Durational and spectral differences in American English vowels: dialect variation within and across regions

Valerie Fridland[1]

[1]Department of English

University of Nevada, Reno

Reno, Nevada 89557-0098

Tyler Kendall[2]

Charlie Farrington[2]

[2]Department of Linguistics

University of Oregon

Eugene, OR 97403-1290

Running Head: Durational and spectral differences in vowels

ABSTRACT

Spectral differences among varieties of American English have been widely studied, typically recognizing three major regionally diagnostic vowel shift patterns (Labov, Ash, and Boberg, 2006). Durational variability across dialects, on the other hand, has received relatively little attention. This paper investigates to what extent regional differences in vowel duration are linked with spectral changes taking place in the Northern, Western, and Southern regions of the U.S. Using F1/F2 and duration measures, the durational correlates of the low back vowel merger, characteristic of Western dialects, and the acoustic reversals of the front tense/lax vowels, characteristic of Southern dialects, are investigated. Results point to a positive correlation between spectral overlap and vowel duration for Northern and Western speakers, suggesting that both F1/F2 measures and durational measures are used for disambiguation of vowel quality. Our findings also indicate that, regardless of region, a durational distinction maintains the contrast between the low back vowel classes, particularly in cases of spectral merger. Surprisingly, Southerners show a negative correlation for the vowel shifts most defining of contemporary Southern speech, suggesting that neither spectral position nor durational measures are the most relevant cues for vowel quality in the South.

PACS numbers: 43.70.Fq, 43.72.Ar

I. INTRODUCTION

Spectral differences in vowel realization across American dialects have been widely documented by sociolinguists over the last 25 years (e.g. Eckert, 1988, 2000; Feagin, 1986; Fridland, 1999, 2001; Fridland and Bartlett, 2006; Gordon, 2005; Labov, 1991, 1994; Labov, Ash, and Boberg, 2006; Thomas, 1989, 1997, 2001). Regional variation in duration, in contrast, has been much less intensively studied. A few papers have examined the relationship between duration and regional vowel variation in American English (Clopper, Pisoni, and de Jong, 2005; Jacewicz, Fox, and Salmons, 2007; Labov and Baranowski, 2006; Tauberer and Evanini, 2009) with most suggesting that Southerners are significantly different from other regions both in duration and in spectral position (Clopper *et al.*, 2005; Jacewicz *et al.*, 2007). An intriguing – but relatively unexamined – question is whether this durational variation correlates with the spectral variation found across and within regions. The current study investigates this relationship, looking both at community level and individual level differences in the spectral and durational characteristics involved in regional vowel shifts.

A. Spectral differences in U.S. dialects

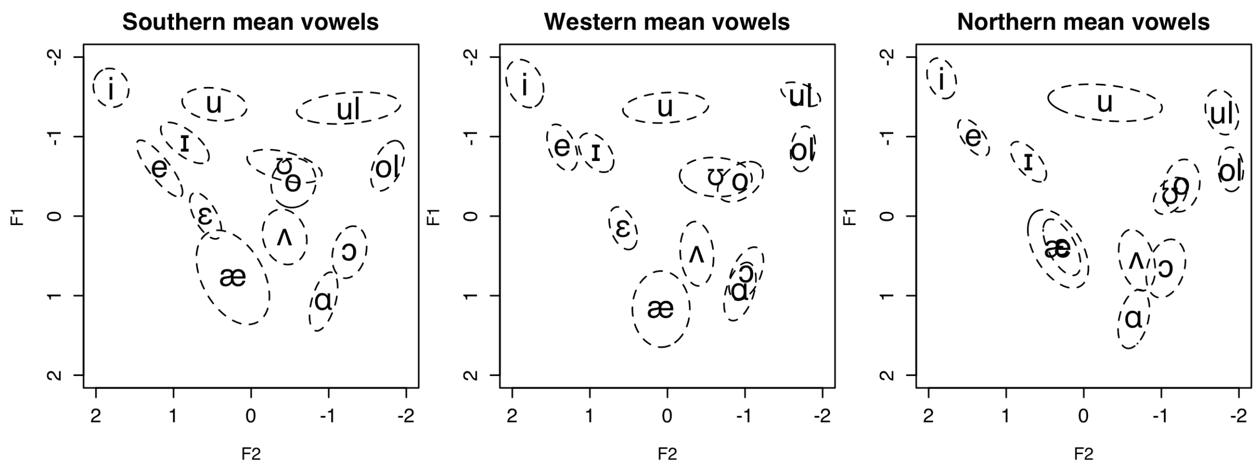
The majority of recent work examining vowel differences in U.S. dialects focuses on three major shifts, the North Cities Shift (NCS), the Southern Vowel Shift (SVS), and the California, or Canadian, Vowel Shift (CVS), affecting vowel realization in the Inland North, the South, and the West, respectively. (See Gordon, 2013; Labov, 1994; or Labov *et al.*, 2006 for greater detail and schematizations of these shifts.)

In Northern speech, NCS related shifts result in extremely raised /æ/ nuclei (often accompanied by breaking) and greatly backed and lowered /ɛ/ (and /ɪ/) nuclei (sometimes accompanied by breaking). The low back vowels remain separate categories, although they lower and front somewhat acoustically. In Southern speech, SVS related shifts lead to centralized /e/ and fronted /ɛ/, often creating spectral overlap between these two classes. Some areas of the South show a similar spectral trend for the high front classes, though this shift is typically not as extensive (Fridland, 1999; Labov *et al.*, 2006). While the NCS pattern shows advancement in younger speakers, SVS features appear to be receding in most urban areas (Baranowski, 2008; Dodsworth and Kohn, 2012; Feagin, 1986; Fridland, 2001; Labov *et al.*, 2006; Prichard, 2010) with younger speakers showing weaker participation in the front vowel tense/lax shift. As in the North, Southern low vowels remain distinct, in part due to an upglided /ɔ/ class, though younger speakers show evidence of greater nucleus overlap (Labov *et al.*, 2006) and decreasing glide differentiation (Jacewicz, Fox, and Salmons, 2011).

In Western speech, the most commonly cited vowel feature is the low back vowel merger, a merger typically absent in the other regions though thought to be rapidly expanding (Labov *et al.*, 2006). While not as well documented as the shift patterns in the North and South, some research suggests that Westerners show retraction in the front lax system, somewhat similar to that affecting Northern speech (Clopper *et al.*, 2005; Kennedy and Grama, 2012; Labov *et al.*, 2006). The greatest spectral contrast in Western vs. Northern speech is found within the low vowel system. Western speech lacks generalized /æ/ raising, a key feature of the NCS, and Northern speech lacks the low back merger, a key feature of the Western dialect.

In recently published work investigating production and perception differences among U.S. dialects, Kendall and Fridland (2012) and Fridland and Kendall (2012) found clear evidence of these three shift patterns. Figure 1 displays the (Lobanov normalized, see Watt, Fabricius, and Kendall, 2011) F1 and F2 means for the vowel onsets for 11 vowel phonemes across three regional groups (South (N = 14 speakers), West (N = 10), and North (N = 14)) for the examined speakers. In addition to the primary vowel classes presented (which do not include pre-lateral, pre-rhotic, or pre-nasal environments), the plots include pre-lateral means for the high and mid back vowels to help illustrate the overall shape of the regional vowel spaces, as pre-liquid tokens lag behind in back vowel fronting and thus define the back periphery. Ellipses depict one standard deviation around the mean for each vowel category.

Figure 1. Mean F1 x F2 vowels by region, Lobanov normalized.



Overall, these regional patterns match expectations based on the distinguishing vowel shifts affecting each region. Such phonetic differences in the speech of regionally diverse speakers seem to suggest a degree of supra-regional orientation in vowel production. An interesting question is what other phonetic cues beyond F1 and F2 might also be sensitive to regional variability. If other less-commonly studied cues, such as vowel duration, are in fact regionally variable, do they co-vary with the shifts described above?

B. Beyond spectral differences

Moving beyond the traditional formant measurements, Faber and Di Paolo (1995) found that multidimensional analysis incorporating a number of acoustic parameters beyond F1/F2 measures showed that speakers "produce and perceive complex sounds that vary simultaneously along a number of different dimensions rather than producing or perceiving any one dimension independently of the others" (60). Though they did not include duration, this work showed both that acoustic parameters beyond F1/F2 structure were important in classifying vowels and that speakers in the same location varied in the weighting of features used to distinguish among nearly merged sounds.

Similarly, in a multidimensional analysis incorporating both formant structure and duration, Wassink (2006) showed that Jamaican Creole, Jamaican English, and American English speakers varied in the degree to which duration disambiguated acoustic overlap among vowels, though all speaker groups showed decreased overlap when duration was included in the analysis.

In terms of the role of duration in the perception of vowel quality, most studies of English have suggested that duration differences across vowel categories play a minor to moderate role in vowel identity, but predominately in concert with (and subordinate to) the frequencies of the first three formants (Ainsworth, 1972; Bennett, 1968; Hillenbrand and Clark, 2000; Mermelstein, 1978; Pols, Van der Kampe and Plomp, 1969; Sawusch, 1996; Stevens, 1959). Duration has been found to have the greatest perceptual effect where spectral distinctions between vowels are small (Ainsworth, 1972; Bennett, 1968; Stevens, 1959), but most studies suggest a phonetic, not phonological, role for duration in English.

Within American dialects, Byrd (1994) looked at a number of differences related to acoustic reduction and found duration of central vowels varied significantly across regions, primarily due to longer duration in the South and South Midland. More recent studies also clearly point to regional differences in duration. Clopper *et al.* (2005) found that Southerners had longer lax vowels than speakers in other U.S. dialect regions. Jacewicz *et al.* (2007) found systematic differences in duration across regional dialects, with Southerners (from Western

North Carolina) showing significantly longer duration than Midland or Northern speakers. This pattern was true of all vowels studied (/ɪ/, /ɛ/, /e/, /æ/ and /aɪ/), not just lax vowels. In a later study, Fox and Jacewicz (2009) found that, along with durational differences, vowel inherent spectral change is a part of a complex of cues sensitive to dialect variation. In particular, measures such as vector length and trajectory change over time appear to be particularly relevant distinctions in Southern dialects, suggesting the longer duration of Southern vowels may be a related phenomenon. While the current paper does not specifically investigate spectral trajectory, it is likely that both formant movement and duration are important and related aspects of dialect variability.

A few studies have also looked at the relationship of spectral and durational features within dialects (Labov and Baranowski, 2006; Langstrof, 2009; Tauberer and Evanini, 2009). While these studies focused on different aspects of this relationship, Labov and Baranowski's work indicates that duration helps speakers differentiate vowels (such as /ɛ/ and /ɑ/ in Northern speech, the focus of their paper) where spectral overlap might cause perceptual confusion. Perception experiments involving manipulations to nucleus duration for both token-types found that duration distinctions of 50 milliseconds or more produced changes in category identification, with /ɛ/ associated with shorter durations than /ɑ/. The authors suggest that NCS-based sound change has lead to a phonological reanalysis of duration in dialects affected. Similarly, Langstrof (2009), looking at shifts in New Zealand English over time, suggests that duration may have served as a transitional cue to vowel quality when the shift involved greater spectral overlap among front lax vowels (which resolved as the shift progressed). In other words, typically an enhancing cue in English dialects, the import of duration may shift toward one of contrast, when other cues (such as F1 and F2) are no longer fully disambiguating.

Looking more at the mechanical relationship between vowel duration and spectral position, Tauberer and Evanini (2009) found that changes in vowel openness caused by regional shift patterns did not show changes in duration that would suggest a solely physiological relationship (with close vowels shorter than lower open vowels, e.g. Crystal and House, 1988; Jacewicz *et al.*, 2007; O'Shaughnessy, 1981; Peterson and Lehiste, 1960). Similarly, Langstrof's (2009) work found no relationship between shifts in F1 and vowel duration in the front lax vowel categories (though he did find significant F2/duration correlations). As that paper discusses, such results suggest the constraints proposed are not mechanically invariable intrinsic constraints and can be overridden by language or dialect specific developments.

This body of work suggests that different dialects may employ duration differently, for example, being used to enrich linguistic contrast as in the North (as suggested by Labov and Baranowski, 2006) or to reference social affiliation (as found in the South by Clopper *et al.*, 2005 and Jacewicz *et al.*, 2007). Also, duration appears to be strongly linked to other linguistic features such as vowel class and consonantal context, and in its relationship with spectral trajectory and change over time. Considering this suggestive previous work, it is surprising that duration has not been studied more extensively as a sociolinguistic variable.

II. METHODS

A. Participants

The data presented here come from a larger study investigating both production and perception across regional U.S. dialects (Fridland and Kendall, 2012; Kendall and Fridland, 2012). A subset of participants from this larger study were recruited from universities in each of the three dialect areas studied to provide speech data from reading passage and word list recitation. For the present inquiry, we examine 14 Southern participants recruited from the

University of Memphis in Tennessee, 10 Western participants recruited from the University of Nevada Reno in Nevada, and 14 Northern participants recruited from the State University of New York at Oswego. Across the field sites, thirty-four of the thirty-eight participants were under the age of 30.

Subjects' regional classifications are based on the dialect regions of the *Atlas of North American English* (Labov *et al.*, 2006). Because data were collected at universities, subjects varied somewhat in terms of where they were raised within each region. Our Northern participants were primarily from New York state and proximal areas in the Inland North, all of our Southern participants were from Tennessee, and our Western participants were from both Nevada and Northern/Central California. Thus, while all subjects showed regional variation in line with the larger shift patterns depicted above, there was some intra-regional variation in the degree of participation in these shifts, as discussed in Fridland and Kendall (2012) and Kendall and Fridland (2012). Regional mean F1 and F2 values for these speakers were plotted in Figure 1, above.

B. Data collection and analysis

The speakers were recorded with a Tascam digital recorder and a Shure WH30XLR head-mounted microphone (or, in a few cases, a Sony MZ-R70 digital recorder and an ATR 410 head-mounted condenser microphone) in a quiet university office (with just the fieldworker and participant present). This approach is typical of the fieldwork-based sociolinguistic interview style employed by most sociolinguists to elicit natural speech samples.

Speakers were recorded reading a reading passage and word list prepared by the authors designed to get a broad sample of linguistic environments for each vowel category. All vowel measurements were made using Praat (Boersma and Weenink, 2012). For all formant values reported in this paper, formant measurements were taken at 1/3 of each vowel token's duration and, as seen in Figure 1, above, normalized using the Lobanov method (Lobanov, 1971) as implemented in the Vowels.R package for the statistical computing language R (Kendall and Thomas, 2012). Formant normalization allows for cross-speaker comparisons (cf. Watt *et al.*, 2011), and the Lobanov technique is consistently found to perform well for sociophonetic purposes (e.g., Flynn, 2011). Normalization by the Lobanov method also has the added benefit of bringing F1 and F2 to similar scales, making measures like Euclidean distance (which is used below) appropriately applied to formant data. All measurements were taken from fully stressed vowels, primarily in monosyllabic words (although eight two-syllable words were used from the reading passage). In total, 95 distinct words were selected for analysis for the data presented here, with additional tokens for 18 of those word types also measured.[1] Due to mispronunciations, omissions, or, occasionally, problems with the audio, the specific number of tokens per subject varied to some small extent.

Vowel duration is normalized in the analyses below by dividing the individual duration measurement for each token by its speaker's mean vowel duration. Thus, in the discussion below, a normalized duration value of 1.0 reflects a duration equivalent to a speaker's mean duration across all vowels. Normalized durations below 1.0 reflect durations shorter than a speaker's mean and normalized durations above a 1.0 reflect longer than mean durations. Normalization is necessary to decouple vowel durational differences from articulation rate differences between speakers.[2]

Table I presents a summary of the data used in this study. In order to facilitate cross study comparisons, the table includes raw duration values, although the following analyses report results for the normalized values only.

TABLE I. Summary of duration data examined (mean values with one standard deviation reported in parentheses)

	/ɑ/			/e/			/i/		
	South	West	North	South	West	North	South	West	North
Norm'd Duration	0.87 (0.31)	0.87 (0.29)	0.95 (0.29)	1.07 (0.35)	1.07 (0.36)	1.10 (0.30)	0.90 (0.40)	0.91 (0.37)	0.89 (0.29)
Raw Duration, *ms*	178 (66)	156 (57)	184 (61)	220 (76)	192 (70)	212 (63)	185 (91)	163 (74)	172 (63)
N	154	103	146	126	91	125	92	68	95
	/ɔ/			/ɛ/			/ɪ/		
	South	West	North	South	West	North	South	West	North
Norm'd Duration	1.07 (0.44)	1.11 (0.43)	1.10 (0.37)	0.93 (0.35)	0.83 (0.31)	0.84 (0.28)	0.97 (0.39)	0.82 (0.30)	0.79 (0.24)
Raw Duration, *ms*	218 (93)	199 (82)	213 (74)	192 (77)	150 (62)	162 (58)	196 (86)	145 (60)	153 (52)
N	165	117	153	112	81	112	100	81	106

III. RESULTS

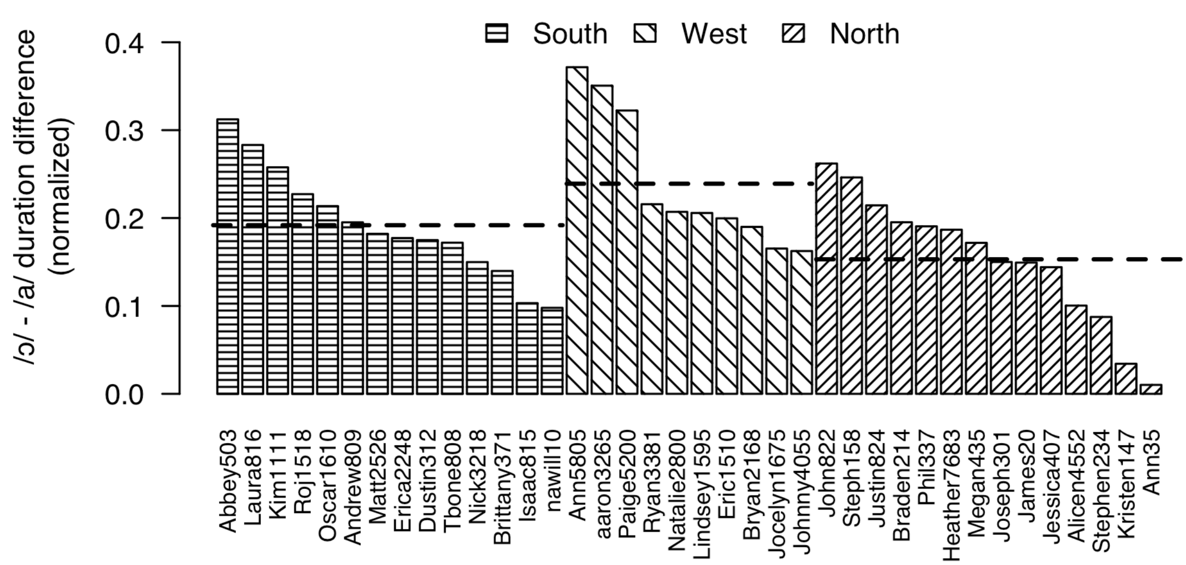
A. The low back merger

We begin our empirical analysis by examining the relationship between duration and spectral position in the low back vowels, /ɑ/ and /ɔ/. Our own data, Figure 1, confirms that the low back merger, in terms of F1/F2 space, is mainly a trait of our Western speakers, with Southerners and Northerners maintaining distinct vowel categories. However, most accounts of the low back merger (cf. Labov *et al.*, 2006) are entirely based on the spectral positions of the vowels (in terms of F1 and F2) and the role that duration plays in the merger is basically unknown.

Table I displayed the mean normalized durations for /ɑ/ and /ɔ/ for each region (South: /ɑ/ = 0.87, /ɔ/ = 1.07; West: /ɑ/ = 0.87, /ɔ/ = 1.11; North: /ɑ/ = .95, /ɔ/ = 1.10). All three regions maintain a length difference between the two vowel classes, with /ɑ/ shorter than /ɔ/, suggesting that duration helps maintain a contrast between the low back vowels for all three regions.

While the West has slightly longer normalized /ɔ/ vowels than the other regions, the difference is not significant by an ANOVA test ($F(2, 35) = 1.49$, $p = 0.24$; all ANOVAs presented here examine speaker means and not individual tokens). Northern /ɑ/ is longer than for the other regions and here the difference is significant (ANOVA: $F(2, 35) = 12.37$, $p < 0.0001$; Tukey HSD: North v. West, $p < 0.001$, North v. South, $p < 0.001$, West v. South, $p = 0.99$). This greater duration for /ɑ/ in the North is most likely related to the NCS pattern and Labov and Baranowski's (2006) finding that increased spectral overlap with /ɛ/ resulted in lengthening of /ɑ/. As a result, this difference in duration seems to be an effect of a specifically local (Northern) shift pattern where increased acoustic overlap leads to length playing a phonemic role. The low back merger in the West presents a similar situation of spectral overlap and we explore here whether duration is likewise playing an increased role in maintaining contrast for the low back vowels in the West.

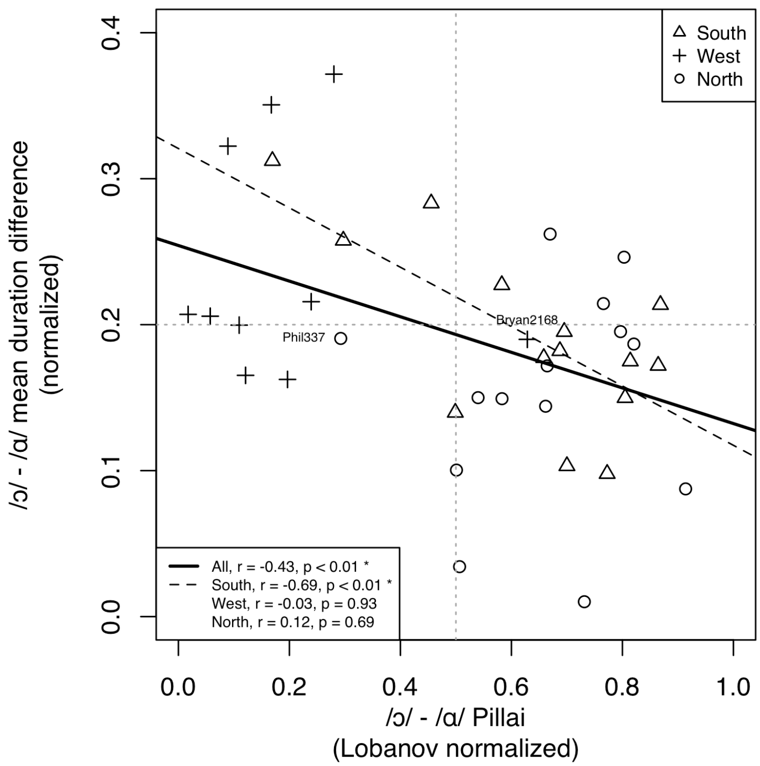
Figure 2. Individual duration differences between /ɔ/ and /ɑ/. (Dashed lines indicate mean duration difference for each region.)



To examine this more closely, Figure 2 plots these durational differences between /ɔ/ and /ɑ/ at the individual level, ordered first by region and then, within region, from greatest to least difference. Regional mean durational differences are depicted with dashed lines. Here rather than a difference between the North and the other regions it is the West that stands out. Western speakers have larger duration differences between /ɔ/ and /ɑ/ than the other regions (although the difference is only significant between the West and the North; ANOVA: $F(2, 35) = 4.28$, $p < 0.05$; Tukey HSD: West v. North, $p < 0.05$, West v. South, $p = 0.26$, North v. South, $p = 0.33$). It is clear from the bar plot that this aggregate difference is driven by the leftmost and rightmost rank-ordered speakers in the West who show much more distinction between these two vowel classes than the most distinct and least distinct speakers in each of the other regions. That not all the Westerners show this pattern is not surprising, as Faber and Di Paolo's (1995) work suggests speakers use different combinations of acoustic parameters in disambiguating near-mergers and that this may shift over time, depending on the direction and ultimate outcome (e.g. if eventual full merger is achieved). In addition, speakers from within the same region and even within the same location are not uniform in the degree to which they conform to variable aspects of local dialect patterns (see Kendall and Fridland, 2012 for expanded discussion of intra-regional variation in shift advancement in these regional vowel patterns).

Overall, Western speakers do show significantly more durational contrast between the low vowel classes than the other regions. Such a finding leads to the question of whether there is a correlation between degree of low back vowel merger, in terms of spectral position, and duration among these speakers. Figure 3 plots the normalized duration difference between /ɔ/ and /ɑ/ against the Pillai score for the two vowels' distributions. Pillai is a useful measure of overlap/non-overlap in distributions (Hay, Warren, and Drager, 2006; Hall-Lew, 2010), which arises as a resulting statistic from a MANOVA analysis (here of F1 and F2 as a function of vowel class). Pillai scores closer to 1 indicate non-overlapping distributions and scores closer to 0 indicate overlapping distributions.[3]

Figure 3. /ɔ/ - /ɑ/ Pillai against /ɔ/ - /ɑ/ duration difference.



Two main "clusters" of speakers emerge from this view – a merged cluster in the upper-left quadrant where speakers have mostly overlapping spectral positions for their low back vowels coupled with a normalized duration difference of greater than 0.15 and an unmerged cluster on the lower-right. Interestingly these unmerged speakers have much smaller duration differences between the low back vowels, with several speakers approaching no difference between the vowels' durations. The correlation between duration difference and Pillai score is significant, with an overall Pearson's r of -0.43 and $p < 0.01$.

In terms of regional patterns, there is a clear distinction between Western and Northern speakers, with only one Western speaker (Bryan2168, labeled in the figure) falling in the main unmerged cluster and only one Northern speaker (Phil337, also labeled in the figure) appearing in the merged cluster. At the same time, while many Southern speakers appear in the unmerged, lower-right cluster as would be expected from previous findings on the South and the separation depicted in the mean Southern vowel plot in Figure 1, there are merged Southern speakers and these spectrally merged Southerners follow the general patterns of having longer duration differences than their unmerged peers. The correlation between duration difference and Pillai is also significant for just the Southern speakers (r = -0.69, p < 0.01). For the North and West, the correlation is not significant, a finding inline with the generally clustered nature of those groups.

In summary, it appears from these data that in American English, regardless of region, there is in fact a durational distinction that actually maintains the contrast between the low back vowel classes. Speakers with greater spectral differentiation of /ɑ/ and /ɔ/ show smaller durational distinctions between the two vowels than speakers showing merger. This pattern is more prominent in the Western and Southern data where spectral overlap is more likely. While F1/F2 measures might suggest full merger in the West, as we see from this data, the underlying contrast between the vowels is actually maintained by duration. This, of course, may be relevant to language change over time, as such 'near-merger' has been suggested to underlie the unmerging of other apparent mergers in the historical record (see Labov, 1994 for more expansive discussion).

B. The mid front vowels

Movement in the relative position of the mid front vowels, /e/ and /ɛ/, is a primary reflex of the SVS and, perhaps, the major way in which speakers affected by the SVS and NCS differ in their vocalic configuration. Though this shift is showing retraction in the South, even young Southerners still show more proximate F1/F2 position for the mid vowel classes (as seen in

Figure 1, above). As these vowels fall into separate tense and lax categories, some associated differences in duration might be expected, as earlier research suggests tense/lax distinctions in American English are also typically associated with shorter durations for lax vowels (Klatt, 1976; Peterson and Lehiste, 1960). In addition, the increased spectral overlap often found for mid front vowels in the South suggests that other acoustic cues might enhance the tense/lax distinction. Given the role found for duration in maintaining distinctions between the low vowels discussed above, the spectral proximity of the tense and lax vowels in the South leads to the question of whether duration may play a similar disambiguating role in the Southern /e/-/ɛ/ reversal.

Table I presented the normalized duration means for /e/ and /ɛ/ for each region (South: /e/ = 1.07, /ɛ/ = 0.93; West: /e/ = 1.07, /ɛ/ = 0.83; North: /e/ = 1.10, /ɛ/ = 0.84). Here, the data show that Southern /ɛ/ is significantly longer than Northern or Western /ɛ/ (ANOVA: $F(2, 35) = 9.28$, $p < 0.001$; Tukey HSD: South v. West, $p < 0.01$, South v. North, $p < 0.01$, West v. North, $p = 0.99$). This lengthening might be related to the tendency toward lax vowel centralization and diphthongization in the SVS pattern, which essentially results in a "tenser" realization of the front lax vowel in the South as noted by Labov (1994; and in New Zealand English by Langstrof, 2009). No significant regional duration differences appear for /e/ (ANOVA: $F(2, 35) = 0.87$, $p = 0.43$). Overall, as a result of longer /ɛ/ and equivalent /e/ durations, the South shows a significantly smaller duration difference in the tense/lax pair than the other regions (ANOVA: $F(2, 35) = 8.24$, $p < 0.01$; Tukey HSD: South v. West, $p < 0.05$, South v. North, $p < 0.01$, West v. North, $p = 0.68$).

Figure 4. Individual duration differences between /e/ and /ɛ/. (Dashed lines indicate mean duration difference for each region.)

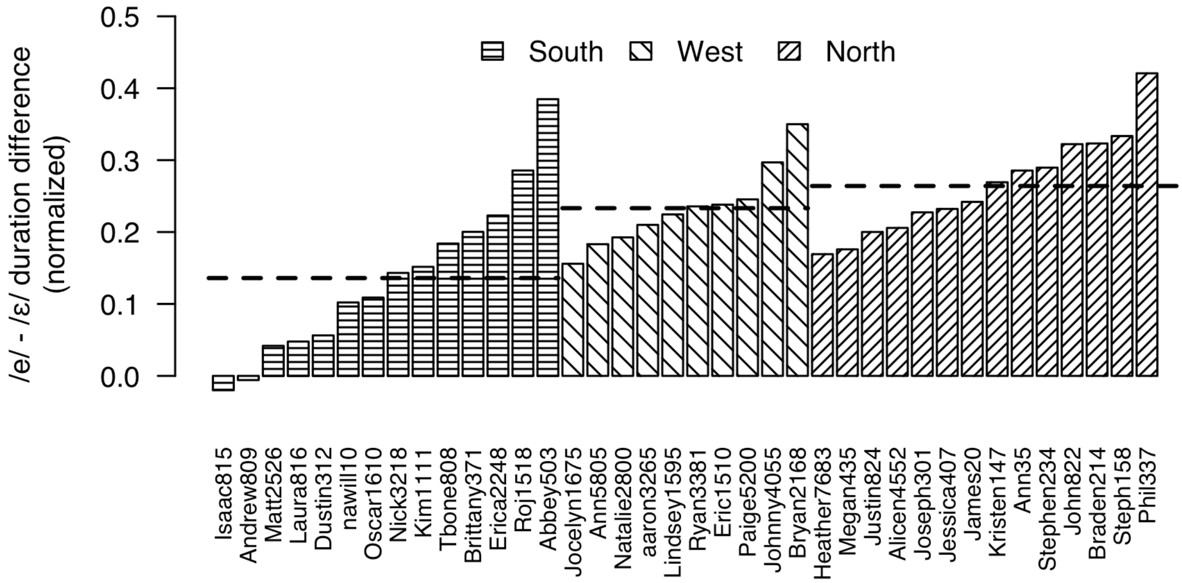
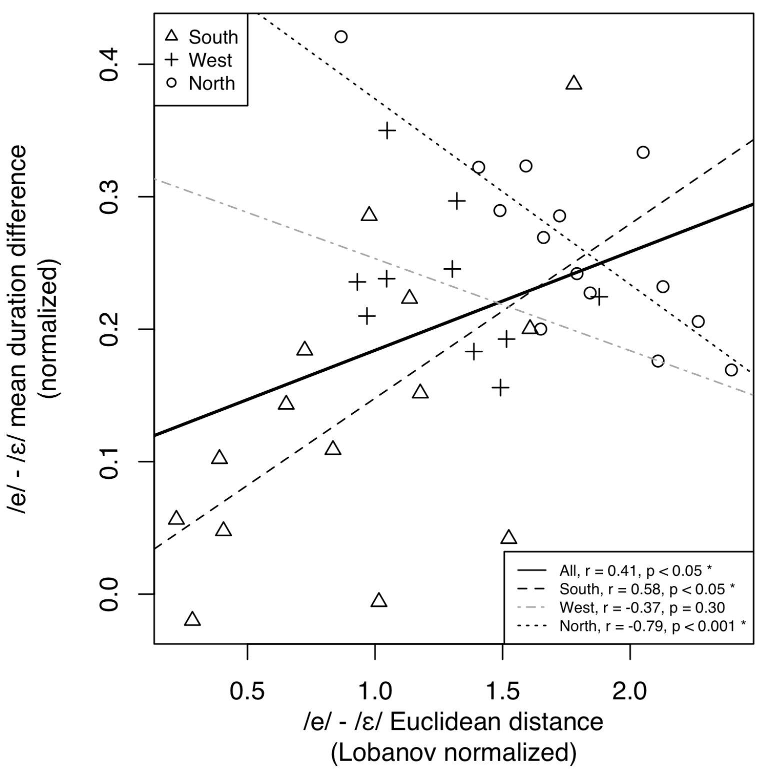


Figure 4 displays the duration differences for the individuals within each region. Here, as with the regional means in Table I, the South stands out as having a smaller durational difference overall. We also notice that the difference is primarily at the left edge of the distribution; half of the Southerners have durational differences below any of the Northerners or Westerners, although the remaining Southerners are on par with the non-Southern groups. This variability seems to be related to the variable degree to which younger speakers participate in SVS shifts more generally (Kendall and Fridland, 2012), a relationship explored in Figure 5.

To examine whether greater spectral overlap resulting from Southern shifted /e/ and /ɛ/ contributes to durational differences, Figure 5 presents a similar plot as was shown for the low back vowels in the previous section. Here, the y-axis displays the normalized duration difference between /e/ and /ɛ/ while the x-axis displays the Euclidean distance between these two vowels' mean points for each speaker. Euclidean distance is used here, rather than Pillai, since the overall distance between the category means is of interest, not the extent to which the distributions overlap (recall the relative lack of direct overlap for /e/ and /ɛ/ from Figure 1, in comparison to /ɑ/ and /ɔ/). The North (Pearson's $r = -0.79$, $p < 0.001$) shows an inverse relationship for the spectral distance (between the onsets of /e/ and /ɛ/) and duration differences. For the Northerners increased distances between /e/ and /ɛ/ (a result of NCS shifts) correspond with lesser duration differences. Though non-significant ($p = 0.30$), we also find a similar tendency in the Western data (this is displayed in the plot through a grayed out regression line). For the South, however, this pattern is strikingly reversed (Pearson's $r = 0.58$, $p < 0.05$). The Southerners with more proximate vowels are also the speakers with the least difference between those vowels' durations. This pattern is so pronounced for the Southerners that it leads to a significant positive overall correlation across all of these speakers (Pearson's $r = 0.41$, $p < 0.05$), as depicted by the solid line in Figure 5.

Figure 5. /e/ - /ɛ/ spectral Euclidean distance against /e/ - /ɛ/ duration difference.



This result contrasts with that found for the low vowels in the South, where decreased spectral overlap correlates with increased durational differences, and seems surprising, as earlier research has suggested that duration tends to become more important as a disambiguating cue when spectral distinctions are small (e.g. Ainsworth, 1972; Bennett, 1968; Labov and Baranowski, 2006; Stevens, 1959), a pattern supported by the low back vowel analysis of III.A. Based on these results, Southern tense and lax vowels would appear to be heading toward F1/F2 and durational merger.

Nonetheless, despite the significantly greater length of Southern lax vowels compared to those in other regions, Southerners still recognize tense and lax vowel differences. In a vowel identification task performed by the same subjects (Fridland and Kendall, 2012; Kendall and Fridland, 2012) participants clearly identified separate tense and lax front vowel categories. In other words, there is something beyond F1 and F2 nucleus position and duration that maintains these vowels' distinctiveness. It seems likely that the distinctions might be tied to vowel trajectory, not just onset position, as breaking is considered a characteristic feature of both Southern tense and lax vowels, with glide directionality traditionally being considered the distinguishing feature. As first noted by Sledd (1966), the front vowels in Southern speech are often typified by triphthongization and breaking with upglided tense vowels and inglided lax vowels. The common stereotype of the Southern drawl may, in fact, be an indicator of what in Southern speech in particular maintains contrasts among vowels that do not exhibit spectral onset or durational differences such as /e/ and /ɛ/. Such an indication is supported by preliminary work by Fox and Jacewicz (2009) that suggests that vector length and total trajectory change across time are distinguishing for a number of vowels in North Carolina and Ohio. Our results also suggest that other less studied acoustic parameters such as spectral slope and trajectory, in concert with durational differences, need to be explored as potential disambiguating cues, particularly in Southern speech.

C. The high front vowels

The high front vowels, /i/ and /ɪ/, are generally seen as following a similar spectral trajectory as the mid front vowels for the SVS, although this tendency has often been found to be less strong/advanced than for the mid front vowels (Fridland, 1999, 2001; Labov *et al.*, 2006). This is the case for the Southern speakers examined here, who have somewhat less proximate high front vowels than mid front vowels, as seen in Figure 1, above. Thus, Southern high front vowels are not as spectrally different from the other regions, at least in terms of F1/F2 onset position. Given the regional differences found for the mid front vowels as just discussed, it is worth asking what role, if any, does duration play in differentiating the regional productions of the high front vowels?

Table I provided /i/ and /ɪ/ normalized durations across regions (South: /i/ = 0.90, /ɪ/ = 0.97; /i/ = 0.91, /ɪ/ =0.82; North: /i/ = 0.89, /ɪ/ = 0.79). Here, the same general relationship for /i/ and /ɪ/ is apparent as for the mid front vowels, namely that there are significant differences in lax vowel length in the South compared to the other regions. As with the mid front vowel data, the durations of the tense category, /i/, are not significantly different across regions (ANOVA: $F(2, 35) = 0.55$, $p = 0.58$), but the lax category, /ɪ/, is highly significantly different (ANOVA: $F(2, 35) = 27.33$, $p < 0.0000001$; Tukey HSD: South v. West, $p < 0.0001$, South v. North, $p < 0.000001$, West v. North, $p = 0.71$). In fact, these differences are more pronounced for the high vowels, with most of the Southerners in these data having longer /ɪ/ vowels than /i/. This is shown in Figure 6, which displays the individual differences between /i/ and /ɪ/ durations. The much longer lax vowels of the South lead to much lower (mostly negative) durational differences between /i/ and /ɪ/ (ANOVA: $F(2, 35) = 17.82$, $p < 0.00001$; Tukey HSD: South v. West, $p < 0.0001$, South v. North, $p < 0.0001$, West v. North, $p = 0.99$).

Figure 6. Individual duration differences between /i/ and /ɪ/. (Dashed lines indicate mean duration difference for each region.)

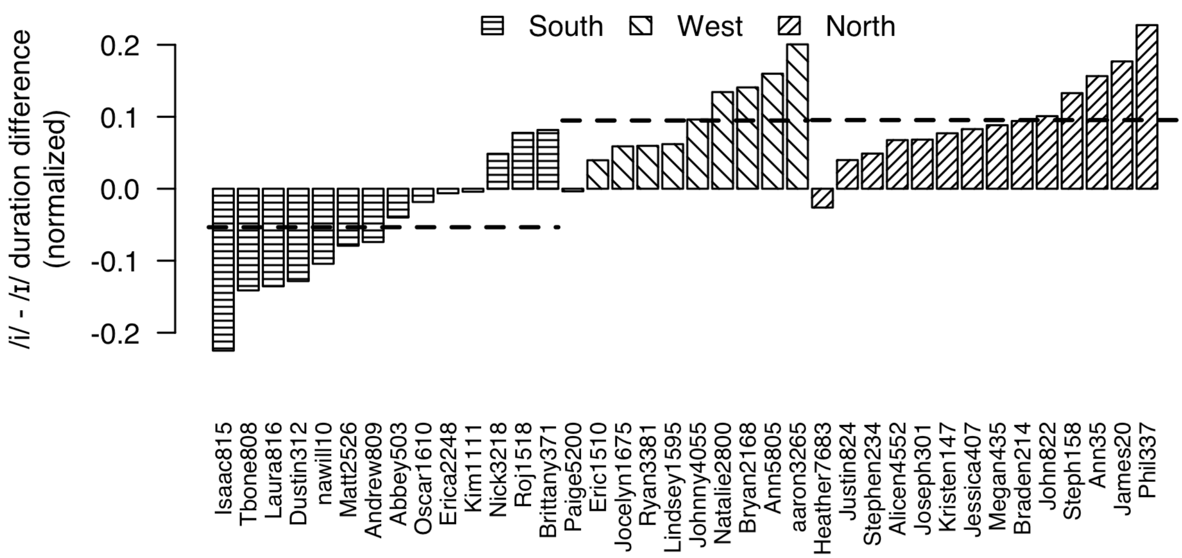
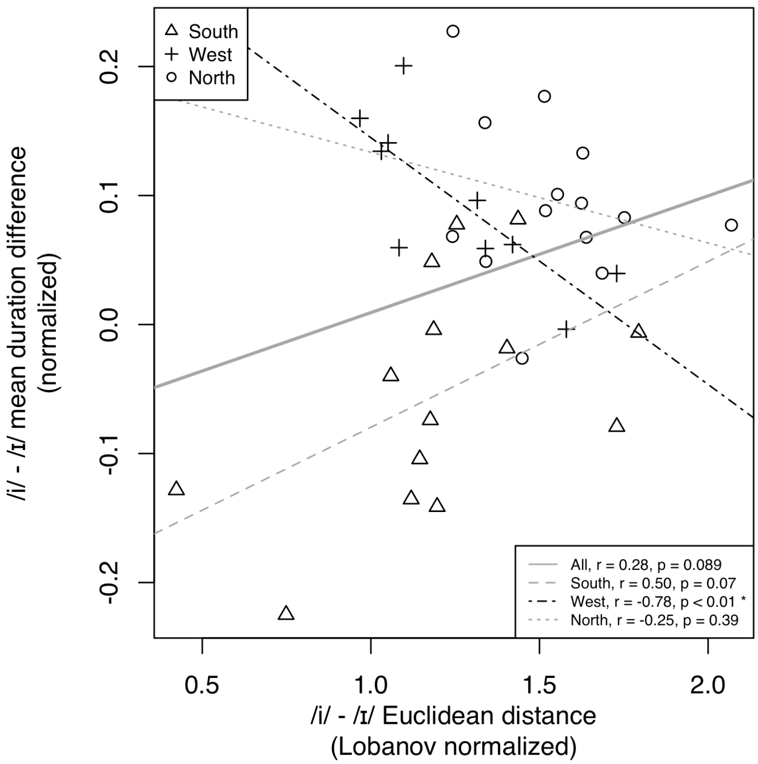


Figure 7 displays a plot identical to that for the mid front vowels for /i/ and /ɪ/, showing duration plotted against the Euclidean distance between vowel classes. Although the correlations are weaker here than for the mid front vowels, the same general trends emerge. Northerners and Westerners have an inverse relationship between spectral distance and durational difference with durational differences decreasing as vowels become farther apart spectrally. The correlation for the West is significant (Pearson's $r = -0.78$, $p < 0.01$), but for the North it is not (Pearson's $r = -0.25$, $p = 0.39$). Southerners on the other hand show the same positive relationship seen for mid front vowels, although here it does not reach significance (Pearson's $r = 0.50$, $p = 0.07$). Due to the distribution of the Southern speakers the overall correlation is also positive, though it too fails to reach significance (Pearson's $r = 0.28$, $p = 0.09$). As before, in order to better highlight the trends in the data, the figure displays fit lines for each group, as well as the combined speakers, with the non-significant correlations grayed out.

Figure 7. /i/ - /ɪ/ spectral Euclidean distance against /i/ - /ɪ/ duration difference.



These findings suggest that, like the mid front vowel tense/lax relationship, duration is simply not a particularly useful cue in the Southern high front system, perhaps due to other, more primary acoustic differences (such as trajectory) that work in concert with duration. In contemporary Northern and Western speech, we find backing and lowering of the high and mid front vowels while Southern speech shows a completely opposite trend. Our results comparing duration and spectral shift reflect the fact that the distinctions between these vowels are fundamentally different than those between the front tense and lax vowels in Northern and Western speech.

IV. CONCLUSION

Speech production data from Southern, Western, and Northern speakers display differences in vowel duration both across and within regions. For the low back vowels, though all three regions maintain durational contrasts that may aid in disambiguation between the vowels, Northerners have significantly longer /ɑ/ vowels, suggesting /ɑ/ lengthening resulting from NCS related spectral overlap with /ɛ/ (as reported by Labov and Baranowski, 2006). For the mid and high front vowels, the data indicate that Southerners have on average smaller (or negative for /i/ - /ɪ/) durational contrasts between their tense and lax vowels than the other regions. In particular, Southern speakers display significantly longer lax vowels for these classes compared to other regions. These results bolster the findings of Clopper *et al.* (2005) and Jacewicz *et al.* (2007) that show that duration is systematically variable by region, and suggest that duration functions differently in Southern compared to Northern and Western speech.

What these previous studies did not examine is how duration differences may be tied to spectral differences in production regionally and how variable a role duration plays depending on dialect. In the work here, systematic duration differences were found both between regions and between speakers showing greater and lesser degrees of spectral overlap as a result of regional shift patterns. Interestingly, although reduced spectral shift correlated with greater duration differences for the low back vowels inline with some other recent findings (Labov and Baranowski, 2006; Langstrof, 2009; Wassink, 2006), this relationship was reversed for the tense/lax vowel pairs for Southern speakers. Lax vowels in the South were significantly longer than those in other dialects (even reversing the typical length relationship for some speakers) and individual speakers with increased tense/lax spectral overlap in the South actually showed less durational contrast as well. An explanation for this result may rest in the unique front vowel system of the South, where lax vowel length and breaking is a long-standing feature. Instead of being a primary cue for these vowels, duration, likely along with spectral change over time, may be part of a package of acoustic distinctions that signals both dialect and vowel category information.

What is clear from our work is that while all three regions are undergoing vowel formant shifts that lead to great distinctions across dialects, Southerners are much more distinct from other dialects, both in formant position and duration, suggesting there is much beyond F1 and F2 measures that may be socially-relevant. In addition, our findings for the low back vowels suggest that duration can be utilized by speakers to maintain contrast in otherwise merged vowel classes, reiterating that considerations of vowel merger cannot stop at overlapping spectral positions alone.

ACKNOWLEDGMENTS

This research has been supported by National Science Foundation grants # BCS-0518264 & BCS-1123460 (PI Fridland), and BCS-1122950 (PI Kendall). We also thank Craig Fickle at the University of Oregon, and Sohei Okamoto at the University of Nevada, Reno, for support with various aspects of this research. Portions of this work were presented at the 2013 International Congress on Acoustics (Fridland, Kendall, and Farrington, 2013) and at the 2011 Annual Meeting of the American Dialect Society. We thank the audiences at those presentations as well as the editor and three anonymous reviewers for helpful comments on earlier versions of this work.

NOTES

[1] The following words were used for vowel measurements.

/i/: *BEAD, DEED, DEEP* (x2)*, eat(ing), PEEK, teas(ing), SEAT,* & *she*.

/ɪ/: *BID, big, BIT, DID* (x2)*, DIP, DISH, KID, kid, sit,* & *TICK*.

/e/: *BADE, BAIT, cake, DATE* (x2)*, GATE, GAVE, haz(y), Kate,* & *TAKE*.

/ɛ/: *BED* (x2)*, BEG, BET, DEAF, DEBT, head(first),* & *Peg* (x3).

/æ/: *aft(er), back, black, chat, DAD, had, HALF, PAD,* & *SAD*.

/ɑ/: *COP, cop, COT, DOCK, fath(er), HOCK, hot* (x3)*, POD, POT, pot,* & *stop*.

/ɔ/: *BOUGHT, CAUGHT, caught, CAUSE, COUGH, DOG, dog, dogs, HAWK, PAW, slosh, talk,* & *toss*.

/ʌ/: *BUT, But, DOES, DUTCH, fun, stuck, sudden, TUCK, UP,* & *up* (x2).

/o/: *BOAT, boat, DOZE, GHOST, POKE, Post, showed (x2),* & *SOAP*; /ol/: *COAL*.

/ʊ/: *Book(ie), brook, FOOT, foot(ing), good* (x2)*, HOOD, put,* & *TOOK*.

/u/: *BOOED, BOOT, dew, DUDE, Duke* (x2)*, shoes, SOUP, soup,* & *SUE*; /ul/: *SCHOOL*.

Words in all-caps are from the word list, while words in sentence case were taken from the reading passage (two-syllable words are listed with the non-measured vowel in parentheses). As noted in the list, in some cases more than one token of a few word types was taken from additional locations in the reading materials. The reading passage itself is available in Kendall (2013: 56-57). The words appeared on the word list in a random order, interspersed with additional word stimuli not used here, but all subjects received identical word lists to minimize prosodic differences due to word order.

[2] Kendall (2013) examined variation in articulation rate for these data. That study found significant articulation rate differences by region with the Westerners speaking (i.e. reading the

passage) significantly faster than the South and North. The South was the slowest, although Southern rates were not significantly different than the North (Kendall, 2013: 75-78). No sex differences were found for articulation rate. Normalizing duration helps to control for these regional articulation rate differences. The normalized durations do not correspond exactly linearly to actual, raw differences. However, a linear model predicting raw duration differences from these normalized differences indicates that they are highly significantly correlated (adjusted $R^2 = 0.88$, $p < 0.0001$) with each 0.1 increment in normalized duration difference corresponding to 18.5 ms in raw duration difference.

[3] Hall-Lew (2010) provides a discussion of generating Pillai scores through MANOVA in the programming language R.

REFERENCES

Ainsworth, W.A. (**1972**). "Duration as a cue in the recognition of synthetic vowels," J. Acous. Soc. Am. **99**, 2350-2357.

Baranowski, M. (**2008**). "The Southern Shift in a marginally Southern dialect," U. Penn. Working Papers in Ling **14**, 35-43.

Bennett, D.C. (**1968**). "Spectral form and duration cues in the recognition of English and German vowels," Lang. and Speech **11**, 65-85.

Boersma, P., and Weenink, D. (**2012**). "Praat: Doing phonetics by computer," Version 5.3.05.

Byrd, D. (**1994**). "The relation of sex and dialect to reduction," *Sp. Comm.* 15, 39-54.

Clopper, C., Pisoni D., and de Jong, K. (**2005**). "Acoustic characteristics of the vowel systems of six regional varieties of American English," J. Acous. Soc. Am. **118**, 1661-76.

Crystal, T., and House, A. (**1988**). "Segmental durations in connected-speech signals: Current results," J. Acous. Soc. Am. **83**, 1553-1573.

Dodsworth, R., and Kohn, M. (**2012**). “Urban rejection of the vernacular: The SVS undone,” Lang. Var. Change **24**, 221-245.

Eckert, P. (**1988**). “Adolescent social structure and the spread of linguistic change,” Lang. Var. Change **1**, 245-208.

Eckert, P. (**2000**). *Linguistic Variation as Social Practice* (Blackwell, Oxford), pp. 85-101.

Faber, A., and M. Di Paolo (**1995**). “The discriminability of nearly merged sounds,” Lang. Var. Change **7,** 35-78.

Feagin, C. (**1986**). “More evidence for vowel change in the South,” in *Diversity and Diachrony*, edited by D. Sankoff (John Benjamins, Amsterdam and Philadelphia), pp. 83-95.

Flynn, N. (**2011**). “Comparing vowel formant normalisation procedures,” York Working Papers in Linguistics **11**, 1-28.

Fridland, V. (**1999**). “The Southern Vowel Shift in Memphis, TN,” Lang. Var. Change **11**, 267-285.

Fridland, V. (**2001**). “Social factors in the Southern Shift: Gender, age and class,” J. Sociolinguistics **5**, 233-53.

Fridland, V., and Bartlett, K. (**2006**). “The social and linguistic conditioning of back vowel fronting across ethnic groups in Memphis, TN,” Eng. Lang. and Ling. **10**, 1-22.

Fridland, V., and Kendall, T. (**2012**). “Exploring the relationship between production and perception in the mid front vowels of U.S. English,” Lingua **122**, 779-93.

Fridland, V., Kendall, T., and Farrington, C. (**2013**). “The role of duration in regional U.S. vowel shifts,” POMA **19**. pp. 060296.

Fox, R. A., and Jacewicz, E. (**2009**). “Cross-dialectal variation in formant dynamics of American English vowels,” J. Acoust. Soc. Am. 126, 2603-2618.

Gordon, M. (**2005**). “The Midwest and West,” in *Handbook of Varieties of English: The Americas and Caribbean, Vol I: Phonology*, edited by E. Schneider (Mouton de Gruyter, Berlin), pp. 338–350.

Gordon, M. (**2013**). “Investigating chain shifts and mergers,” in *The Handbook of Language Variation and Change*, 2nd ed., edited by J.K. Chambers and N. Schilling (Wiley-Blackwell, Oxford), pp. 203-219.

Hall-Lew, L. (**2010**). “Improved representation of variance in measures of vowel merger,” POMA **9**, pp. 060002.

Hay, J., Warren, P., and Drager, K. (**2006**). “Factors influencing speech perception in the context of a merger-in-progress,” J. Phon. **34**, 458-484.

Hillenbrand, J., and Clark, M. (**2000**). “Some effects of duration on vowel recognition,” J. Acous. Soc. Am. **108**, 3013-3023.

Jacewicz, E., Fox, R., and Salmons, J. (**2007**). “Vowel duration in three American English dialects,” Am. Speech **82**, 367-85.

Jacewicz, E., Fox, R., and Salmons, J. (**2011**). “Cross-generational vowel change in American English,” Lang. Var. Change **23**, 45-86.

Kendall, T. (**2013**). *Speech Rate, Pause, and Sociolinguistic Variation: Studies in Corpus Sociophonetics* (Palgrave Macmillan, Houndmills Basingstoke), pp. 64-80.

Kendall, T., and Fridland, V. (**2012**). “Variation in perception and production of mid front vowels in the U.S. Southern Vowel Shift,” J. Phonetics **40**, 289-306.

Kendall, T., and Thomas, E.R. (**2012**). Vowels.R: Vowel manipulation, normalization, and plotting. R package version 1.2. http://CRAN.R-project.org/package=vowels (date last viewed 2/6/14).

Kennedy, R., and Grama, J. (**2012**). “Chain shifting and centralization in California vowels: An acoustic analysis,” Am. Speech **87**, 39-56.

Klatt, D. (**1976**). “Linguistic uses of segmental duration in English: Acoustic and perceptual evidence,” J. Acous. Soc. Am. **59**, 1208-1221.

Labov, W. (**1991**). “The three dialects of English,” in *New Ways of Analyzing Variation*, edited by P. Eckert, (Academic Press, New York), pp. 1-44.

Labov, W. (**1994**). *Principles of Linguistic Change: Internal Factors* (Blackwell, Malden, MA), pp. 113-292, 293-390.

Labov, W., Ash, S., and Boberg, C. (2006). *The Atlas of North American English: Phonetics, Phonology and Sound Change*, (De Gruyter, Berlin).

Labov, W., and Baranowski, M. (**2006**). “50 msec,” Lang. Var. Change **18**, 1-18.

Langstrof, C. (**2009**). “On the role of duration in the New Zealand English front vowel shift,” Lang. Var. Change **21**, 437-53.

Lobanov, B. M. (**1971**). “Classification of Russian vowels spoken by different speakers.” J. Acous. Soc. Am. **49**, 606-608.

Mermelstein, P. (**1978**). “On the relationship between vowel and consonant identification when cued by the same acoustic information,” Percept. Psychophys. **23**, 331-336.

O'Shaughnessy, D. (**1981**). “A study of French vowel and consonant durations,” J. Phonetics **9**, 385-406.

Peterson, G., and Lehiste, I. (**1960**). “Duration of syllable nuclei in English,” J. Acous. Soc. Am. **32**, 693-703.

Pols, L. C. W., van der Kamp, L. J. Th., and Plomp, R. (**1969**). “Perceptual and Physical Space of Vowel Sounds,” J. Acous. Soc. Am. **46**, 458-467.

Prichard, H. (**2010**). “Linguistic Variation and Change in Atlanta, Georgia,” U. Penn. Working Papers in Linguistics **16**, 141-149.

Sawusch, J.R. (**1996**). “Effects of duration and formant movement on vowel perception,” ICSLP **4**, 2482-2485.

Sledd, J.H. (**1966**). “Breaking, Umlaut, and the Southern drawl,” Language **42**, 18-41.

Stevens, K.N. (**1959**). “Effect of duration on identification,” J. Acous. Soc. Am. **31**, 109.

Tauberer, J., and Evanini, K. (**2009**). “Intrinsic vowel duration and the post-vocalic voicing effect: Some evidence from dialects of North American English,” Proceedings of Interspeech 2009*,* 2211-2214.

Thomas, E.R. (**1989**). “The Implications of /o/ Fronting in Wilmington, North Carolina,” Am. Speech **64**, 327-333.

Thomas, E.R. (**1997**). “A rural/metropolitan split in the speech of Texas Anglos,” Lang. Var. Change **9**, 309-332.

Thomas, E.R. (**2001**). *An acoustic analysis of vowel variation in New World English* (Duke University Press, Durham, NC), pp. 15-160.

Wassink, A. (**2006**). “A geometric representation of spectral and temporal vowel features: Quantification of vowel overlap in three linguistic varieties,” J. Acous. Soc. Am. **119**, 2334-2350.

Watt, D., Fabricius, A., and Kendall, T. (**2011**). “More on vowels: Plotting and normalization,” in *Sociophonetics: A Student's Guide*, edited by M. Di Paolo and M. Yaeger-Dror (Routledge, London), pp. 107-118.