Vowel production in a spoken Arapaho corpus

Christian T. DiCanio
dicanio@haskins.yale.edu
D. H. Whalen
whalen@haskins.yale.edu

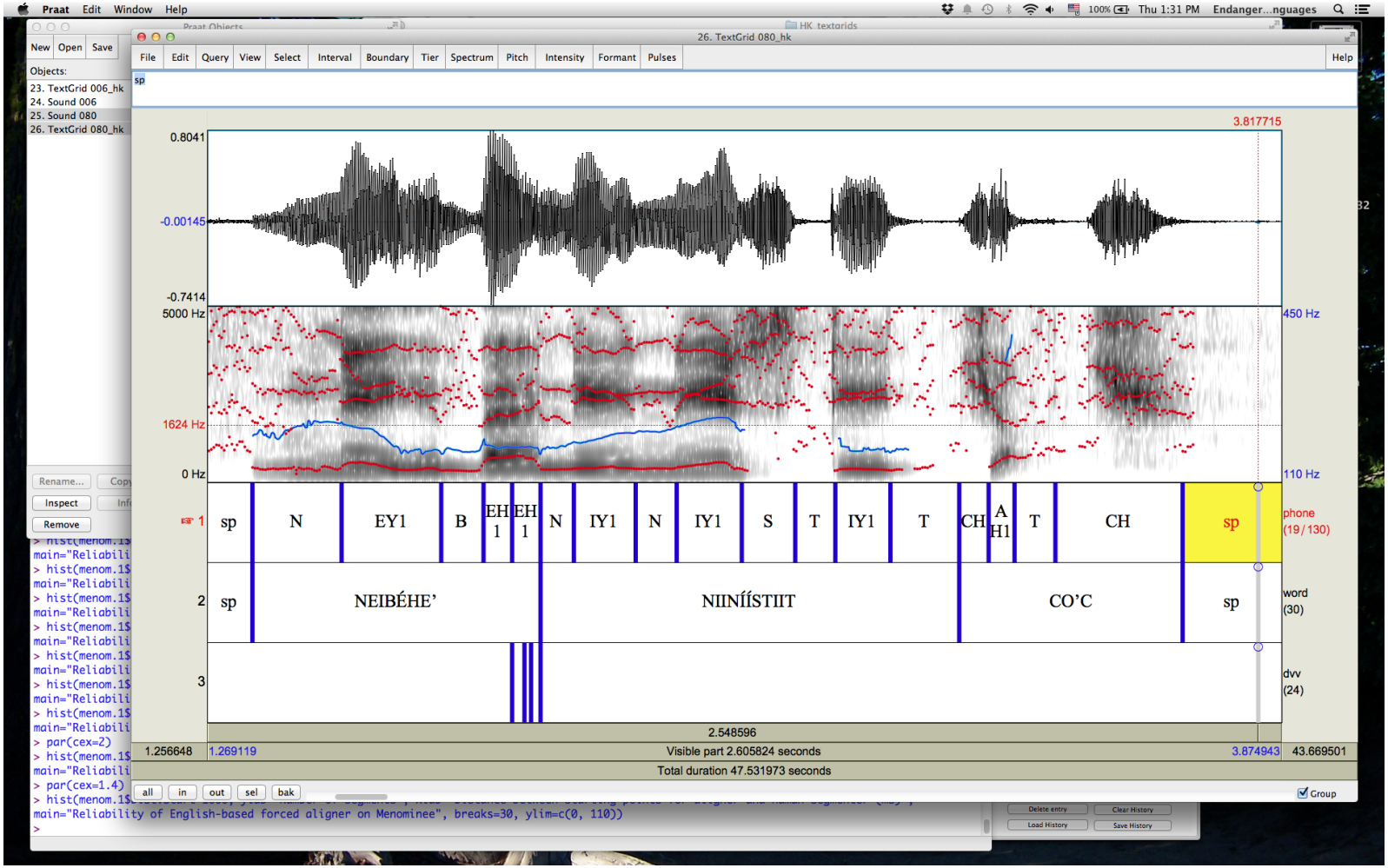
Haskins Laboratories
http://linguistics.berkeley.edu/∼dicanio

10/24/14
46^{th} Algonquian conference

Applying an aligner for phonetic purposes

- Endangered language corpora are the most useful for linguistic analysis when we can easily extract linguistic data from them.
- Computational tools such as *forced alignment* allow us to more easily segment existing recordings.
- Segmentation allows us to automatically extract word and speech sounds for phonetic and phonological research.

What is a segmented corpus?



What can you do once your corpus is segmented?

1. Easily extract recordings of words/sentences that can be included in a speaking dictionary.
2. Easily find out how rare or obscure words are pronounced.
3. Study how the pronunciation of words or morphemes changes across contexts and speakers.
4. Examine how certain sounds are produced for phonetic research.

Vowel production in Arapaho

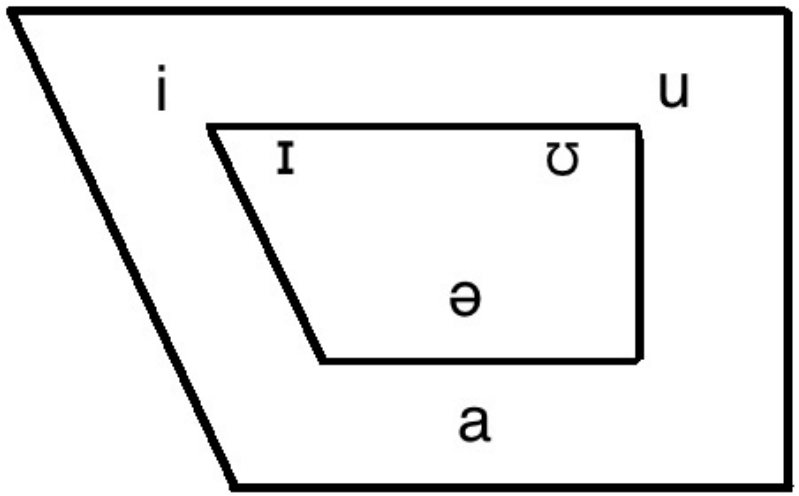
- Our goal here was to examine how short and long vowels in Arapaho were produced in both elicited speech and narrative speech.
  - While there are also *extra* long vowels in Arapaho, these are rather infrequent in texts and we did not analyze them here.
- This task was aided by forced alignment, which was used for the initial segmentation of the narrative corpus.

Why study Arapaho vowels?

- Arapaho, like many Algonquian languages, has a series of short and long vowels. How are these pronounced in different contexts?
- Languages vary in how vowel length is phonetically realized. In some languages, short vowels are reduced compared to long vowels, as in Western Apache (Gordon et al., 2001) and Creek (Johnson and Martin, 2001).
- However, languages differ in just *which* vowel qualities are reduced and in some, like Ndumbea (Gordon and Maddieson, 1999) and Norwegian (Behne et al., 1996), no reduction is found.
- Is Arapaho more like Creek in this regard or more like Norwegian? What about vowels in running speech?

Vowel quality and quantity

- "As against their short counterparts, the long high vowels are closer and tenser." (Salzman, 1956, 53).
- "When the vowels occur long, they show changes in quality as well as quantity. In particular, /ii/ and /uu/ have values approaching [i] and [u]." (Cowell and Moss Sr., 2008, 13).
- Neither Salzman nor Cowell & Moss mention any quality difference among low vowels. Impressionistically, low vowels seem to differ in quality too, e.g. [ɛ] vs. [æː], [ʌ] vs. [ɔː].



- Typically, the expectation is that when vowels are shorter, speakers lack the duration to reach a vowel target. This is called *vowel undershoot* (Lindblom, 1963). Over time, these patterns become part of the phonology of a language, so vowels of different *quantity* also differ in *quality*.

Arapaho vowels

	Front	Back
High	ɪ (i) i (ii)	ʊ (u) u (uu)
Low	ɛ (e) ɛɛ (ee)	ɔ (o) ɔɔ (oo)

Diphthongs	eɪ (ei) ɪɬ (io)	oʊ (ou) aɪ (oe)

Examples

Length	Word	Gloss	Word	Gloss
short	his	'*liver*'	hɔ́θɔʔ	'*star*'
long	níís	'*two*'	hɔhɔ́ɔ́t	'*tree*'
short	hɛθ	'*dog*'	θɔɔnoʊ́hut	'*He's lazy.*'
long	tʃɛ́ɛθ	'*accidentally*'	tɔʔúút	'*hammer*'

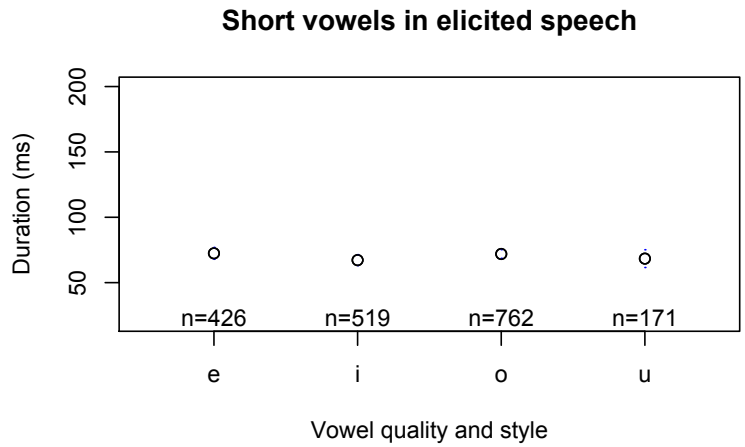
In addition to length, these vowels seem to be different in quality too.

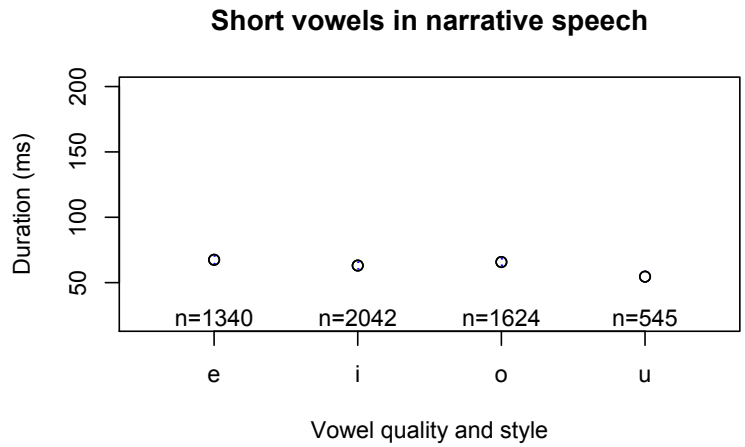
Method

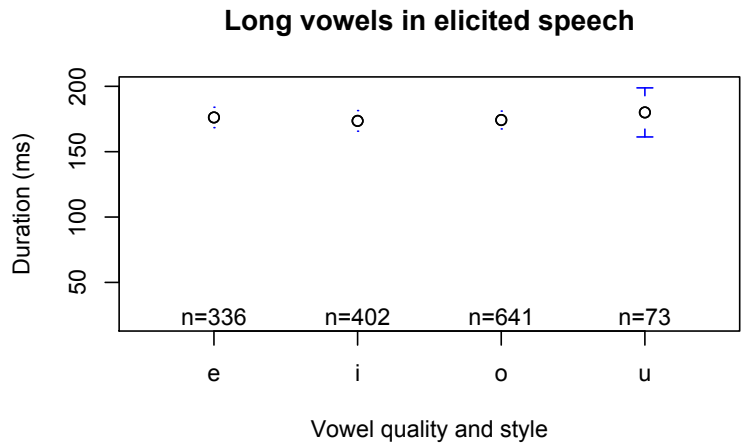
- Corpus consisted of elicited data (6 hours) to narrative data (1 hour, 36 minutes). A subset consisting of four speakers' data was used.
- A total of 12,426 vowels were analyzed.
- Used forced alignment for initial segmentation of narrative speech. This tool allows us to automatically segment the corpus with transcriptions.
- Hand-correction of the forced alignment was done for the narrative speech data. Hand-labelling was done for the elicited speech.
- Labeled vowels and analyzed formant values on vowels of different lengths using a script written for Praat (Boersma and Weenink, 2013).

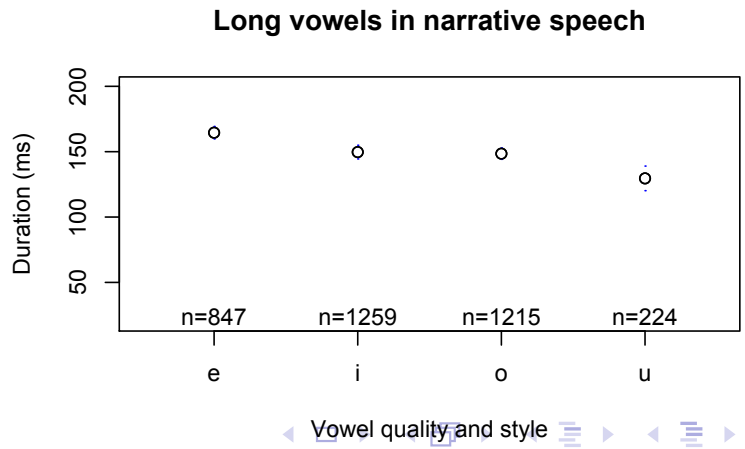
Results: duration

Across speech style, long vowels are more than twice as long as short vowels.









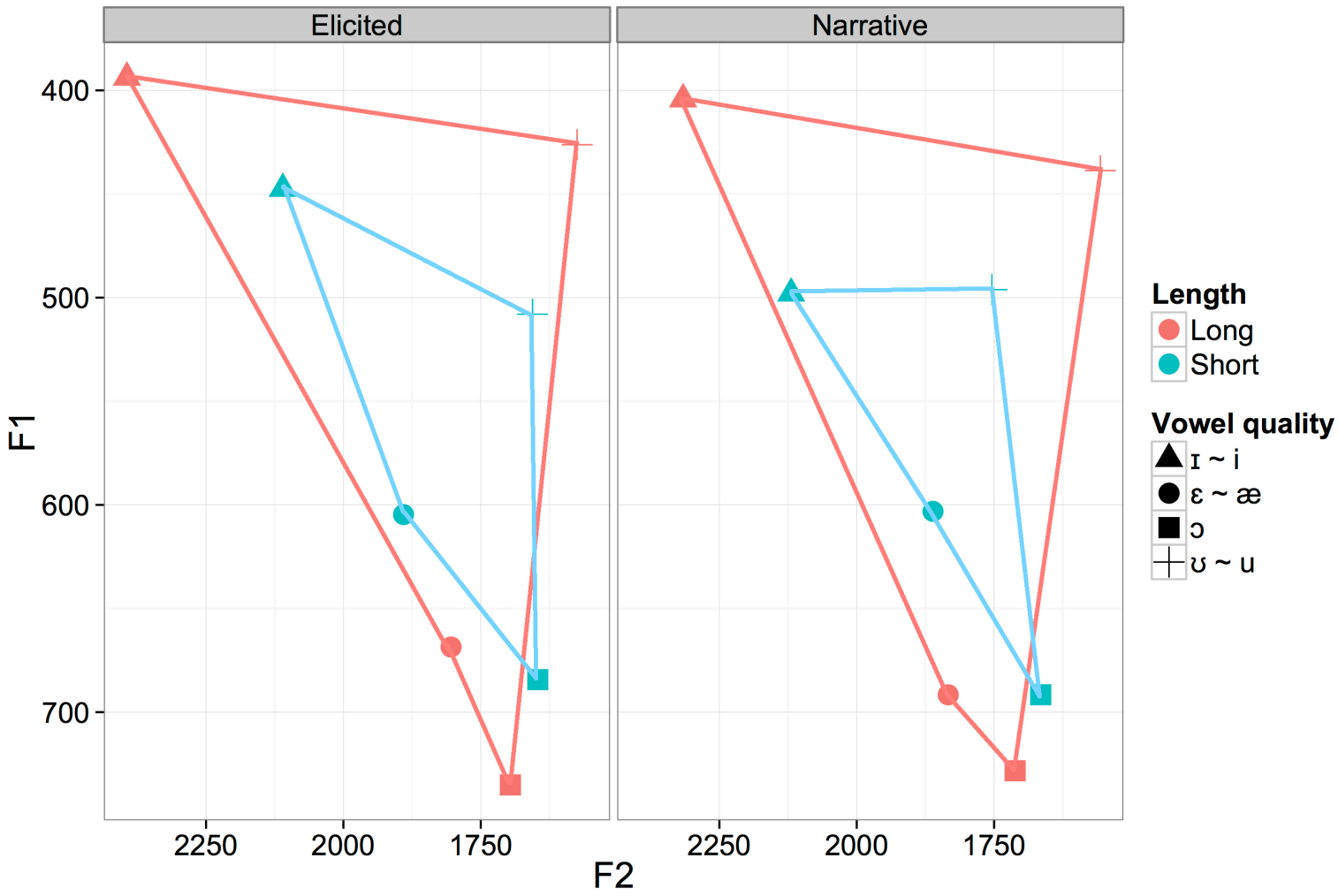
Results: duration

Long vowels are shorter when produced in narrative, running, speech, than when produced in elicited speech.

	Short	Long
Elicited	71 ms	176 ms
Narrative	65 ms	152 ms

It is fairly typical for vowels to be shorter in running speech than in elicited speech. This duration compression only significantly affects the long vowels here.

Results: Length and quality



Results: Length and quality

- All short vowels are reduced compared to long vowels, but the back vowel ‘o’ /ɔ/ is not reduced as much as the others.
- Even though the duration of long vowels changes across speech styles, their quality stays the same.
- Yet, while the duration of short vowels is similar across speech styles, their quality changes more across speech styles; short vowels undergo more reduction in narratives than in elicited speech.
- The acoustic distance (measured in formants) between long and short vowels is greater in narrative speech than in elicited speech.

Discussion

- Findings fit with the more general notion that short vowels are *centralized*, or undershot, while long vowels are more peripheral in the vowel space. Arapaho is more like Creek than Norwegian in this regard.

- While Arapaho vowels clearly contrast in vowel length, it appears that vowel quality is emphasized more in narrative speech than in elicited speech.

- Since narrative speech tends to be faster than elicited speech, this observation appears to be a case of phonetic compensation. As speech rate increases, speakers and listeners may rely more on the quality of the vowels than just their duration.

Conclusions

- The relationship between vowel length and quality holds up across different speech styles, but it is stronger in narrative speech.
- When teaching Arapaho pronunciation, one must both emphasize the duration of the vowels *and* their differences in quality. Quality may be more important in natural, connected speech.
- While forced alignment saved time in the segmentation of Arapaho speech here, its use was also revealing.
  1. It was not useful for the elicited speech, as the English here was not transcribed.
  2. Substantial human correction of alignment was still needed for narrative speech because of the mismatch between the transcription and the surface forms.

Acknowledgements

Thank you!

This work was supported by NSF Grant No. 0966411 to Haskins Laboratories: "From endangered language documentation to phonetic documentation."

National Science Foundation
WHERE DISCOVERIES BEGIN

We would also like to thank the Arapaho community and Lisa Conathan.

Behne, D. M., Moxness, B., and 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.

Boersma, P. and Weenink, D. (2013). Praat: doing phonetics by computer [computer program]. www.praat.org.

Cowell, A. and Moss Sr., A. (2008). *The Arapaho Language*. University Press of Colorado.

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

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

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

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35:1773–1781.

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