SPECTRAL COARTICULATION IN HAWAIIAN /aV/ AND /aCV/ SEQUENCES

Thomas Kettig

University of Hawaiʻi at Mānoa
tkettig@hawaii.edu

ABSTRACT

This study provides an acoustic analysis of the Hawaiian stressed low short vowel /a/ in /aV/ and /aCV/ positions, testing whether the height and backness of /a/ differ systematically based on the following vowel. F1 and F2 trajectories were measured in 1,550 /a(C)V/ tokens in over 41 minutes of spontaneous speech from one speaker, Larry Kimura. In /aV/ clusters, the F1 of /a/ was reduced preceding a high vowel (p<.001) and F2 was reduced preceding a back vowel (p=.001). Compared to /aCa/ sequences, /aCi/ exhibited significantly reduced F1 (p=.004) and larger F2 (p<.001) values. Overall, there were no significant differences in coarticulatory effects between clusters and /aCV/ contexts.

The results are discussed in relation to debates on the representation of Hawaiian vowel clusters as being comprised of two distinct phonemes or one. This study represents one of the first acoustic studies of spontaneous speech in the under-documented Hawaiian language.

Keywords: Hawaiian; vowels; [a]; diphthongs; coarticulation

1. INTRODUCTION

Hawaiian (ʻŌlelo Hawaiʻi) is an endangered Polynesian language indigenous to the islands of Hawaiʻi. Hawaiian is currently undergoing a renaissance after a sharp decline in speaker numbers in the 20th century.

There exists a small literature on the vowel inventory of Hawaiian, which has relied on auditory transcriptions [4, 6, 10, 15, 16] or acoustic analysis of small data sets [12, 13]. The present study seeks to expand on this literature with acoustic measurements of the pronunciation of /a/ in /aV/ and /aCV/ contexts.

2. PREVIOUS DESCRIPTIONS OF HAWAIIAN /aV/ AND /aCV/

Parker Jones [12] describes the vowels of Standard Hawaiian as /i e a o u/ and /iː eː aː oː uː/, plus the short diphthongs /ae ai ao au ei eu iu oi ou/ and the long diphthongs /aːe aːi aːo aːu eːi oːu/. He notes that diphthongs in Hawaiian are not ‘unit phonemes’, and analyzes them as simply sequences of individual vowels. Some sequences of vowels may give rise to multiple syllables (e.g. disyllabic *kia* /ˈki.a/ ‘pillar’) while those previously classified as diphthongs [4], including /aV/ sequences, can serve together as the peak of a single syllable (e.g. *‘aina* /ˈʔai.na/ ‘meal’) [15, 16].

In a short (<2 min.) passage of read speech from one male Hawaiʻi Island speaker, Parker Jones [12] tested differences in the realization of /a/ in /ai/ and /ae/, finding F1 to be significantly lower and F2 to be significantly higher in /ai/. His diagrams of the trajectories of the short diphthongs depict /ai au/ surfacing as [ɐi ɐu], while /ae ao/ surface as [ae ao].

Piccolo [13] also describes the vowel system of two female speakers, one Hawaiian L1 from Niʻihau (speaking the distinctive Niʻihau dialect [9]) and one L2 from Hawaiʻi Island, based on recordings of four repetitions of nineteen words. She found that for the Niʻihau speaker, /a/ had a notably greater F2 in /ai/ clusters than in /ae/ clusters, while the Hawaiʻi Island speaker did not show robust differentiation between the /a/ of /ai/ and of /ae/. Both speakers exhibited a large amount of overlap in the starting point of /ao/ and /au/, and both exhibited a notably larger F2 in the /a/ of /ai/ and /ae/ compared with /au/ and /ao/. However, stress was not fully controlled for (cf. [11]), and only a single word per vowel was analyzed.

It has also been reported that /a/ raises toward /ɐ/ when followed by /i/ or /u/ in the next syllable, and sometimes when /a/ is in the following syllable [17]. This raising is not universally noted in descriptions of Hawaiian, though; for instance, Parker Jones [12] narrowly transcribes *makani* /ma.ˈka.ni/ ‘wind’ as [mã.ˈka.nĩ], without raising to [mã.ˈkɐ.nĩ] or [mã.ˈkʌ.nĩ].

This study seeks to describe the acoustic characteristics of the Hawaiian short vowel /a/ in stressed /aCV/ and /aV/ positions in order to clarify its phonetic and phonological patterning.

If acoustic evidence indicates that an unstressed vowel has the same coarticulatory effect on a preceding stressed /a/ regardless of intervening consonant, this would strengthen the case against considering Hawaiian diphthongs as unitary phonemes, in concert with previous arguments based on the metrical system [15, 16]. These coarticulatory phenomena may not result from the vowels being in direct segmental contact, but rather from occurring within the same metrical foot.

3. METHOD

3.1. Source of data

The Hawaiian speaker for whom there exists the largest amount of publicly-accessible recorded and transcribed data is Dr. Larry Kimura, one of the most well-respected leaders in the Hawaiian revitalization movement. Kimura learned Hawaiian from his grandmother while growing up in Waimea, Hawaiʻi Island (though like most in his generation, Hawaiʻi Creole English was his dominant L1 [18]). His voice is familiar to many Hawaiian learners given the sheer number of his recordings and his many decades of revitalization work. As one of the founders of the ʻAha Pūnana Leo language nest organization, his voice has been a model, directly and indirectly, for many L2 and L1 speakers.

The speech data investigated here comprises a subset of the Kaniʻāina corpus of the *Ka Leo Hawaiʻi Radio* show, first produced and presented on KCCN by Kimura in 1972 [5]. Speech from six episodes from 1972 and 1973 make up the subset used in this study. In total, just under 42 minutes of Kimura's continuous speech was analyzed.

3.2. Segmentation

The .wav files for each episode were paired with transcriptions aligned roughly at the utterance level, excluding any sections of overlapping speech or segments of the show conducted over a telephone line. A model for the Montreal Forced Aligner [7] was trained by inputting paired sound and transcription TextGrids. In addition to the 42 minutes of speech from Larry Kimura that is analyzed here, training data for the aligner model also included just under 41 minutes of total speech from five other in-studio guests featured in the six episodes [5]. Output TextGrids containing the start and end points of individual words and phones was created for each recording.

3.3. Vowel measurements

Using Praat [3], measurements of F1 and F2 were extracted at the point of maximum F1 in /a/ using a script adapted for this purpose. This measurement location was chosen because regardless of coarticulatory effects resulting from contact with the previous or following segment, the point at which F1 reaches its maximum will consistently be the inflection point at the lowest point in the vowel space. For /aV/ clusters, measurements of the second vowel were taken at the point of maximum F2.

In order to control for effects of stress, observations were filtered such that only syllables bearing primary stress (cf. [11]) are analyzed here.

The words *ʻae* /ˈʔae/ 'yes' (the most frequent word in the subset and often used in back-channelling) and *laila* /ˈlei.la/ 'there, then' (also high frequency, with the spelling not reflecting its current pronunciation [6]) were excluded from measurement. Outliers were identified and excluded based on either F1 or F2 of /a/ being greater than two standard deviations away from the mean; in the case of /aV/ clusters, this outlier exclusion process was also applied to observations based on the F1 and F2 of the second vowel.

There was wide variation in the observed frequency of the different environments after outlier removal, ranging from just 7 tokens each of /ae/ and /ao/ to 202 tokens of /au/ and 306 tokens of /ai/; /aCV/ environments had higher counts but were also unbalanced, with 62 tokens of /aCu/, 101 of /aCe/, 104 of /aCo/, 186 of /aCi/, and 575 of /aCa/. The statistical limitations of this dataset should be noted, and the results are, naturally, reflective only of this one speaker.

4. RESULTS

Linear mixed effects models were constructed in R [14] using the lme4 package [1]. Vowel plots were created using phonR package [8].

4.1. Stressed /aV/ clusters

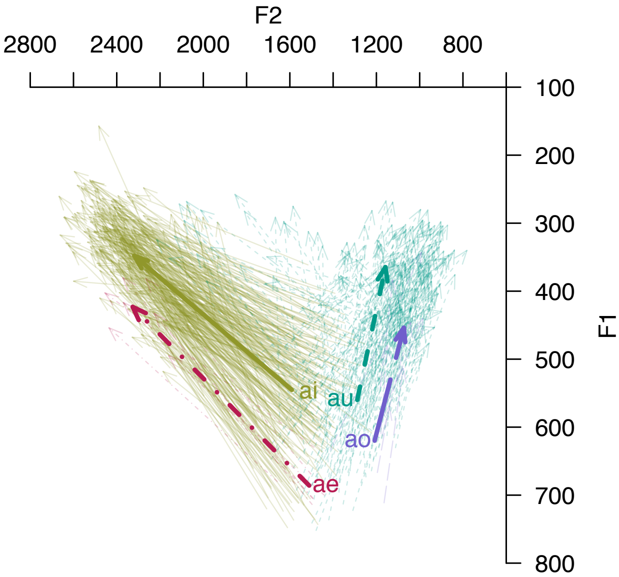
In the first set of models for /aV/ observations, F1 or F2 of /a/ was set as the dependent variable, with the height (high vs. mid) and backness (front vs. back) of the following vowel and their interactions entered as fixed effects. The identity of the segment following /aV/ was also entered as a fixed effect; possible following segments in non-word final contexts were /h ʔ k l m n p w/ as well as /a/ (in words like *aia* /ˈai.a/ 'there' and *kaua* /ˈkau.a/ 'war'); in word-final /aV/ contexts, following /i e o u/ or a prosodic break were also possible. Word was included as a random intercept. All fixed effect variables were sum contrast coded.

Tokens included word-final /aV/ (e.g. *kai* /ˈkai/ 'sea') and penultimate /aV/ (e.g. *kaona* /ˈkao.na/ 'hidden meaning'). Including stress placement as a fixed effect was not found to improve the models' fit.

Fig. 1 plots the trajectories of /ai ae au ao/ from the F1 maximum of /a/ to the F2 maximum of the second vowel. F1 was found to be significantly reduced preceding a high vowel compared to a mid vowel (β=-110, *SE*=31.4, *t*=-3.52, *p*<.001). F2 was found to be significantly reduced preceding a back vowel compared to a front vowel (β=-241, SE=72.7, t=-3.32, p=.001). A classic coarticulation effect is found: the tongue is higher in articulatory space preceding /i/ and /u/, lower preceding /e/ and /o/, backer preceding /o/ and /u/, and fronter preceding /i/ and /e/.

A following prosodic silence was set as the reference level for segment following /aV/. /m/ predicted a reduced F1 (β=-77.6, SE=29.9, t=-2.58, p=.010) and /e/ predicted a greater F2 (β=145.7, SE=57.4, t=2.54, p=.012).

Figure 1: /aV/ trajectories (speaker: Larry Kimura). Individual trajectories (light) and means (bold) shown.



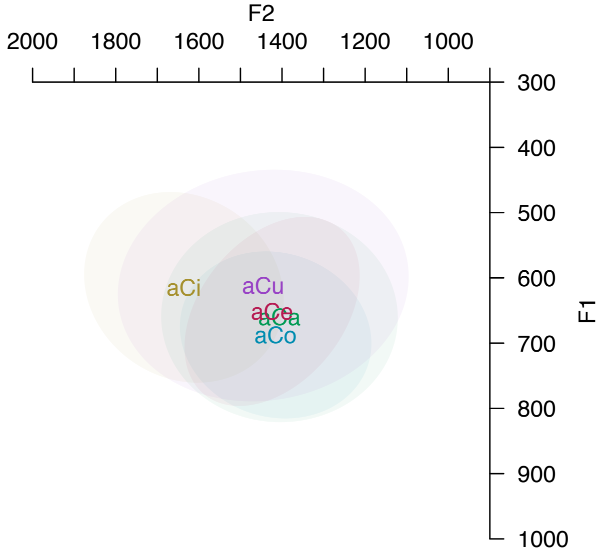
4.2.1. Pairwise comparisons

In a second set of models, F1 or F2 of /a/ was set as the dependent variable, with the identity of the second vowel (/i e o u/) and the following segment (/h ʔ k l m n p w i e a o u/ or prosodic break) included as fixed effects. Word was entered as a random intercept. Tukey's post-hoc pairwise analysis, correcting for multiple comparisons with the Benjamini-Hochberg procedure [2], showed that F1 was significantly reduced in /ai/ compared to /ae/ (β=141, SE=41.6, t=3.39, p=.004) and /ao/ (β=108, SE=47.7, t=2.28, p=.045). F1 was also significantly reduced in /au/ compared to /ae/ (β=112, SE=41.0, t=2.73, p=.018). F2 was significantly greater in /ai/ compared to /ao/ (β=-372, SE=110, t=-3.39, p=.002) or /au/ (β=-270, SE=47.2, t=-5.68, p<.001).

4.3. Stressed /a/ in /aCV/

Fig. 2 plots the location of /a/ in /aCV/ contexts. For these linear mixed effects models, F1 or F2 of /a/ was the dependent variable, with the identity of the second vowel (/i e a o u/) and intervening consonant (/h ʔ k l m n p w/) included as fixed effects. Word was also entered as a random intercept.

Figure 2: /a/ by /aCV/ context (speaker: Larry Kimura). Means at F1 maximum shown with ellipses drawn ±1 sd over F1 and F2 values.



4.3.1. Effects of following V

The reference level for second vowel was set as /aCa/, the environment hypothesized to be least likely to be affected by vowel-to-vowel coarticulation. F1 was significantly reduced in /aCi/ (β=-53.2, SE=18.2, t=-2.92, p=.004) and marginally reduced in /aCu/ (β=-41.6, SE=22.1, t=-1.88, p=.062). F2 was significantly greater in /aCi/ (β=220, SE=33.0, t=6.67, p<.001). In sum, a following /i/ exerts a significant coarticulatory effect on the preceding stressed /a/.

4.3.2. Effects of intervening consonant

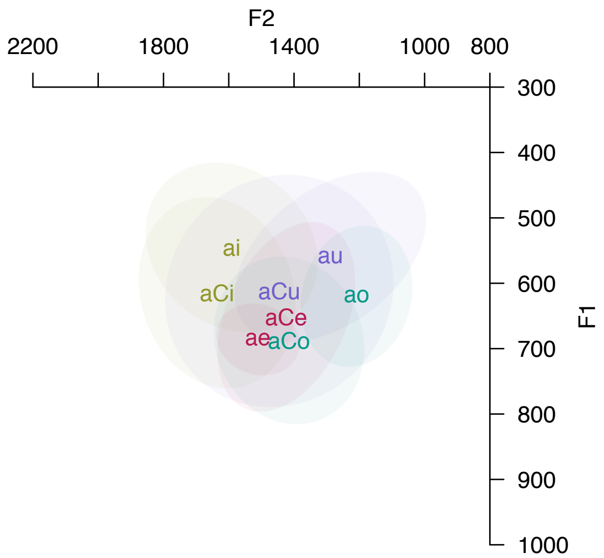
The reference level for intervening consonant was /h/. An intervening /ʔ/ (β=74.1, SE=34.3, t=2.16, p=.033) or /k/ (β=87.0, SE=30.5, t=2.85, p=.005) predicted a significantly greater F1, while an intervening /p/ predicted a reduced F1 (β=-64.7, SE=28.0, t=-2.31, p=.023). F2 was also greater with an intervening /p/ (β=290, SE=52.4, t=5.54, p<.001) and reduced with an intervening /w/ (β=-223, SE=109, t=-2.05, p=.042).

4.4. Comparison of /aV/ and /aCV/

Fig. 3 plots the relative locations of all /a/ in both /aV/ and /aCV/ contexts; /aCa/ is excluded from this analysis. (It is conceptually possible to regard /a:/ to be equivalent to /aa/, but this consideration of vowel length is beyond the scope of the current investigation.) A linear mixed effects model was constructed with F1 or F2 as the dependent variable; fixed effects were entered as the height (mid vs. low) and backness (back vs. front) of the second vowel, cluster status (/aV/ vs. /aCV/), and all two- and three-way interactions. In addition, following segment was entered as a fixed effect and word was entered as a random intercept.

F1 was significantly predicted by the height of the following vowel, with following high vowel corresponding to a reduced F1 (β=-70.2, SE=19.3, t=-3.63, p<.001). No effects of /aV/ vs. /aCV/ status or its interactions were significant. A following /l/ also significantly reduced F1 (β=-56.1, SE=26.7, t=-2.10, p=.036). While it appears in Fig. 3 that /ai/ and /au/ are especially raised in articulatory space in comparison to their /aCi/ and /aCu/ counterparts, the interaction of height and cluster status on F1 was marginal (β=-72.7, SE=38.2, t=-1.90, p=.059).

Figure 3: Relative /a/ positions in /aV/ and /aCV/ contexts (speaker: Larry Kimura). Means at F1 maximum shown with ellipses drawn ±1 sd over F1 and F2 values.



A following high vowel predicted a greater F2 (β=115, SE=38.5, t=2.99, p=.003), and a following back vowel predicted a reduced F2 (β=-169, SE=38.3, t=-4.42, p<.001). No effects of cluster status or its interactions were significant. Following segments /m/ (β=-116, SE=52.0, t=-2.23, p=.026) and /i/ (β=-165, SE=81.8, t=-2.02, p=.044) were found to predict a reduced F2, while a following /e/ (β=142, SE=54.9, t=2.60, p=.010) predicted a greater F2.

5. DISCUSSION

The present study provides evidence that Larry Kimura, a speaker of Standard Hawaiian, exhibits robust coarticulation of stressed /a/ to the following vowel in /aV/ clusters (Fig. 1). Vowel-to-vowel coarticulation also occurs in /aCi/ environments (Fig. 2), with /a/ being higher and fronter than in the /aCa/ context. Comparison of /aV/ and /aCV/ environments (Fig. 3) demonstrates that overall, /a/ before a high vowel is higher and fronter in articulatory space, while /a/ before a back vowel is backer. The presence of an intervening consonant is associated with only slightly less coarticulation than when the vowels abut.

These acoustic findings are consistent with a view of /aV/ sequences as phonemic clusters of monophthongs, in harmony with the previous literature on Hawaiian and other Polynesian languages [12, 15, 16]. Though the syllable and metrical structure of Hawaiian indicates that certain combinations of vowels, including /aV/ clusters, can jointly comprise the peak of a single syllable, the second vowel appears to exert a similar coarticulatory effect on /a/ whether or not it is in the same syllable. This suggests that in Hawaiian, such allophony may operate at the level of the foot (perhaps Schütz's [16] 'measure') rather than at the level of the syllable.

Several implications for the description of the Hawaiian vowel system can be identified. Phonetically, it is notable that /aV/ clusters exhibit not only coarticulatory raising based on the height of the following vowel, but robust backness effects based on the vowel's backness, too. Phonologically, it is tentatively suggested that tautosyllabic vowel clusters may be better described as containing two vowels of equal status [15], rather than as containing a peak and an offglide [4, 16]. Further investigation comparing the spectral qualities of the second vowel in these clusters with their corresponding stressed and unstressed singletons would help to clarify this picture. Accordingly, the unitary vowels of Hawaiian are suggested to be simply /i e a o u/ and /iː eː aː oː uː/, with the locus of functional complexity located at the syllabic and metrical levels (as suggested by Schütz [16]).

Whether or not a consonant intervenes does not make a significant difference in these patterns of observed coarticulation; however, the trend that the /a/ in /ai/ and /au/ raises more than in /aCi/ and /aCu/ may indicate a slightly greater coarticulatory effect when the two vowels are in direct contact. Further augmentation of the data under investigation – with more words and speakers included, as well as considerations of word frequency – will provide a more complete picture of the acoustic phonetics and phonological representations of Hawaiian vowels.

6. REFERENCES

[1] Bates, D., Mächler, M. Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *J. Statistical Software* 67(1), 1–48.

[2] Benjamini, Y., Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Statistical Society, Series B*. 57(1), 289–300.

[3] Boersma, P., Weenink, D., 2018. *Praat: Doing Phonetics by Computer*. Computer program, version 6.0.31.

[4] Elbert, S. H., Pukui, M. K. 1979. *Hawaiian Grammar*. Honolulu: University Press of Hawai‘i.

[5] Kimura, L. L K. 1972–1973. *Ka Leo Hawai‘i*. Kani‘āina, ulukau.org Digital Repository of Ka Haka ‘Ula O Ke‘elikōlani Coll. of Hawaiian Lang., U. Hawai‘i Hilo, L. Kimura (producer, presenter). 1972-04-11, HV24.004, http://hdl.handle.net/10125/59150; 1972-10-31, HV24.013, http://hdl.handle.net/10125/59156; 1972-11-14, HV24.015, http://hdl.handle.net/10125/59157; 1972-12-05, HV24.018, http://hdl.handle.net/10125/59158; 1973-02-06, HV24.021, http://hdl.handle.net/10125/59159; 1973-05-01, HV24.033, http://hdl.handle.net/10125/59164.

[6] Kinney, R. K. 1956. A non-purist view of morphomorphemic variations in Hawaiian speech. *J. Polynesian Soc.* 65(3), 282–286.

[7] McAuliffe, M., Socolof, M, Mihuc, S., Wagner, M., Sonderegger, M. 2018. *Montreal Forced Aligner*. Computer program, version 1.1.

[8] McCloy, D. 2016. *PhonR: R tools for Phoneticians and Phonologists.* R package, version 1.0-7.

[9] NeSmith, R. K. 2002. Tūtū's Hawaiian and the emergence of a Neo Hawaiian Language. Master's thesis, U. Hawai‘i at Mānoa.

[10] Newbrand, H. L. 1951. A phonemic analysis of Hawaiian. Master's thesis, U. Hawai‘i at Mānoa.

[11] Parker Jones, ‘Ō. 2010. A computational phonology and morphology of Hawaiian. Dissertation, U. Oxford.

[12] Parker Jones, ‘Ō. 2017. Hawaiian. *J. Int. Phon. Assoc.*: online.

[13] Piccolo, F. 2005. Where is the Hawaiian langauge headed? *U. Hawai‘i Working Papers in Linguistics* 36(1), 1–24.

[14] R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

[15] Rehg, K. 2007. Does Hawaiian have diphthongs? And how can you tell? In: Siegel, J., Lynch, J., Eades, D. (eds), *Language Description, History and Development*. Amsterdam: John Benjamins, 119–131.

[16] Schütz, A. 1981. A reanalysis of the Hawaiian vowel system. *Oceanic Linguistics* 20(1), 1–43.

[17] Schütz, A., Kanada, G. N. K., Cook, K. W. 2005. *Pocket Hawaiian Grammar*. Waipahu: Island Heritage.

[18] Wilcox, L. 2016. Larry Kimura. *Long Story Short*. Television Program, PBS Hawaii, 31-05-2016, http://pbshawaii.org/wp-content/uploads/2016/06/LSS_921_Larry_Kimura_Web_Transcript.pdf.