Interspeech 2024
1-5 September 2024, Kos, Greece

Affricates in Lushootseed

Ted Kye[1]

[1]University of Washington
kyeted20@gmail.com

Abstract

In this study, I conduct the first acoustic analysis of affricates in Lushootseed (Coast Salish branch). The findings reveal that several acoustic measurements, such as VOT, release frication duration, intensity, spectral moments (Center of Gravity), and voice onset quality, characterize affricate contrasts with respect to their place of articulation and laryngeal type (i.e., voiced, voiceless, ejective). This study raises questions concerning (a) the acoustic-articulatory correlates of affricates, (b) typological classification of ejective affricates, and (c) the diachronic status of voiced affricates in Coast Salish languages.

Index Terms: affricates, ejectives, typology, Coast Salish

1. Introduction

Like other Salish languages, Lushootseed (a Coast Salish language spoken in the Puget Sound region of the Pacific Northwest) is known for having many obstruents in its phonemic inventory. Of the 37 consonants in Lushootseed, 31 are obstruents (i.e., 84% of consonants). Seven of these obstruents are affricates: Namely, /dz ts ts' dʒ tʃ tʃ' tɬ'/ ‹dᶻ c c̓ j č č̓ λ̓›. An interesting feature of Lushootseed is that Lushootseed has a set of voiced affricates /dz dʒ/ ‹dᶻ j›, which is not found in many Coast Salish languages. Although some research has examined obstruents for other Salish languages [1], [2], [3], only a few conducted a rigorous acoustic analysis of obstruents in these languages [2], [3]. Affricates are of particular interest to document because, despite being a ubiquitous class of sound across Salish languages, very little is known about affricates phonetically in Coast Salish. In this paper, I provide the first acoustic phonetic analysis of affricates in Lushootseed.

Cross-linguistically, ejectives are known to differ in several acoustic dimensions [3], [4], [5], [6], [7], [8], [9], [10], [11], [12]. According to Lindau [13] and Kingston [4], [5], there are typological differences in the realization of ejectives. Lindau [13] and Kingston [4], [5] proposed a model that classifies distinct types of ejectives, where ejectives can be classified as either stiff or slack. Stiff ejectives are characterized by a silent period between the consonant release and voice onset, which results in a long Voice Onset Time (VOT). Moreover, stiff ejectives are achieved with increased longitudinal tension and medial compression of the vocal folds, which results in a raised f0 at voice onset and a tense or modal voice onset quality. Other characteristics of stiff ejectives include a sharp rise in the amplitude of the following vowel and a relatively intense burst. Slack ejectives, on the other hand, are characterized by a short VOT and little longitudinal tension, resulting in a depressed f0 and creaky quality at voice onset. Other characteristics of slack ejectives include a steep rise in amplitude of the following vowel and a normal burst intensity. Table 1 provides a summary of the Lindau-Kingston model of stiff and slack ejectives.

Table 1: *Proposed ejective typology following the Lindau-Kingston model of ejectives.*

Acoustic Dimensions	Stiff ejective	Slack ejective
VOT	Long	Short
f0 perturbation	Raised	Depressed
Voice quality	Modal or tense	Creaky
Rise time	Fast	Slow
Burst intensity	Intense	Normal

Although typological research focuses on ejective stops, few studies have focused on ejective affricates. Percival [11], [12] found that ejective affricates have acoustic correlates that differ from those of ejective stops. This leads to the question of how the acoustic dimensions used to characterize ejective types apply for affricates. Moreover, it is typologically rare to find languages with a three-way laryngeal contrast in affricates. Another goal of this paper is to address these questions by conducting an acoustic analysis of the three-way laryngeal affricate contrast (voiced, voiceless, ejective) in Lushootseed. Addressing these questions in this study can contribute towards our understanding of the typology of ejectives (and affricates in general) across languages and how the acoustics of the three-way laryngeal contrast of affricates is realized.

2. Methods

2.1. Recordings

Because there are no first language speakers of Lushootseed remaining, recordings from the archives were examined. These recordings come from the Metcalf Collection (part of the Burke Museum's Special Collections), which are recordings that date to the 1950's and were produced by the musicologist Leon Metcalf. These recordings were digitized with a sampling rate of 44.1kHz with a 32-bit depth. 11 of these recordings were examined (total duration of approximately 2 hours and 3 minutes), which are recordings of traditional Salish myths and private correspondences (i.e., these are recordings of connected (spontaneous) speech).

2.2. Speakers

Two speakers were examined: Annie Jack (AJ) and Martha Lamont (ML). Annie Jack was an elder speaker of Southern Lushootseed. She was born around the 1870's or 1880's and lived in the Muckleshoot Tribal reservation. Her living descendants include Denise Bill (great-granddaughter), Willard Bill, Jr. (great-grandson), Elise Bill-Gerrish (great-great-granddaughter and daughter of Denise Bill), and Justice Bill (great-great-grandson and son of Willard Bill, Jr.). Seven recordings (duration of approximately 67 minutes) of Annie Jack were examined. Martha Lamont was an elder speaker of

10.21437/Interspeech.2024-1607

Northern Lushootseed who was born around the 1880's or 1890's. She was a Snohomish speaker who lived in Tulalip. Her living descendants include Hank Williams (grandson) and Hank's descendants. Four recordings (duration of approximately 51 minutes) of Martha Lamont were examined.

2.3. Measurements and Sampling Procedure

Duration, intensity, spectral moments, and voice onset quality measurements were extracted using Praat [14]. For duration measurements, following [15], [16], Voice Onset Time (VOT) and the duration of the affricates' release frication were measured. Decibels of Sound Pressure Level (dB SPL) was measured along the affricates release frication at five time points (10%, 30%, 50%, 70%, 90%).

For spectral moments, Center of Gravity (CoG) was measured. The method used comes from Shadle [17] and Forrest et al. [18]. Following Shadle [17], time-averaging was used by extracting multiple Discrete Fourier Transform (DFT) power specturms along the release frication duration, where each window was extracted with a length of 5ms. Each of these windowed DFTs were then averaged through a matrix of intensity and sampling frequencies for each token. CoG was then calculated from the time-averaged spectrums.

To examine the voice onset quality of affricates, fundamental frequency (f0), jitter, and amplitude rise (i.e., intensity difference) were examined. Average f0 was obtained at the voice onset (10% point) and the vowel midpoint (50% point). f0 was measured by interpolating the pitch track and then extracting the f0 from the voicing component of the vowel. Jitter was obtained from 30ms windows at voice onset and the vowels midpoint. Voiced period marks (pulses) were generated in Praat. These marks were inspected for errors and corrected by hand if necessary. Jitter is the variation in the duration of successive fundamental frequency cycles [19]. The mean percent jitter ratio was calculated by taking the average absolute difference between consecutive pulses divided by the average period and multiplied by 100 [20]. Intensity (in dB SPL) was obtained at the voice onset (10% point) and at the vowels' peak amplitude. Jitter, f0, and intensity were normalized, as shown in Table 2. f0 perturbation was calculated by subtracting the mean f0 at vowel onset from the mean f0 at vowel midpoint. Similarly, jitter perturbation was calculated by subtracting the mean jitter at voice onset by the mean jitter at the vowel midpoint. RMS amplitude difference was calculated by subtracting the intensity (in dB SPL) at the vowel's peak amplitude from the intensity at the vowel onset.

Table 2: *Formulas for normalized measurements.*

Measurement	Formula
f0 perturbation	(f0 at onset) – (f0 at midpoint)
jitter perturbation	(jitter at onset) – (jitter at midpoint)
RMS amplitude difference	(RMS energy at vowel peak) – (RMS energy at vowel onset)

All seven affricates /dz ts ts' dʒ tʃ tʃ' tɬ'/ were examined word-initially and word-medially (intervocalically). Figure 1 illustrates the ejective alveolar affricate /ts'/ ‹c̓› in (a) word-initial position, and (b) word-medial position. At least 20 or more observations of each affricate (10 or more from each speaker) were used for the analysis (see Table 3). However, fewer than 20 tokens of /dʒ/ ‹ǰ› were obtained because /dʒ/ ‹ǰ› occurred rarely in the data.

Table 3: *Number of tokens for each affricate.*

Speaker	dz	ts	ts'	dʒ	tʃ	tʃ'	tɬ'
Annie Jack	13	76	16	6	65	31	36
Martha Lamont	10	30	11	5	59	23	54
Total	23	106	27	11	124	54	90

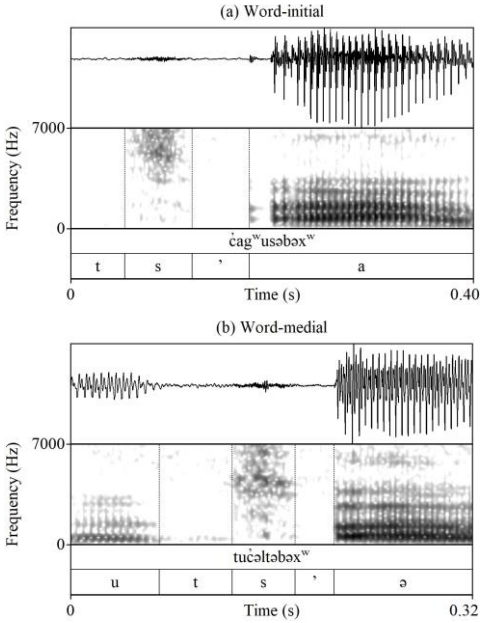


Figure 1: *Ejective alveolar affricate /ts'/ ‹c̓› in (a) word-initial position, and (b) word-medial position.*

2.4. Analysis

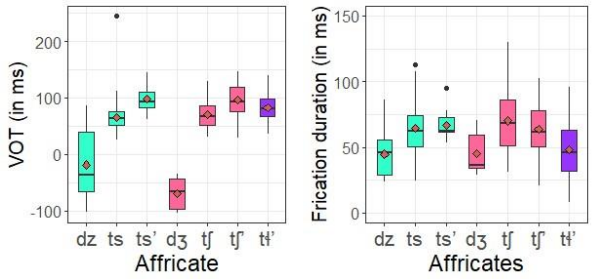
The data was fit into a linear mixed effects model to analyze the contrast of each affricate with respect to its place of articulation and laryngeal type (i.e., voiced, voiceless, ejective). The package *lme4* was used [21], where a *t*-value of 2 or greater was considered statistically significant. Each acoustic measure was used as dependent variables, while place of articulation (i.e., alveolar, post-alveolar, lateral) and laryngeal type (i.e., voiced, voiceless, ejective) were used as fixed effects. Speakers and words were treated as random effects, where place of articulation was used as a random slope for measures of duration, intensity, and spectral moment; while laryngeal type was used as a random slope for measures of duration and voice onset quality. The following equation was used to analyze the contrast across affricates with respect to its acoustic dimensions:

(1) *Measurements ~ Place + Laryngeal Type + (Place, Laryngeal Type | Speaker) + (1 | Word)*

3. Results

3.1. Duration measurements

The VOT and release frication duration for each affricate is illustrated in Figure 2. For VOT across each laryngeal type, ejective affricates are significantly longer than voiced (β = 124.753, *SE* = 7.219, *t* = 17.281) and voiceless (β = 26.118, *SE* = 4.939, *t* = 5.289) affricates. It should be noted that there were instances where /dz/ had a positive VOT, which can be observed in Figure 2a (where the VOT for /dz/ spans across both positive and negative values along the y-axis). The realization of /dz/ ‹dᶻ› as devoiced (i.e., positive VOT) usually occurred in word-initial position.



(a) VOT (b) Frication duration

Figure 2: (a) *VOT (in ms), and (b) release frication duration (in ms), for affricates in Lushootseed, where green is alveolar, red is post-alveolar, and purple is lateral.*

Voiced affricates have a significantly shorter release frication duration than voiceless (β = -21.615, *SE* = 4.816, *t* = -4.488) and ejective (β = -11.852, *SE* = 5.283, *t* = -2.243) affricates. Although the release frication duration for ejective affricates is significantly shorter than voiceless (β = -9.763, *SE* = 3.626, *t* = -2.693) affricates, this was not always observed with respect to place of articulation and across speakers, where the ejective alveolar affricate /ts'/ ‹c̓› for the speaker ML appeared to be (on average) greater than the voiceless alveolar affricate /ts/ ‹c›, while they are relatively the same for the speaker AJ.

3.2. Intensity characteristics

Figure 3 illustrates intensity at five time points for each affricate with respect to their laryngeal type and place of articulation. As Figure 3 illustrates, there is a rising intensity at the 90% time point for voiced affricates, a falling intensity at the 90% time point for ejectives, and in-between for voiceless affricates. Voiced affricates have a significantly greater intensity at 90% than voiceless affricates (β = 3.867, *SE* = 0.879, *t* = 3.867) and voiceless affricates signficantly greater than ejectives (β = 6.325, *SE* = 1.366, *t* = 4.632). Interestingly, the lateral ejective affricate is characterized by a falling intensity throughout the entire frication period (from 10% to 90%). This contrasts from alveolar and post-alveolar affricates, where there is a rising intensity from 10% up to about 50% or 70%, and then a fall.

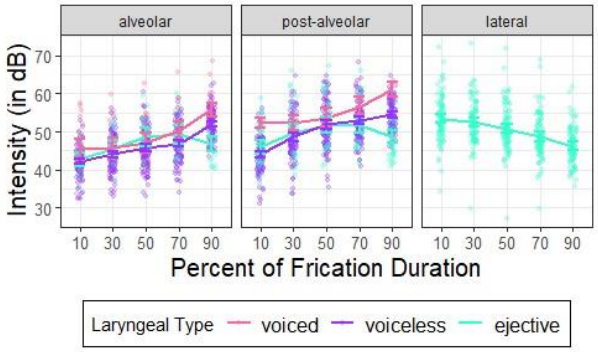


Figure 3: *Intensity of affricate's release frication across five time points for each laryngeal type with respect to their places of articulation.*

3.3. Spectral moment: Center of Gravity (CoG)

CoG for each affricate can be observed in Figure 4. The CoG for the alveolar place of articulation was significantly greater than post-alveolar (β = 645.306, *SE* = 99.337, *t* = 6.496), and post-alveolar greater than lateral (β = 1061.082, *SE* = 295.226, *t* = 3.594).

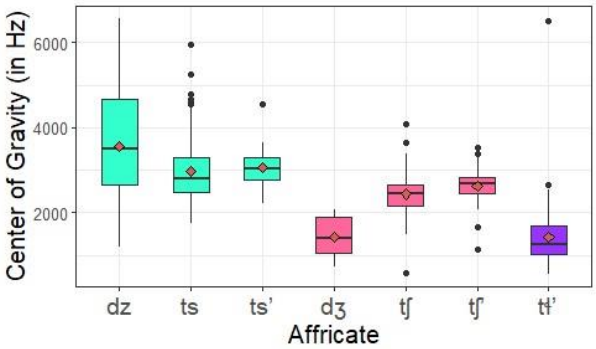


Figure 4: *Center of Gravity (in Hz) for each affricate.*

3.4. Voice onset quality

3.4.1. f0 perturbation

Figure 5 illustrates f0 perturbation for the affricate's laryngeal types, where ejectives showed depressed f0, voiceless raised, and voiced static. The f0 perturbation for voiceless affricates was significantly greater than voiced (β = 12.699, *SE* = 3.383, *t* = 3.754), and voiced significantly greater than ejectives (β = 9.092, *SE* = 3.283, *t* = 2.770).

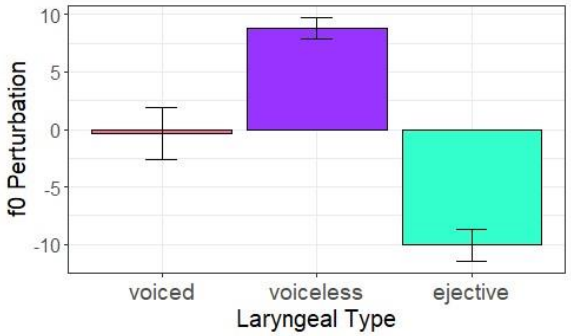


Figure 5: *f0 perturbation across each affricates laryngeal type.*

3.4.2. Jitter perturbation

Figure 6 illustrates that ejective affricates have the highest jitter perturbation compared to voiced and voiceless affricates, which suggests that the voice onset quality for ejectives is creaky voiced. Moreover, the jitter perturbation for voiceless affricates is greater than voiced affricates.

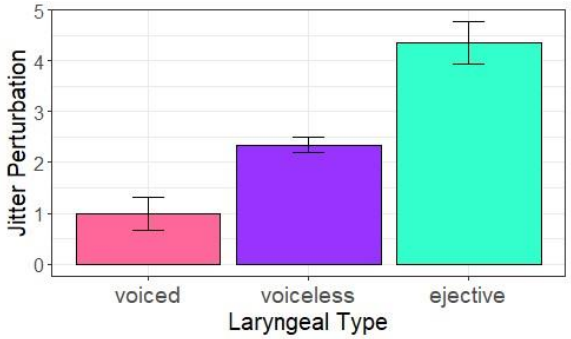


Figure 6: *Jitter perturbation across each affricates laryngeal type.*

Ejectives have a significantly greater jitter perturbation than voiceless affricates (β = 2.179, *SE* = 1.039, *t* = 2.098), and voiceless affricates greater than voiced (β = 1.306, *SE* = 0.585, *t* = 2.232).

3.4.3. RMS amplitude difference

Figure 7 reveals that the intensity difference (i.e., the difference between the amplitude at voice onset and the vowels peak amplitude) is greater for ejectives. The intensity difference for ejective affricates was significantly greater than voiced (β = 1.665, *SE* = 0.6621, *t* = 2.515) and voiceless (β = 1.364, *SE* = 0.4203, *t* = 3.245) affricates, but voiceless affricates did not significantly differ from voiced.

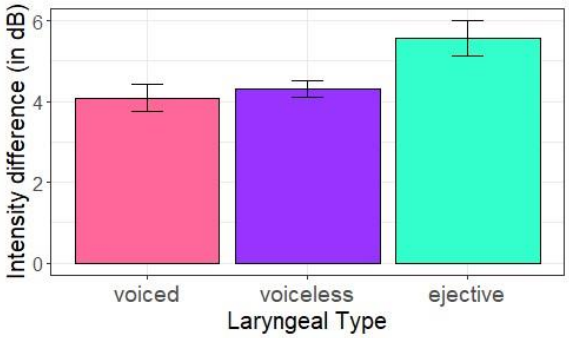


Figure 7: *Intensity difference across each affricates laryngeal type.*

4. Discussion

The current findings reveal that there are several acoustic dimensions that characterize affricate contrasts in Lushootseed. Duration, Center of Gravity (CoG), intensity, and voice onset quality are good measures that characterize the affricate contrast with respect to their place of articulation and laryngeal type. The findings pose interesting problems towards (a) the acoustic-articulatory correlates of affricates, (b) typological classification of ejectives, and (c) diachronic status of voiced affricates in Coast Salish.

4.1. Acoustic-articulatory correlates of affricates

CoG was found to differentiate affricates in Lushootseed, suggesting differences in place of articulation as is typically found across languages [22], where alveolar > post-alveolar > lateral. CoG for post-alveolars occurred in the low-mid frequency range. However, CoG for the alveolars is relatively lower than CoG of alveolars typically observed in other languages [18], [22], [23]. Alveolars tend to show peaks well above 7kHz, [24], [25], [26]. Unfortunately, these peaks were not traceable in the speech signal from some of these recordings because of the upper frequency cutoff of the microphone signal, where the cutoff could be observed near 7kHz. This may have obscured measures of CoG for the voiceless alveolar affricate. However, relative differences in the affricates' place of articulation can be inferred.

The current findings reveal that the lateral place of articulation differs in CoG and intensity when compared with laterals in other languages, where CoG for lateral affricates and fricatives tend to be closer to the post-alveolar place of articulation [11], [12], [22], [23], [27]. However, the CoG for the lateral ejective affricate was significantly lower than post-alveolars in Lushootseed. It is possible that the oral release for the lateral ejective affricate was made posteriorly along the sides of the palate. The more posterior the constriction, the lower the frequency of the main peak [17], [22]. However, it is also possible that the low CoG was due to the length of the side channels during the release. If the constriction was made more posteriorly along the sides of the palate, this would create a longer channel along the buccal cavity for air to flow through following the release, which may correlate with the low CoG. This suggests that the lateral ejective affricate in Lushootseed may involve a different articulatory target when compared with other languages.

4.2. Typological classification of ejectives

The current findings support the view that there are more variable realizations of ejectives across languages and that ejectives cannot simply be divided into either stiff or slack, [3], [6], [7], [8], [9], [10]. There are several problems with classifying ejective affricates under the Lindau-Kingston model in Lushootseed. The Lindau-Kingston model assumes that VOT is longer for stiff ejectives and shorter for slack ejectives. The VOT's of ejective affricates in Lushootseed is longer than plain (voiceless pulmonic) affricates, which may suggest that ejective affricates in Lushootseed are stiff. However, the voice onset quality of ejective affricates has properties of slack ejectives (i.e., depressed f0, creaky voice onset, and greater intensity difference at voice onset). This mixture of stiff and slack ejectives reveals that Lushootseed ejective affricates cannot be classified exclusively as either stiff or slack.

Given the consistently falling intensity of the release frication across ejectives in the data, intensity characteristics of the affricates release frication across laryngeal types do not seem to vary along stiff vs. slack parameters. Rather, the consistency suggests physiological constraints in the glottal gestural timing of affricates when transitioning into a voiced source vs. a silent source. Ejectives involve a closed glottis during the articulatory release, which subsequently causes a delay in voicing when transitioning from the release frication to the following vowel. This leads to a period of silence following the release of ejectives. The intensity of the affricates' release frication is being pulled towards the silent period, which subsequently leads to a drop in intensity. The fact that this has also been observed in other non-Salish languages [11], [12] suggests that intensity characteristics of the affricates frication is not a typologically unique feature in Lushootseed.

4.3. Diachronic status of voiced affricates in Coast Salish

In Coast Salish, there are only three languages that have voiced affricates: Lushootseed, Twana, and Comox-Sliammon. The historical origins of voiced affricates can be inferred based on cognates found across Coast Salish languages and within Lushootseed. It is likely the case that the voiced alveolar affricate /dz/ ‹dᶻ› developed from the palatal approximant /j/. For example, cognates of the Lushootseed word /dzaq'/ ‹dᶻaq̓› 'fall' in other Coast Salish languages include /jaq'/ ‹yaq̱'› 'fall' in Squamish and /jeq'/ ‹yeq'› in Hul'q'umi'num'. This suggests that the sound change might be */j/ > /dz/. This may have been phonetically motivated by domain-initial strengthening. Evidence of domain-initial strengthening can be observed when /dz/ occurs in word-initial position. As the current findings reveal, /dz/ is sometimes realized with a positive VOT in this position, which suggests that domain-initial strengthening might play a factor in the diachronic shift in the language. Domain-initial strengthening has also been observed in other Salish languages such as Quinalt [28] and Klallam [29].

5. Acknowledgements

Many thanks go to the Burke Museum for making these recordings available. I would also like to thank Laurel Sercombe of UW's Ethnomusicology Archives, who gave me the privilege to access these recordings. Special thanks go to the late Leon Metcalf, who spent about five years tape-recording elder speakers of Lushootseed during the 1950's. I would also like to acknowledge the family members of the speakers in this study: Denise Bill (great-granddaughter of Annie Jack), Willard Bill, Jr. (great-grandson of Annie Jack), Elise Bill-Gerrish (great-great-granddaughter of Annie Jack and daughter of Denise Bill), Justice Bill (great-great-grandson of Annie Jack and son of Willard Bill, Jr.); and Hank Williams (grandson of Martha Lamont), his daughter, and his (and Martha's) descendants. Most of all, I am strongly in debt to the speakers themselves: Annie Jack and Martha Lamont, renowned storytellers whose legacies will never be forgotten and forever be preserved in these recordings. May the language and spirit of Annie Jack and Martha Lamont continue to live on within their descendants.

6. References

[1] S. Bird, D. Gerdts, J. Leonard, "The realization of obstruents across speakers of Hul'q'umi'num'," *Canadian Acoustics,* vol. 44, no. 3, pp. 134-135, 2016.

[2] E. Flemming, P. Ladefoged, S. Thomason, "Phonetic structures of Montana Salish," *Journal of Phonetics,* vol. 36, no. 3, pp. 465-491, 2008.

[3] M. Percival, "Contextual variation in the acoustics of Hul'q'umi'num' ejective stops." In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences,* vol. 19, pp. 3270-3274, 2019.

[4] J. Kingston, *The phonetics and phonology of the timing of oral and glottal events.* Doctoral Dissertation, Department of Linguistics, University of California, Berkeley, 1985.

[5] J. Kingston, "The phonetics of Athabaskan tonogenesis," *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, pp. 131-179, 2005.

[6] N. Warner, "Acoustic characteristics of ejectives in Ingush." *Proceeding of Fourth International Conference on Spoken Language Processing,* vol. 3, pp. 1525-1528, 1996.

[7] R. Wright, S. Hargus, K. Davis, "On the categorization of ejectives: data from Witsuwit'en." *Journal of the International Phonetic Association,* vol. 32, no. 1, pp. 43-77, 2002.

[8] M. Stevens, J. Hajek, "Positional Effects on the Characterization of Ejectives in Waima'a," in *Proc. INTERSPEECH 2008,* Brisbain, Australia Sept. 22, 2008.

[9] S. Hargus. Witsuwit'en grammar: Phonetics, phonology, morphology. UBC Press.

[10] M. Percival, "Dene stop contrasts: Data from Déljnę Slavey," in The Scottish Consortium of ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences,* pp. 976-981.

[11] M. Percival, "Affricate contrasts in Eastern Oromo," in *Acoustics Week in Canada 16,* 2016.

[12] M. Percival, "Affricate contrasts in Déljnę Slavey," in *171st Meeting of the Acoustical Society of America,* May 27, 2016.

[13] M. Lindau, "Phonetic differences in glottalic consonants," in *Journal of Phonetics*, vol. 12, no. 2, pp. 147-155, 1984.

[14] P. Boersma, D. Weenink, "Praat: doing phonetics by computer [Computer program]", Version 6.4.06, Feb. 25, 2024.

[15] A. S. Abramson, "Laryngeal timing in Karen obstruents," In F. Bell-Berty, L. Raphael (eds.), *Producing speech: Contemporary issues. For Katherine Safford Harris*. Woodbury, NY: American Institute of Physics Press; pp. 155-165, 1995.

[16] A. S. Abramson, D. H. Whalen, "Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions," in *Journal of Phonetics,* vol. 63, pp. 75-86, 2017.

[17] C. Shadle, "Acoustics and aerodynamics of fricatives," in A. Cohn, C. Fougeron, M. Huffman (eds.), *Handbook of Laboratory Phonology,* pp. 511-526, Oxford University Press, Oxford, UK, 2012.

[18] K. Forrest, G. Weismer, P. Milenkovic, R. N. Dougall, "Statistical analysis of word-initial voiceless obstruents: Preliminary data," in *Journal of the Acoustical Society of America*, vol. 84, no. 1, pp. 115-123, 1988.

[19] M. Gordon, P. Ladefoged, "Phonation types: A cross-linguistic overview," in *Journal of Phonetics,* vol. 29, no. 4, pp. 383-406, 2001.

[20] Y. Koike, "Application of Some Acoustic Measures for the Evaluation of Laryngeal Dysfunction," in D. Runia, A. Mendelson, D. Winston (eds.), *Studia phonologica,* vol. 7, pp. 17-23.

[21] D. Bates, M. Mächler, B. Bolker, S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," in *Journal of Statistical Software,* vol. 67, no. 1, pp. 1-48, 2015.

[22] M. Gordon, P. Barthmaier, K. Sands, "A cross-linguistic acoustic study of voiceless fricatives," in *Journal of the International Phonetic Association,* vol. 32, no. 2, pp. 141-174, 2002.

[23] S. Hargus, G.-A. Levow, R. Wright, "Acoustic Characteristics of Deg Xinag Fricatives", in *University of Washington Working Papers in Linguistics,* pp. 1-53.

[24] C. Shadle, C. Scully, "An articulatory-acoustic-aerodynamic analysis of [s] in VCV sequences," in *Journal of Phonetics,* vol. 23, no. 1-2, pp. 53-66, 1995.

[25] C. Shadle, *The Acoustics of Fricative Consonants* (Doctoral Dissertation, Massachusetts Institute of Technology).

[26] L. Koenig, C. Shadle, J. L. Preston, C. R. Mooshammer, "Toward improved spectral measures of /s/: Results from adolescents," in *Journal of Speech and Language Hearing Research*, vol. 56, pp. 1175-1189.

[27] P. Ladefoged, I. Maddieson, *The Sounds of the World's Languages*, Blackwell Publishing.

[28] M. Swadesh, "Salish phonologic geography", in *Language,* vol. 28, no. 2, pp. 232-248, 1952.

[29] L. C. Thompson, M. T. Thompson, "Language universals, nasals, and the Northwest Coast," in M. E. Smith (ed.) *Studies in Linguistics in Honor of George L. Trager,* pp. 441-456, The Hague: Mouton.