each of the latter. We eliminated Rolloff90 since it was correlated with Rolloff75. Of the correlated F2 and F3 transition measures, we kept only F2LongSlope and F3LongSlope, since these two measures provide a more complete temporal profile of formant transitions compared to the short slopes. It was more or less an arbitrary decision to keep the long slope measures in the analysis in place of the long delta measures. The initial set of 52 measures was thus reduced to a set of 13: NormInt, Centroid, Kurtosis, Rolloff25, Rolloff75, F2LongSlope, F3LongSlope, FricDur, MaxPos, MFCC0, MFCC1, MFCC2, and MFCC3.

### *5.2.2 Cluster analysis*

To begin to visualize the relationships among the fricatives, we performed a cluster analysis. We represented each fricative as the multidimensional vector corresponding to the centroid of the instances of that class (for the 13 non-correlated measures). We then clustered the resulting representations, using Ward's method for agglomerative clustering (Ward 1963, Everitt 1993). The dendrograms appear in Figure 5-Figure 6. The shortness of bar height indicates the closeness of the relationship, and the vertical axis represents the within-cluster sum-of-squares distance at a particular level.

For fricatives before [a] (Figure 5), the three pairs of fricatives which are most similar to each other in the set of non-correlated measures are (a) [χ] and [h], then (b) [ɬ] and [θ], then (c) [ʃ] and [ʂ]. Group (d), coronal fricatives other than [s], forms the next level of grouping. The next level, (e), consists of [s] and the group (d) coronal fricatives.

Figure 5. Cluster analysis using Ward's method: fricatives before [a]



The groupings after [a] (Figure 6) show some similarities and differences to the pre-[a] context. The most similar group, (a), consists of two sibilants, [ʃ] and [ʂ]. The next most similar group, (b), consists of [ɬ] and [ç], which then cluster to form group (c), with [θ]. As [h] does not occur in this position, [χ] forms the next group (d) with the group (a) fricatives. [s] is the most dissimilar fricative in this position.

Figure 6. Cluster analysis using Ward's method: fricatives after [a]

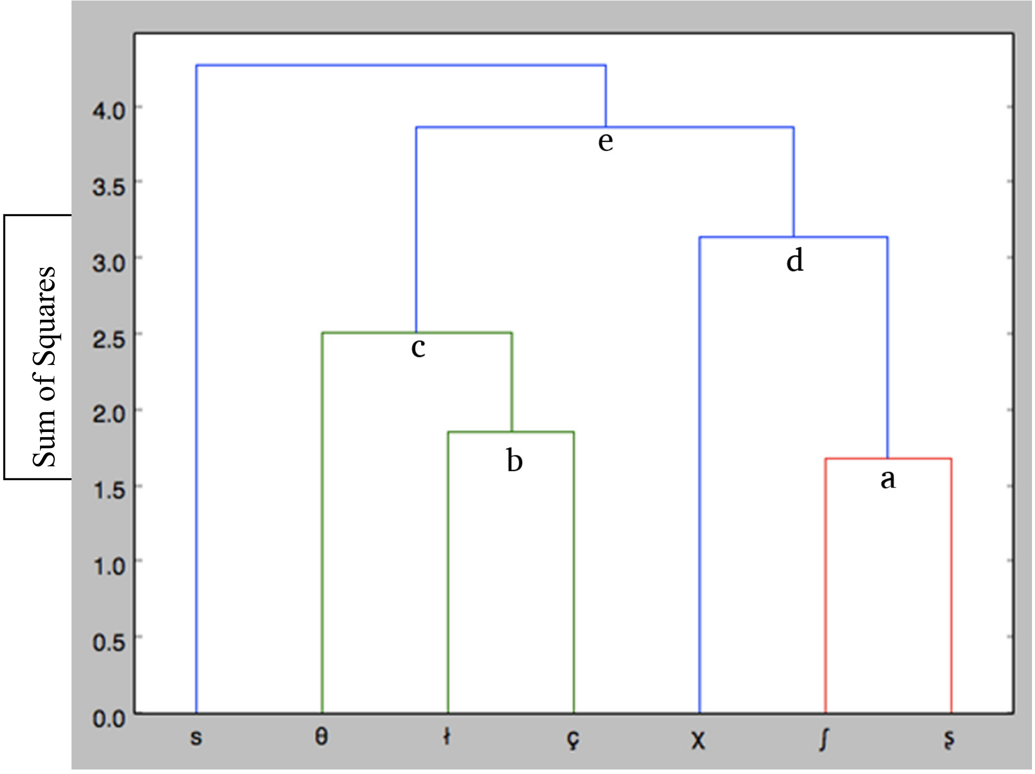


Table 6 provides another way of looking at the groupings in the cluster analysis. Across positions it can be seen that [ʃ ʂ] and [ɬ θ] are two of the most consistent groupings.

Table 6. Groupings across positions

| | a | b | c | d | e |
|---|---|---|---|---|---|
| pre-[a] | χ h | ɬ θ | ʃ ʂ | ɬ θ ʃ ʂ | ɬ θ ʃ ʂ s |
| post-[a] | ʃ ʂ | ɬ ç | ɬ ç θ | ʃ ʂ χ | ɬ ç θ ʃ ʂ χ |

### *5.2.3 Linear mixed effects model*

Next, we investigated the relative contribution of the non-correlated measures by statistical modelling using a linear mixed effects model. For this statistical analysis, a z-score normalization of each measure was performed for each speaker, in order to minimize the impact of between-speaker variation. Specifically, each measured value $x$ was replaced by $(x - \mu)/\sigma$, where $\mu$ and $\sigma$ are each speaker's mean value and standard deviation, respectively, for that measure.

We employed a linear mixed effects model using the R package MCMCglmm (Hadfield 2010) with fricative identity as a categorical dependent variable, the measures described above as fixed effects, and speaker and word repetition as random effects. Separate analysis was conducted for pre- and post-vocalic instances.

The deviance information criterion (DIC) (Spiegelhalter et al. 2002) measures the goodness of fit of the model, where a lower value represents a better fit. Table 7 (pre-[a] fricatives) and Table 8 (post-[a] fricatives) present the fixed effects in decreasing order of DIC, comparing the original model and the model built when that fixed effect is excluded. The values represent the DIC of the model that results from removing the specified fixed effect. A higher DIC for the model which excludes a particular fixed effect indicates that removing that factor leads to a worse model. Thus, a higher DIC indicates that the factor had a greater importance to the original model. Delta DIC, also shown in the tables, is computed with respect to the DIC of the full model, to indicate information loss when a feature is removed.

Table 7. Deviance information criterion results for pre-[a] fricative measures

| *Measure* | *DIC* | *Delta DIC* |
|---|---|---|
| MFCC0 | 814.24 | 58.33 |
| NormInt | 770.57 | 14.65 |
| MFCC1 | 768.63 | 12.71 |
| Centroid | 765.32 | 9.41 |
| Kurtosis | 764.39 | 8.48 |
| MFCC3 | 764.26 | 8.35 |
| F2LongSlp | 763.10 | 7.19 |
| F3LongSlp | 761.80 | 5.89 |
| MFCC2 | 760.15 | 4.24 |
| Rlloff75 | 758.14 | 2.23 |
| FricDur | 756.76 | 0.85 |
| Rlloff25 | 754.97 | -0.94 |
| Maxpos | 754.54 | -1.37 |

Table 8. Deviance information criterion results for post-[a] fricative measures

| *Measure* | *DIC* | *Delta DIC* |
|---|---|---|
| F2LongSlp | 831.79 | 66.96 |
| MFCC0 | 780.89 | 16.07 |
| MFCC3 | 772.29 | 7.46 |
| MFCC1 | 771.56 | 6.74 |
| Maxpos | 771.28 | 6.45 |
| NormInt | 770.34 | 5.51 |
| Centroid | 764.81 | -0.02 |
| Kurtosis | 764.74 | -0.09 |
| MFCC2 | 764.71 | -0.12 |
| Rlloff75 | 764.33 | -0.50 |
| Rlloff25 | 761.50 | -3.33 |
| F3LongSlp | 760.01 | -4.82 |
| FricDur | 756.00 | -8.83 |

Our interpretation of Table 7-Table 8 is deferred to 5.3.1.

### *5.2.4 Analysis by classification*

We continued our investigation of the relative contribution of the non-correlated measures with an analysis-by-classification experiment employing a machine learning classifier, another methodology which has been used in phonetics (e.g. Styler 2015 for vowel nasalization, Panfili 2018 for voice quality) to detect which of various measures of a large initial measurement set perform the most work in distinguishing a set of contrasts. As with the linear mixed effects modelling in 5.2.3, we performed this experiment using a z-score normalization of each measured value.

To assess the relative contribution of different acoustic features to the identification of fricative categories, we trained a machine learning classifier to distinguish among the eight fricative classes (seven in each position) and then compared the results of classification based on different subsets of features. We employed a Support Vector Machine (SVM) (Boser et al. 1992) classifier as implemented in the Weka package (Hall et al. 2009), trained using Sequential Minimum Optimization with default parameter settings. SVMs learn the separating hyperplane between two categories in high dimensional space such that the margin between instances of different categories is maximized. Multi-class classification is achieved by training multiple pairwise classifiers and selecting the most frequently assigned label. In our experiments, we first group instances by position and then perform 10-fold cross-validation on the resulting subset, iteratively training on 90% of the data and testing on the remaining 10%. We report results accumulated over all folds.

Overall, we achieved 71.8% accuracy on pre-vocalic fricatives and 72.3% on post-vocalic tokens, relative to a 15% most common class baseline. We then performed an ablation analysis, comparing the results of performing leave-one-feature-type-out classification experiments and removing the feature type with the largest impact on classification accuracy. This process was then repeated iteratively. The ranking of the measures by this ablation analysis appears in Table 9 for pre-vocalic fricatives and in Table 10 for post-vocalic fricatives:

Table 9. Ablation analysis results for classification experiment, pre-[a] measures

| *Measure* | *Accuracy* | *Reduction from all* |
|---|---|---|
| All | 71.8% | --- |
| NormIntens | 66.8% | -5.0% |
| Rlloff25 | 62.7% | -9.1% |
| MFCC1 | 58.2% | -13.6% |
| F3LongSlp | 57.3% | -14.6% |
| Rlloff75 | 55.9% | -15.9% |
| MFCC0 | 50.5% | -21.4% |
| Centroid | 45.5% | -26.4% |
| MFCC3 | 36.4% | -35.5% |
| Kurtosis | 28.6% | -43.2% |
| MFCC2 | 25.5% | -46.4% |
| Maxpos | 14.1% | -57.7% |
| F2LongSlp | 9.6% | -62.3% |

Table 10. Ablation analysis results for classification experiment, post-[a] measures

| *Measure* | *Accuracy* | *Reduction from all* |
|---|---|---|
| All | 72.3% | |
| F2LongSlp | 60.7% | -11.6% |
| MFCC0 | 55.8% | -16.5% |
| MFCC3 | 50.5% | -21.9% |
| Centroid | 47.3% | -25.0% |
| NormIntens | 43.8% | -28.6% |
| MFCC2 | 36.2% | -36.2% |
| Kurtosis | 33.9% | -38.4% |
| Rlloff25 | 29.0% | -43.3% |
| Rlloff75 | 25.0% | -47.3% |
| MFCC1 | 17.9% | -54.5% |
| F3LongSlp | 15.2% | -57.1% |
| Maxpos | 13.0% | -59.4% |
| FricDur | 9.4% | -63.0% |

Our interpretation of the results of Table 9-Table 10 is given in 5.3.1.

As described earlier, the creation of the classifier involves the training of pairwise classifiers to distinguish between each pair of fricatives. The final classifier combines the outputs of these pairwise classifiers. In our framework, we employ linear classifiers such that the output of the classifier $y$ is computed as $y = wx + b$, where $x$ is the input feature vector, $b$ is a threshold value, and $w$ is the trained weight vector. Classification is determined by whether or not the output is greater than 0. Given this model, we can use the weight vectors learned for each pairwise classifier to determine which feature receives the most weight and thus plays the most important role in the classification decision. Table 11 displays the feature receiving the highest weight (absolute value) for the corresponding pairwise classifier in prevocalic position. The measure which distinguished the greatest number of fricative pairs was MFCC3 (6 pairs), followed by MFCC2 and Centroid (5 pairs each), then Rlloff25 (2 pairs), and finally F2LongSlp, F3LongSlp, MFCC0 at 1 pair apiece. Strikingly, Centroid distinguished /s/ from every other fricative, but only in that set of contrasts.

Table 11. Most important feature (measure) distinguishing pairs of fricatives before [a]

| | /s/ | /ɬ/ | /ʃ/ | /ʂ/ | /χ/ | /h/ |
|---|---|---|---|---|---|---|
| /θ/ | Rlloff25<br>2.1 | MFCC3<br>2.0 | MFCC2<br>1.9 | MFCC3<br>1.6 | MFCC2<br>2.1 | MFCC2<br>1.9 |
| /s/ | | Centroid<br>2.1 | Centroid<br>2.0 | Centroid<br>2.5 | Centroid<br>1.6 | Centroid<br>1.7 |
| /ɬ/ | | | MFCC2<br>1.9 | Rlloff25<br>2.4 | F2LongSlp<br>1.8 | MFCC2<br>1.7 |
| /ʃ/ | | | | F3LongSlp<br>2.1 | MFCC3<br>2.7 | MFCC3<br>2.2 |
| /ʂ/ | | | | | MFCC3<br>1.9 | MFCC3<br>1.7 |
| /χ/ | | | | | | MFCC0<br>2.9 |

Table 12 shows the result of similar pairwise comparisons for post-vocalic fricatives. Here the measure which distinguished the greatest number of pairs of fricatives was F2LongSlp (7 pairs), followed by MFCC0 (4 pairs), then MFCC2 (3 pairs), MFCC3 and Centroid (2 pairs each), and finally MFCC1 and Rlloff25 (1 pair each).

Table 12. Most important feature (measure) distinguishing pairs of fricatives after [a]

| | /s/ | /ɬ/ | /ʃ/ | /ʂ/ | [ç] | /χ/ |
|---|---|---|---|---|---|---|
| /θ/ | MFCC0<br>2.4 | MFCC2<br>2.2 | MFCC3<br>2.4 | MFCC0<br>1.6 | F2LongSlp<br>3.5 | MFCC2<br>2.6 |
| /s/ | | MFCC1<br>1.8 | MFCC2<br>1.7 | Centroid<br>2.5 | MFCC0<br>2.0 | Centroid<br>1.9 |
| /ɬ/ | | | MFCC3<br>2.3 | Rlloff25<br>1.8 | F2LongSlp<br>3.9 | F2LongSlp<br>2.7 |
| /ʃ/ | | | | F2LongSlp<br>2.4 | F2LongSlp<br>3.4 | MFCC3<br>2.8 |
| /ʂ/ | | | | | MFCC0<br>-2.6 | F2LongSlp<br>2.4 |
| [ç] | | | | | | F2LongSlp<br>2.9 |

## *5.3 Discussion*

### *5.3.1 Effect of place on measures*

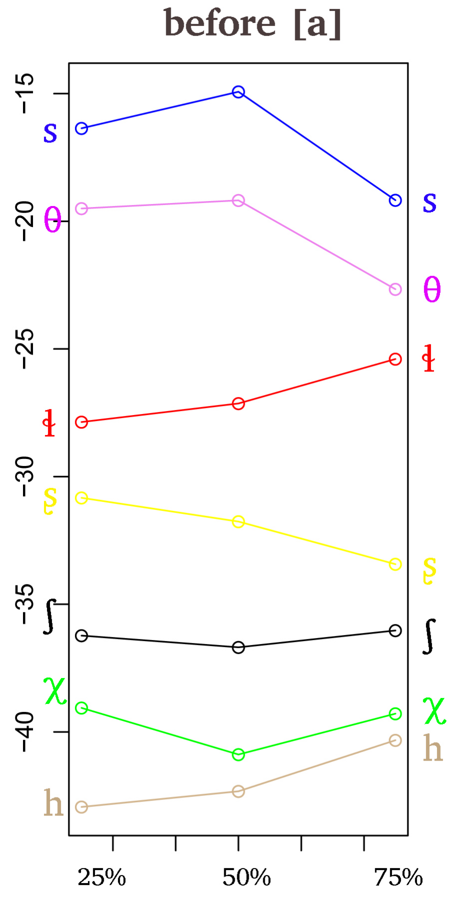
Research question (1) sought to determine the acoustic correlates of fricative place in Deg Xinag, and also to determine which acoustic measures are the most reliable indicators of place of articulation in Deg Xinag.

#### 5.3.1.1 Before /a/

As reported in 5.2.3, for pre-[a] fricatives, the Deviance Information Criterion test indicated that a relatively large set of measures (even in the reduced set of 13 non-correlated acoustic measures) had a significant effect distinguishing the group of 7 fricatives from each other. In this position, the measures which did the most work are those related to spectral shape (MFCC0, MFCC1, Centroid, Kurtosis, and MFCC3) and energy (NormIntensity), while F2 and F3 transitions were also meaningful. Each of these measures captures a slightly different dimension of the fricatives, so the groupings of the fricatives on the graphs presented below will be seen to be different.

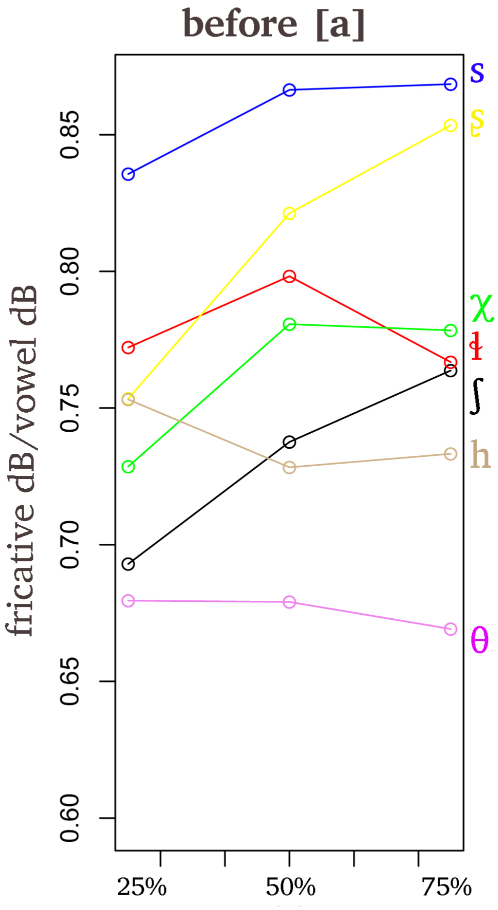
The MFCC metrics help to characterize the distribution of energy in the spectrum, while more conventional measures of spectral energy distribution, centroid and kurtosis, capture the overall spectral shape. Importantly, two of these energy metrics, NormIntensity and MFCC1, also contributed significantly to the classification accuracy in the ablation analysis (5.2.4). Figure 7-Figure 8 illustrate the ways in which the two metrics that were at the top of both the DIC and the ablation analyses, MFCC1 and NormInt, respectively, differentiate the spectral and energy characteristics of the different fricatives. As seen in Figure 7 (MFCC1), focusing on midpoint measures, /h/ and /χ/ have a larger ratio of energy at lower frequencies versus higher frequencies relative to other fricatives. Then in decreasing order are /ʃ ʂ ɬ θ/ and finally /s/. Figure 7 also illustrates how /h χ ɬ ʃ/ exhibit an increase in this ratio of energies from the midpoint to the offset before /a/. This change results in a different grouping of fricatives (into two sets) at fricative offset: /h χ ʃ ʂ/ and /ɬ θ s/.

Figure 7. MFCC1 of pre-vocalic Deg Xinag fricatives, at three positions within each fricative



Results for Normalized Intensity (Figure 8), were similar to our initial observations about intensity from FFT spectra (see §10), where we observed pre-vocalically that the sibilant fricatives [s ʂ ʃ] had the highest intensity of frication, [ɬ χ ç] had intermediate frication energy, and [θ h] had low intensity frication. For our NormIntens measure, at midpoint, the sibilants /s ʂ/ are the loudest fricatives, with /s/ louder than /ʂ/ at most points sampled. Then there is an intermediate group of fricatives consisting of /ʃ ɬ χ h/. The quietest fricative is /θ/. Two of these are fairly different at offset, with /ʂ/ becoming louder, comparable to /s/, and /ɬ/ becoming quieter, even more like /χ/ and /ʃ/.

Figure 8. Intensity of pre-vocalic Deg Xinag fricatives relative to following [a] (NormIntens), at three positions within each fricative



### 5.3.1.2 After /a/

In post-vocalic position, the most important distinguishing factors in the DIC analysis are F2 transitions (F2LongSlp), spectral shape (MFCC0, MFCC3, MFCC1, and MaxPos), and energy (NormInt). In the ablation analysis, three of these contributed the most to classification accuracy: F2 transition (F2LongSlp) and two spectral shape metrics (MFCC0 and MFCC3). Figure 9-Figure 11 illustrate the ways in which the F2 transition and the spectral shape measures separate the fricatives.

Figure 9 is a graphical display of post-fricative F2 transitions (slope measures) within [a]. The greatest F2 perturbations were found before [ç] and [ʂ], then the remaining coronals [s ɬ ʃ θ], with [χ] having the least effect on F2 transition (as might be expected since F2 for /a/ and uvulars is similar, cf. Alwan 1989, Elmazouzi et al. 2014).

Figure 9. F2ShrtSlp (left boxes) and F2LongSlp (right boxes) for [a] before fricatives

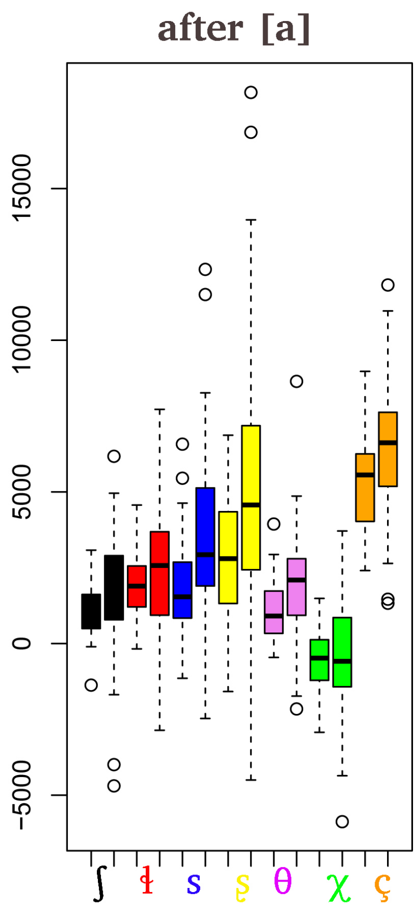
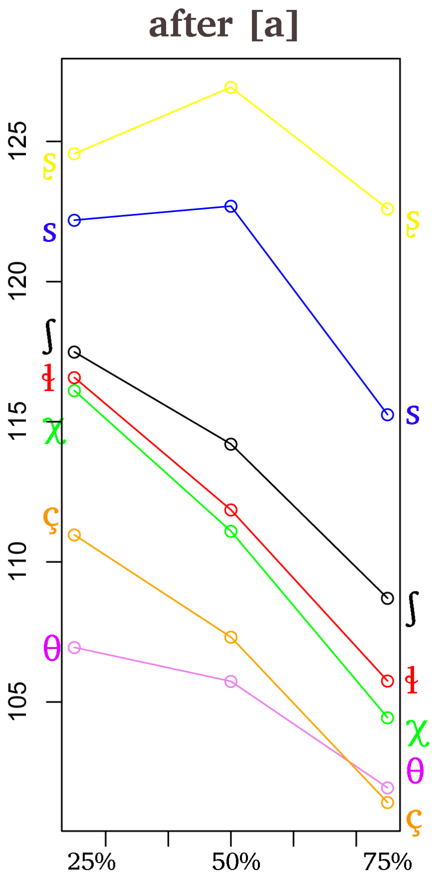


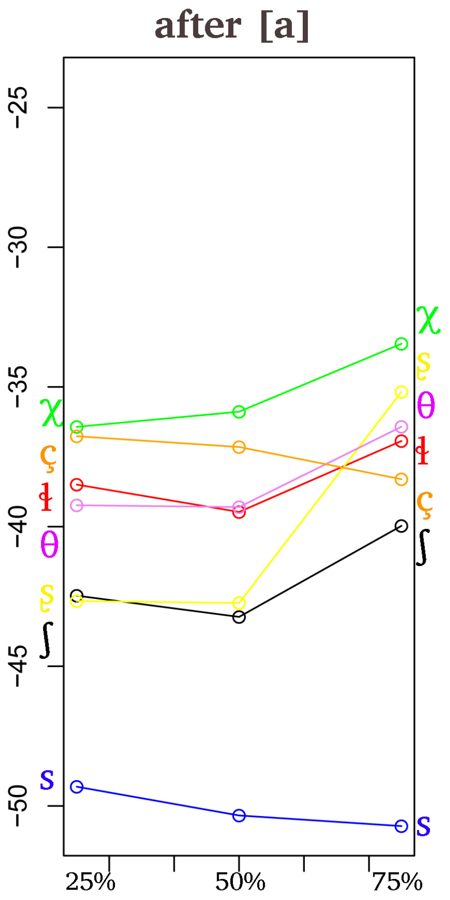
Figure 10 shows the degree of differentiation among the fricatives in MFCC3. There are three basic groupings: [ʂ ʃ ç], [s ɬ], and [θ χ].

Figure 10. MFCC3 of post-vocalic Deg Xinag fricatives, at three positions within each fricative



With respect to MFCC0, seen in Figure 11, at 25% and 50% [s] has the greatest energy, followed by [ʂ ʃ], then [θ ɬ], with [χ ç] having the least amount of energy for fricatives after [a].

Figure 11. MFCC0 of post-vocalic Deg Xinag fricatives, at three positions within each fricative



5.3.1.3 Summary

In comparing the metrics across prevocalic and postvocalic positions, recall that six of the seven places of articulation are the same in both positions. However, the measures that do the most work in distinguishing the fricatives are a somewhat disjoint set. In postvocalic position it is the F2LongSlp formant metric and two different spectral energy metrics (MFCC3 and MFCC0). Two of these, F2LongSlp and MFCC0, also distinguished a large number of fricatives in pairwise comparisons (cf. Table 12). In prevocalic position, the NormInt and MFCC1 energy metrics best distinguish the fricatives in the results reviewed in this section, but in pairwise comparisons (cf. Table 11), MFCC3, MFCC2 and Centroid distinguished the most pairs.

### *5.3.2 Comparison with other languages*

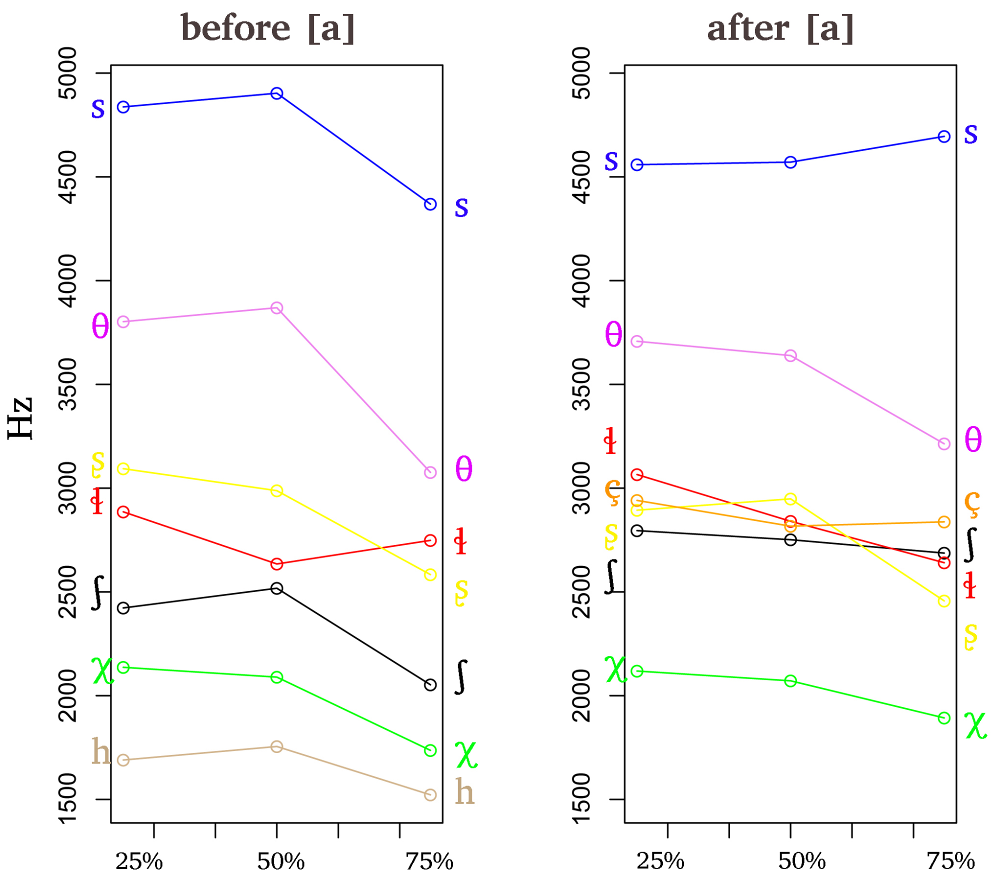
In this section we address our second research question, how the Deg Xinag fricatives compare with fricatives from other languages, particularly those in the same family as Deg Xinag, in terms of measures which distinguish them from each other. Despite differences in statistical analyses across studies, we strive to compare our results to those of other studies which investigated spectral moments, formant transitions, duration and intensity.

5.3.2.1 Spectral moments

<u>Centroid</u>: Our Deg Xinag data are roughly comparable to Hupa, Western Apache and Navajo, discussed in §2. As shown in Figure 12, /s/ has the highest centroid, just like

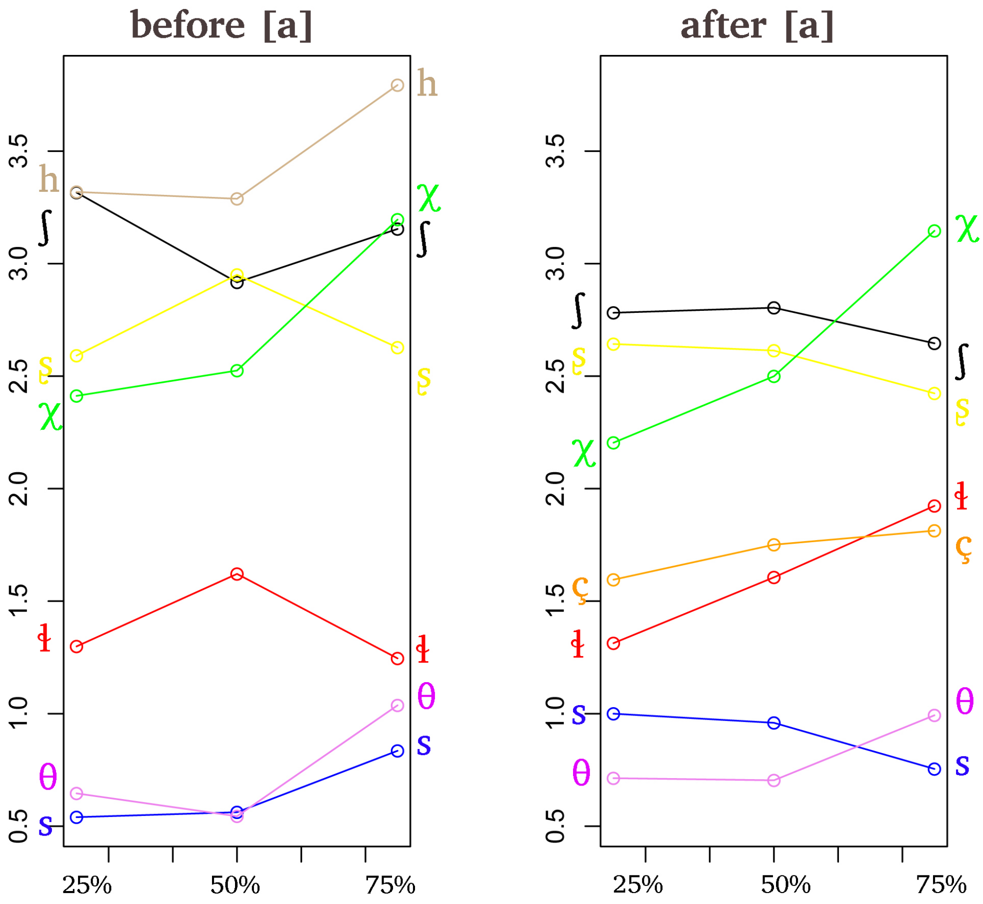
Hupa, W. Apache and Navajo (and also /s/ in Western Aleut (relative to /ç ł x χ/) (Gordon et al. 2002), and /s z/ in English (Jongman et al. 2000)). Deg Xinag /θ/ has the second highest centroid in both positions. In Deg Xinag, /ʃ ł/ are close in centroid, as in Hupa, but the relative height varies slightly with /ł/ having a slightly higher centroid before /a/ but with virtually identical heights after /a/. After [a], there are really only four settings for centroid, with [ç ş ł ʃ] forming a tight cluster. Also like the Hupa velar fricatives, Deg Xinag /χ/ has a low centroid, with only /h/ lower than /χ/, similar to Gaelic /x/, which had a significantly lower center of gravity than /fʲ ʃ ç s/ (Gordon et al. 2002).

Figure 12. Line graph of centroid of Deg Xinag fricatives at three time points, for fricatives before [a] and after [a]



<u>Skewness</u>: Graphical results for Deg Xinag skewness are shown in Figure 13. The fricatives with highest values on the Y-axis (/h χ ʃ ş/) have a greater concentration of energy in the lower frequencies. Before /a/, /ł θ s/ all have the lowest skew, whereas after /a/ /ł/ groups with [ç] in the middle of the range. Our results for skewness are reminiscent of English (Jongman et al. 2000), where /s z/ had the least amount of skew, and /θ ð/ the second least.

Figure 13. Line graph of skewness of Deg Xinag fricatives at three time points, for fricatives before [a] and after [a]. Y-axis units are arbitrary numbers.



Skewness is inversely correlated with Centroid (see Table 5) in Deg Xinag (just like Johnson 1993 found for Xhosa) and for that reason was not selected as one of our final measures, as discussed in 5.2.1.

Variance and kurtosis: Graphical results for Deg Xinag variance[14] are shown in Figure 14, which we discuss in conjunction with our results for kurtosis (Figure 15). Although these two measures were not correlated, the results are very similar.

Our results for Deg Xinag show a higher variance for /ɬ/ than for /ʃ/, similar to the findings of McDonough 2003 for Navajo, with the difference being more striking before /a/ than after it. /θ/ has the flattest spectrum as well as the most variance, followed by /s/, then /ɬ/. Next are /χ ʃ ʂ/ and [ç], with relatively peaked shapes and low variance. Finally, /h/ has the most peaked shape, as well as the least amount of variability around the mean.

---

[14] We use variance in the same way that other studies use standard deviation.

Figure 14. Line graph of variance of Deg Xinag fricatives at three time points, for fricatives before [a] and after [a]

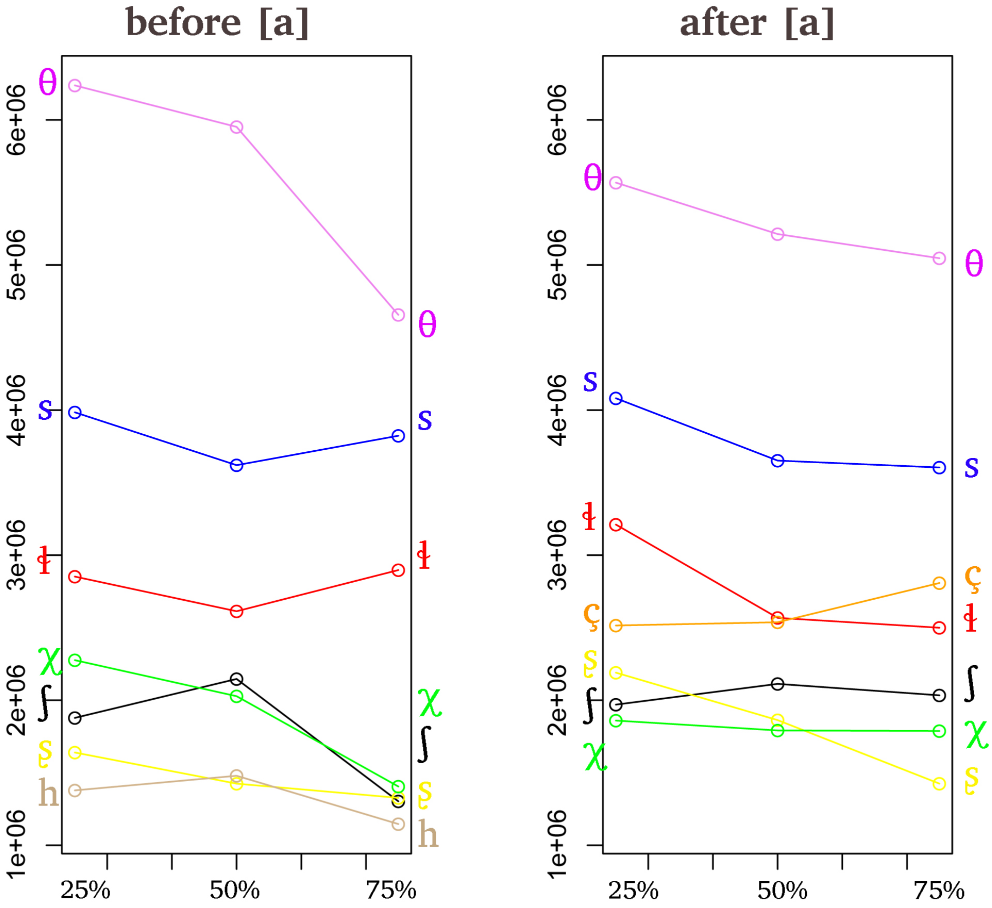
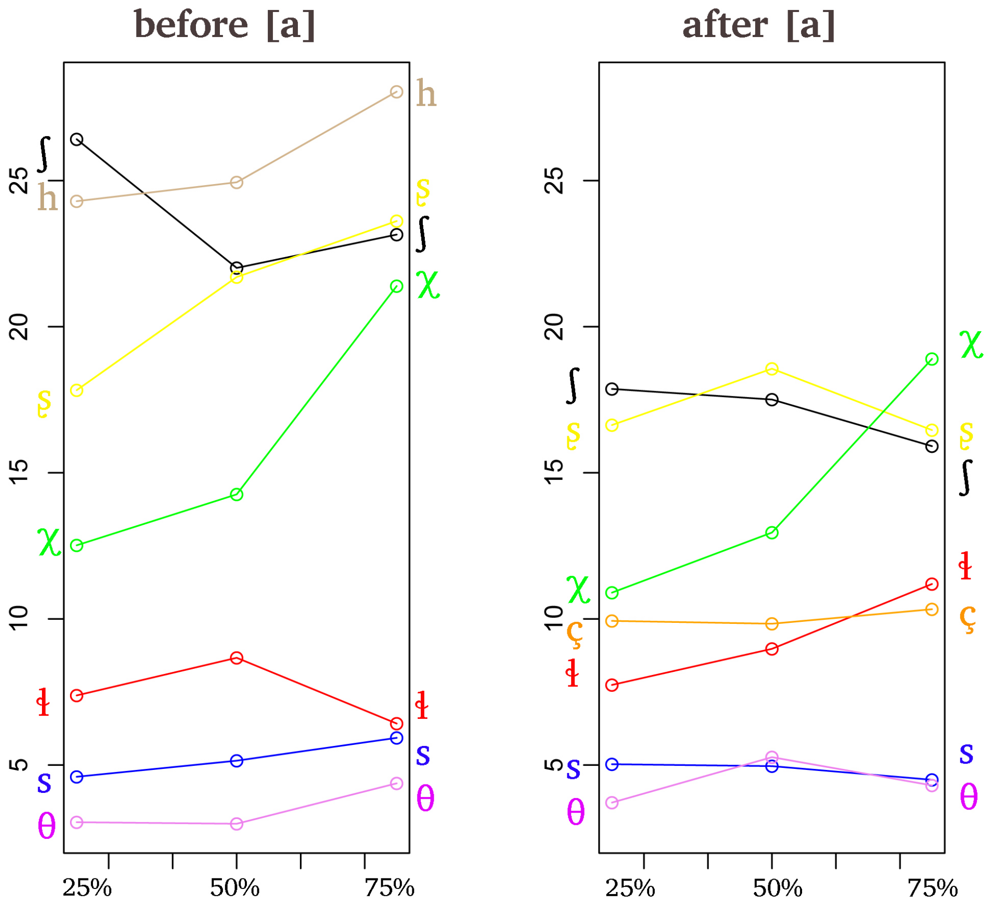


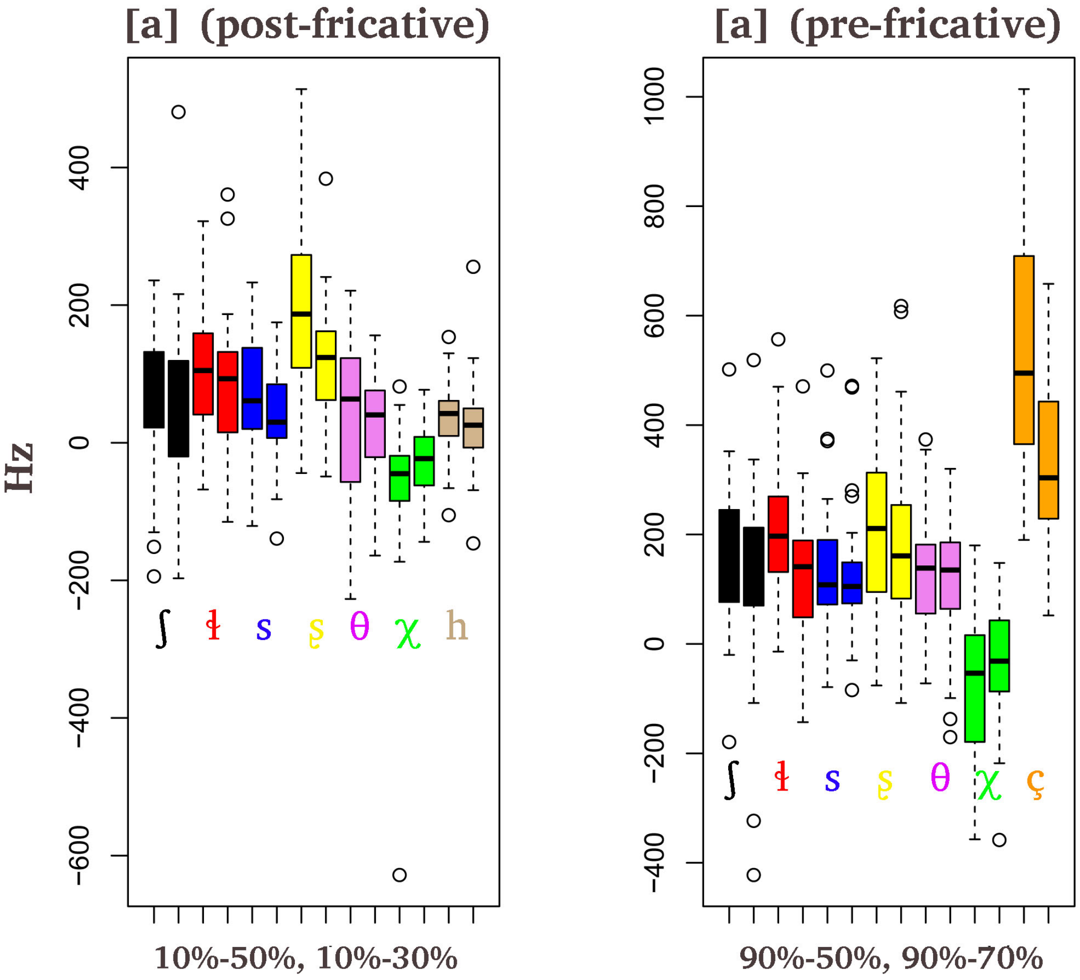
Figure 15. Line graph of kurtosis of Deg Xinag fricatives at three time points, for fricatives before [a] and after [a]. Y-axis units are arbitrary numbers.



5.3.2.2 Formant transitions

Deg Xinag F2 transition data is shown in Figure 16 for our difference measures, F2LongΔ and F2ShortΔ. The rightmost graph ([a] before fricative) is most comparable to the Hupa velar fricative data. For the vowel before /χ/, F2 lowers at the point in the vowel closest to the fricative. Compared to the other fricatives, /χ/ shows the most F2 lowering, similarly to [x] in Hupa. For the vowel following /χ/, F2 lowers less compared to the post-vocalic /χ/, but of the DX fricatives only /χ/ perturbs F2 of following [a] downwards.

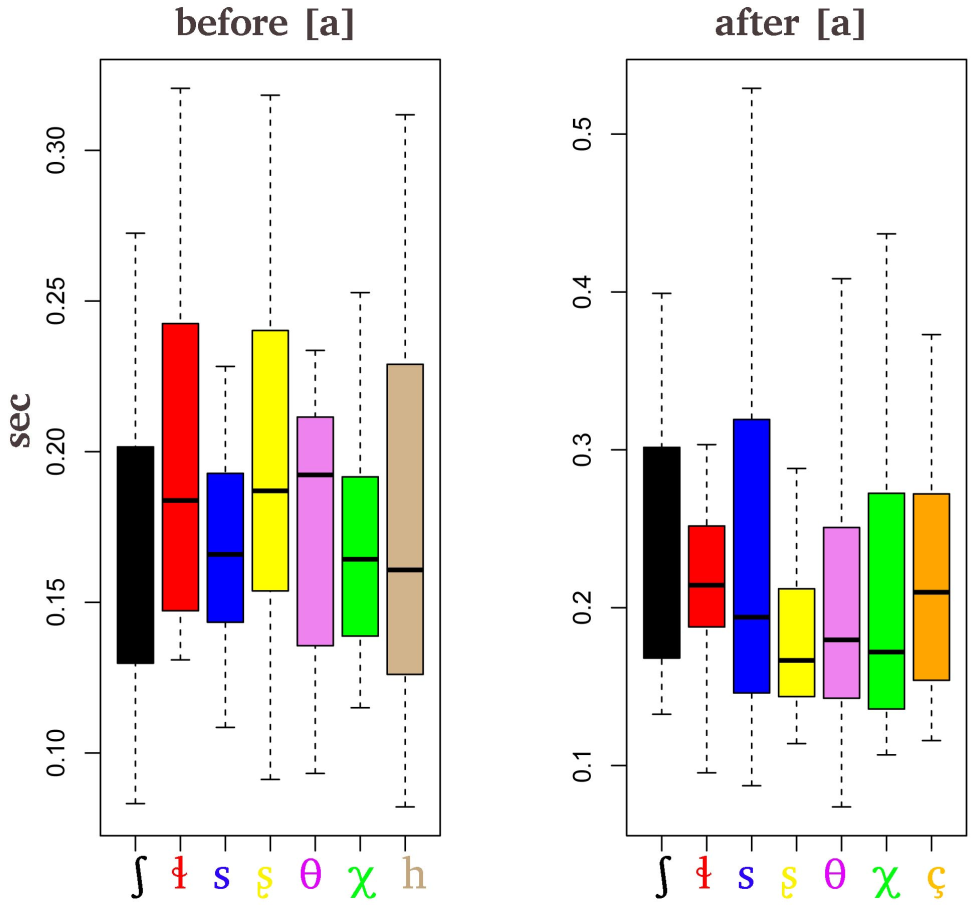
Figure 16. Box plots of F2 differences at two points within each vowel, 10%-50% (F2LongΔ) and 10%-30% (F2ShortΔ) (for [a] after fricatives), and 90%-50% (F2LongΔ) and 90%-70% (F2ShortΔ) (for [a] before fricatives)



### 5.3.2.3 Duration

Deg Xinag durational data are provided in Figure 17. For fricatives after [a], [ç] is the longest fricative and [ʂ] the shortest, but for fricatives before [a], the longest fricatives were [θ] and [ʂ]. Fricatives before [a] are slightly shorter on average than fricatives after [a].

Figure 17. Duration of Deg Xinag fricatives, in two positions

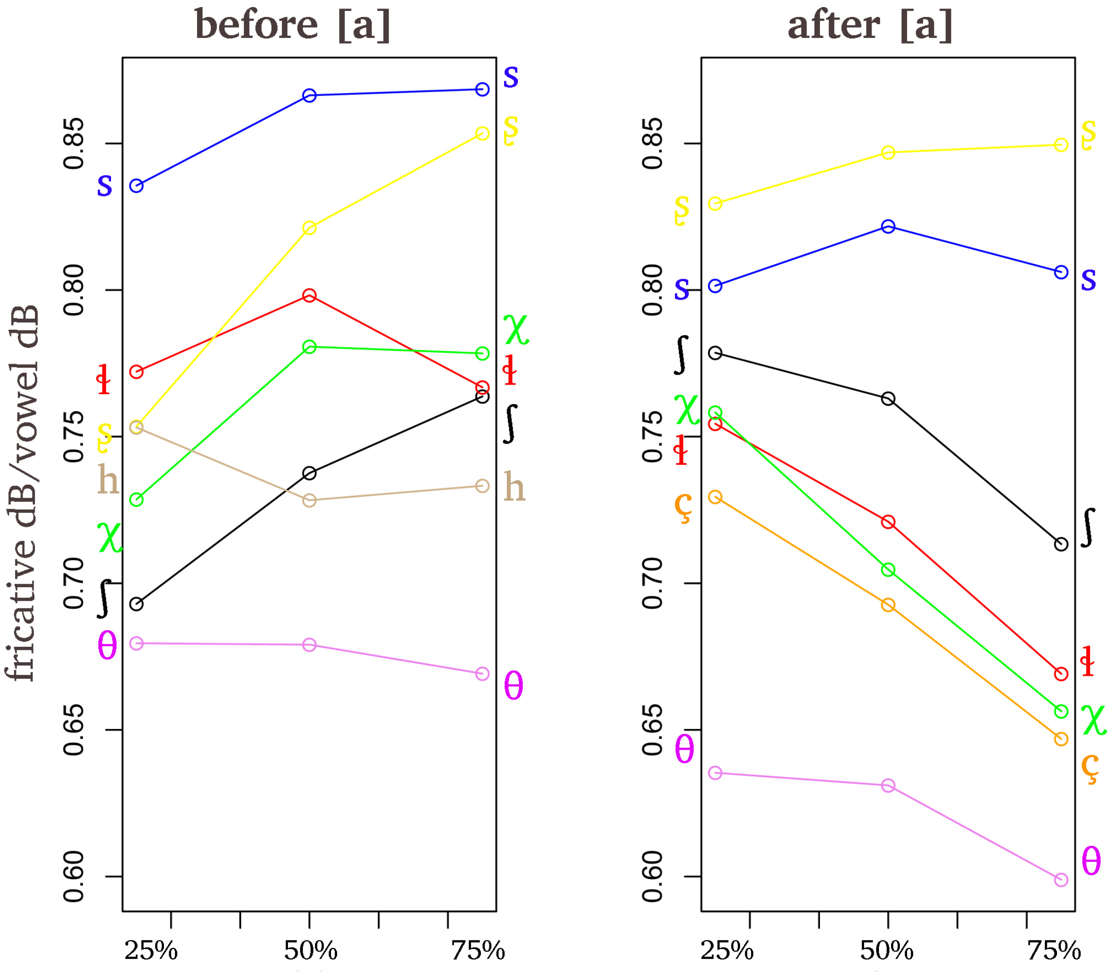


As seen in Table 7-Table 10 duration is not a highly ranked measure distinguishing fricatives before or after [a] in either the linear mixed effects model, the ablation analysis, or pairwise comparisons. Our results are in accord with those of Gordon et al. 2002, who noted that duration in general did not distinguish voiceless fricatives in their cross-linguistic study.

### 5.3.2.4 Intensity

Graphical results for normalized intensity, NormInt, were provided above in Figure 8 for fricatives before [a]. That graph is reproduced in Figure 18 with fricatives after [a] for comparison. It can be seen that in both positions the loudest fricatives are the sibilants /s ʂ/ and the quietest fricative is /θ/. The remaining fricatives, including the sibilant /ʃ/ and lateral /ɬ/, have intermediate values in both positions, but like Navajo, /ʃ/ has higher normalized intensity after /a/ than /ɬ/.

Figure 18. Intensity of Deg Xinag fricatives relative to adjacent [a], at three positions within each fricative



### 5.3.2.5 Summary

Deg Xinag, even with its larger fricative inventory, is generally comparable to the other Athabaskan languages for which there are instrumental studies of fricatives: Hupa, Navajo, and Western Apache. In all four languages, the highest centroid fricatives are /s/, followed by /ɬ ʃ/ with the dorsal fricatives [x χ] having the lowest centroid. /ɬ/ is more variable in Navajo than /ʃ/, also in Deg Xinag. In Navajo, /ʃ/ and /ɬ/ have intensity distributions skewed towards the lower frequencies relative to /s/, which we also found in Deg Xinag. Duration did not play a clear role in distinguishing the fricatives in Deg Xinag, just as it did not in Hupa or Western Apache. In Hupa and Navajo, /ʃ/ is of higher intensity than /ɬ/, but in our study, using a normalized measure of intensity, /ʃ/ is higher in intensity than /ɬ/ only after [a]. In Hupa, F1 and F2 were lower on vowels before velar than non-velar fricatives. In Deg Xinag, the least amount of F2 perturbation was found in [a] adjacent to [χ], which already has a low frequency concentration of energy.

## 6 Experiment 2: Towards understanding **θ, *ɬ* > /ɬ/ in Koyukon

In 5.3.2 we saw that the fricatives of Deg Xinag have many of the same properties as the corresponding sounds of Navajo, Western Apache and Hupa. Moreover, in the cluster analysis given in 5.2.2, we saw that before [a], [θ ɬ] are two of the most similar fricatives to each other, and after [a], [θ ɬ] form another tight grouping that includes [ç].

Recall from §3 that **θ *ɬ* have merged as /ɬ/ in Koyukon. If we use Deg Xinag as an approximation of the ancestor to Koyukon-Holikachuk-Deg Xinag (Goddard 1996), can we understand why **θ* and **ɬ* have merged in Koyukon, and why they have merged as /ɬ/ and not /θ/?

### *6.1 Materials and methods*

The data for Experiment 2 are a subset of the data used in Experiment 1. Here we provide statistical analyses of the subset pertinent to /ɬ/ and /θ/, 138 tokens across the eight speakers.

### *6.2 Results*

Deviance information criteria for /θ/ vs. /ɬ/ are given in Table 13-Table 14 for these fricatives at midpoint. Our approach here is to consider the measures at the bottom of the rankings and ask which measure performs relatively little work distinguishing /θ ɬ/ in both positions. In this case the answer is F2LongSlp, one of the formant transition measures.

Table 13. Deviance information criterion results for /θ/and /ɬ/ measures (pre-[a])

| *Measure* | *DIC* | *Delta DIC* |
|---|---|---|
| MFCC2 | 10.88 | 6.88 |
| Kurt | 8.80 | 4.81 |
| FricDur | 8.41 | 4.41 |
| Rlloff25 | 7.93 | 3.93 |
| Rlloff75 | 7.84 | 3.84 |
| MFCC3 | 6.83 | 2.84 |
| MaxPos | 6.47 | 2.48 |
| MFCC1 | 6.43 | 2.44 |
| NormInt | 5.92 | 1.93 |
| MFCC0 | 5.78 | 1.78 |
| Centroid | 5.76 | 1.77 |
| F3LongSlp | 5.38 | 1.39 |
| F2LongSlp | 4.62 | 0.62 |

Table 14. Deviance information criterion results for /θ/and /ɬ/ measures (post-[a])

| *Measure* | *DIC* | *Delta DIC* |
|---|---|---|
| Centroid | 8.57 | 5.74 |
| NormIntens | 6.32 | 3.49 |
| MFCC1 | 5.86 | 3.02 |
| MaxPos | 5.18 | 2.35 |
| Rlloff75 | 5.15 | 2.32 |
| Kurt | 5.14 | 2.31 |
| MFCC0 | 5.11 | 2.28 |
| Rlloff25 | 3.69 | 0.85 |
| F3LongSlp | 3.61 | 0.77 |
| MFCC3 | 3.55 | 0.72 |
| F2LongSlp | 3.52 | 0.69 |
| MFCC2 | 2.69 | -0.15 |
| FricDur | 2.66 | -0.18 |

### *6.3 Discussion*

Comparing Deg Xinag to its close relative Koyukon allows us to speculate on some of the forces at work in the merger of */θ/and */ɬ/, and furthermore to speculate on why the two fricatives merged into the lateral rather than into the dental fricative. Non-sibilant fricatives, especially quiet ones like [θ], rely on F2 transitions for the perception of place of articulation (e.g. Malécot 1956). The quieter the fricative, the more easily the spectral cues in the frication noise are lost due to environmental masking and the more reliant the listener becomes on cues in the formant transitions (for a discussion see Wright 1999 and Wright 2001). The lateral fricative [ɬ], on the other hand, is louder (see Figure 18), and therefore its internal spectral cues should be more robust to environmental masking, making it easier for a listener to recover without reliance on the formant transitions. Thus, there is presumably an asymmetry in the likelihood that the frication cues to place of articulation for the two fricatives are lost. As seen in Figure 19-Figure 20, the F2 transitions for [θ] and [ɬ] in preceding [a] are similar. At vowel midpoint, F2 is around 1500-1600 Hz, raising to around 1800 at vowel offset.

Figure 19. Wide-band spectrogram and lowest 5 formants of portion of [vəjəɬ ətəʧaɬ] 'he/she is busy' (speaker HM)

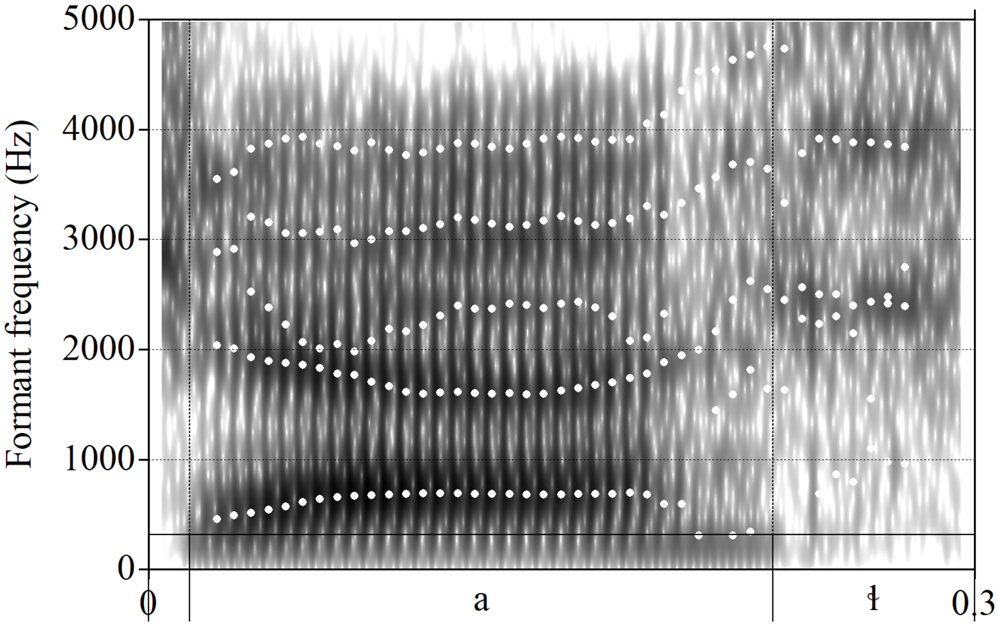
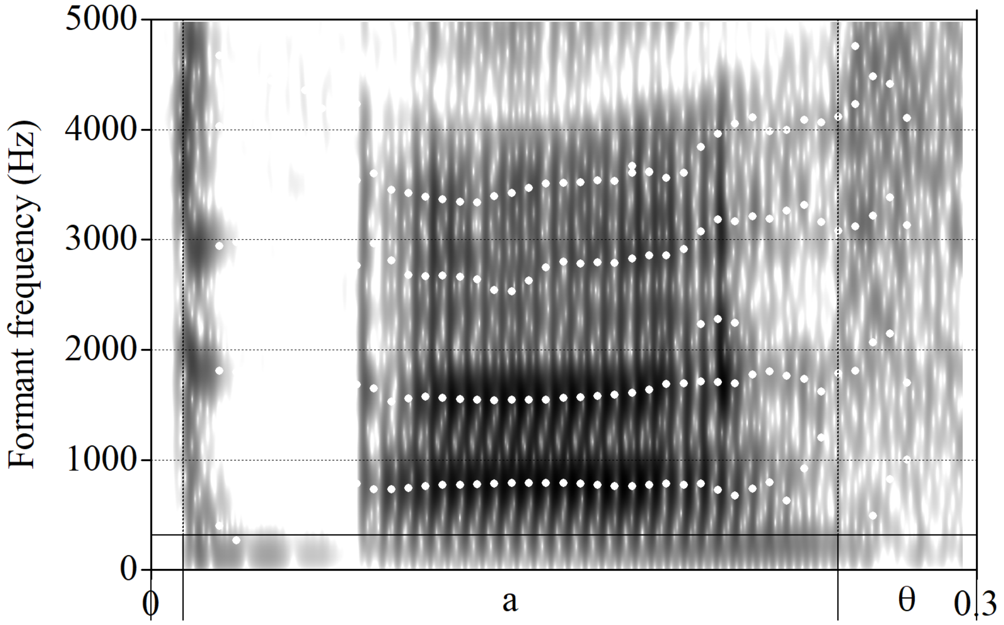


Figure 20. Wide-band spectrogram and lowest 5 formants of portion of [ŋəɬt'aθ] 'fry it' (speaker HM)



This similarity in F2 transitions may have lain the groundwork for a potential merger, which occurred in Koyukon in the direction of the louder fricative, /ɬ/. This process would probably have begun in the post-vocalic position where formant transitions carry less reliable place cues (e.g. Fujimura et al. 1978).[15]

## 7 Experiment 3: Clarifying the identity of prefixal "x"

As mentioned in §0, the stem-initial contrast between /χ/ and /h/ (e.g. /voˈχo/ 'for him', /ʁʊˈhoɬ/ 'he is walking') is fairly clear. In this context the fricative is followed by a stressed vowel. However, when the following vowel is unstressed, there is no contrast between /χ/ and /h/ in this context, and there has been inter-transcriber disagreement about the place of articulation of this fricative, "x". For example, the "areal" verbal prefix (Thompson 1993) has been transcribed both /h/- (as reflected in the practical orthography of one version of the language name *Deg Hit'an* (lit. 'people of this area'), Krauss 1974) and /χ/- (as reflected in the spelling *Deg Xinag* (lit. 'this language'), Kari 1978). In this section we report on the acoustic properties of "x", as attested in two productive prefixes, the areal and also the third person plural subject prefix.

### *7.1 Materials and methods*

The participants and recording equipment were the same as in Experiment 1. The target word list for Experiment 3 is shown in Table 15. The quality of the following vowel was balanced within the word list.

---

[15] The difference in intensity combined with the similarity in formant transitions may also contribute to the skewed distribution of /θ/ and /ɬ/ crosslinguistically. Both /θ/ and /ɬ/ are relatively rare and they contrast in a very small number of languages, but /ɬ/ is more common than /θ/ overall. In the genetically balanced 317 language sample of UPSID (Maddieson 1984), [ɬ] occurs in 9.4% of languages and [θ] in 5% of languages. In the 1672 language sample of Phoible (Moran et al. 2014), 6% of languages have [ɬ] while 4% have [θ], and only three languages (0.001%) have both: Tanacross and Dene Sųłiné (a.k.a. Chipewyan) (both Athabaskan), and Halkomelem (Salish).

Table 15. Word list for Experiment 3

| | /χ/ | /h/ | prefixal “x” |
|---|---|---|---|
| ____ a | χa ‘grease’ | “x”ənehaç ‘(you sg.) talk’ | “x”aθʈʂət ‘they’re lying down’ |
| ____ a | vaχa ‘with, by means of it’ | kag enahan̥ ‘you started to pick berries’ | “x”aθtɬo ‘they are (in position)’ |
| ____ə | χən ‘river’ | kəteʁ enohəɬ ‘(3s) will camp out’ | “x”ənehaç ‘(you sg.) talk’ |
| ____o | χot ‘slowly’ | kag ʊhon̥ ‘he’s eating berries’ | “x”otəɬ ‘they’re walking’ |

The words in Table 15 were recorded at the same time as those for Experiment 1, thus under the same conditions (see 5.1.1).

In Experiment 3, fricatives from words on the target word list which were shorter than 70 ms were not discarded: since “x” occurs before unstressed vowels, this fricative was often shorter than the cut-off duration used in Experiment 1.

The same measures were used as in Experiment 1.

### *7.2 Results*

We performed statistical and classification experiments analogous to those in 5.2 to analyse this particular contrast on a z-score transform of each measured value. There were 162 instances each of /h/, /χ/, and “x”.

#### *7.2.1 Linear mixed effects model*

Here we created a linear mixed effects model comparable to that for the experiment in §5, but in this case built only with the data specified in Table 15. We computed the deviance information criterion value (DIC) for each model resulting from exclusion of one of the fixed effects. The measures whose removal yielded the greatest degradation in model fit appear highest in Table 16.

Table 16. Deviance information criterion results for measures (/h/, /χ/, and “x”)

| *Measure* | *DIC* | *Delta DIC* |
|---|---|---|
| MFCC0 | 936.98 | 77.77 |
| NormInt | 926.35 | 67.14 |
| MFCC2 | 919.16 | 59.94 |
| FricDur | 889.36 | 30.14 |
| MFCC3 | 881.28 | 22.06 |
| Rlloff25 | 877.04 | 17.82 |
| Centroid | 869.91 | 10.69 |
| Maxpos | 862.00 | 2.78 |
| F2LongSlp | 861.54 | 2.32 |
| Rlloff75 | 859.92 | 0.71 |
| MFCC1 | 859.69 | 0.47 |
| F3LongSlp | 859.07 | -0.15 |
| Kurtosis | 858.66 | -0.56 |

For ease of comparison with results of the ablation analysis performed in 7.2.2, a DIC model was also created with measures for just /h/ and /χ/:

Table 17. Deviance information criterion results for measures (/h/ vs. /χ/)

| *Measure* | *DIC* | *Delta DIC* |
|---|---|---|
| MFCC2 | 270.64 | 58.79 |
| MFCC0 | 266.19 | 54.34 |
| FricDur | 257.05 | 45.20 |
| MFCC3 | 236.15 | 24.30 |
| Rlloff25 | 235.26 | 23.41 |
| Centroid | 221.79 | 9.94 |
| NormInt | 219.74 | 7.89 |
| Maxpos | 212.83 | 1.03 |

It can be observed that five of our measures (Kurtosis, Rlloff75, MFCC1, F2LongSlp, F3LongSlp) are absent from Table 17. This is because they were not useful for discriminating between /h/ and /χ/: there was no difference (or only slight improvement of the model) when those measures were removed.

### *7.2.2 Analysis by classification*

We performed a pairwise classifier analysis as described above for Experiment 1 (5.2.4) to determine which feature receives the most weight in distinguishing the three entities /h/, /χ/ and "x". Results are shown in Table 18:

Table 18. Most important feature (measure) distinguishing /h/, /χ/ and "x"

| | /χ/ | "x" |
|---|---|---|
| /h/ | FricDur<br>3.2 | MFCC1<br>3.0 |
| /χ/ | | NormInt<br>3.2 |

Next we trained a classifier only on the /h/ and /χ/ instances, to distinguish those two classes. In comparable cross-validation experiments, we found the classifier achieved 89.8% accuracy, using all acoustic measures, indicating that these measures distinguish these two fricatives with high accuracy. We then performed ablation experiments with our focussed 13 measures as in Experiment 1, to assess the relative contribution of the different measures for discriminating between these categories. Table 19 presents the results of these ablation experiments.

Table 19. Ablation analysis results for classification experiment, /h/ vs. /χ/

| *Measures* | *Accuracy* | *Reduction from all* |
|---|---|---|
| All | 80.3% | |
| FricDur | 75.6% | -4.6% |
| MFCC2 | 72.5% | -7.7% |
| MFCC3 | 70.1% | -10.2% |
| MFCC1 | 68.5% | -11.7% |
| Centroid | 64.2% | -16.1% |
| Kurt | 60.5% | -19.8% |
| F2LongSlp | 56.2% | -24.1% |

Here too a number of measures (MFCC0, F3LongSlp, Maxpos, NormInt, Rlloff25, Rlloff75) are not listed in Table 19. These measures resulted in a classifier that was little better than chance (<56%).

FricDur was the most important measure differentiating /h/ and /χ/, as seen in Table 18, with /χ/ longer in duration than /h/ on average. FricDur, along with MFCC2, MFCC3, were among the top four measures distinguishing /χ/ from /h/ in both the ablation (Table 19) and DIC analyses (Table 17) for /h/ vs. /χ/.

Given the classifier that distinguishes /h/ and /χ/ with high accuracy as above, we then applied it to the instances labelled as "x". Of the 162 instances of "x" presented to the classifier, 119 were classified as /χ/ while the remaining 43 were classified as /h/.

We then analyzed the data by speaker to see if this variability was the result of consistent differences between speakers. Results are shown in Table 20.

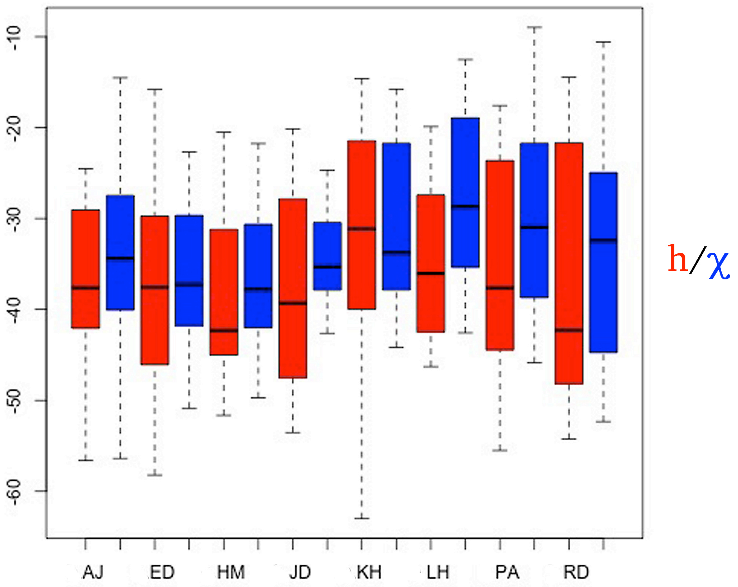
Table 20. Classification of "x" as /h/ or /χ/, by speaker

| | χ | h |
|---|---|---|
| AJ | 18 | 4 |
| ED | 20 | 0 |
| HM | 16 | 4 |
| JD | 18 | 2 |
| KH | 10 | 10 |
| LH | 15 | 5 |
| PA | 6 | 14 |
| RD | 16 | 4 |

It can be seen from Table 20 that all speakers produced some instances of "x" that were interpretable as /χ/. Most speakers also produced some instances of "x" that were interpreted as /h/, except for ED, who produced no /h/. Most speakers produced more instances of /χ/ for "x" than /h/, with the exceptions of KH, who produced equal numbers of /χ/ and /h/, and PA, who produced more instances of /h/ than /χ/.

Some of this uncertainty in the classification of "x" might be attributable to variation in the production of /h/ and /χ/. Figure 21 is a bar graph showing the between-speaker variability seen with one of our measures, MFCC2. It can be seen that for 6 speakers /h/ has lower MFCC2 values than /χ/. However, the two fricatives have roughly equal MFCC2 values for one speaker (ED), and /χ/ has a higher MFCC2 value than /h/ for another (KH).

Figure 21. Bar graph of MFCC2 for /h/ vs. /χ/, by speaker. The horizontal line within each bar represents the mean.



It must be kept in mind that although MFCC2 was the highest ranking measure in the DIC analysis (Table 17) and the second highest measure in the ablation analysis (Table 19), it was only one of the measures we used to compare /χ/ and /h/. Thus /χ/ and /h/ must be distinct for a speaker like ED by some other measure(s).

### *7.3 Discussion*

The 3PL subject prefix is reconstructed as **qʰə-* (Leer 2000) and the areal prefix as **qʰʊ-* (Thompson 1993, Leer 2005), prefixes whose Deg Xinag reflexes are our mystery fricative “x”. Note that the place of articulation of the 3PL and areal prefixes is reconstructed as uvular, not laryngeal. We interpret the variability in our data as a sound change in progress. In Table 21 we show the reflexes of these two prefixes in the Athabaskan languages of western Alaska (see Figure 1). Their place of articulation is uvular or velar in most Alaskan languages but has shifted to laryngeal in Koyukon. Additionally, there is variation between uvular and laryngeal in Holikachuk, the language immediately upriver to Deg Xinag along the Yukon River.

Table 21. Reflexes of 3PL and areal prefixes in Alaskan Athabaskan languages near Deg Xinag

| | source | areal | 3pl subject |
|---|---|---|---|
| Koyukon | Kwaraceius & Jones 2000 | hɔ̆-[16] | hə- |
| Tanana | Tuttle 1998, Tuttle & Hargus 2004 | xʷ- | xə- |
| Dena’ina | Tenenbaum 1978 | qʰə- | qʰə- |
| Ahtna | Kari 1990 | qʰo-, hʷ- | qʰə- |
| Holikachuk | Kari et al. 1978 | χ- ~ h-[17] | χ- ~ h- |
| Upper Kuskokwim | Collins & Petruska 1979 | x(ʊ)- | x- |

---

[16] Thompson et al. 1983 describe ɔ̆ as “similar to the u in ‘but’ only more rounded”. Marlow 2000 describes this vowel as “low back rounded reduced vowel similar to *ou* in English *tough*”.

[17] In Kari et al. 1978, some lexical items contain <x> (=/χ/) and some contain /h/ as Holikachuk reflex of the areal prefix.

It appears that the place of articulation of the 3PL and areal prefixes in Deg Xinag is shifting from uvular to laryngeal[18]. We view this debuccalization as a type of lenition. Recall that there is no contrast between /χ/ and /h/ in the verb prefixes. Root-initial consonants /χ/ and /h/ are consistently stressed, whereas prefix-initial “x” is not stressed. /χ/ is thus leniting to /h/ before an unstressed vowel.

## 8 Conclusions

Our primary goal in this article was to determine which acoustic measure(s) best distinguish(es) the large number of fricatives found in Deg Xinag. Our approach was to start with 52 measures, which we reduced to 13 non-correlated measures. We found that no single measure distinguished all fricatives in pairwise comparisons. Moreover, the top measures distinguishing fricative pairs depended on whether the fricatives were pre- or postvocalic. Prevocalic fricatives were best distinguished by overall energy and energy ratio (our NormInt and MFCC1 measures), whereas the postvocalic fricatives were best distinguished in formant transitions (F2LongSlp) and two different energy measures, MFCC0 and MFCC3.

Statistical analyses on the /θ/ and /ɬ/ subset of data shed some light on a merger of **θ* and **ɬ* as /ɬ/ in closely related Koyukon. These fricatives (and the affricates which they are part of) merged as the louder of the two, /ɬ/. We speculated that this is because the formant transitions for these two fricatives are similar, rendering them perceptually confusable. We further suggested that the directionality of the merger may have been due to the relative intensities of **θ* vs. **ɬ*, and therefore the recoverability of the frication noise, favoring survival of the lateral.

Finally, we employed the methods of the main study to address an outstanding problem in Deg Xinag linguistics, the identity of a fricative which occurs before unstressed vowels that has been variously identified as /χ/ or /h/ by previous Deg Xinag field linguists. Our analyses showed that the mystery fricative has more characteristics of /χ/ than /h/ but that there is variation within and between speakers. We posited that /χ/ is in the process of changing to /h/ in this context, as has occurred in some other Athabaskan languages.

Although we did not provide a simple answer the question of which acoustic measure(s) of fricatives are best, akin to finding the equivalent of F1 and F2 for vowel quality, our results confirm that fricatives are a multi-dimensional phonetic entity, like voice quality. Recall that Garrelek 2020, in his study of the six voice quality distinctions in !Xóõ, began with 18 measures and could only identify eight of these with the highest absolute correlations with three functions returned by a linear discriminant analysis.

Other key methodological aspects of the current study include use of real words, data obtained from native speakers (not trained phoneticians), and fricatives measured in two positions (before and after [a]). Because the Deg Xinag speakers were producing fricative noise that was less distinct than the formant transitions in post-vocalic position, we caution against comparing measures taken from different positions without taking into account a variety of contextual factors that may influence a talker’s pronunciation. We recommend that future studies of fricatives employ a large set of metrics rather than relying on only a handful of (possibly correlated) measures. We also recommend that the measurement set include MFCCs, since like Spinu & Lilley 2016 for Romanian, we

---

[18] These prefixes have also shifted from uvular to laryngeal outside Alaska, in Witsuwit’en (Hargus 2007) and Navajo (Leer 1987).

found that the cepstral coefficients actually did a better job distinguishing the Deg Xinag fricatives than the more widely used spectral moments. At the same time, however, we acknowledge that a limit of our study is that we only sampled measured properties of fricatives from three points within each fricative. A more dynamic approach to measurement as in Reidy 2016 might yield different results.

## 9 Acknowledgements

We thank the following native speakers of the Yukon dialect of Deg Xinag for their participation in this study: Edna Deacon, James Dementi, Raymond Dutchman, Phillip Arrow, the late Katherine Hamilton, the late Lucy Hamilton, the late Alta Jerue, and the late Hannah Maillelle.

This research was supported by NSF (OPP-0137483, to Hargus) and by Conoco Phillips (2006 grant to Michael Krauss).

We thank an anonymous reviewer of *UWWPL* and editor Anna Moroz for comments on previous versions of this article.

## 10 Appendix

Composite FFTs of each fricative, generated as discussed in 5.1.2, (separated by position) are provided in this appendix. Each graph consists of 9 FFTs (one for each speaker in the study, plus an average across speakers) separated by an offset of 10-15 dB on the (unlabeled) vertical axis. The three male speakers (RD, PA, and JD) are near the top of each graph.

Figure 22. /θ/ before /a/

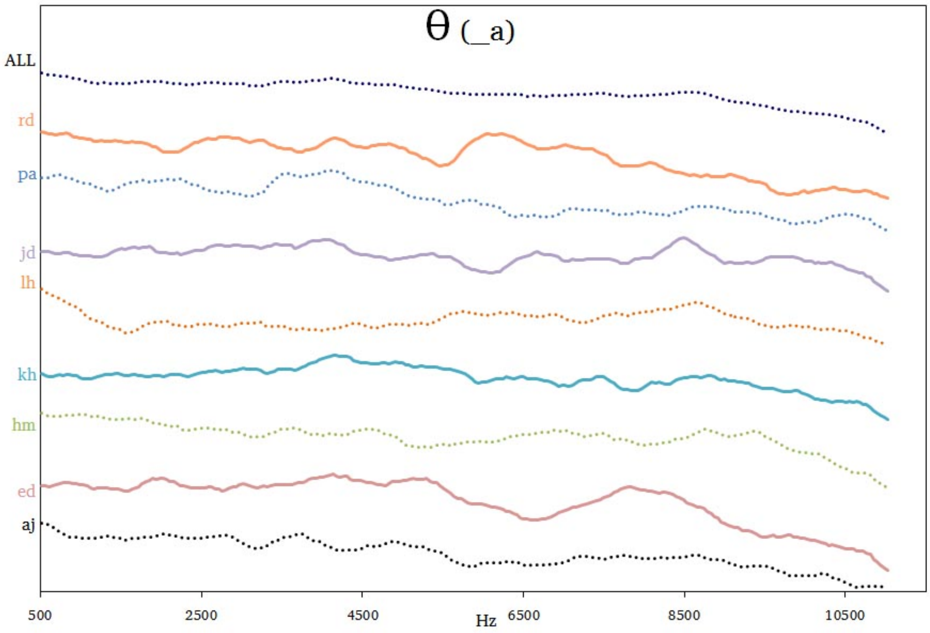


Figure 23. /θ/ after /a/

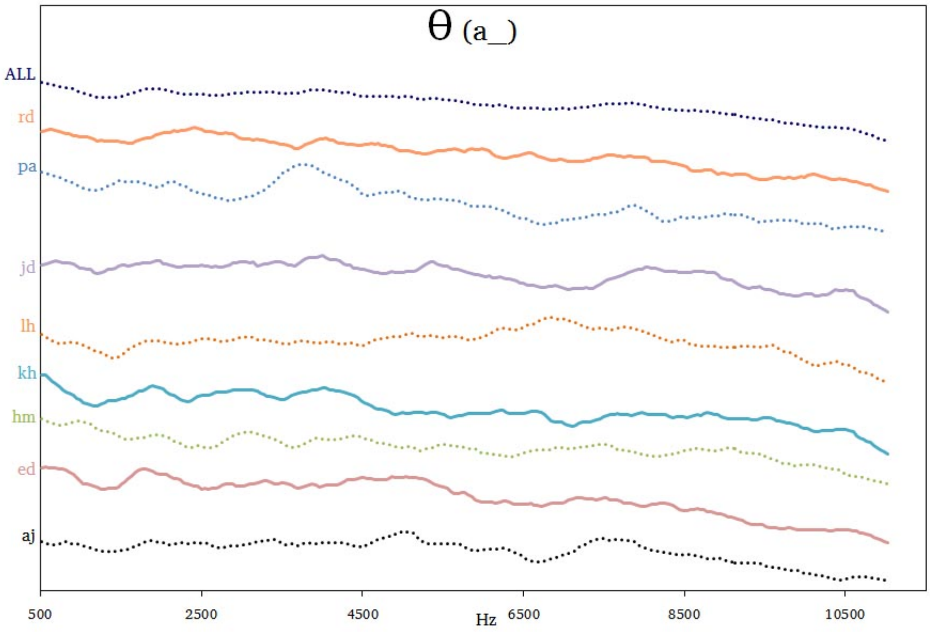


Figure 24. /ɬ/ before /a/

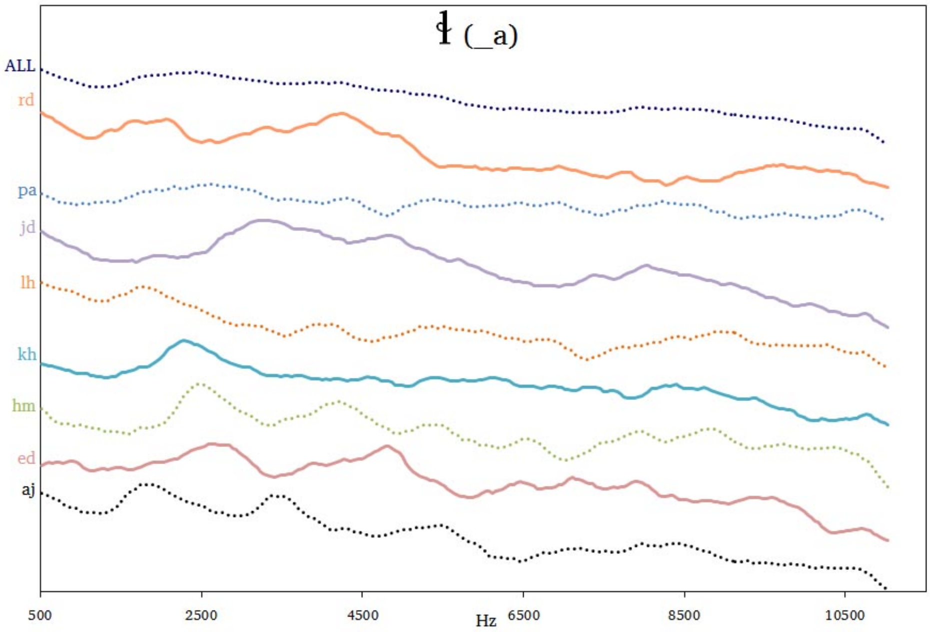


Figure 25. /ɬ/ after /a/

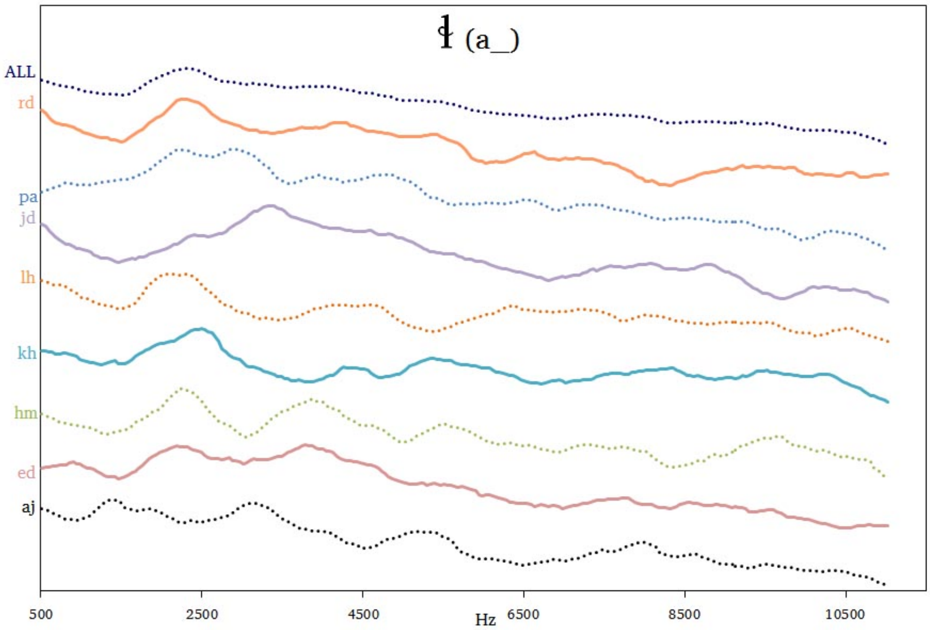


Figure 26. /ʃ/ before /a/

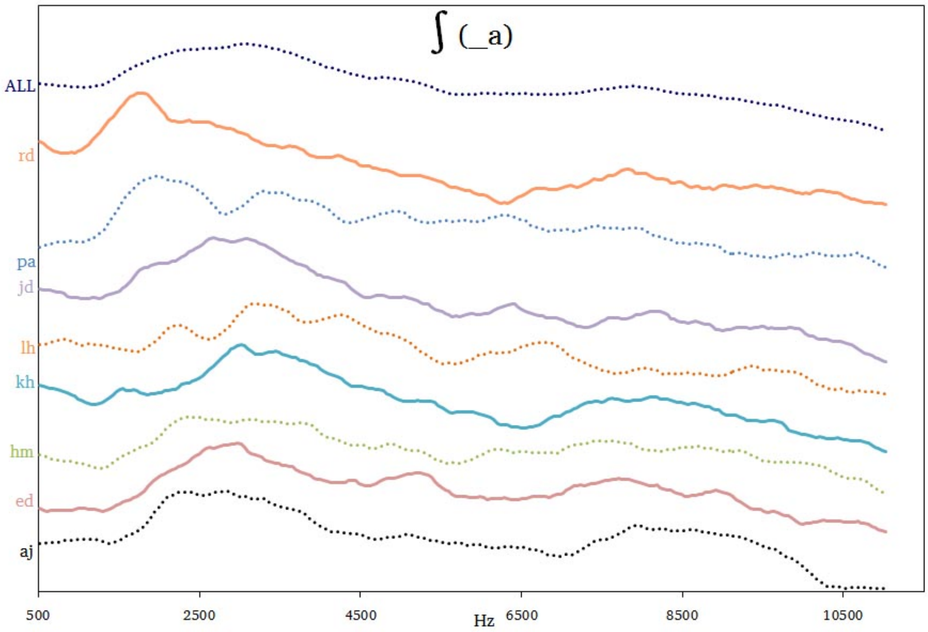


Figure 27. /ʃ/ after /a/

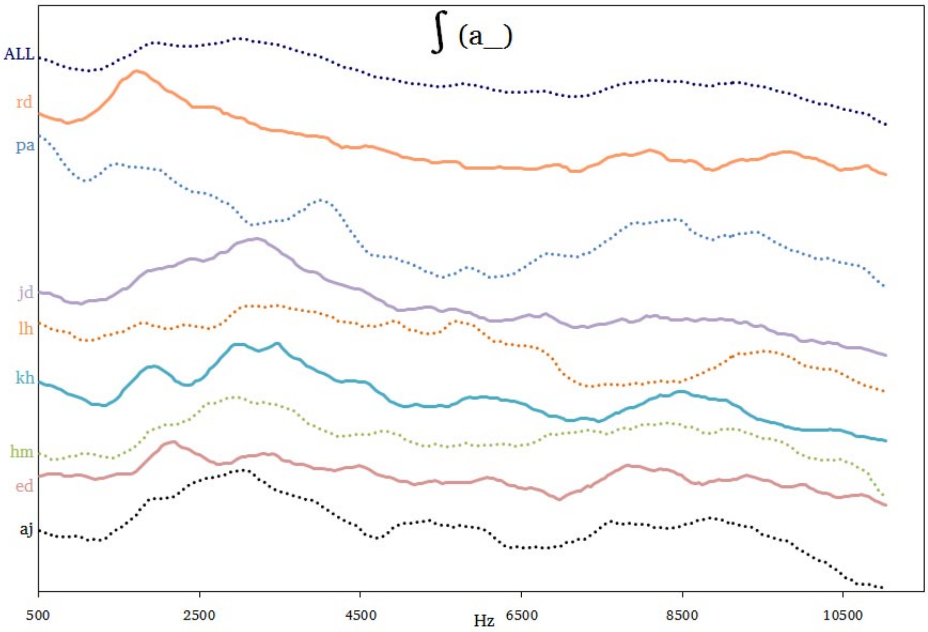


Figure 28. /ʂ/ before /a/

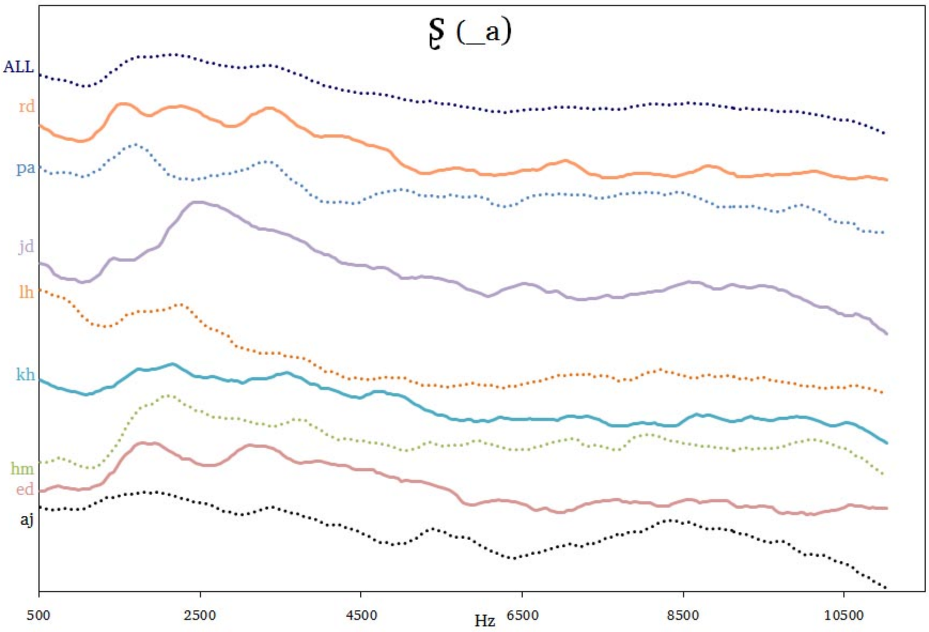


Figure 29. /ʂ/ after /a/

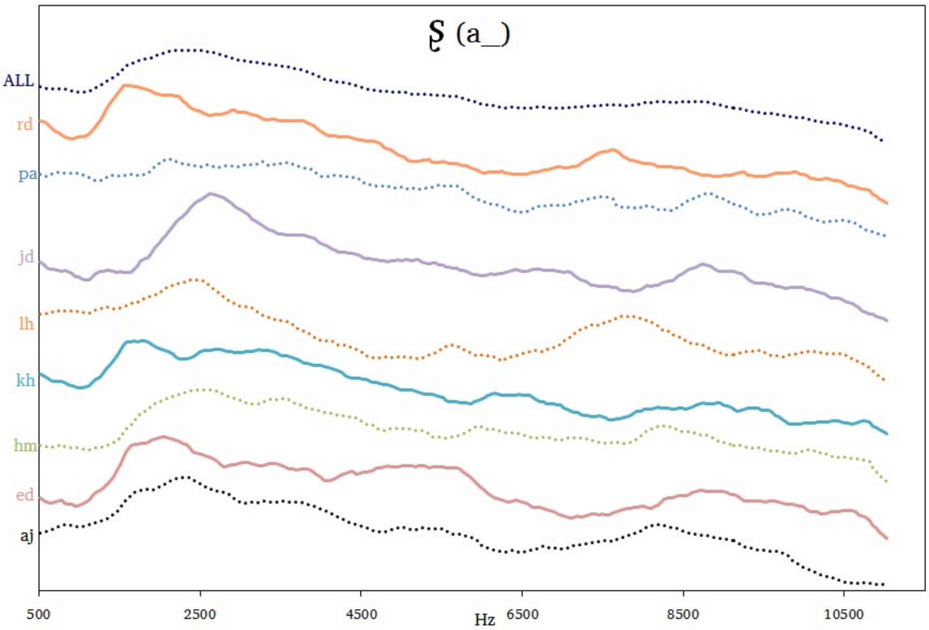


Figure 30. /s/ before /a/

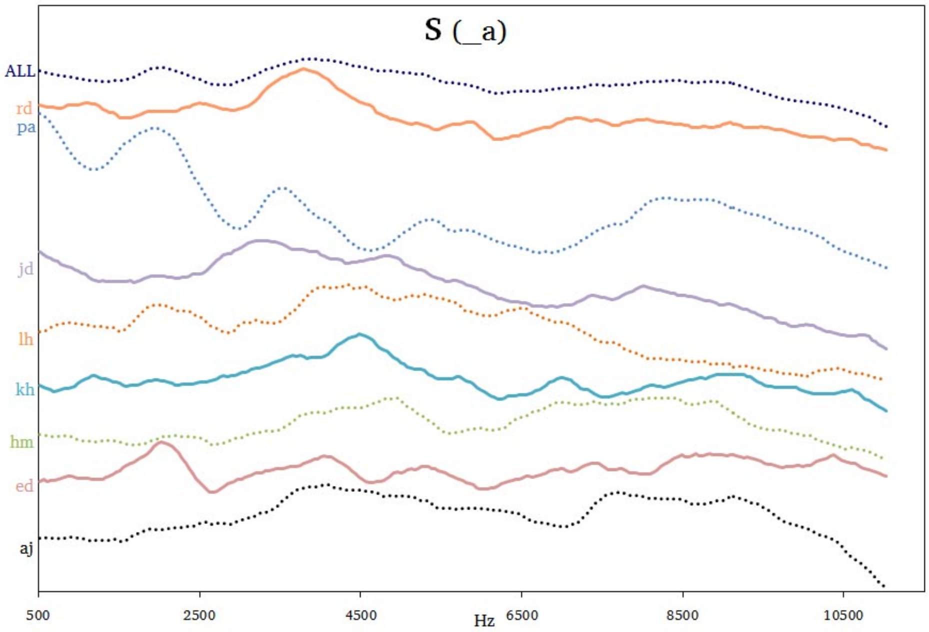


Figure 31. /s/ after /a/

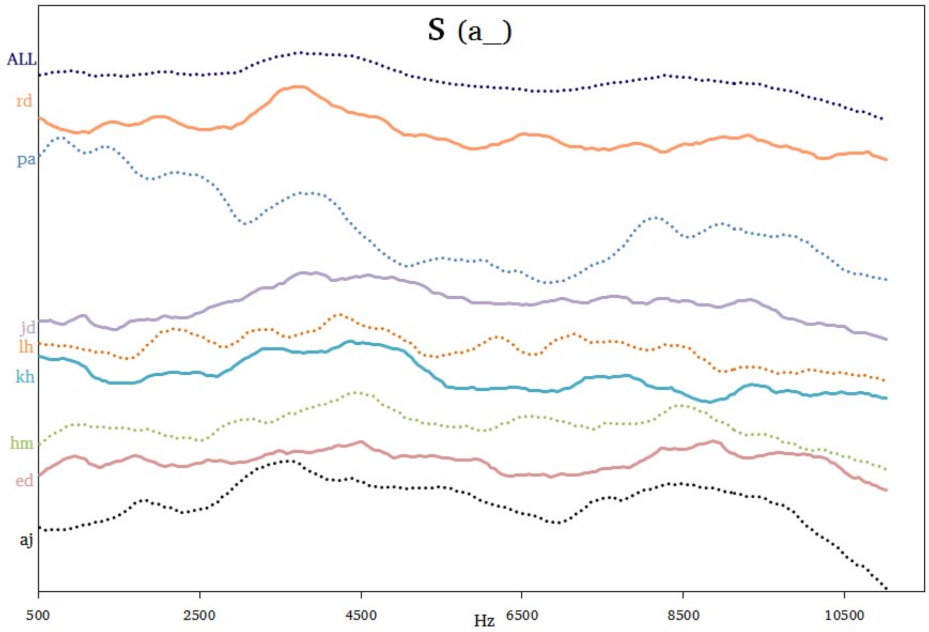


Figure 32. /χ/ before /a/

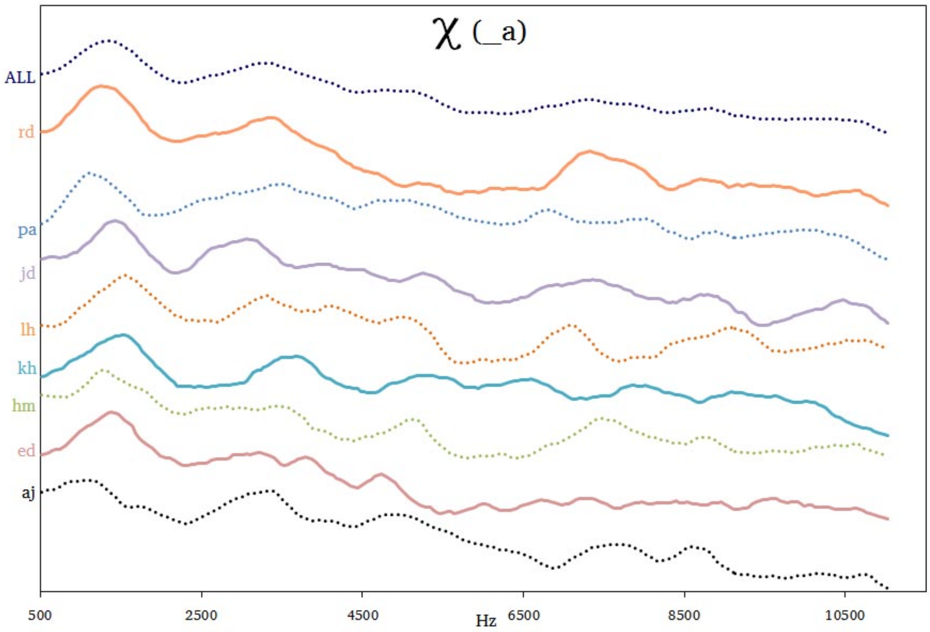


Figure 33. /χ/ after /a/

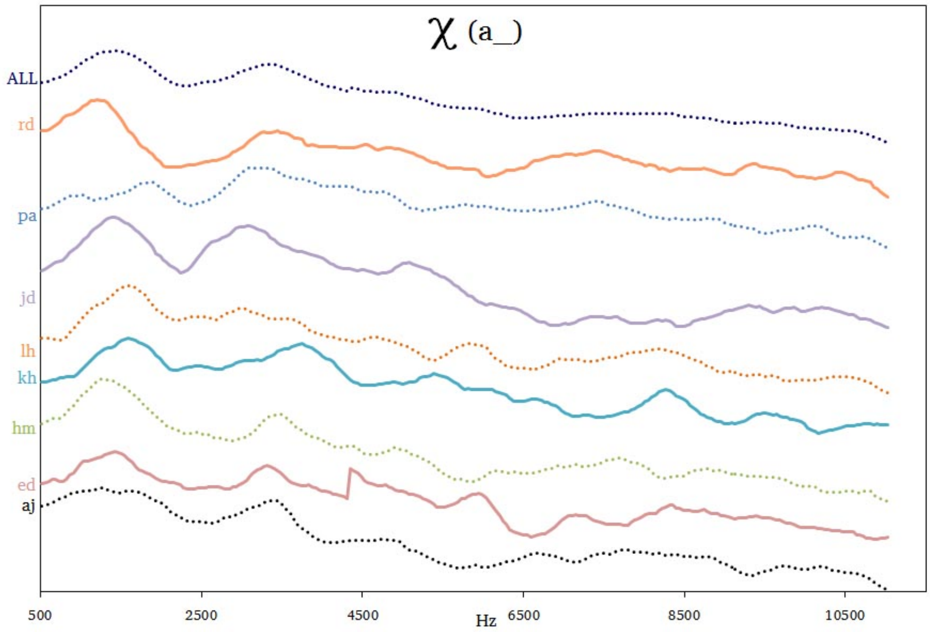


Figure 34. /h/ before /a/

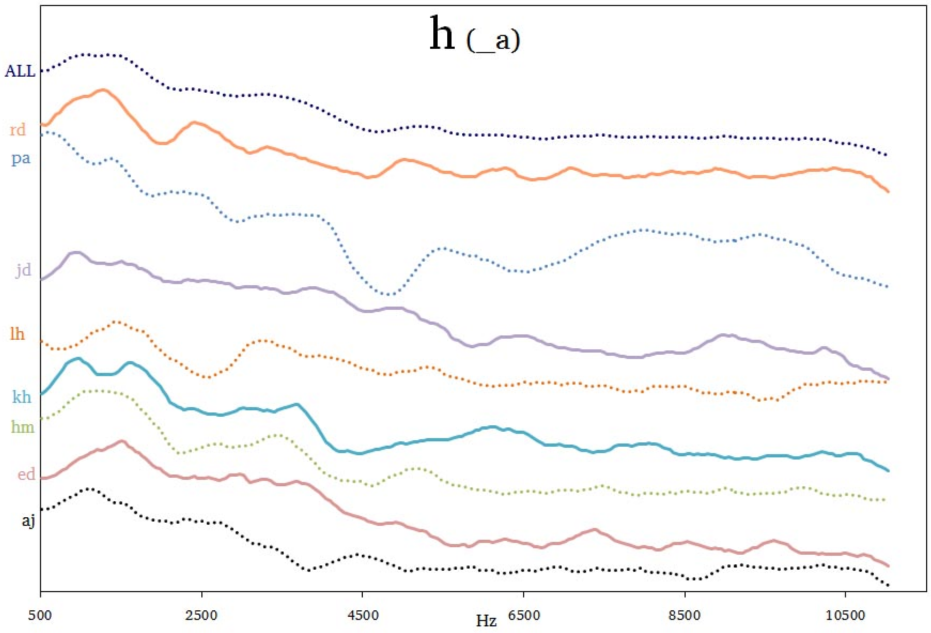
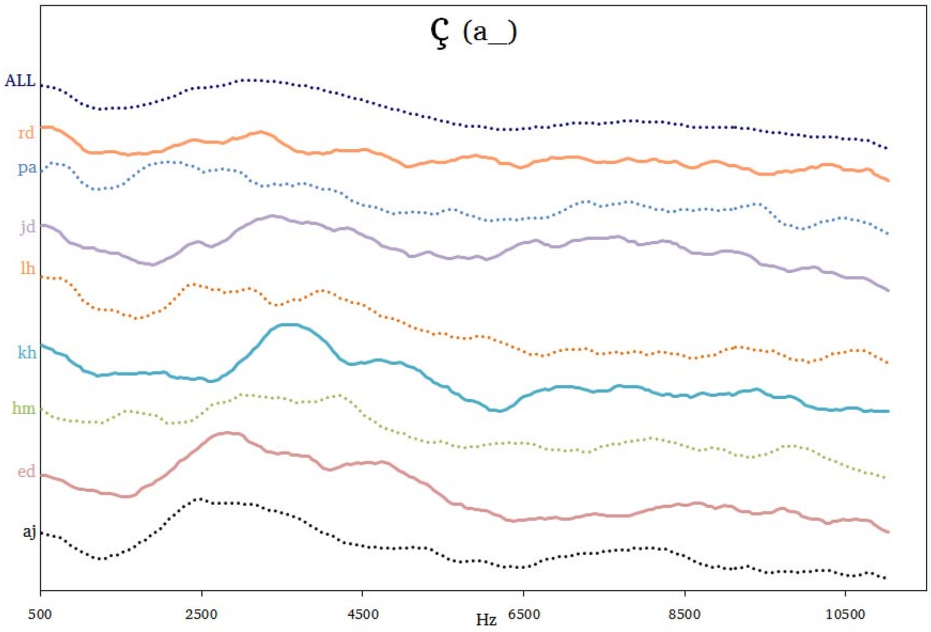


Figure 35. [ç] after /a/



## 11 References

Alwan, Abeer Abdul-Hussain. 1989. Perceptual cues for place of articulation for the voiced pharyngeal and uvular consonants. *Journal of the Acoustical Society of America* 86(2). 549-56.

Behrens, Susan J. & Sheila E. Blumstein. 1988. Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *Journal of Phonetics* 16(3). 295-98.

Bladon, Anthony, Christopher Clark & Katrina Mickey. 1987. Production and perception of sibilant fricatives: Shona data. *Journal of the International Phonetic Association* 17(1). 39-65.

Boersma, Paul & David Weenink. 2013. Praat: Doing phonetics by computer.

Boser, Bernhard E., Isabelle M. Guyon & Vladimir N. Vapnik. 1992. A Training Algorithm for Optimal Margin classifiers. In ACM Special Interest Group for Automata and Computability Theory (ed.) *Proceedings of the Fifth Annual ACM*

*Workshop on Computational Learning Theory: July 27-29, 1992, Pittsburgh, Pennsylvania*, 144-52. New York: Association for Computing Machinery.

Collins, Raymond L. & Betty Petruska. 1979. *Dinak'i (Our Words): Upper Kuskokwim Athabaskan Junior Dictionary*. Anchorage: National Bilingual Materials Development Center, University of Alaska.

Davis, Steven B. & Paul Mermelstein. 1980. Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28(4). 357-66.

Elmazouzi, Leila, Ilham Mounir, Soufyane Mounir, Abdelmajid Farchi & Jamal Elabbadi. 2014. Acoustic Cues to Fricatives Place of Articulation Produced in Moroccan Dialect. *International Journal of Electronics Communication and Computer Engineering* 5(6). 1357-61.

Everitt, Brian S. 1993. *Cluster Analysis*. London: Arnold.

Eyben, Florian, Martin Wöllmer & Björn Schuller. 2010. openSMILE - The Munich Versatile and Fast Open-Source Audio Feature Extractor. *ACM Multimedia 2010. International Conference 25-29th October 2010, Firenze, Italy*. 1459-62.

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(1). 115-23.

Fujimura, Osamu, M.J. Macchi & L.A. Streeter. 1978. Perception of stop consonants with conflicting transitional cues: a cross-linguistic study. *Language and Speech* 21(4). 337-46.

Garrelek, Marc. 2020. Acoustic discriminability of the complex phonation system in !Xóõ. *Phonetica* 77(2). 131.

Goddard, Ives. 1996. Introduction. In Ives Goddard (ed.) *Languages*, 1-16. Washington D.C.: Smithsonian Institution.

Gordon, Matthew. 1995. The Phonetic Structures of Hupa. In Ian Maddieson (ed.) *Fieldwork Studies of Targeted Languages IV (UCLA Working Papers in Phonetics, 93.)*, 1-24. Los Angeles: UCLA Department of Linguistics Phonetics Lab.

Gordon, Matthew, Paul Barthmaier & Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association* 32(2). 141-74.

Gordon, Matthew, Brian Potter, John Dawson, Willem de Reuse & Peter Ladefoged. 2001. Phonetic structures of Western Apache. *International Journal of American Linguistics* 67(4). 415-48.

Hadfield, J.D. 2010. MCMC methods for Multi-response Generalised Linear Mixed Effects Models: The MCMCglmm R package. *Journal of Statistical Software* 33(2). 1-22.

Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann & Ian H. Witten. 2009. The WEKA Data Mining Software: An Update *SIGKDD Explorations* 11(1).

Hargus, Sharon. 2007. *Witsuwit'en Grammar: Phonetics, Phonology and Morphology*. Vancouver: UBC Press.

Hargus, Sharon. 2008. Deg Xinag lateral affricates: Phonetic and historical perspectives. Paper presented to the Society for the Study of the Indigenous Languages of the Americas, Chicago, 2008.

Hargus, Sharon. 2010. Vowel quality and duration in Yukon Deg Xinag. In Siri G. Tuttle & Justin Spence (eds.), *Working Papers in Athabaskan Languages 2009*, 33-73. Fairbanks: Alaska Native Language Center.

Hargus, Sharon. 2011. Effects of morpheme type on Deg Xinag ejectives. *Journal of the Acoustical Society of America* 129(4). 2451.

Hargus, Sharon. 2013. Dene. In Daniel W. Hieber Sharon Hargus, and Edward Vajda (ed.) *Working Papers in Athabaskan (Dene) Languages 2012: Bellingham, Washington August 15-17, 2012*, 13-23. Fairbanks: Alaska Native Language Center.

Hargus, Sharon. 2016. Deg Xinag word-final consonants and voice quality. In Matt Coler, Leo Wetzels & Heriberto Avelino (eds.), *Laryngeal features of Native American languages*, 71-128. Leiden: Brill.

Harris, Katherine S. 1958. Cues for the discrimination of American English fricatives in spoken syllables. *Language and Speech* 1(1). 1-7.

Iskarous, Khalil, Joyce McDonough & D H. Whalen. 2012. A gestural account of the velar fricative in Navajo. *Journal of Laboratory Phonology* 3(1). 195-210.

Jassem, Wiktor. 1968. Acoustic description of voiceless fricatives in terms of spectral parameters. In Wiktor Jassem (ed.) *Speech Analysis and Synthesis*, 189-206. Warsaw: Panstwowe Wydawnictwo Naukowe.

Jesus, Luis M.T. & Christine H. Shadle. 2002. A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics* 30(3). 437-64.

Johnson, Keith. 1993. Acoustic and Auditory Analyses of Xhosa Clicks and Pulmonics. In *UCLA Working Papers in Phonetics*, 37-46. Los Angeles.

Johnson, Keith. 1997. *Acoustic and Auditory Phonetics*. Oxford: Blackwell.

Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108(3). 1252-63.

Kari, James. 1974. Kuskokwim Ingalik Alphabet and Key Words: Alaska Native Language Center, University of Alaska Fairbanks.

Kari, James. 1978. *Deg Xinag: Ingalik Noun Dictionary (Preliminary)*. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks.

Kari, James. 1990. *Ahtna Athabaskan Dictionary*. Fairbanks: ANLC.

Kari, James, Jimmy Alexander, John Deacon & Olga Deacon. 1978. *Holikachuk Noun Dictionary (Preliminary)*. Fairbanks: Alaska Native Language Center.

Kay Elemetrics Corp. 2004. Multi-Speech and CSL Software.

Krauss, Michael. 1962. Ingalik: Alaska Native Language Archive, University of Alaska Fairbanks.

Krauss, Michael. 1964. Proto-Athapaskan-Eyak and the Problem of Na-Dene: The Phonology. *International Journal of American Linguistics* 30(2). 118-31.

Krauss, Michael. 1974. Native Peoples and Languages of Alaska. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks.

Krauss, Michael. 2000. Koyukon Dialectology and Its Relationship to Other Athabaskan Languages. In Jules Jetté & Eliza Jones (eds.), *Koyukon Dictionary*, l-lxv. Fairbanks: Alaska Native Language Center, University of Alaska Fairbanks.

Krauss, Michael & Victor Golla. 1981. Northern Athapaskan Languages. In June Helm (ed.) *Handbook of North American Indians, vol. 6. Subarctic*, 67-85. Washington D.C.: Smithsonian Institution.

Krauss, Michael, Gary Holton, Jim Kerr & Colin T. West. 2011. Indigenous Peoples and Languages of Alaska. Fairbanks and Anchorage: Alaska Native Language Center and UAA Institute of Social and Economic Research.

Kwaraceius, Joe & Eliza Jones. 2000. Koyukon Verb Paradigms. In Jules Jetté & Eliza Jones (eds.), *Koyukon Athabaskan Dictionary*, 777-804. Fairbanks: Alaska Native Language Center.

Leer, Jeff. 1987. Navajo and Comparative Athapaskan. In Robert Young & William Morgan (eds.), *The Navajo Language: A Grammar and Colloquial Dictionary*, 264-301. Albuquerque: University of New Mexico Press.

Leer, Jeff. 1996. The Evolution of the Stem Syllable in Gwich'in. In Eloise Jelinek, Sally Midgette, Keren Rice & Leslie Saxon (eds.), *Athabaskan Language Studies: Essays in Honor of Robert W. Young*, 193-234. Albuquerque, NM: University of New Mexico Press.

Leer, Jeff. 2000. The Negative/Irrealis Category in Athabaskan-Eyak-Tlingit. In Theodore B. Fernald & Paul R. Platero (eds.), *The Athabaskan Languages: Perspectives on a Native American Language Family*, 51-72. Oxford: Oxford University Press.

Leer, Jeff. 2005. How stress shapes the stem-suffix complex in Athabaskan. In Sharon Hargus & Keren Rice (eds.), *Athabaskan Prosody*, 278-318. Amsterdam and Philadelphia: John Benjamins.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

Malécot, André. 1956. Acoustic Cues for Nasal Consonants: An Experimental Study Involving a Tape-Splicing Technique. *Language* 32(2). 274-78.

Maniwa, Kazumi, Allard Jongman & Travis Wade. 2009. Acoustic characteristics of clearly spoken English fricatives. *Journal of the Acoustical Society of America* 125(6). 3962-73.

Marlow, Patrick. 2000. Koyukon Sound System and Orthography. In Jules Jetté & Eliza Jones (eds.), *Koyukon Athabaskan Dictionary*, lxvi-lxxi. Fairbanks: Alaska Native Language Center.

McDonough, Joyce. 2003. *The Navajo Sound System*. Dordrecht: Kluwer.

Moran, Steven, Daniel McCloy & Richard Wright. 2014. *PHOIBLE Online*. http://phoible.org.

Munson, Benjamin. 2001. A method for studying variability in fricatives using dynamic measures of spectral mean. *Journal of the Acoustical Society of America* 110(2). 1203-06.

Nartey, Jonas N.A. 1982. *On Fricative Phones and Phonemes*. Los Angeles: UCLA PhD dissertation.

Norlin, K. 1983. Acoustic analysis of fricatives in Cairo Arabic. *Working Papers (Lund University Department of Linguistics)* 25. 113-37.

O’Shaughnessy, Douglas. 2008. Invited paper: Automatic speech recognition: History, methods and challenges. *Pattern Recognition* 41(10). 2965–79.

Osgood, Cornelius. 1940. *Ingalik material culture*. New Haven: Yale University Press.

Panfili, Laura. 2018. *Cross-linguistic acoustic characteristics of phonation: A machine learning approach*: University of Washington PhD dissertation.

Reidy, Patrick F. 2016. Spectral dynamics of sibilant fricatives are contrastive and language specific. *Journal of the Acoustical Society of America* 140(4). 2518-29.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin & Angelika van der Linde. 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B* 64(4). 583–639.

Spinu, Laura & Jason Lilley. 2016. A comparison of spectral coefficients and spectral moments in the classification of Romanian fricatives *Journal of Phonetics* 57. 40-58.

Strevens, Peter. 1960. Spectra of fricative noise in human speech. *Language and Speech* 3(1). 32-49.

Styler, Will. 2015. *On the Acoustical and Perceptual Features of Vowel Nasality*: University of Colorado PhD dissertation.

Svantesson, Jan-Olof. 1986. Acoustic analysis of Chinese fricatives and affricates. *Journal of Chinese Linguistics* 14(1). 53-70.

Tenenbaum, Joan. 1978. *The Morphology and Semantics of the Tanaina Verb*. New York: Columbia University PhD dissertation.

Thompson, Chad. 1993. The Areal Prefix hu-- in Koyukon Athapaskan. *International Journal of American Linguistics* 59(3). 315-33.

Thompson, Chad, Melissa Axelrod & Eliza Jones. 1983. *Han Zaadlitl'ee: Yukon Koyukuk School District Koyukon Language Curriculum (Koyukon Language Curriculum: Scope and Sequence)*. Nenana, AK: Yukon Koyukuk School District.

Toloşi, Laura & Thomas Lengauer. 2011. Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics* 27(14). 1986–94.

Tronnier, M. & M. Dantsuji. 1993. Some acoustic characteristics of glottal and palatal fricatives in Japanese and German. In Jan-Olof Svantesson (ed.) *Working Papers (Lund University Department of Linguistics)*, 213-28.

Tuttle, Siri G. 1998. *Metrical and tonal structures in Tanana Athabaskan*. Seattle: University of Washington PhD dissertation.

Tuttle, Siri G. & Sharon Hargus. 2004. Explaining variability in affix order: The Athabaskan areal and third person prefixes. In Gary Holton & Siri G. Tuttle (eds.), *Working Papers in Athabaskan Linguistics 2004*, 70-99. Fairbanks: ANLC.

van Dommelen, Wim A. 2003. An acoustic analysis of Norwegian /ç/ and /ʃ/ as spoken by young people. *Journal of the International Phonetic Association* 33(2). 131-41.

Ward, Joe H., Jr. 1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* 58. 236-44.

Wright, Richard. 1999. Tsou consonant clusters and auditory cue preservation. *Selected Papers from the Eighth International Conference on Austronesian Linguistics: Symposium Series of the Institute of Linguistics, Academia Sinica* 1. 277-312.

Wright, Richard. 2001. Perceptual cues in contrast maintenance. In Elizabeth Hume & Keith Johnson (eds.), *The Role of Speech Perception in Phonology*, 251-77. San Diego: Academic Press.