THE INTERACTION OF VOWEL LENGTH AND SPEECH STYLE IN AN ARAPAHO SPEECH CORPUS

Christian DiCanio[1,2], D. H. Whalen[2,3,4]

[1]Department of Linguistics, University at Buffalo, Buffalo, NY, USA

[2]Haskins Laboratories, New Haven, CT, USA
[3]Program in Speech-Language-Hearing Sciences; Linguistics, CUNY Graduate Center, New York, NY, USA
[4]Department of Linguistics, Yale University, New Haven, CT, USA
cdicanio@buffalo.edu, whalen@haskins.yale.edu

ABSTRACT

Cross-linguistically, vowel length contrasts may involve changes in vowel quality. In a different but more gradient way, speech style influences vowel articulation. While both alter vowel acoustics, it remains unclear whether both are byproducts of general processes of vowel undershoot or reflect a modification of articulatory gestures independent of durational constraints. This study investigates the influence of distinctive vowel length and speech style on vowel production in an Arapaho speech corpus. We find that length contributes most strongly to differences in duration and vowel quality, where short vowels are more centralized and long vowels more peripheral. However, the effect of speech style is asymmetrical: long vowels undergo greater durational compression in narrative speech than short vowels do, but the latter undergo greater changes in quality. These findings support the view that speech style produces not only patterns of vowel undershoot, but also active changes in vowel articulation.

Keywords: vowels, acoustics, speech style, length, endangered languages

1. INTRODUCTION

The acoustic properties of vowels are surprisingly variable within and across languages. In addition to coarticulatory variation [2, 11, 13, 23, 27, 31], vowel quality varies with many phonological properties, such as length, stress, and voice register [1, 8, 15, 12, 14, 19, 20, 26]. Such factors are intrinsic to phonological systems and can be contrasted with paralinguistic, or extrinsic factors which also influence vowel quality. The latter include the effects of speech rate and speech style on vowel production [7, 16, 4, 25]; and the sociolinguistic role that vowel variation may play within a speech community [17].

Given how many different types of factors may influence vowel formants, the important considerations for phonetic science are (a) how do specific factors influence vowel production when compared with other ones and (b) *to what extent* are the effects of these factors generalizable across languages as opposed to being language-specific. The current study addresses these concerns by evaluating both the influence of vowel length and speech style on vowel reduction in a corpus of Arapaho (ISO 639-3 *arp*) speech. While these two, independent factors are known to influence vowel quality, there remains an open question as to how they interact with each other not only within a language like Arapaho, but more generally as well.

2. BACKGROUND

2.1. Vowel length and vowel undershoot

Cross-linguistically, vowel quality may occur as an additional cue to indicate a vowel length contrast [18]. In a survey of 56 languages with a vowel length contrast, approximately 30% were reported to possess an accompanying difference in quality [22]. The typical pattern is for short vowels to occupy a more central position within the vowel space while long vowels occupy a more peripheral one.

The centralization of short vowels is argued to reflect a general process of *vowel undershoot*, whereby temporal constraints in producing shorter duration vowels result in the inability of a speaker to reach the intended vowel target [21, 25]. This process has generally been discussed in relation to the influence of stress on vowel production, as unstressed vowels are typically shorter than stressed ones, but it also may account for quality differences which accompany length contrasts. Vowel quality varies with vowel quantity in languages like German [26], Thai [1], Hausa [20], Western Apache [10], and Creek [14]. Short and long vowels, however, do not vary in quality in Norwegian [3] and Ndumbea [9].

Why might such variation across languages occur? After all, spectral changes in vowel production are perceptually salient as a cue to vowel length even for those listeners do not speak a language with contrastive length or use vowel quality differences to cue length. In a study examining durational, spectral, and F_0 cues to length, Lehnert-Lehouillier [19] finds that Spanish speakers are able to use spectral cues to perceive contrastive vowel length even though Spanish does not possess such a distinction. The same result was found for listeners of Japanese, where short vowels do not significantly differ in quality from long vowels [12]. One explanation for this phenomenon is that, while spectral cues may be salient to all listeners, they are only actively recruited as secondary cues within certain languages. Such recruitment may involve actively distinct articulatory gestures independent from those imposed by the temporal constraints of vowel production. In a study of the kinematics and acoustics of German vowels in stressed and unstressed positions, Mooshammer & Geng [26] find that short, unstressed vowels do not substantially differ from their stressed counterparts in duration, but do in fact undergo greater coarticulation with their surrounding consonants when unstressed. Thus, short vowels may involve less articulatory control in certain languages for which quality varies with quantity.

These findings highlight the main question motivating work on the phonetics of vowel length: are changes in quality independent or a byproduct of temporal constraints on speech production? We investigate this question here by looking at how duration and quality vary across speech styles for Arapaho vowels.

2.2. Vowel reduction and speech style

Just as intrinsic factors may influence vowel articulation, factors such as speech rate and style also have an influence. Vowels produced at a slower speech rate tend to be hyperarticulated compared with those produced at a faster rate. In a corpus study of conversational French, Meunier & Espesser [24] found that naturally-occurring shorter duration vowels occupied a more raised and reduced vowel space than those vowels which had longer duration. In a study investigating the influence of rate on vowel reduction for Japanese speakers, Hirata & Tsukada [12] find that short vowels are centralized during fast speech, but not in other contexts. Long vowels undergo no reduction with rate. While this particular finding suggests that vowel quality in Japanese is stable across different vowel lengths, there is reason to suspect that speech style mediates such stability. Earlier work by Keating & Huffman [16] on the influence of style on Japanese vowel production found much greater centralization of short vowels than predicted by Hirata & Tsukada. Japanese speakers in this older work were asked to read passages of natural text while speakers in Hirata & Tsukada's study repeated nonsense words at different speech rates.

For Japanese, differences in speech style may account for these contrasting findings. Elicited speech is characterized by slower speech rate and hyperarticulation in relation to narrative speech [4, 7]. While vowel reduction found in narrative speech may stem from undershoot due to an increased speech rate, it may reflect a distinct articulatory setting used in more spontaneous speech. In a recent study, DiCanio et al. [7] found that while elicited vowels in Yoloxóchitl Mixtec were significantly longer than vowels from narratives, duration alone was not sufficient to capture the observed quality differences. How do differences in speech style influence patterns of vowel reduction in a language with a vowel length contrast like Arapaho?

2.3. Arapaho phonology

Arapaho is an endangered Algonquian language spoken in Wyoming, USA. Like many Algonquian languages, it possesses a relatively simple consonant inventory and a vowel system with contrastive length [6]. In addition to short and long vowels, there also exist surface phonetic sequences that are *extra* long. Such sequences are relatively rare and may be analyzed as a sequence of short and long vowels. There are four diphthongs in the language and these behave like long vowels. Certain diphthongs may also be extra long. The vowel inventory is shown in 1.

Table 1: Vowel length contrasts in Arapaho

Short	ɪ	ɛ	ʊ	ɔ				
Long	iː	ɛː	uː	ɔː	eɪ	oʊ	aɪ	ɪɬ
Extra long	iːː	ɛːː	uːː	ɔːː	eɪː	oʊː		

Earlier work on Arapaho describes vowel quality differences which co-occur with the length contrast. In particular, long high vowels are described as tenser and closer than short high vowels [6, 30]. To our ears, low vowels seem to differ in quality too, e.g. [ɛ] vs. [æː], [ə] vs. [ɔː], though this is not mentioned in previous descriptions of the language. The current study examines to what degree low short vowels undergo a process of phonetic undershoot when compared with high short vowels. Thus, our findings do not only contribute to our knowledge on the phonetics of vowel reduction, but also provide a more general picture of the phonetics of Arapaho.

3. CORPUS STUDY

3.1. Methods

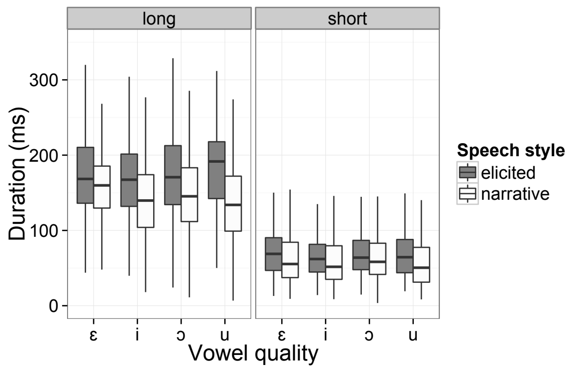
For the present study, we analyzed a subset of a corpus of Arapaho speech collected and transcribed by Lisa Conathan between 2004–2006. The analyzed corpus consists of both elicited and narrative speech from four speakers (3 female, 1 male). The elicited data comprised 5 hours 55 minutes of speech and the narrative data comprised 35 minutes of speech. As we are interested in the production of short and long vowels, only monophthongal vowels were considered in this study. Diphthongs, of which there are no short counterparts, and the extra long vowels were excluded. A total of 9,269 vowels were analyzed (5,232 from elicitations, 4,037 from narratives). Though little narrative speech was evaluated in the current study, the total number of analyzed vowels across speech styles was comparable.

The narrative data was initially force aligned using the Penn forced alignment system (P2FA) [32, 33]. Following alignment, a research assistant corrected faulty boundaries and hand segmented the elicited data. A script was written for Praat [5] which extracted vowel duration and the first three formants at three separate time points for each short vowel and five separate time points for each long vowel. Formant analysis was done using LPC covariance. Though dynamic formant data was collected, only values at midpoints are analyzed in the current work. Data analysis and statistics were done using R [28].

3.2. Results: Duration

Figure 1 shows the effect of vowel length and speech style (elicited vs. narrative) on vowel duration in Arapaho. A linear mixed effects model was constructed with three fixed effects (vowel length, speech style, and vowel quality) and one random effect (speaker). Two-way interactions were also included in this full model. Model comparison was done using a step-wise procedure excluding fixed effects and interactions and then analyzing the variance of each less specified model with the full model using a χ^2 test. A comparison of the full model with subsequent ones found all effects to be statistically significant except for the interaction of vowel length and vowel quality. Individual results are discussed below.

Figure 1: The effect of vowel length and speech style on duration in Arapaho vowels.



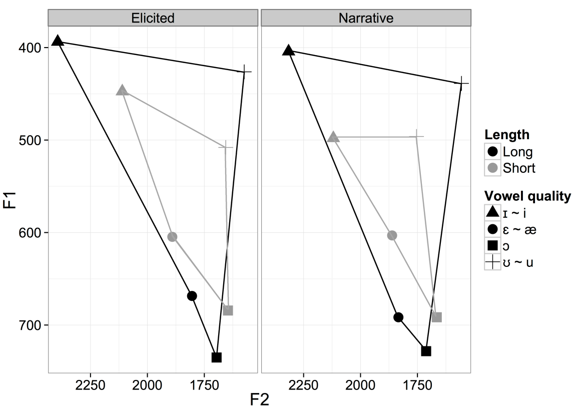
The main effect of vowel length on duration was significant (AIC = -22103 vs. -27640; $\chi^2[2] = 5540$, $p < .001$). The average ratio between short vowels (mean duration = 67.8 ms) and long vowels (mean duration = 165.5 ms) was quite long (1 : 2.4). However, a large amount of overlap in duration was also observed. This overlap was largely asymmetrical, as some long vowels were produced with very short durations but fewer short vowels were produced with long durations. The main effect of speech style was also significant (AIC = -27545 vs. -27640; $\chi^2[5] = 104.5$, $p < .001$). Vowels produced in narratives were, on average, 20% shorter than those in elicited speech. The durational ratio between narrative and elicited vowels was 1:1.2. However, this effect was also asymmetrical, as there was a significant interaction of vowel length and speech style on vowel duration (AIC = -27549 vs. -27640; $\chi^2[5] = 92.3$, $p < .001$). Long vowels in narrative speech were 28 ms shorter than those in elicited speech, but short vowels in narrative speech were only 6 ms shorter than their counterparts in elicited speech. Thus, the durational ratio of short to long vowels was smaller in narrative (1:2.3) than in elicited speech (1:2.5).

There was a significant main effect of vowel quality on duration as well (AIC = -27623 vs. -27640; $\chi^2[6] = 28.1$, $p < .001$). On average, high vowels (/i, u/) were slightly shorter than non-high vowels (/ɛ, ɔ/. This effect was the weaker than the influence of speech style and was more prominent among long vowels than short vowels.

3.3. Results: Formants

Figure 2 shows the effect of vowel length and speech style on vowel quality in Arapaho. The same statistical analysis was applied to the formant data where we treated F1 and F2 as dependent variables in distinct models. There was a significant main effect of vowel quality on F1 (AIC = 117432 vs. 114680; $\chi^2[9] = 2770.2$, $p < .001$) and on F2 (AIC = 129514 vs. 127595; $\chi^2[6] = 1931.4$, $p < .001$). A significant main effect of length on F1 (AIC = 114912 vs. 114680; $\chi^2[5] = 242.8$, $p < .001$) and F2 (AIC = 127790 vs. 127595; $\chi^2[5] = 203.6$, $p < .001$) was also found. This reflected a significant overall lowering of F2 in short vowels.

Figure 2: The effect of vowel length and speech style on the Arapaho vowel space.



A significant vowel x length interaction was found for F1 (AIC = 114695 vs. 114680; $\chi^2[3] = 21.3$, $p < .001$), but this effect was stronger for F2 (AIC = 127730 vs. 127595; $\chi^2[3] = 140.8$, $p < .001$). Short non-high vowels had significantly lower F1 values than long ones, but short high vowels had significantly *higher* F1 values than long ones. F2 was lower for short /i/ when compared with long /iː/, but slightly higher for short /ɛ/ when compared with long /ɛː/. F2 was higher for short /u/ when compared with long /uː/ but did not vary substantially for the short and long variants of /ɔ/.

A significant effect of speech style on F1 and F2 was found (F1: AIC = 114698 vs. 114680; $\chi^2 = 28.4$, $p < .001$; F2: AIC = 127616 vs. 127595; $\chi^2 = 22.8$, $p < .001$). Speech style influenced mainly the short vowels in the language, though the F2 value for /iː/ was significantly raised in elicited speech. F1 was slightly raised for all short vowels in narrative speech, while F2 was raised mainly for the *high* short vowels in narrative speech. In sum, the effect of length on vowel quality was robust across speech styles and involved a significant centralization of all short Arapaho vowels. Such effects varied with speech style, where vowels in narrative speech were produced with a fronted and slightly lowered articulation asymmetrically influencing the short vowels.

4. DISCUSSION & CONCLUSIONS

Like German or Creek, contrastive vowel length in Arapaho involves both significant differences in duration and vowel quality. Short high vowels are lowered and centralized relative to long high vowels. This finding accords with previous descriptions of the quality differences in Arapaho [30, 6], suggesting that the distinction among them is more accurately transcribed as [ɪ] vs. [iː] and [ʊ] vs. [uː]. Short low vowels are raised relative to long low vowels, but more so for the front low vowel than for the back one. This suggests that the distinction among low vowels is more accurately transcribed as [ɛ] vs. [æː] and [ɔ] vs. [ɔː].

Vowel length exerts a stronger effect on vowel quality and duration than speech style does. The effects of speech style are surprisingly asymmetrical. Long vowels underwent greater overall durational compression in more casual, narrative speech, than short vowels did. Yet, style influenced F1 and F2 values *more so* for short vowels in the language than for long ones. If durational compression alone were responsible for changes in vowel production, as predicted with vowel undershoot, we would predict a stronger influence of speech style on long vowels than short vowels. We find the opposite pattern though. This suggests that speech style induces different articulatory gestures for short vowels independent from duration, a finding in accord with work on Mixtec and German [7, 26].

Why might languages show such asymmetries with speech style? Consider that speech rate is increased in running, narrative speech when compared with careful, elicited speech. If speakers are actively aware of patterns of durational compression due to rate, changes in vowel targets in running speech may serve to enhance a length contrast. As long vowels are durationally compressed, short vowels vary more in vowel quality. Thus, for those languages where vowel quality is recruited as a cue to length, it may enter into a trading relation with it [29].

The current work also serves to demonstrate the utility of automatic methods, such as forced alignment, for the analysis of phonetic data from endangered language corpora. These corpora reflect a substantial resource not only for descriptive phonetic research, but also for testing hypotheses regarding the fundamental nature of speech production. As we seek more general explanations for different speech patterns, the continued creation, access, and use of these data will become increasingly more important.

5. ACKNOWLEDGEMENTS

This work was supported by NSF Grant 0966411 to Haskins Laboratories (Whalen, PI).

6. REFERENCES

[1] Abramson, A. S., Ren, N. 1990. Distinctive vowel length: duration vs. spectrum in Thai. *Journal of Phonetics* 18(1), 79–92.

[2] Beddor, S. P., Harnsberger, J. D., Lindemann, S. 2002. Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics* 30, 591–627.

[3] Behne, D. M., Moxness, B., Nyland, A. 1996. Acoustic-phonetic evidence of vowel quantity and quality in Norwegian. *Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm* 2, 13–16.

[4] Beinum, F. Koopmans-van 1980. *Vowel contrast reduction, an acoustic and perceptual study of Dutch vowels in various speech conditions*. PhD thesis University of Amsterdam, The Netherlands. Academische Pers B. V., Amsterdam.

[5] Boersma, P., Weenink, D. Version 5.3.5 2013. Praat: doing phonetics by computer [computer program]. www.praat.org.

[6] Cowell, A., Moss Sr., A. 2008. *The Arapaho Language*. University Press of Colorado.

[7] DiCanio, C., Nam, H., Amith, J. D., Castillo García, R., Whalen, D. H. 2015. Vowel variability in elicited versus spontaneous speech: evidence from Mixtec. *Journal of Phonetics* 48, 45–59.

[8] Gordon, M., Ladefoged, P. 2001. Phonation Types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.

[9] Gordon, M., Maddieson, I. 1999. The phonetics of Ndumbea. *Oceanic Linguistics* 38(1), 66–90.

[10] Gordon, M., Potter, B., Dawson, J., de Reuse, W., Ladefoged, P. 2001. Phonetic structures of Western Apache. *International Journal of American Linguistics* 67(4), 415–448.

[11] Hillenbrand, J. M., Clark, M. J., Nearey, T. M. 2001. Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America* 109(2), 748–763.

[12] Hirata, Y., Tsukada, K. 2009. Effects of speaking rate and vowel length on formant frequency displacement in Japanese. *Phonetica* 66, 129–149.

[13] House, A. S., Fairbanks, G. 1953. The influence of consonantal environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25, 105–113.

[14] Johnson, K., Martin, J. 2001. Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica* 58, 81–102.

[15] de Jong, K., Zawaydeh, B. 2002. Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration. *Journal of Phonetics* 30, 53–75.

[16] Keating, P. A., Huffman, M. K. 1984. Vowel variation in Japanese. *Phonetica* 41, 191–207.

[17] Labov, W. 2001. *Principles of Linguistic Change, Vol. 2: Social Factors*. Blackwell.

[18] Lehiste, I. 1970. *Suprasegmentals*. Cambridge, MA.: MIT Press.

[19] Lehnert-Lehouillier, H. 2010. A cross-linguistic investigation of cues to vowel length perception. *Journal of Phonetics* 38, 472–482.

[20] Lindau-Webb, M. 1985. Hausa vowels and diphthongs. *Studies in African Linguistics* 16(2), 161–182.

[21] Lindblom, B. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35, 1773–1781.

[22] Maddieson, I. 1984. *Patterns of Sounds*. Cambridge University Press.

[23] Maddieson, I., Emmorey, K. 1985. The relationship between semivowels and vowels: cross-linguistic investigations of acoustic difference and coarticulation. *Phonetica* 42(4), 163–174.

[24] Meunier, C., Espesser, R. 2011. Vowel reduction in conversational speech in French: The role of lexical factors. *Journal of Phonetics* 39, 271–278.

[25] Moon, S.-J., Lindblom, B. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *Journal of the Acoustical Society of America* 96(1), 40–55.

[26] Mooshammer, C., Geng, C. 2008. Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association* 38(2), 117–136.

[27] Ohde, R. N., Sharf, D. J. 1975. Coarticulatory effects of voiced stops on the reduction of acoustic vowel targets. *Journal of the Acoustical Society of America* 58, 923.

[28] R Development Core Team, A., Vienna 2013. R: A language and environment for statistical computing [computer program], version 3.0.2. http://www.R-project.org, R Foundation for Statistical Computing.

[29] Repp, B. H. 1982. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin* 92(1), 81–110.

[30] Salzmann, Z. 1956. Arapaho I: Phonology. *International Journal of American Linguistics* 22(1), 49–56.

[31] Stevens, K. N., House, A. S. 1963. Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research* 6(2), 111–128.

[32] Yuan, J., Liberman, M. 2008. Speaker identification on the SCOTUS corpus. *Proceedings of Acoustics - 2008*.

[33] Yuan, J., Liberman, M. 2009. Investigating /l/ variation in English through forced alignment. *Interspeech - 2009* 2215–2218.