Intrinsic vowel duration and the post-vocalic voicing effect

Some evidence from dialects of North American English

Joshua Tauberer and Keelan Evanini

University of Pennsylvania

September 4, 2009

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Description of the corpus

The Atlas of North American English

- A telephone survey of 762 speakers (439 analyzed acoustically)
- All dialect regions of North America represented: at least two speakers sampled from every Metropolitan Statistical Area with more than 50,000 residents
- Most comprehensive corpus of dialect variation in North America available
- Speech data consists of free conversation, minimal pair tests, and other elicitation methods (style not recorded consistently in log files)
- All word tokens analyzed for the ANAE were extracted into individual sound files: 134,000 tokens total

Dialect regions in ANAE

Dialect Region	Speakers	Tokens
North	124	26,299
South	76	21,814
Midland	63	15,335
West	41	9,123
Canada	28	7,089
Western PA	13	3,685
Mid-Atlantic	12	2,950
Eastern New England (ENE)	10	2,141
Southeast	10	2,838
New York City (NYC)	5	1,706
Total	*382*	*92,980*

Forced alignment

To obtain duration measurements for the vowels, phoneme-level segmentation was created by forced alignemnt.

We used the P2FA forced alignment system (Yuan and Liberman 2008) (*http://ling.upenn.edu/phonetics/p2fa/*):

- Acoustic models trained on 25.5 hours of speech from the SCOTUS corpus (Supreme Court oral arguments)
- Monophone HMM models
- 32 Gaussian mixture components
- 39 PLP coefficients

Duration normalization: Motivation

Many of the variables manipulated in the ANAE affect duration.
Vowels are longer:

- before fricatives than stops
- before velars than labiodentals
- before voiced than voiceless consonants
- in the last syllable of a word

and some vowels are longer (AY) than others (UH).

These effects increase the variability in the data — when we don't take them into account. If the sample is not balanced, they can also confound the results.

Duration normalization: Method I

As a result, we used a linear model over log durations to estimate the effects of the variables we were not interested in, and then subtracted out those effects before further analysis.

We used a multiplicative model of duration where each factor increases or decreases the expected duration of the vowel by some percentage:

$$\text{duration} = c \cdot x_1^{m_s} \cdot x_2^{m_f} \cdot x_3^{p_l} \cdot x_4^{m_a} \cdot x_5^{p_v} \ldots$$

where c is some constant, m_s is 1 if the vowel precedes a stop and 0 otherwise, m_f is 1 if the vowel precedes a fricative and 0 otherwise, etc. x_i are the percent changes in duration caused by each factor.

On a log scale, this becomes a simple linear model so we can estimate the effects x_i with linear regression.

Duration normalization: Method II

Percent change in duration relative to a vowel preceding a voiceless labial stop (P) in word-medial position, according to linear regression:

affricate	5.6%
fricative	12%
labiodental	-11%
velar	8.6%
final syllable	27%
voiced	19%

We essentially factored out the effects we were not interested in by applying the appropriate factors to each token.

Duration normalization: Method III

The normalization procedure is replacing each duration with its residual according to the model.

Our model is somewhere between a simple linear model and the more complex model proposed by Klatt (1973) based on the notion of “incompressibility”.

For our study on post-vocalic voicing, the model contained vowel class and post-vocalic place and manner of articulation. (Only word-final syllables were used.) For our study of intrinsic vowel duration, the model contained post-vocalic place and manner of articulation, voicing, and number of following syllables.

Outline

Previous studies of North American English

Previous studies of duration variation by region have usually focused on North vs. South

- Freiman (1979): no difference in WPM between Southerners and Northerners
- Deser (1990): longer vowel durations among Southern AAE speakers than Northern ones
- Wetzell (2000): speakers from North Carolina have longer vowel durations than a group from the Mid-Atlantic and North
- Clopper et al. (2005): EH and AH longer in the South than the North
- Jacewicz et al. (2007): IH, EH, and AE longer in the South than the North or Midland

Duration by region results

Region	Duration (msec)
NYC	133
ENE	140
Canada	142
Mid-Atlantic	146
North	149
Western PA	150
West	153
Midland	154
South	156
Southeast	159

Explanation of dialect variation

- Regional variation in vowel duration likely not due to overall speaking rate differences
- For example, no difference between Southerners (N=1,421) and Northerners (N=445) in the Fisher corpus: both average 193 WPM
- Rather, the differences are likely due to the different natures of the sound changes in progress in each region
- The Soutern Shift causes tensing and diphthongization of high, front lax vowels
  - /ɪ/ → [iə] as in *sit*
  - /ɛ/ → [eə] as in *set*

Outline

Effect of voicing of following consonant

Durations of vowels are longer before a voiced consonant (versus voiceless).

- Schematically it can be thought of as:
  /bæt/ *bat* versus /bæːd/ *bad*, or
  V → [+long] / ___ [+voice]
- But the underlying mechanism is far from clear.

The duration difference can be quite significant but depends greatly on context. (Umeda 1975)

- In pre-pausal position, the ratio of vowel duration voiced-to-voiceless (of following consonant) is around 1.5-to-1, i.e. 50% increase.
- Word and utterance-medially the ratio is just 1.1 (10% increase), or less.

Effect of voicing of following consonant

- This phenomenon goes by many names, including the 'post-vocalic consonant voicing effect', and sometimes the 'vowel length effect', or "VLE" for short.
- Voewls affected by voicing of stops as well as fricatives
- The different is *magnified* when there are other factors that are increasing vowel duration, such as position, stress, and overall speech tempo.
    - This means that when studying the VLE controlling other variables is very important.
- Why is it interesting? We don't know what the mechanism is behind it.
    - Physiological?
    - Phonological rule?
    - Part of the phonetic implementation of [voice]?

Is the VLE universal?

Cross-linguistic comparisons are not very reliable because there are so many factors that affect vowel duration. Nevertheless, it appears English is unique at least in degree.

Language	Ratio	Study
English	1.63	Chen (1970)
French	1.15	Chen (1970)
Catalan	1.17	Charles-Luce (1992)
Arabic	1.03	Flege and Port (1981)
Polish	1.0	Keating (1980)

(A subsequent study of English and French made the difference look much smaller (Laeufer 1992)).

What about within dialects of (American) English?

Veatch (1991):

- 4 dialects. 1–3 speakers per dialect.
  - Alabama, Chicago white, Jamaican Creole: One speaker from each dialect showed a weak but significant duration difference.
  - Los Angeles Chicano: No statistically significant difference.

Jacewicz et al. (2007):

- Madison, Wisconsin; Columbus, Ohio; and western North Carolina. 18 speakers per dialect.
- Each dialect showed a duration difference by voicing, but an interaction by dialect was found.

VLE results: By dialect region

We report the first comprehensive survey across all dialects of North American English.

Dialect Region	Ratio
NYC	1.13
South	1.16
Canada	1.19
West	1.19
Midland	1.21
North	1.23
ENE	1.24
Southeast	1.24
Mid-Atlantic	1.25
Western PA	1.27

- **Overall ratio for all speakers: 1.2**

VLE results: By dialect region

- The VLE appears to exist in all dialect regions.
- The ratios are lower than what has been reported elsewhere, but this is probably due to the nature of the tasks.
- What do we make of the range from NYC (1.13) to Western PA (1.27)? NYC vowels are shorter, so we expect a smaller VLE. But this makes the South unexpected. We don't know.

VLE results: By dialect

Of the more narrow ‘dialect’ classification of speakers, there were two notable outliers:

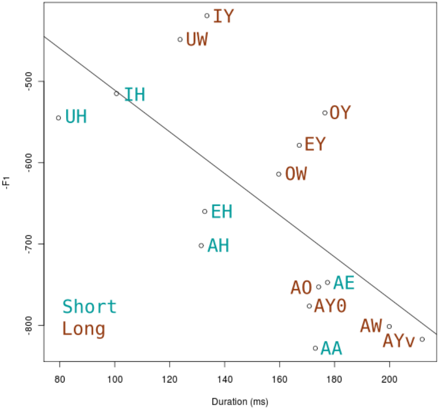
Dialect	Ratio
Boston	1.33
Maine	1.02

Does Maine lack a VLE?

This data suggests further study. There are too few speakers in the samples of these dialects (6 from Boston, 2 from Maine) to draw any conclusions now.

Outline

Intrinsic vowel duration



Mean durations of vowels vary from vowel to vowel and these means are correlated with vowel height. (House and Fairbanks 1953, Peterson and Lehiste 1960)

In our data:
r = .68 (N=15)
+18ms for every +100 Hz F_1

Intrinsic vowel duration: Why?

The lower the vowel the greater its “intrinsic” duration.

Explanations:

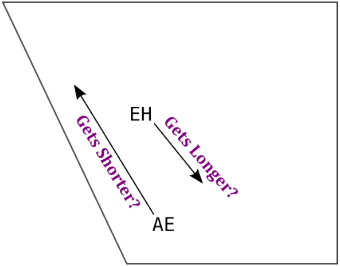
1. Differences are **physiological** (Lehiste 1970)
2. Differences are **phonological**: Each vowel has a duration target specified in the grammar (Lisker 1974)

Arguments against physiology:

- Low vowels have longer steady states, not transitions (Lisker 1974)
- Duration can be distinctive (Labov and Baranowski 2006)
  - Pittsburgh: AH and monophthongal AW overlap in formant space but differ by approximately 100ms in duration
  - Inland North: EH and AA overlap in formant space but differ by 50ms

Intrinsic vowel duration: Our hypothesis

What happens to intrinsic duration as a vowel undergoes sound change?



As a vowel is lowered, e.g., through a chain shift:

- the physiological explanation predicts the vowel's intrinsic duration will increase
- the phonological explanation predicts that the duration target will remain unchanged

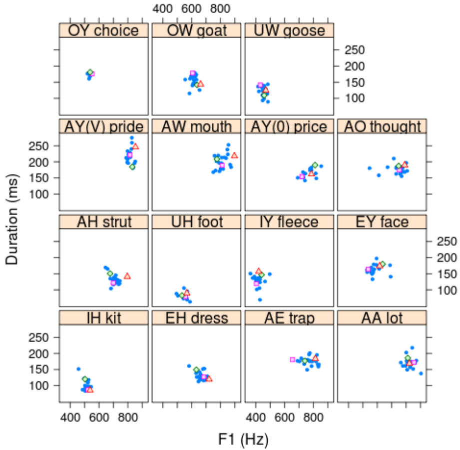
— other things being equal.

Intrinsic vowel duration: Method

Hypothesis: As a vowel undergoes sound change (i.e. change in F_1), its duration changes accordingly.

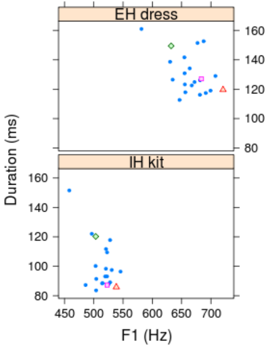
Prediction: If we take a single vowel and trace it *across dialects*, we should see a positive correlation between F_1 and duration.

Intrinsic vowel duration: Results



Intrinsic vowel duration: Results

What happens to intrinsic duration as a vowel undergoes sound change?



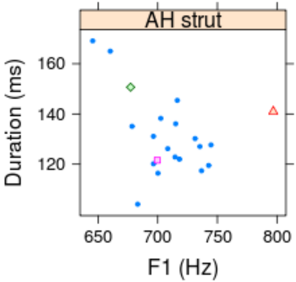
Two vowels had significant correlations: IH and EH.

But correlations were negative! e.g. IH was longer in dialects where the vowel is higher (lower F_1).

Caveats

The correlations (or lack thereof) are difficult to interpret here: formant trajectories undergo sound change as well. Monophthongization and diphthongization processes might increase or decrease intrinsic vowel duration for reasons independent of vowel height, though this has yet to be shown empirically.

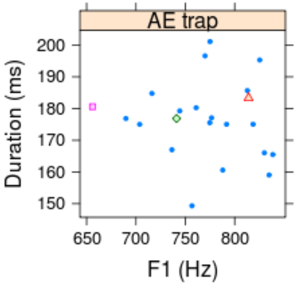
Test Case 1: Pittsburgh shift



AH has been lowered as part of the Pittsburgh shift; F_1 mean is 785 Hz, 50–100 Hz higher than for all other dialects.

Nevertheless, Pittsburgh AH is clearly located in the middle of the distribution of durations across dialects. This would tend to support the explanation that intrinsic duration is largely phonological.

Test Case 2: AE



F_1 variability in AE is quite large, but the variability in duration is not.

Does AE have a phoneme-specific duration target that is not influenced by regional differences in vowel height?

But many dialects with raised AE also develop a schwa-like off-glide which could obscure the change in duration due to height.

Summary

- Survey of vowel duration in a comprehensive corpus of North American dialect variation (ANAE)
- Regional differences in overal duration not due to different rates of speech, but can be explained by sound changes in progress
- The effect of post-vocalic voicing on duration is relatively stable across all dialect regions
- Results from individual vowels across regions suggest that “intrinsic” vowel length is not purely physiological, but also partly phonological

References I

Charles-Luce, Jan. 1992. The effects of semantic context on voicing neutralization. *Phonetica* 50:28–43.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22:129–159.

Clopper, Cynthia G., David B. Pisoni, and Kenneth de Jong. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America* 118:1661–1676.

Deser, Toni. 1990. Dialect transmission and variation: An acoustic analysis of vowels in six urban Detroit families. Doctoral Dissertation, Boston University.

Flege, J.E., and R. Port. 1981. Cross-language phonetic interference: Arabic to English. *Language and Speech* 24:125–146.

References II

Freiman, Howard A. 1979. Speech rate as a function of dialect geography. In *Perspectives on applied sociolinguistics: From the language medium of education to the semiotic language of the media*, ed. Robert N. St. Clair, 128–136. Lawrence, Kan.: Coronado.

House, Arthur S., and Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *J. Acoust. Soc. Am.* 25:105–113.

Jacewicz, Eva, Robert A. Fox, and Joseph Salmons. 2007. Vowel duration in three American English dialects. *American Speech* 82:367–385.

Keating, Patricia Ann. 1980. A phonetic study of a voicing contrast in Polish. Doctoral Dissertation, Brown University.

Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *J. Acoust. Soc. Am.* 54:1102–1104.

References III

Labov, William, and Maciej Baranowski. 2006. 50 msec. *Language Variation and Change* 18:223–240.

Laeufer, Christiane. 1992. Patterns of voicing-conditioned vowel duration in French and English. *Journal of Phonetics* 20:411–440.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.

Lisker, Leigh. 1974. On 'explaining' vowel duration variation. *Glossa* 8:233–246.

Peterson, Gordon E., and Ilse Lehiste. 1960. Duration of syllable nuclei in english. *The Journal of the Acoustical Society of America* 32:693–703.

Umeda, Noriko. 1975. Vowel duration in American English. *J. Acoust. Soc. Am.* 58:434–445.

Veatch, Thomas Clark. 1991. English vowels: Their Surface Phonology and Phonetic Implementation in Vernacular Dialects. Doctoral Dissertation, University of Pennsylvania.

References IV

Wetzell, Brett. 2000. Rhythm, dialects, and the Southern drawl. Master's thesis, North Carolina State University.

Yuan, Jiahong, and Mark Liberman. 2008. Speaker identification on the SCOTUS corpus. In *Proceedings of Acoustics '08*.