WORD POSITION AND STRESS COMBINE TO AFFECT VOICE ONSET TIME IN HAWAIIAN

Lisa Davidson[1] and ʻŌiwi Parker Jones[2]

New York University[1], University of Oxford[2]
lisa.davidson@nyu.edu, oiwi.parkerjones@eng.ox.ac.uk

ABSTRACT

Hawaiian, or ʻŌlelo Hawaiʻi, is an Eastern Polynesian language spoken on the islands of Hawaiʻi. This study examines voice onset time (VOT) in Hawaiian as produced by speakers on the 1970s-80s radio program Ka Leo Hawaiʻi. Hawaiian has only two primary oral stops /p k/ ([t] is a rare variant in these speakers' dialects). First, we show that /p/ and /k/ are unaspirated in this generation of Hawaiian speakers. Second, we examine whether syllable stress and word position affect VOT duration. Results show an interaction indicating that word initial position leads to longer VOT in unstressed and secondarily stressed syllables, but not in syllables with primary stress. In comparison to previous research on aspirated stops, which are often affected by prosodic prominence, and unaspirated stops, which are much less sensitive to prosody, stress and word position work together for a nuanced effect on the realization of VOT in Hawaiian.

Keywords: voice onset time, stops, Hawaiian, stress, word position

1. INTRODUCTION

Hawaiian, or ʻŌlelo Hawaiʻi, is an Eastern Polynesian language spoken on the islands of Hawaiʻi. After a precipitous decline in the intergenerational transmission and use of Hawaiian following the overthrow of the Hawaiian Kingdom and the banning of Hawaiian in schools in 1890s [1], a revitalization movement began in the 1970s. This movement has been successful in creating Hawaiian-medium schooling opportunities that have led to a new generation of native speakers of Hawaiian [2-4]. Around the same time that the Hawaiian-medium schooling began, a Hawaiian language radio program called Ka Leo Hawaiʻi (KLH) was broadcast that included on-air interviews with bilingual speakers from the last generation who had learned Hawaiian from birth from their families and communities [5]. This source of data has been used in previous studies of the phonetics of Hawaiian [6, 7], and is used in the current study.

The focus of this study is on the voice onset time (VOT) of Hawaiian, as it was produced by the KLH speakers in the 1970s-80s. Hawaiian contains 2 primary oral stops, /p/ and /k/ (as well a [t] variant of /k/ in a small number of words)[1], such as [papa] 'class', and [kokoke] 'near'. No instrumental study of VOT has previously been carried out for this generation of speakers, and previous impressionistic descriptions have not always agreed. Elbert and Pukui [8] comment that these stops are less aspirated than English, but Newbrand [9] observes that her speakers seem to produce aspiration at least word initially. A related finding is that Hawaiʻi Creole English (HCE) is reported to have VOT values that are significantly shorter than English aspirated stops [10], perhaps because HCE is influenced by Hawaiian.

The first goal of this study is to establish whether the Hawaiian spoken by the speakers who appeared on KLH contains aspirated or unaspirated stops. A related study on Māori found that greater use of English did correlate with longer VOT in Māori voiceless stops over time [11], so it may be the case that contact with English had already begun to lengthen VOT in Hawaiian among the generation of KLH speakers.

The second goal is to examine how the VOT of the single-series stops of Hawaiian interacts with stress and word boundaries. A typical finding for a language like English or (Jutland) Danish with (contrastive) aspirated stops is that VOT is longer in stressed position [12, 13], while [14] found an effect of phrase-initial position but not stress for stops that were unaccented in the phrase in English. For single series (non-contrastive) stops, the effect of prosodic variables on VOT seems to be minimal. For example, in Plains Cree, which has only voiceless unaspirated stops, syllable position (used as a proxy for either stress or pitch accent, which is still contested) is not significant [15]. Sierra Norte de Puebla Nahuatl, which has only one stop series that is consistently realized with positive VOT, word-initial vs. word-medial position has no consistent effect [16]. As for languages which contrast voiceless unaspirated (fortis) stops and voiced (lenis) stops, the role of prosodic prominence is less clear; prominence further shortens short lag VOT in Dutch [17], but has no effect in Spanish [18].

Expectations for the effect of stress and word position on Hawaiian VOT may depend on the degree of aspiration found for /p, k/, but the lack of contrast

with another stop series may also affect the findings. On one hand, if the KLH speakers produce stops in the aspirated range, then it may be expected that prosodic factors associated with prominence like syllable stress and word initial position would lead to longer VOT. If stops are instead unaspirated, then prominence could lead to shorter VOT, as in Dutch. Another possibility is that Hawaiian could show an intermediate level of aspiration, perhaps similar to HCE; previous research is agnostic about the prosodic effects on a "semi-aspirated" voiceless stop. On the other hand, since other languages without a contrast are generally insensitive to prosodic prominence for VOT length, Hawaiian may also follow this pattern too, regardless of the degree of VOT found for the stops.

2. METHOD

2.1. Participants

This analysis examines spontaneous speech collected from 7 Hawaiian elders (4M, 3F) who were interviewed on the radio program Ka Leo Hawai'i in the 1970s. These interviewees were bilingual speakers of Hawaiian and English, but the program was carried out entirely in Hawaiian.

2.2. Language background and materials

We report only on /p, k/, since use of [t] is very limited. Oral stops can appear in stressed and unstressed syllables in Hawaiian. The words in (1) illustrate the six possible combinations of heavy and light syllables that have been proposed for words that only have one stress in Hawaiian [19-21]. Syllables with diphthongs and long vowels are heavy, and all heavy syllables are stressed. If there are 2 or more light syllables at the end of the word, the penultimate syllable will contain stress [20, 22, 23]. Words containing more than 3 syllables or multiple heavy syllables will also have secondary stress, with primary stress always occurring on the rightmost stressed syllable. Examples of words with multiple stress are in (2) (this list of types is not exhaustive).

(1) a.	ˈσL σL	[ˈpa.ka] 'raindrops'
b.	σLˈσL σL	[pa.ˈke.ʔo] 'to project'
c.	ˈσH	[ˈkai] 'ocean', [ˈpoː] 'night'
d.	σLˈσH	[ku.ˈkui] 'candlenut tree'
e.	ˈσH σL	[ˈpai:.ka] 'fight'
f.	σLˈσH σL	[pa.ˈlao.a] 'bread'
(2) a.	ˌσHˈσH	[ˌpi:.ˈkoi] 'ache'
b.	ˌσL σLˈσH	[ˌpi.ha.ˈʔuː] 'jam-packed'
c.	ˌσH ˈσL σL	[ˌko:.ˈʔa.la] 'to broil'
d.	ˌσL σLˈσL σL	[ˌpa.ka.ˈla.na] 'Chinese violet'

2.3. Data analysis

Using episode transcripts and audio, each stop produced was segmented in Praat. Stops were not included if they were in a word containing an error or a hesitation. Each stop closure and burst was measured separately, as shown for [k, p] in Figure 1. The burst was marked as beginning where aperiodic noise started after the closure, and ended when F1 of the following vowel became visible.

The stops that were extracted from the KLH transcripts were annotated for word position and stress of the syllable using Parker Jones' [21] computational prosodic grammar of Hawaiian. For example, the grammar parses the word *pakeʻo* [pakeʔo] 'to project' as (WORD (PRWD (SYLL_L (CONSONANT p) (SHORT_VOWEL a)) (FOOT (SYLL_L (CONSONANT k) (SHORT_VOWEL e)) (SYLL_L (CONSONANT ʔ) (SHORT_VOWEL o))))). Regular expression matching was then used to detect whether /p/ or /k/ appeared in a primary stressed, secondary stressed, or unstressed syllable, or in a word-initial or word-medial syllable. For [pa.ˈke.ʔo], /p/ is word-initial, /k/ is word-medial, and /k/ is stressed (see [24] for details on the method). Word frequencies were also obtained from an independent Hawaiian text corpus of about 1.6M words [25].

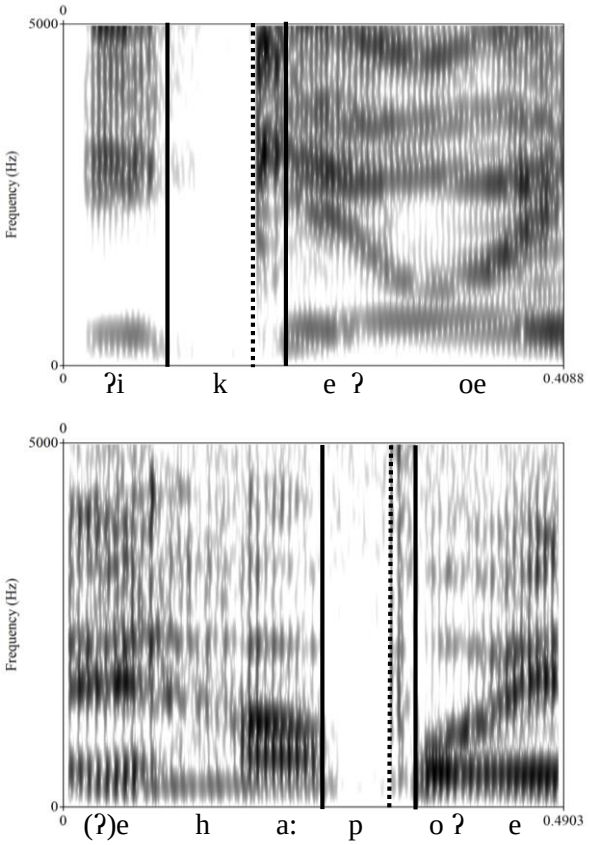

Figure 1: Multiple realizations of oral stops in [ʔike ʔoe] 'you see', 32ms, and [ʔeha: poʔe] 'four people', 23ms. Burst onset is demarcated with a dotted line.

In addition to the type of words in (1) and (2), Hawaiian also has 3 monosyllabic light syllable particles, [ka/ke] (allomorphs of the definite article) and [ko] (possessive marker), which are written as individual words in the orthography (e.g. *ka hale* 'the house'). These are typically unstressed, and are coded as such in this study. There are 5692 stops with a burst analysed in this study. In addition to the stops with closures and bursts, 1802 more are produced with frication, which will not be further examined here because of space limitations.

3. RESULTS

The histogram in Figure 2 indicates the range of values for VOT for /p/ and /k/ over all speakers and contexts in Hawaiian. The mean for /p/ is 24ms (s.d.=14ms, N=1547) and for /k/ is 39ms (s.d.=16ms, N=4145).

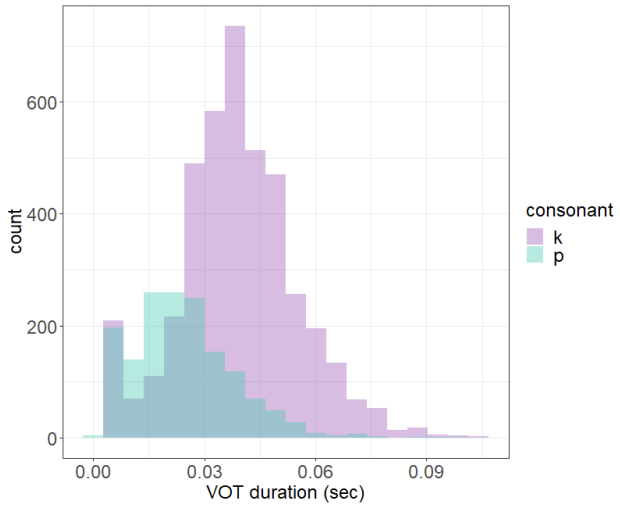

Figure 2: Histogram of VOT durations

The effects of consonant (/p, k/), word position (initial, medial), syllable stress (primary stressed, secondary stressed, unstressed), and lexical frequency (log transformed) on VOT duration are examined with a mixed effects linear regression in R. All of the categorical predictors were sum coded. A 3-way interaction between consonant, word position, and stress was included, as well as random intercepts for speaker and word. The results are in Table 1.

These results show that there is an effect of consonant: the 15ms difference between /k/ and /p/ is significant. Moreover, the histogram in Figure 2 indicates that some tokens do appear to have values that are in the range of aspirated stops. While individual differences are not the focus of this study, one concern is that these tokens may be attributable to one or two speakers. However, the plot of each of the talkers in Figure 3 demonstrates that despite some individual variability, all of the talkers are within the range of unaspirated stops, and that all of them produce some especially long tokens. This suggests that no individual speaker is producing most of the stops in the aspirated range.

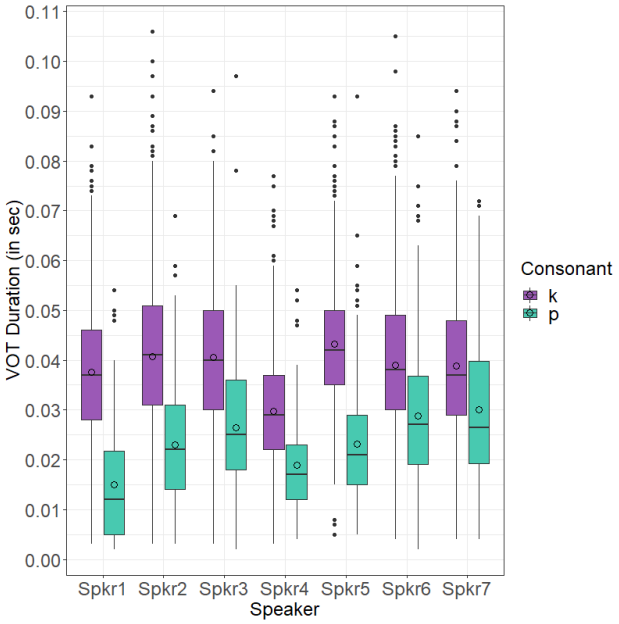

Figure 3: Average VOT durations for individual speakers

	Estim	Df	t val.	Pr(>|t|)
(Intercept)	0.0385	14.5	23.848	**<0.001**
consonant(p)	-0.0153	954.5	-14.05	**<0.001**
wd_pos(initial)	0.0017	1322	3.067	**0.002**
stress(primary)	0.0013	567.9	1.769	0.08
stress(secondary)	-0.0001	771.2	-0.103	0.92
log_word_freq	-0.0004	3501	-1.72	0.09
consonant(p): wd_pos(init)	-0.0003	1015	-0.256	0.79
consonant(p): stress(primary)	-0.0003	993.5	-0.282	0.78
consonant(p): stress(secondary)	0.0014	1515	0.898	0.37
wd_pos(init): stress(primary)	-0.0014	1435	-2.302	**0.021**
wd_pos(init): stress(secondary)	-0.0008	1654	-0.924	0.36
cons(p):wd_pos (init):str(prim)	0.0005	1058	0.413	0.68

Table 1: Linear mixed effects regression. Bold indicates a significant result.

There is no main effect for the levels of stress. Word position is significant, though the overall difference between the averages is small (word initial: 36ms, word medial: 33ms). There is no effect of word frequency. The negative interaction between word position and stress for primary stressed syllables occurs because there is a smaller

difference between word initial and word medial position than there is for secondary and unstressed syllables, both of which reflect the main effect of stress with longer VOT in word initial position. This effect is illustrated in Figure 4.

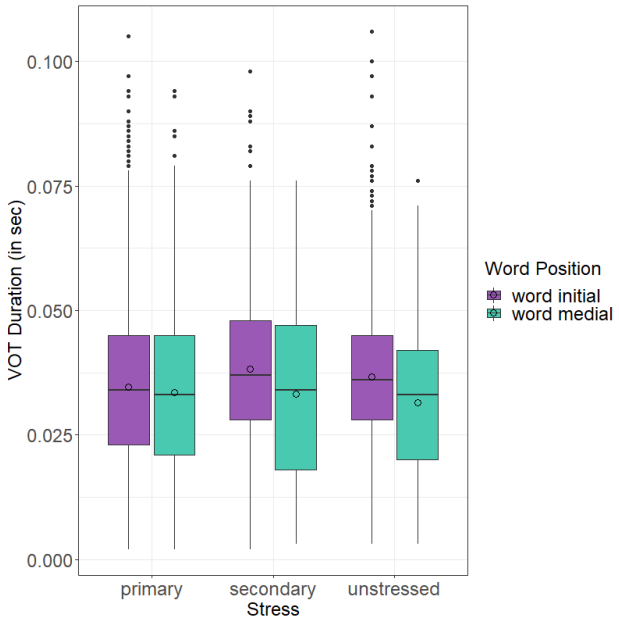

Figure 4: VOT durations by stress level (primary, secondary, unstressed) and word position (word-initial, word-medial)

4. DISCUSSION

Results indicate that for the generation of Hawaiian speakers who appeared as guests on Ka Leo Hawai‘i, the oral stops /p, k/ can be considered unaspirated. The values found here can be compared to previous cross-linguistic surveys [26, 27], which together report VOT values in more than 20 languages. The bilabial stops measured in [26, 27] range from 10-22ms, so while the Hawaiian average value of 24ms for [p] is on the high end of the range, it is consistent with other languages. As for velars, the Hawaiian value of 39ms for [k] fits squarely within the range of 27-56ms reported in these papers. Though we have not measured English speech for these same speakers, a recent study examining the VOT duration of aspirated stops in American English from 600 talkers reported a mean of 51ms for [p^h] and 54ms for [k^h] (across all contexts and averaged over all stress types) [28]. While only z-scored comparisons are reported in [29] for different contexts, VOT of stressed syllables is longer than unstressed syllables in English.

One limitation of this study is that we did not account for broader intonational patterns or prosodic elements like focus or accent, since this data comes from a corpus of spontaneous speech, and an intonational grammar of Hawaiian has not yet been developed. Impressionistic examination of the interviews suggests that especially long VOTs may be attributable to phenomena like focused utterances and animated speech, but a full accounting of these tokens awaits future research.

Previous research has shown two outcomes for the effect of prosodic variables on unaspirated stops. In Spanish, which has a voicing contrast, and in Plains Cree, which does not, stress had no effect on the duration of VOT [15, 18]. In Dutch, prosodically stronger positions including syllable stress, phrasal accent, and prosodic boundary strength all led unaspirated stops to have shorter VOTs than stops in weaker positions [17]. The authors concluded that in Dutch, VOT shortening is a fortition process that occurs to demarcate stronger prosodic positions.

The word position and stress results for Hawaiian show some significant effects, though their magnitude is small. Overall word initial VOT is 3ms longer, and the main effect of stress is not significant. The interaction of word position and stress indicates that longer word initial VOT is attributable to secondary and unstressed syllables (5ms difference), while the difference between word initial and medial positions for primary stress is nearly non-existent (1ms). Moreover, the numerical values for word-initial VOT in secondary and unstressed syllables are higher than for primary stress (Figure 4). This finding is not consistent with the literature for aspirated stops (e.g., [12, 29]), which usually reports that word initial position or primary stress lead to increased VOT duration (though in Scottish English spontaneous speech, for example, the word position difference for voiceless stops in stressed syllables is not significant [30]). If the effect for Hawaiian stops in secondary and unstressed syllables is perceptible, perhaps it could be an acoustic cue that a less prominent syllable belongs to the beginning of a word, instead of being the final unstressed syllable of a preceding word.

The picture that emerges for the interacting effects of stress and word position on Hawaiian VOT is more nuanced than most previous results for either aspirated or unaspirated stops. While the values are small, a 5ms difference is approximately the just noticeable difference for a 30ms burst [31]. Future work should examine whether this acoustic information is accessible for speech segmentation, as well as how it interacts with other cues to stress levels in Hawaiian.

5. ACKNOWLEDGMENTS

Thanks to Aidan Katson and JoJo Yang for their assistance with segmentation in Praat.

6. REFERENCES

[1] Wilson, W. 2013. Assessing Hawaiian. In: Kunnan, A. J. (ed.), *The Companion to Language Assessment* (pp. 1748-1758). Malden, MA: Wiley.

[2] Wilson, W., and Kamanā, K. 2001. Mai Loko Mai O ka ʻIʻini: Proceeding from a dream. In: Hinton, L. and Hale, K. (eds.), *The Green Book of Language Revitalization in Practice* (pp. 147-176). Leiden: Brill.

[3] Warner, S. L. N. E. 2001. The Green Book of Language Revitalization in Practice. In: Hinton, L. and Hale, K. (eds.), *The Movement to Revitalize Hawaiian Language and Culture* (pp. 133-146): Brill.

[4] Kawai'ae'a, K., Houseman, A. K., and Alencastre, M. 2007. Pūʻā i ka ʻŌlelo, Ola ka ʻOhana: Three generations of Hawaiian language revitalization, *Hūlili: Multidisciplinary Research on Hawaiian Well-Being*. 4: 183-237.

[5] Kimura, L. L. K. 2020. Kaniʻāina, ulukau.org Digital Repository of Ka Haka ʻUla O Keʻelikōlani College of Hawaiian Language, University of Hawaiʻi at Hilo, Ka Leo Hawaiʻi. 2020.

[6] Davidson, L. 2021. Effects of word position and flanking vowel on the implementation of glottal stop: Evidence from Hawaiian, *J. Phon.* 88, 101075.

[7] Kettig, T. 2021. Haʻina ʻia mai ana ka puana: The vowels of ʻōlelo Hawaiʻi. PhD dissertation, University of Hawaii at Mānoa.

[8] Elbert, S., and Pukui, M. K., *Hawaiian Grammar*, Honolulu: University of Hawaii Press, 1979.

[9] Newbrand, H. L. 1951. A phonemic analysis of Hawaiian. MA thesis, University of Hawaiʻi at Mānoa.

[10] Drager, K., Grama, J., Gonzalez, L. *et al.* 2016. Voice onset time and closure duration of word-initial voiceless plosives in Pidgin conversation, . *J. Acoust. Soc. Am.* 140: 4, 3052-3052.

[11] Maclagan, M., and King, J. 2007. Aspiration of plosives in Māori: Change over time, *Australian Journal of Linguistics*. 27: 1, 81-96.

[12] Lisker, L., and Abramson, A. S. 1967. Some effects of context on voice onset time in English stops, *Lang. Speech*. 10: 1, 1-28.

[13] Puggaard, R. 2019. Modeling regional variation in voice onset time of Jutlandic varieties of Danish. In: Van de Velde, H., Haug Hilton, N. and Knooihuizen, R. (eds.), *Selected papers from the Tenth International Conference on Language Variation in Europe* (pp. 79-110). Amsterdam: John Benjamins.

[14] Cho, T., and Keating, P. 2009. Effects of initial position versus prominence in English, *J. Phon.* 37: 4, 466-485.

[15] Hodgson, M. 2021. Voice onset time in Plains Cree. MA Thesis, University of Calgary.

[16] Kakadelis, S. 2018. Phonetic Properties of Oral Stops in Three Languages with No Voicing Distinction. PhD Dissertation, CUNY Graduate Center.

[17] Cho, T., and McQueen, J. M. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress, *J. Phon.* 33: 2, 121-157.

[18] Simonet, M., Casillas, J., and Diaz, Y. 2014. The effects of stress/accent on VOT depend on language (English, Spanish), consonant (/d/, /t/) and linguistic experience (monolinguals, bilinguals). In: Campbell, F., Gibson, M. and Hirst, D. (eds.), *Proceedings of Speech Prosody 7* (pp. 202-206).

[19] Schütz, A. 2010. Measures and morphemes: A functional approach to Hawaiian accent. In: Bowden, J., Himmelmann, N. P. and Ross, M. (eds.), *A journey through Austronesian and Papuan linguistic and cultural space: Papers in honor of Andrew Pawley* (pp. 405-422). Canberra: Pacific Linguistics.

[20] Schütz, A. 1981. A reanalysis of the Hawaiian vowel system, *Oceanic Linguistics*. 20: 1, 1-43.

[21] Parker Jones, O. 2010. A computational phonology and morphology of Hawaiian. DPhil. thesis, University of Oxford

[22] Parker Jones, O. 2018. Hawaiian, *J. Int. Phon. Assoc.* 48: 1, 103-115.

[23] Schütz, A. 1978. Accent in two Oceanic languages, *Anthropological Linguistics*. 20: 4, 141–149.

[24] Anonymous. submitted. Prosodic effects on Hawaiian glottal stop realization, *Phonetica*.

[25] Parker Jones, O., and Shillingford, B., 2018. Composing RNNs and FSTs for small data: Recovering missing characters in old Hawaiian text, in 32nd Annual Conference on Neural Information Processing Systems IRASL Workshop, Montreal.

[26] Cho, T., and Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages, *J. Phon.* 27, 207-229.

[27] Cho, T., Whalen, D. H., and Docherty, G. 2019. Voice onset time and beyond: Exploring laryngeal contrast in 19 languages, *J. Phon.* 72, 52-65.

[28] Chodroff, E., Godfrey, J., Khudanpur, S. *et al.* 2015. Structured variability in acoustic realization: A corpus study of voice onset time in American English stops. In: *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: University of Glasgow.

[29] Sprinkle, M. 2022. In search of phonetic evidence for prosodically-motivated aspiration. BA Thesis, Linguistics, College of William and Mary.

[30] Stuart-Smith, J., Sonderegger, M., Rathcke, T. *et al.* 2015. The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian, *Lab. Phon.* 6, 505-549.

[31] Stevens, K., *Acoustic Phonetics*, Cambridge, MA: MIT Press, 1998.

[1] While [t] is rare in the dialects of the KLH speakers, it does occur more commonly in the dialect of Niʻihau, a small island that maintained a continuously speaking community during the 20th century to today.