



# Exploring the relationship between production and perception in the mid front vowels of U.S. English

Valerie Fridland<sup>a,\*</sup>, Tyler Kendall<sup>b</sup>

<sup>a</sup> Department of English, MS 0098, University of Nevada, Reno, NV 89557, United States

<sup>b</sup> Department of Linguistics, 1290 University of Oregon, Eugene, OR 97403, United States

## ARTICLE INFO

### Article history:

Received 1 April 2011

Received in revised form 1 August 2011

Accepted 21 December 2011

Available online 31 January 2012

### Keywords:

U.S. English

Sociophonetics

Speech perception

Southern Vowel Shift

Northern Cities Shift

## ABSTRACT

A combined vowel perception/vowel production study was designed to address the question of how variability in speech production relates to variability in speech perception. For the study, participants from three different U.S. dialect regions (North, South and West) completed an online vowel identification experiment. A subset of participants from each region also provided speech samples. In this paper, we acoustically analyze these participants' vowel productions for the mid front vowel classes, and compare these productions to the same speakers' performance on a mid front vowel categorization task, asking what kinds of links exist between speakers' actual speech production and their perception of vowel categories. We are particularly interested in the role that regional vowel shifts currently affecting speech in the North and the South of the U.S. (the Northern Cities Shift and the Southern Vowel Shift, respectively) play in both production and perception. In addition to examining differences across regions, we investigate intra-regional variation in terms of individuals' productive realization of shift and the extent to which this correlates with the same individuals' perceptions. Results show that both regional affiliation and individual participation in regional shifts in production play a role in perception of the /e/-/ɛ/ continuum.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

A great deal of work in sociophonetics has made clear the central role of variability in displaying social distinctions, such as those associated with region, gender or ethnicity. In order for variation to mark such distinctions requires, however, not just its realization in production, but, crucially, the recognition of such variation as meaningful within its communities of use. Yet, research on the sociolinguistic aspects of perception has lagged far behind that of production studies. In addition, much sociolinguistic work on variability, particularly vowel variation, focuses on the dialects of communities as abstractions, where language resides not with the individual but within the shared system of the larger community (Labov et al., 1972; Weinreich et al., 1968; but see Johnstone, 1996; Macaulay, 1991). This paper attempts to look at both the level of the individual and the level of the community in examining vowel production and vowel perception. The central question driving this research is *how does variability in speech production relate to variability in speech perception?*

A number of recent studies have examined how changes in the vowel systems in American dialects are leading toward more regional distinction in the production of vowels phonetically (e.g. Eckert, 1988, 2000; Labov et al., 2006; Thomas, 2001). Three main patterns of regional shift affecting U.S. dialects have been widely documented: The Northern Cities Shift

\* Corresponding author. Tel.: +1 775 784 7546.

E-mail addresses: [fridland@unr.edu](mailto:fridland@unr.edu) (V. Fridland), [tsk@uoregon.edu](mailto:tsk@uoregon.edu) (T. Kendall).

(NCS), the Southern Vowel Shift (SVS) and the California (or Canadian) Vowel Shift (CVS). These distinctions primarily seem to affect front vowels, resulting in greater regional divergence in the front system while more similar tendencies (e.g. back vowel fronting) affect the back system. For speakers in the North, NCS affected systems typically show backed and/or lowered short front (/ɪ/ and /e/) vowels, raised /æ/ and peripheral front tense (/i/ and /e/) classes (Gordon, 1997, 2002; Labov, 1994; Labov et al., 2006). In contrast, in the South, systems affected by the SVS show strong /e/ centralization and /ɛ/ peripheralization which sometimes results in the acoustic overlap or reversal of these vowels (Feagin, 1986; Fridland, 1998, 2000, 2001; Labov, 1991, 1994, 2001; Labov et al., 2006; Thomas, 2001). In some areas of the South (mainly Appalachia), there is similar reversal found for the high front vowels (/i/ and /ɪ/), though this is less common. Finally, for areas of the West participating in a shift known as the California or Canadian vowel shift (CVS), we find the falling of the short front classes making Western /ɪ/ and /e/, in such cases, somewhat similar to the lowered variants found in the North (Clarke et al., 1995; Luthin, 1987; Thomas, 2001). However, the presence of widespread low-back vowel merger in the West and /æ/ raising in the North uniquely identifies these regions (Labov et al., 2006).<sup>1</sup> Presumably, alignment with these shift patterns identifies speakers as members within a regional speech community, though not all community members will show the same degree of participation in shifts. Yet, as we explore in the next section, very little research has explicitly studied differences that may exist in vowel perception as a result of or in relation to these active vowel shifts in production.

The current study attempts to explicitly tease out the link between regionally diverse speakers' own production of variable vowel categories and their perception of these same categories. To do this, participants raised in the Northern, Western or Southern U.S. performed a vowel identification task along a continuum between /e/ and /ɛ/, and a subset of these listeners were then subsequently recorded reading a passage and word list containing a number of realizations of each vowel. In this paper, we acoustically analyze these participants' vowel productions and compare these productions to the same speakers' performance on the vowel identification task, asking what kinds of links exist between speakers' actual speech production and their perception of vowel categories. A main interest of the research is to explore the role that differential participation in regional vowel shifts currently affecting speech in the North and the South of the U.S. plays in both production and perception at the individual level.

## 2. Background on regional differences in perception

The majority of work on regional differences in perception has focused on the perception of social meaning rather than linguistic form. For example, it is clear from Dennis Preston's work on perceptual dialectology that regional differences are highly salient and that speakers show little hesitation assigning wide-ranging social traits such as differing levels of pleasantness and correctness to speakers they perceive as regionally divergent (Preston, 1989, 1993). Similarly, a number of other studies have also illustrated how social meaning and language forms are related (Baker et al., 2009; Baugh, 2000; Bayard et al., 2001; Campbell-Kibler, 2007, 2009, 2010; Fridland et al., 2004, 2005; Graff et al., 1986; Hay and Drager, 2010; Plichta and Preston, 2005; Purnell et al., 1999). Beyond such studies on social perceptions, there has also been some work investigating the perception of linguistic form, with most of these suggesting that social differences across listeners and/or speakers influence the way they perceive stimuli of varying kinds (e.g. Clopper and Pisoni, 2004, 2007; Di Paolo and Faber, 1990; Foulkes and Docherty, 2006; Graff et al., 1986; Hay et al., 2006; Jannedy and Hay, 2006; Janson, 1986; Koops et al., 2008; Labov and Ash, 1997; Niedzielski, 1999; Plichta and Preston, 2005; Purnell et al., 1999; Strand, 1999; Strand and Johnson, 1996; Willis, 1972; Wolfram et al., 1999). Foundational work by Johnson et al. (1999), Strand (1999) and Strand and Johnson (1996) showed that the perception of both fricative and vowel continua were affected by perceived speaker gender. Niedzielski (1999) found that changing the national identity of speakers as indicated on an answer sheet altered listeners' perception of the vowel category itself. Indeed, both listener and speaker attributes have been found to affect vowel perception, as suggested by the work of Hay et al. (2006), examining the perception of a near-merger. In addition, Clopper and Pisoni's (2007) work on dialect classification suggests dialect exposure (listener mobility) affected the perception of dialect variation.

Less literature has examined how individuals vary perceptually within a community or what link exists between individuals' production and perception. Experimental phoneticians have periodically examined these questions, but from a more general speech processing,<sup>2</sup> rather than sociolinguistic, perspective (Bell-Berti et al., 1979; Dahan et al., 2008; Frieda et al., 2000; Johnson and Mullennix, 1997; Johnson et al., 1993; Newman, 2003; Nygaard et al., 1994). Recently, a few studies, with interests in sociolinguistically related questions, such as Evans and Iverson (2004, 2007) and Sumner and Samuel (2009), have looked at how variable norms among individuals relate to their perception of those variables. In Sumner and Samuel's work, for example, three groups of subjects were tested on their perception of /r/-less variants for short-term and long-term priming effects on word recognition. The subject groups differed on whether they produced /r/-less variants themselves (New York /r/-less participants), regularly heard such variants but did not exhibit them productively (New York /r/-ful participants), or never used nor regularly heard /r/-less variants ("General American" /r/-ful participants). Their

<sup>1</sup> For sake of space, we have not included schematics of these shifts. Diagrams and more information about these shifts can be found in publications like Gordon (2002), Labov et al. (2006), and Thomas (2011:280–283). Within this paper, Figs. 1–3 roughly illustrate the main tendencies of each of these shifts.

<sup>2</sup> There is a large body of speech science literature which investigates speech perception, such as that examining categorical perception or cross-language perception which, for the sake of space, we do not mention here. The works cited are primarily those most relevant to the current paper, such as those addressing the link between individual perception and production and talker contingent adaptation.

results suggested that individuals' own production as well as local community production affected processing. Our research similarly poses the question of how both one's own use of a vowel variant and one's familiarity with a particular community-based variant affects how individuals' perceive category boundaries for that vowel. In the sections that follow we will examine the regional variation in production data from a subset of our subjects and then move to the results of the perception test for all subjects. In the final section, we consider how the variation in production in terms of shift participation within each region translates into perceptual variation as well.

### 3. Experiment method

#### 3.1. Perception test design

The vowel perception test was designed using a forced-choice vowel identification task (Hillenbrand et al., 1995; Strange, 1995; Thomas, 2002), which measured vowel category judgments for an /e/ to /ɛ/ continuum synthesized between natural speech endpoints, which were provided by a single middle-age male talker born in New England but raised in the West.

The continuum range for the /e/~/ɛ/ vowel pair was determined based on the sample speaker's production values for each selected endpoint categories. (The identification task also included other vowel continuum pairs which are not treated in this paper). These values were compared with those provided by earlier work (summarized in Kent and Read, 2002) on average vowel frequency ranges and contemporary regional vowel shifts to ensure they fell within the average frequency range for American speakers. Based on these endpoints, the stimuli were created through a vowel synthesis program designed by Bartek Plichta. More extensive software design details and examples, along with the program itself, are available online at <http://bartus.org>. Vowel duration was neutralized across each vowel pair by means a PSOLA-based computational method so that all tokens, regardless of original differences in vowel inherent spectral change, exhibited a duration intermediate the two original endpoints.<sup>3</sup> The vowel continuum was embedded into two different consonant contexts (preceding labial and preceding alveolar) resulting in two seven-step continua between *bait*~*bet* and between *date*~*debt*.

The perception test was administered over the Internet. Subjects participated in the experiment either in a university lab or were instructed to take the test in a quiet location with external noise sources minimized. Prior to the task itself, listeners responded to a series of demographic and hearing/listening questions and were given a practice run using a non-tested continuum. Participants were instructed to wear headphones, but, due to the online nature of the experiment, we were not able to enforce this requirement. Participants were asked in the preliminary questions about their listening environment, i.e. whether they were wearing headphones. Then, listeners heard a single vowel-continuum step one token at a time (played once) and were asked to select which of the endpoint tokens they believed they had heard from two relevant choices, e.g. *bait* or *bet*. Each step in each vowel continuum was played 4 times randomized over the course of the study.

#### 3.2. Perception study participants

Participants were recruited from universities in each of the three dialect areas studied. Northern participants were recruited from the State University of New York at Oswego in New York state, Southern participants were recruited from the University of Memphis in Tennessee and from Virginia Tech in Virginia. Finally, Western participants were recruited from the University of Nevada Reno in Nevada. Participants were asked where they were raised from age 4 until adulthood and this information was used to separate participants into regional groups. Only those participants who self-identified as from one of the relevant regions were included in the analyzed data. For the present study, we limit our investigation of the perception data to the 217 subjects from these three regions who self-identified as Anglo-American (i.e. "White"). Table 1 shows the breakdown by listener region for the perception study. The majority of subjects reported wearing headphones as was requested.

**Table 1**  
Participants from each region.

	South		West	North
	Tennessee (TN)	Virginia (VA)	Nevada (NV)	New York (NY)
Full perception sample	20	48	90	58
Perception–production subsample	13	–	10	9

#### 3.3. Speech sample data

In addition to participating in the vowel identification tasks, a subset of participants were also recorded reading a reading passage and word list so that we would have production data to correlate with the perceptual findings. As outlined in Table 1,

<sup>3</sup> The continuum between these two vowel category endpoints involves contrast between vowels with different spectral properties. However, since we are comparing participants' perceptual behavior in terms of how principled change in F1/F2 affects category change and this same monophthongal/diphthongal contrast is present for all U.S. English speakers tested, we believe this is appropriate for the present study. (Readers interested in the dynamics of the synthesized stimuli can email the authors or refer to the Appendix of Kendall and Fridland, 2012.)

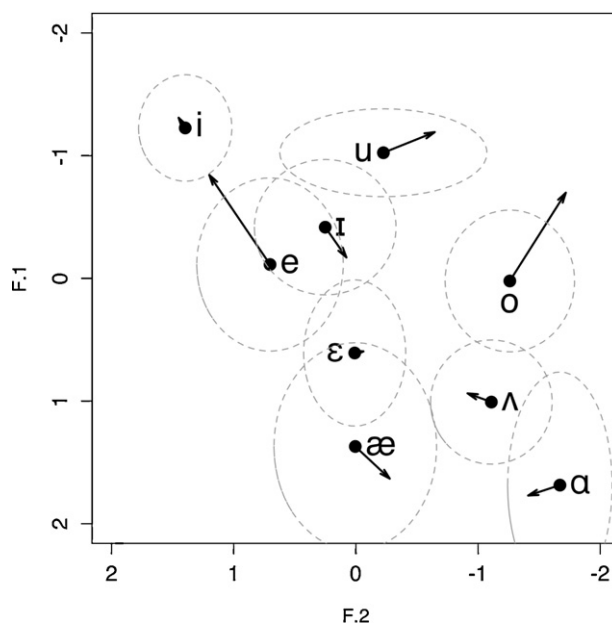


Fig. 1. Mean vowels for 13 Southern speakers (Lobanov normalized).

for the present study we focus on 13 participants from the Memphis, TN (South) field site, 10 participants from the Reno, NV (West) field site, and 9 participants from the Oswego, NY (North) field site who contributed speech data to the project.

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). All speakers read the same reading passage and word list with the same instructions (to read the passage over before recitation and to pause briefly between each word list recitation).

In order to obtain general vowel space information and detailed data on the mid front vowels for each speaker, thirty-nine vowels were measured from the word list and nine vowels were measured from the reading passage. All vowels were measured from monosyllabic words, with the exception of three disyllabic words in the reading passage, where we measured the stressed, first syllable. Praat (Boersma and Weenink, 2010) was used for all acoustic measurements. Vowel data, including  $f_0$ ,  $F_1$ ,  $F_2$ , and  $F_3$  for two points in time – at the 1/3 point and at the 2/3 point of the vowel's temporal duration – as well as each vowel's duration were measured using a Praat script designed for this study<sup>4</sup>. In the remaining discussion, all vowel data are manipulated and (as needed) normalized using the Vowels package (Kendall and Thomas, 2009; <http://cran.r-project.org/web/packages/vowels/>) for R (R Development Core Team, 2010).

## 4. Results

### 4.1. Production across regions

Figs. 1–3 display the production data averaged across speakers in each region, normalized via the Lobanov method (cf. Watt et al., 2011). The light gray ellipses indicate one standard deviation around the mean, computed from all tokens of each vowel class for all relevant (i.e. within-region) speakers. Each vowel class is depicted as a vector with the mean of the 1/3 temporal measurement point as the filled dot and the mean of the 2/3 temporal measurement point as the arrow head. Vowel classes are labeled using IPA notation.

Looking first at generalized community behavior productively, the tendencies in each region clearly show reflexes of the vowel shifts discussed previously. Fig. 1 illustrates the overall vowel productions for the Southerners. As is evident in the figure, the Southerners' mid front vowel classes are quite close together and, in fact, the standard deviations slightly overlap. This proximity between /e/ and /ɛ/ is especially apparent in contrast to the mean plots for Northerners (Fig. 2) and Westerners (Fig. 3). The Southerners' high front classes, acoustically quite separate, do not appear to be affected by SVS-related shift. Such a finding is not surprising considering that such shift is typically only found in a much smaller area of the South (Labov et al., 2006), one that excludes Western Tennessee where all of these Southerners were raised. Back vowel fronting is also quite advanced in this sample, with /u/ tokens overlapping some high front lax tokens.

<sup>4</sup> The script is available for public download at [http://ncslaap.lib.ncsu.edu/tools/scripts/vowel\\_capture\\_aug09.praat](http://ncslaap.lib.ncsu.edu/tools/scripts/vowel_capture_aug09.praat).

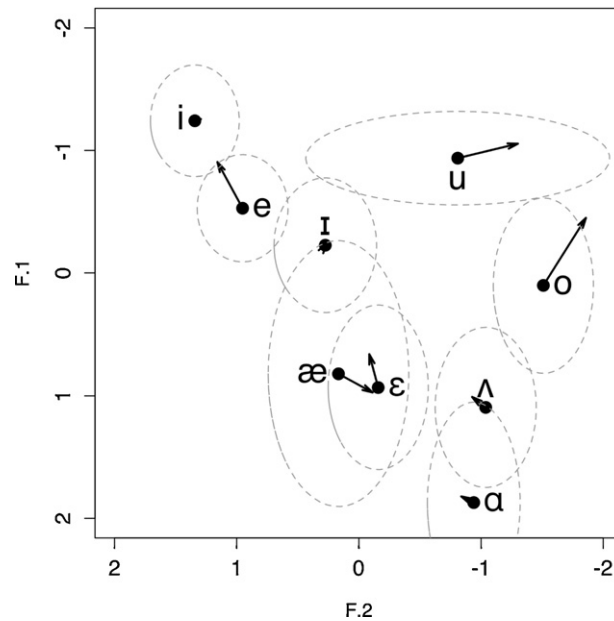


Fig. 2. Mean vowels for 9 Northern speakers (Lobanov normalized).

In Fig. 2, which shows the Northerners' vowels, a much wider gap between mid front classes is evident than in the Southern plot. The /e/ class remains quite peripheral while the short front classes /ɪ/ and /ɛ/ appear backed and lowered, suggesting NCS related shifts. In addition, /æ/ raising is evident, resulting in substantial overlap between the low front and mid front lax vowel classes. While back vowels do appear to be fronting, there is a large spread in the distribution of tokens, suggesting more tokens remain backed than in the Southern sample. Such results confirm those reported by Labov et al. (2006) suggesting that Southerners are more advanced in back vowel fronting than Northerners.

Fig. 3 illustrates the mean distribution of the Western sample speakers, who show a system somewhat intermediate that of the other regions. In general, Westerners maintain more separation between the tense and lax vowels than found for the Southern system, but do not seem to be as separated as Northerners. While the high and mid tense vowels are similar acoustically to those in the North, the front lax vowels clearly distinguish the systems, with less evidence of lowering in the high

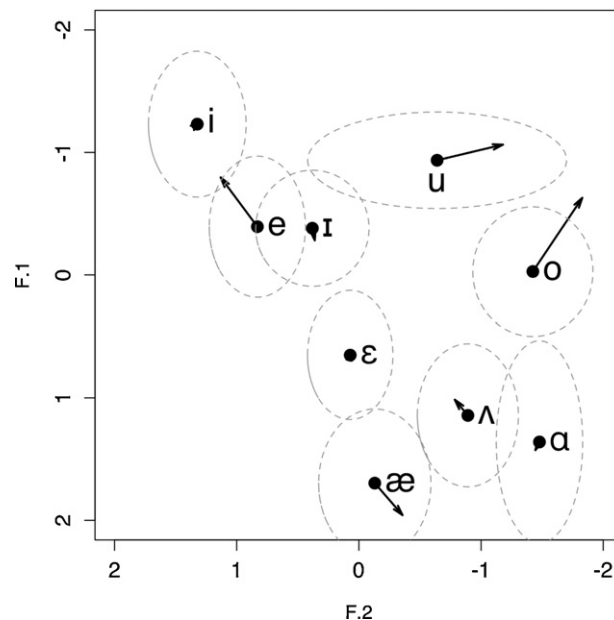


Fig. 3. Mean vowels for 10 Western speakers (Lobanov normalized).

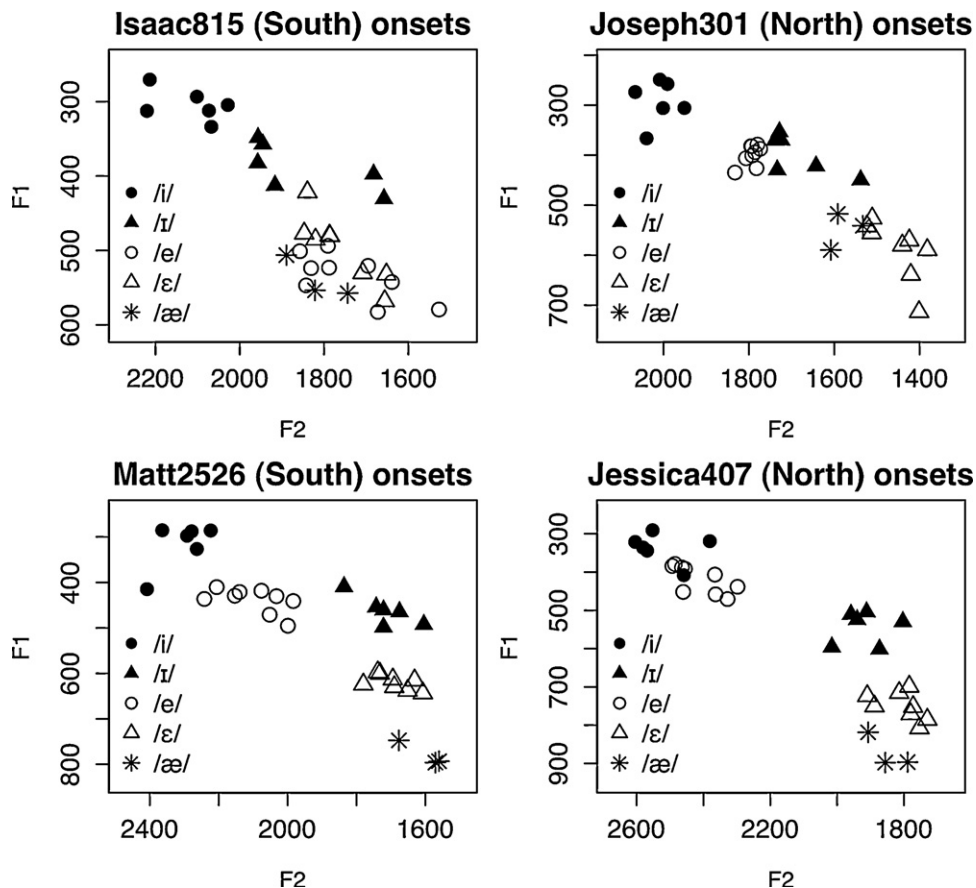


Fig. 4. Front vowel onsets for two Southerners and two Northerners.

and mid front classes and no low front raising. Back vowel fronting also appears to affect the region to a degree intermediate that in the South and North. Viewing these figures as indicating general regional tendencies, it is clear there are differences among these speakers in community-based production norms in line with their dominant regional shift patterns.

However, as implied by the ellipses in Figs. 1–3 and confirmed when we look at individual speakers' productions within each region, we also see there is quite a bit of variation among speakers from the same communities. To illustrate this briefly, Fig. 4 displays front vowel onsets for two of the Southerners and two of the Northerners. As can be seen from these plots, although there may be central tendencies toward community or regional norms when we aggregate across speakers, there is also quite a bit of variability across the speakers productively. Some southern speakers, like Isaac815, show a great deal of SVS related /e/-/ɛ/ reversal, while other Southerners, like Matt2526, clearly do not. Similarly, in the North, Jessica407 shows much more NCS backing and lowering of /ɛ/ compared to, for instance, Joseph301. At the same time, Joseph301 shows a much more raised /æ/ than Jessica407. The difference between sociolects and idiolects, of course, comes as no surprise to any researcher who has looked at community-based production data.<sup>5</sup> But, the central question here is, essentially, how does such variability affect speakers' perceptual systems? In the remaining sections we will explore how this variability in production globally and individually is related to variability in perception.

#### 4.2. Perception data results for all listeners

Figs. 5 and 6 show the mean results over the seven steps of the continua for the vowel identification task in the /b/ and /d/ contexts, respectively, for the 217 perception study participants (introduced in Section 3.2) from the three regions. Community-based production differences such as those involved in the SVS and the NCS do appear to create perceptual distinctions among the regions. From these figures, it is apparent that Southerners show different perceptual responses

<sup>5</sup> Closer examinations of the individuals' vowel spaces than we have room to fully explore in this paper indicate that there may be social correlates to some of the production differences across individuals. For instance, Fig. 7, later in the paper, indicates that females have larger /e/-/ɛ/ distances than males. However, we find no sex effects for any of the perception data in Sections 4.2 and 4.3. So, while there are some social explanations/correlations for these production differences, they do not end up playing a role in the perception data or the perception–production link, particularly since we are looking at such a link per individual. We thank an anonymous reviewer for pointing out the importance of our clarifying this aspect of the data.



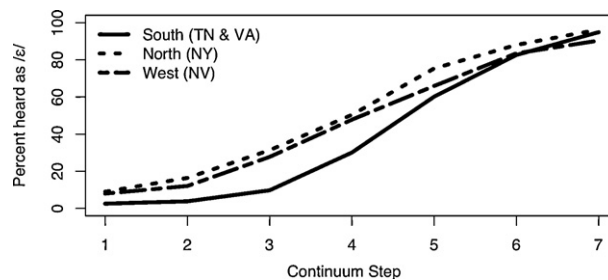


Fig. 5. Identification results for /e/-/ɛ/ in /b/ context (i.e. bait~bet).

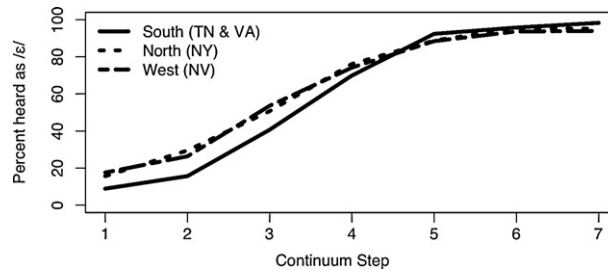


Fig. 6. Identification results for /e/-/ɛ/ in /d/ context (i.e. date~debt).

compared to the other groups for this vowel class. Southerners, whose productive means in Fig. 1 above indicate the presence of SVS related /e/ centralization, also show a significantly longer perception of /e/ along the continuum. In the /b/ context, the shift toward /ɛ/ perception – the cross over point at which each region's subjects hear 50% or more tokens as /ɛ/ – does not occur until between steps 4 and 5 for Southern listeners compared to between steps 3 and 4 for Northern and Western listeners. This difference in cross over points is significant through a one-way ANOVA ( $F(2, 214) = 5.21, p < 0.01$ ). Post hoc Tukey comparisons indicate that the cross over differences between Southerners and both other regions are significant (South vs. North, 95% CI  $[-1.51, -0.20], p < 0.01$ ; South vs. West, 95% CI  $[-1.27, -0.04], p < 0.05$ ). For the /d/ context, the perceptions are more similar across all three regions and an ANOVA indicates the differences are not significant ( $p = 0.3$ ), but the same general pattern is seen – the Southerners cross over between steps 3 and 4, while the Northerners and Westerners cross over between steps 2 and 3. Not surprisingly, the Northerners and Westerners, whose systems are similarly unaffected by Southern Vowel Shift tendencies, do not show the same prolonged /e/ perception as Southerners.

Mixed-effect regression models were built on the entire data for each continuum to test the likelihood of subjects hearing /ɛ/ rather than /e/. These models include random intercepts for subject and random slopes by subject for continuum step and fixed effects for the continuum step and each subject's regional background. The models also test, by including as a factor in the regression, whether the use of headphones plays a (systematic) role in the subjects' perceptions, and, for our purposes, this acts a control for the influence of that factor on the factors of interest. Subject sex was also tested in the models, but did not come out as significant for any of the perception data. Of direct interest here is the effect of region. The models show that the difference in perception between the Southerners and the Northerners and Westerners is significant (/b/: North vs. South, log-odds 2.98,  $p < 0.001$ , West vs. South, log-odds 3.58,  $p < 0.0001$ ; /d/: North vs. South, log-odds 1.72,  $p < 0.05$ , West vs. South, log-odds 1.82,  $p < 0.01$ ), but that the difference between the Northerners and Westerners is not (/b/:  $p = 0.4$ ; /d/:  $p = 0.9$ ).<sup>6</sup>

Considering the data presented so far, it appears that as a group the Southerners both hear more /ɛ/ and productively realize /e/ at more central F1/F2 positions. Given the effects of the Southern Vowel Shift in production, these results can be interpreted as indicating that the SVS alters perception in similar ways to production. Interestingly, we do not see a similar effect on the Northern listeners. They realize, as a group, similar perceptions to the Western listeners, despite the fact that their mid front vowel productions, via the NCS, move as well, although in the opposite direction as the Southerners. This lack of distinction from the Westerners may be due to the fact that for these vowels the NCS and CVS shifts are not greatly dissimilar, with both resulting in fairly similar movement of the lax system (backing/lowering of high and mid lax vowels in NCS vs. lowering in CVS) that result in greater distance between /e/ and /ɛ/ acoustically. SVS shifts, on the other hand, create much greater contrast to the North and West in production, with spectral overlap resulting from /e/ centralization while Northern and Western /ɛ/ shift processes (backing/lowering) create greater separation between the two vowel classes. It seems that the maintenance of a peripheral tense system and this tendency toward increased /e/-/ɛ/ distance stemming from Northern and Western shift processes results in similar perception. This lack of contrast may suggest that the mid front vowels are not very regionally distinctive in production or perception outside of the South. Such a finding conforms to expectations

<sup>6</sup> Full discussions of the statistical models for the overall perception data are provided in Kendall and Fridland (2012).

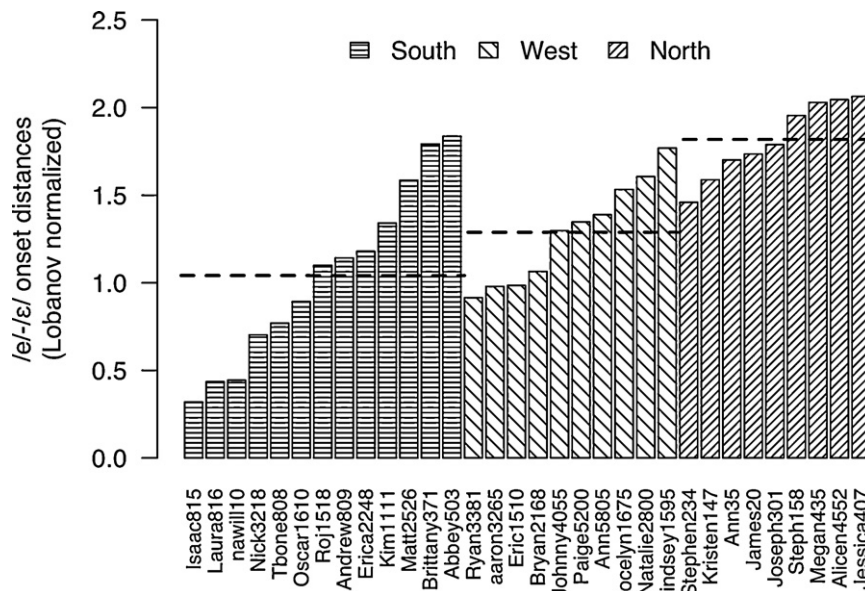


Fig. 7. Euclidean distances for /e/-/ε/ for subsample speakers, arranged by region.

about dialect saliency that suggest Southern speech differences are most regionally marked to non-linguists (Clopper and Pisoni, 2004, 2007; Niedzielski and Preston, 2003; Preston, 1989, 1993). However, Northern and Western speech contrasts are not absent as they are still maintained through other differences that are promoting greater distinctiveness (such as /æ/ raising in the North and the low back vowel merger in the West) among speakers in those regions.

While these results point to differences and similarities in general community norms perceptually, we also want to consider individual variation and the relationship between production and perception at the individual level, which we move to in the next section.

#### 4.3. Individual production/perception comparison

A primary interest of this paper is to examine how engagement with locally based vowel variation is reflected in the relationship between production and perception at the level of the individual, not just through community norms. In order to do this we use a Euclidean distance measure, which calculates the relative distance between each individual's /e/ and /ε/ tokens in production in F1 and F2 space (Kendall and Fridland, 2010).<sup>7</sup> Using the relevant vowel shifts as a basis, Southern speakers with closer /e/-/ε/ distance measures can then be indexed as greater 'shifTERS' in the SVS mid front shift compared to Southerners showing more distant /e/ and /ε/ classes. Similarly, a hallmark of the Northern Cities Shift is the backing and lowering of the mid front lax class /ε/, while /e/ remains at the front periphery (as evidenced in our sample (Fig. 2)). Thus, a greater Euclidean distance between /e/ and /ε/ for Northerners could be considered as greater participation in this shift compared to Northern speakers showing closer mid front classes. Such a shift measure in the North (separation) dovetails nicely with the contrasting measure that results from shift participation in the South (closeness). So, instead of just ranking speakers by simple regional orientation, this Euclidean distance measure can, in essence, provide a measure of degree of shift participation that can be used to rank speakers beyond simply binning them based on their region of origin or other impressionistic (or even acoustic) shift measures.

Fig. 7 provides a summary of the Euclidean distances for the speakers in the production subsample according to region. As can be seen, the overall mean (dashed line) for each group illustrates the overall effect of the various shift processes, with Southerners as a group showing the least distance between these classes and Northerners as a group, not surprisingly, showing the most. Regional affiliation clearly separates these groups and the regression results for all 217 listeners (in Section 4.2) certainly suggest that region is a significant factor in predicting perception behavior.

Still, as is seen in Fig. 8, which shows the same Euclidean distance measures by rank-order rather than by region, the most distinct relative positioning of these vowel classes comes from the contrast between those most actively participating in the Southern shift and in Northern shift, rather than simple regional orientation. Several Southerners (those whom we would classify both instrumentally and impressionistically as the most Southern shifted) cluster on the left-hand side of the plot, with the most proximate /e/ and /ε/ vowels, and several Northerners (those most Northern Cities shifted) cluster on the right,

<sup>7</sup> Euclidean distance is calculated via the standard formula:  $\sqrt{(F2_e - F2_\epsilon)^2 + (F1_e - F1_\epsilon)^2}$ , where  $F1_v$  and  $F2_v$  are the mean Lobanov normalized first and second formants for vowel  $v$  for each speaker. In addition to making formant values more comparable across speakers, Lobanov normalization has the added benefit of converting raw F1 and F2 to roughly equivalent scales, a requirement for the (meaningful) use of Euclidean distance.



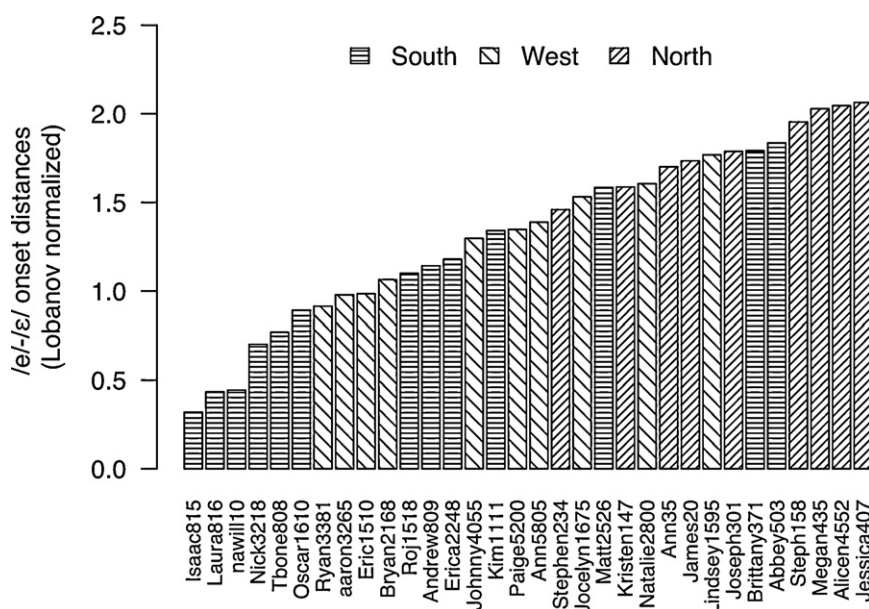


Fig. 8. Euclidean distances for /e/-/ε/ for subsample speakers, rank ordered by Euclidean distance.

with the least proximate /e/ and /ε/ vowels. Using this Euclidean distance measure, it is clear that, within each region, there is quite a bit of variability in degree of shift evidenced by different speakers and, echoing the individual plots shown earlier (in Fig. 4), only some of the speakers are fully engaged in the production shifts.

How specifically does this individual variability in production relate to the way that the speakers perceive the vowel continuum? Ordered by Euclidean distance along the horizontal axis, Figs. 9 and 10 display each participant's perceptual response (% ε) for each step of the *bait~bet* and *date~debt* continuum, respectively. The final panel, at the bottom right, shows the mean across all steps for these subjects. All plots show, in a gray dashed line, the actual response for each listener and, in a solid black line, a smoothed lowess line that captures the main trend in the data (Cleveland, 1981). When comparing perceptual decision behavior at each step for each listener–talker, we find that the speakers with the smallest distance between /e/ and /ε/ (Southern shifters) and the speakers with the greatest distance (Northern shifters) have, surprisingly, somewhat similar perceptions of the /e/-/ε/ continua, particularly in the early to middle steps of the continuum. This appears to occur despite the fact that the shifts affecting the individuals on opposite ends of the plots are making production of these two vowel classes comparatively *more* distinct (i.e. *less* similar across regions). This relationship is visible in the figures by attending to the perceptual behavior of the speakers at each end of the rank ordered plots, as they display the least (leftmost) to greatest (rightmost) difference between /e/ and /ε/ productively, and is especially apparent through the smoothed lowess lines.

To determine if these relationships are significant, the distance between each speaker's /e/ and /ε/ classes was included in mixed-effect statistical models of the perception data. The Euclidean distance between /e/ and /ε/ as a single, linear term did not arise as significant in exploratory models. In other words, the existence of a linear relation between Euclidean distance and perception of the /e/-/ε/ continuum (i.e. increases in Euclidean distance altering perception in a simple linear fashion) was not supported by statistical analysis. However, based on the curvilinear pattern which arose in Figs. 9 and 10 suggesting those with the least and greatest Euclidean distance shared similar perception, Euclidean distance was tested as a polynomial predictor with a squared term and this was found to be significant, though regional origin remained significant in the model, and as the stronger of the two predictors (likelihood ratio tests indicate that a model with only region is better than a model with only Euclidean distance). Table 2 and Fig. 11 display fixed-effects from the mixed-effect logistic regression model for the /b/ context continuum. Table 3 and Fig. 12 display the effects from the model for the /d/ context. The two models are almost identical, except that the /b/ context model has a significant interaction between step and region, while the /d/ context does not. As can be seen especially by comparing the figures (Figs. 11 and 12), the /b/ context data exhibits a larger range of (accounted for) variation. Both models fit the data quite well, as indicated by the Somers' Dxy and C statistics provided under the tables.<sup>8</sup> In essence, the models confirm what seemed apparent in the individual step-by-step plots in Figs. 9 and 10 – that the participants with the greatest distance between the mid front vowels and those with the least show similar perception along much of the continuum, and are significantly different from those in their communities with less shift.

<sup>8</sup> These Somers' Dxy and C statistics are computed following advice by Baayen (2008:281–282) and values so close to 1 indicate that the models' fits are quite good.

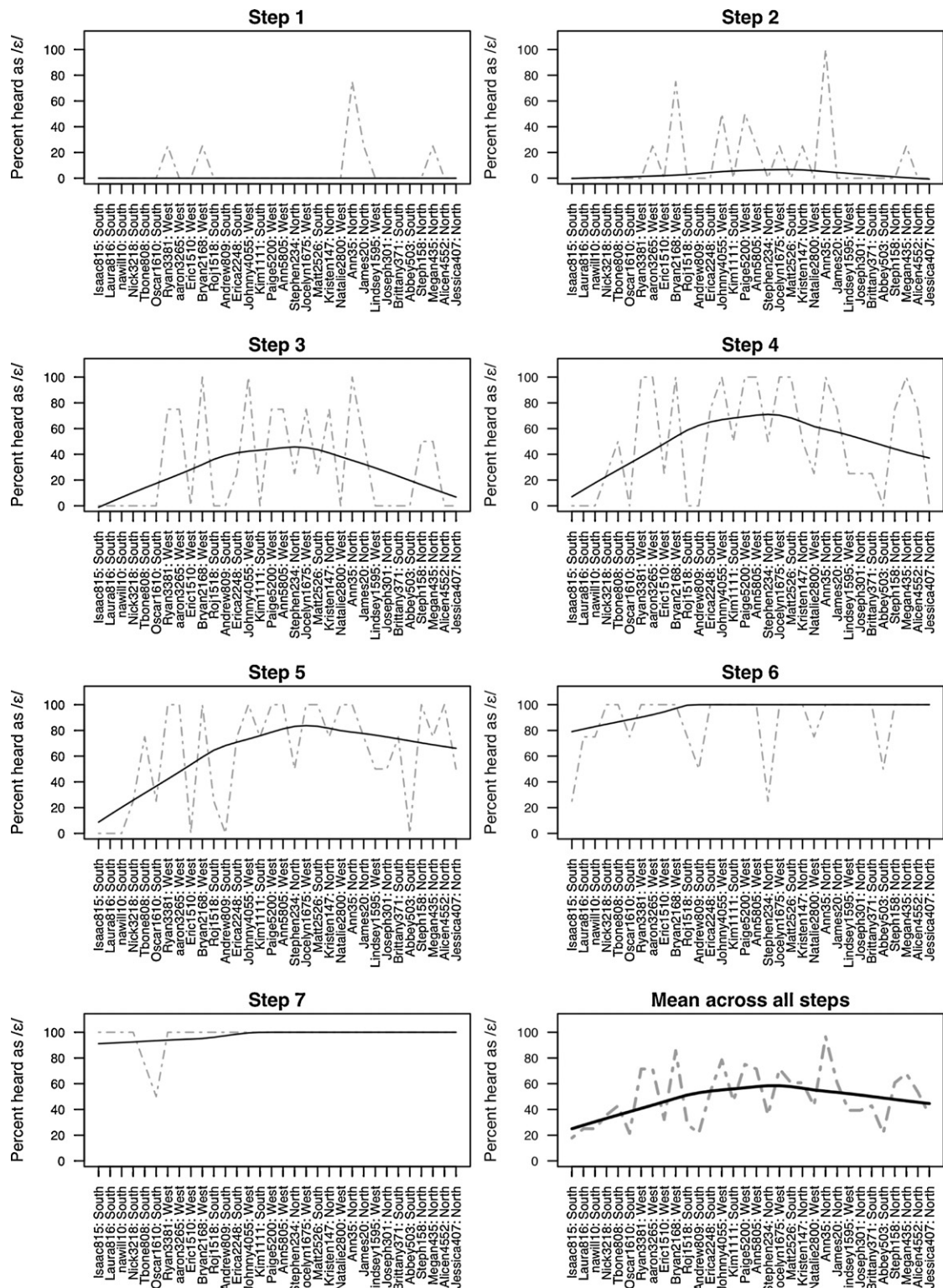
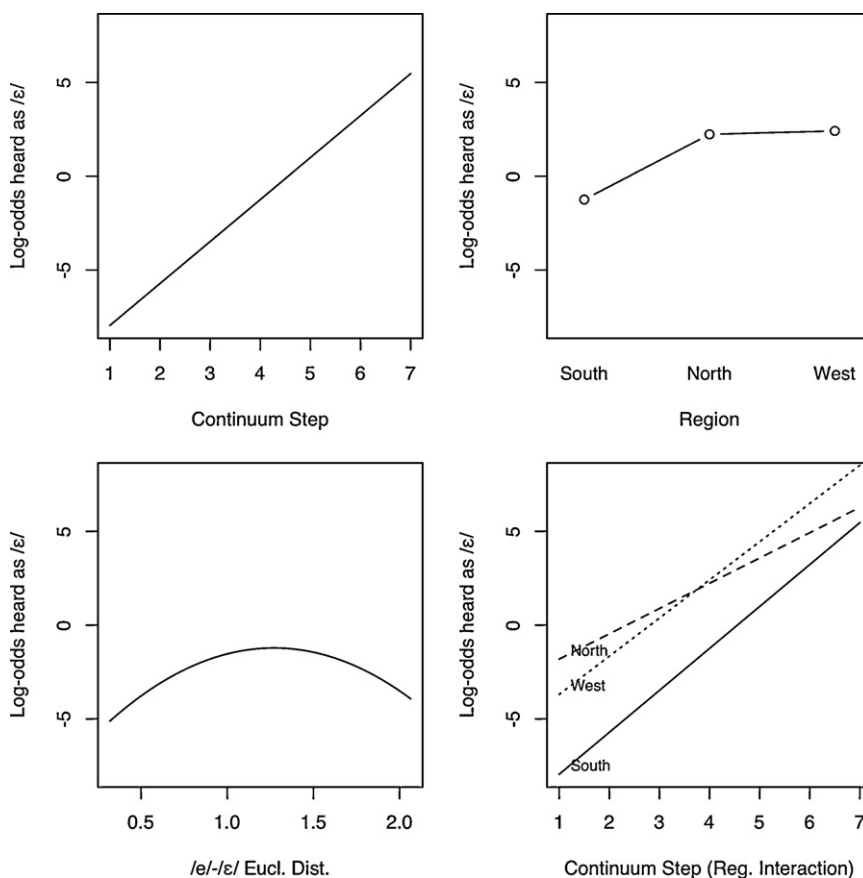


Fig. 9. Perception for /b/ context /e/-/ε/ continuum for each step, and mean perception, for each listener ranked by Euclidean distance.



**Table 2**Mixed-effect model results for /b/ context (*bait~bet*) for subsample listeners.

Factor	Log-odds Est.	Std. Err.	Z value	p
(Intercept)	−17.118	2.995	−5.716	<0.00001
Continuum Step	2.236	0.260	8.588	<0.00001
South vs. North	7.028	1.787	3.933	<0.0001
South vs. West	4.459	1.685	2.646	<0.01
eEdist	10.947	4.452	2.459	=0.01
eEdist <sup>2</sup>	−4.307	1.794	−2.400	<0.05
Step × South vs. North	−0.888	0.311	−2.855	<0.01
Step × South vs. West	−0.201	0.363	−0.554	[=0.58]

Somers'  $D_{xy}$  = 0.938,  $C$  = 0.969.**Fig. 11.** Model result plots for /b/ context (*bait~bet*) for subsample listeners.**Table 3**Mixed-effect model results for /d/ context (*date~debt*) for subsample listeners.

Factor	Log-odds Est.	Std. Err.	Z value	p
(Intercept)	−10.228	1.966	−5.203	<0.00001
Continuum Step	1.718	0.168	10.221	<0.00001
South vs. North	1.504	0.864	1.741	[=0.08]
South vs. West	2.348	0.669	3.513	<0.001
eEdist	7.287	3.251	2.242	<0.05
eEdist <sup>2</sup>	−2.656	1.297	−2.407	<0.05

Somers'  $D_{xy}$  = 0.914,  $C$  = 0.957.

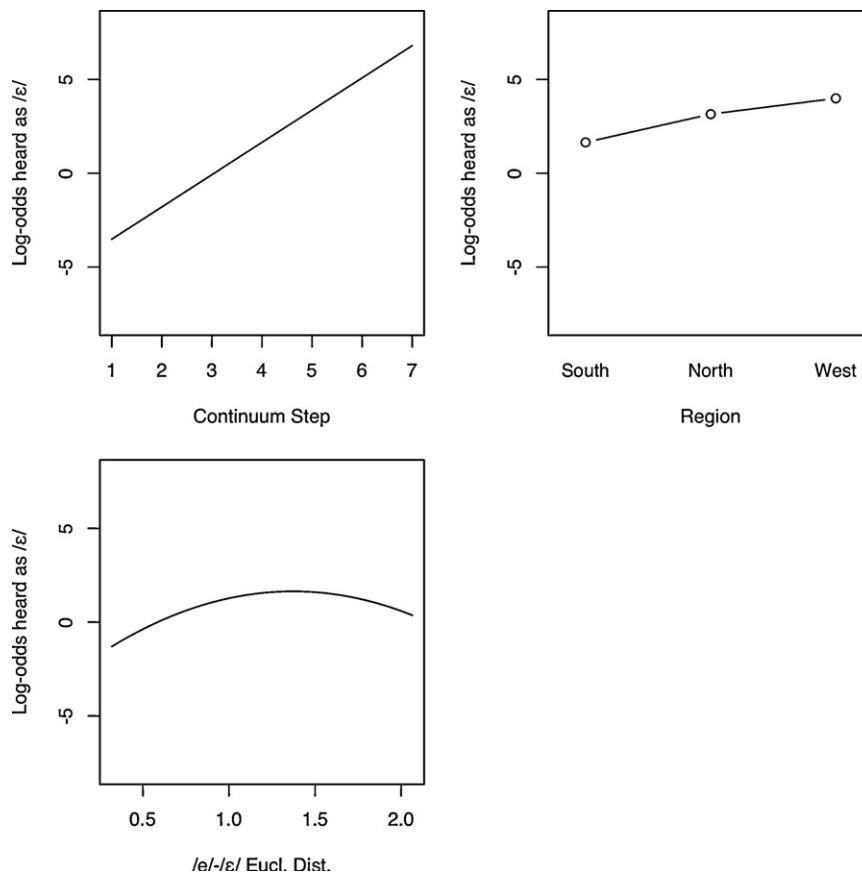


Fig. 12. Model result plots for /d/ context (*date~debt*) for subsample listeners.

To understand what, at first glance, seems a strange relationship, it is helpful to look back at the production data for these groups. In SVS affected Southerners, as seen in Isaac815's system in Fig. 4, /e/ is centralized, moving down and back in spectral space toward /ε/. Thus, /e/ and /ε/ become more similar spectrally. On the other hand, looking at, for instance, Jessica407 in Fig. 4, it can be seen that in the NCS /ε/ moves down and back, away from /e/, with the result that /e/ and /ε/ become less similar in spectral space. Yet, both kinds of speakers are seen to favor backed perceptions of /e/ (i.e. longer perceptions of /e/ as the stimuli moves back and down toward /ε/).

How do these productive patterns help illuminate the perception data results? An explanation, we believe, can be found in exemplar-based models of grammar (e.g. Goldinger, 1996, 1997, 1998; Johnson, 1997, 2007; Hay et al., 2006; Pierrehumbert, 2001, 2006). The individual variability we find productively within each region (and, more generally, community) suggests that, in each speech community, there is a range of exemplars available to each listener. As found in recent work by Sumner and Samuel (2009) and Evans and Iverson (2004, 2007), however, not all exemplars are ranked equally and individuals' own productions influence how their representations are stored. When a community is engaged in shift, we suggest that 'best' exemplars or targets move in the direction of shift, particularly for those actively engaged in the change, though other norms may be recognized as potential realizations. Yet shifting production targets do not necessarily have a one-to-one relationship with shifting perceptual targets. So, for SVS shifters, an expectation for central /e/ targets based on SVS-related variation creates centralized (i.e. backed) /e/ perception and, for NCS shifters, an expectation for backed and lowered /ε/ targets create backed /ε/ perception, and thus more room for /e/ perception. The overall effect is that both shifts result in the expectation of a backed /ε/ realization relative to non-shifters in their communities, suggesting a larger /e/ perceptual range for both groups of shifters.

## 5. Conclusions

In summary, the investigation of production and perception across regional dialects produced a number of different findings. First, regional differences in production were found in line with previously reported results that clearly indicate distinct regional vowel shift patterns (citations and discussion in Section 1). Southerners showed evidence of SVS related shifts in the mid front vowels while Northerners showed evidence of NCS related shifts, particularly in the front lax subsystem. The West showed a pattern intermediate the South and North. However, beyond these general shift patterns,

each region showed variability at the level of the individual in degree of shift participation. Put simply, some speakers are greater participants in the shifts than others.

Second, region was a significant predictor of perceptual behavior. Region of origin made a difference in how participants heard the vowel continuum, with Southerners maintaining a longer /e/ perception along the continuum than non-Southerners. Northerners and Westerners were not significantly different from each other. Such results are in line with the general sense by scholars and “folk linguists” (Niedzielski and Preston, 2003; Preston, 1989), who note language in the U.S. South to be highly salient and distinct from other regions.

Third, the Euclidean distance between each speaker’s mid front vowel classes, /e/ and /ɛ/, was also a significant predictor of perceptual behavior, suggesting that individual production, and not just generalized community norms, relates to how the continuum was perceived by listeners. These results suggested, however, a complex relationship between shift participation and perception of the vowel continuum in that the greater the degree of shift participation in *either* the Northern Cities or Southern Shift the more similar the perception of the /e/~/ɛ/ boundary, with shift participation leading to generally increased /e/ perception farther down the continuum.

To extrapolate from our findings, perception appears to depend both on what you yourself produce (i.e. as an individual speaker) *and* who you are more generally (i.e. as a member of a specific community). Such results appear consistent with the findings of Sumner and Samuel (2009) who found, in their study of ‘covert’ and ‘overt’ /r/ users in New York, that recognition in primed lexical decision tasks was aided both by a listener’s own system (whether they used /r/-less speech or not) and that of their community (whether they lived in an /r/-less community or not). Like our study, their results suggested that processing is affected by both what you say and what others around you say, which they found reflected differences in how forms were encoded. While we would suggest that our results are most compatible with an exemplar-based model of speech processing (as vowel stimuli appear to be variably processed among speakers with similar dialect exposure), further work is needed to illuminate what such differences suggest about how representations are encoded for speakers. Certainly, what is clear from our results is that variation is not limited to production but is also a perceptual phenomenon, something often overlooked in traditional treatments of sociolinguistic variation.

## Acknowledgements

This research has been supported by a grant from the National Science Foundation Linguistics Program BCS#0518264. This paper builds on the results presented in a short preliminary paper in the *Penn Working Papers in Linguistics* (Kendall and Fridland, 2010) and in Kendall and Fridland (2012), a paper that looks more closely at our data on the Southern Vowel Shift. A few sections of this paper, such as the treatment of the experimental and instrumental analysis methods, are discussed in those other papers. Finally, we would like to thank Teresa Dalle and Sohei Okamoto for their help collecting and managing the data.

## References

- Baayen, H., 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press, Cambridge.
- Baker, W., Eddington, D., Nay, L., 2009. Dialect identification: the effects of region of origin and amount of experience. *American Speech* 84, 48–71.
- Baugh, J., 2000. Racial identification by speech. *American Speech* 362–364.
- Bayard, D., Weatherall, A., Gallois, C., Pittam, J., 2001. Pax Americana?: Accent attitudinal evaluations in New Zealand, Australia, and America. *Journal of Sociolinguistics* 5, 22–49.
- Bell-Berti, F., Raphael, L., Pisoni, D., Sawusch, J., 1979. Some relationships between speech production and perception. *Phonetica* 36, 373–383.
- Boersma, P., Weenink, D., 2010. Praat: Doing Phonetics by Computer (Version 5.1.43) (Computer program).
- Campbell-Kibler, K., 2007. Accent, (ING) and the social logic of listener perceptions. *American Speech* 82, 32–64.
- Campbell-Kibler, K., 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21, 135–156.
- Campbell-Kibler, K., 2010. Sociolinguistics and perception. *Language and Linguistics Compass* 4, 377–389.
- Clarke, S., Elms, F., Amani, Y., 1995. The third dialect of English: some Canadian evidence. *Language Variation and Change* 7, 209–228.
- Cleveland, W.S., 1981. LOWESS: a program for smoothing scatterplots by robust locally weighted regression. *The American Statistician* 35, 54.
- Clopper, C., Pisoni, D., 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32, 111–140.
- Clopper, C., Pisoni, D., 2007. Free classification of regional dialects of American English. *Journal of Phonetics* 35, 421–438.
- Dahan, D., Drucker, S., Scarborough, R., 2008. Talker adaptation in speech perception: adjusting the signal or the representations? *Cognition* 108, 710–718.
- Di Paolo, M., Faber, A., 1990. Phonation differences and the phonetic content of the tense-lax contrast in Utah English. *Language Variation and Change* 2, 155–204.
- Eckert, P., 1988. Adolescent social structure and the spread of linguistic change. *Language Variation and Change* 1, 208–245.
- Eckert, P., 2000. *Linguistic Variation as Social Practice*. Blackwell, Oxford.
- Evans, B., Iverson, P., 2004. Vowel normalization for accent: an investigation of best exemplar locations in northern and southern British English sentences. *Journal of the Acoustical Society of America* 115, 352–361.
- Evans, B., Iverson, P., 2007. Plasticity in vowel perception and production: a study of accent change in young adults. *Journal of the Acoustical Society of America* 121, 3814–3826.
- Feagin, C., 1986. More evidence for vowel change in the South. In: Sankoff, D. (Ed.), *Diversity and Diachrony*. John Benjamins, Amsterdam & Philadelphia, pp. 83–95.
- Foulkes, P., Docherty, G., 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34, 409–438.
- Fridland, V., 1998. *The Southern Vowel Shift: Linguistic and Social Factors*. Dissertation, Michigan State University.
- Fridland, V., 2000. The Southern Vowel Shift in Memphis, TN. *Language Variation and Change* 11, 267–285.
- Fridland, V., 2001. Social factors in the Southern Shift: gender, age and class. *Journal of Sociolinguistics* 5, 233–253.
- Fridland, V., Kathryn, B., Kreuz, R., 2004. Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN. *Language Variation and Change* 16, 1–16.
- Fridland, V., Kathryn, B., Kreuz, R., 2005. Making sense of variation: pleasantness and education ratings of regional vowel variants. *American Speech* 80, 366–387.



- Frieda, E., Walley, A., Flege, J., Sloane, M., 2000. Adults' perception and production of the English vowel /i/. *Journal of Speech, Language, and Hearing Research* 43, 129–143.
- Goldinger, S., 1996. Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22, 1166–1183.
- Goldinger, S., 1997. Words and voices: perception and production in an episodic lexicon. In: Johnson, K., Mullennix, J. (Eds.), *Talker Variability in Speech Processing*. Academic Press, San Diego, pp. 33–66.
- Goldinger, S., 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251–279.
- Gordon, M., 1997. *Urban Sound Change Beyond City Limits: The Spread of the Northern Cities Shift in Michigan*. Dissertation, University of Michigan.
- Gordon, M., 2002. Investigating chain shifts and mergers. In: Chambers, J.K., Trudgill, P., Schilling-Estes, N. (Eds.), *The Handbook of Language Variation and Change*. Blackwell Publishing, Oxford, pp. 244–266.
- Graff, D., William, L., Harris, W., 1986. Testing listeners' reactions to phonological markers of ethnic identity. In: David, S. (Ed.), *Diversity and Diachrony*. John Benjamins, Amsterdam & Philadelphia, pp. 45–53.
- Hay, J., Drager, K., 2010. Stuffed toys and speech perception. *Linguistics* 48, 865–892.
- Hay, J., Warren, P., Drager, K., 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34, 458–484.
- Hillenbrand, J., Laura, G., Michael, C., Wheeler, K., 1995. Acoustic characteristics of English vowels. *Journal of the Acoustical Society of America* 97, 3099–3111.
- Jannedy, S., Hay, J., 2006. Modeling sociophonetic variation. *Journal of Phonetics* 34, 405–408.
- Janson, T., 1986. Sound change in perception: an experiment. In: John, O., Jeri, J. (Eds.), *Experimental Phonology*. Academic, Orlando, FL, pp. 253–260.
- Johnson, K., 1997. Speech perception without speaker normalization: an exemplar model. In: Johnson, K., Mullennix, J. (Eds.), *Talker Variability in Speech Processing*. Academic Press, San Diego, pp. 145–166.
- Johnson, K., 2007. Decisions and Mechanisms in Exemplar-based Phonology. In: Sole, M.J., Beddor, P., Ohala, M. (Eds.), *Experimental Approaches to Phonology*. In Honor of John Ohala. Oxford University Press, pp. 25–40.
- Johnson, K., Mullennix, J. (Eds.), 1997. *Talker Variability in Speech Processing*. Academic Press, San Diego.
- Johnson, K., Flemming, E., Wright, R., 1993. The hyperspace effect: phonetic targets are hyperarticulated. *Language* 69, 505–528.
- Johnson, K., Strand, E., D'Imperio, M., 1999. Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics* 27, 359–384.
- Johnstone, B., 1996. *The Linguistic Individual: Self-Expression in Language and Linguistics*. Oxford University Press, New York/Oxford.
- Kendall, T., Fridland, V., 2010. Mapping production and perception in regional vowel shifts. *Penn Working Papers in Linguistics* 16.2.
- Kendall, T., Fridland, V., 2012. Variation in perception and production of mid front vowel shift in the US Southern Vowel Shift. *Journal of Phonetics*, doi:10.1016/j.wocn.2011.12.002.
- Kendall, T., Thomas, E., 2009. Vowels: Vowel Manipulation, Normalization, and Plotting in R. R Package, Version 1.0-3. , URL:<http://cran.r-project.org/web/packages/vowels/>.
- Kent, R., Read, C., 2002. *The Acoustic Analysis of Speech*. Thomson Learning, Canada.
- Koops, C., Gentry, E., Pantos, A., 2008. The effect of perceived speaker age on the perception of PIN and PEN in Houston, Texas. *Penn Working Papers in Linguistics* 14.2, 93–101.
- Labov, W., 1991. The three dialects of English. In: Penelope, E. (Ed.), *New Ways of Analyzing Variation*. Academic Press, New York, pp. 1–44.
- Labov, W., 1994. *Principles of Linguistic Change: Internal Factors*. Blackwell, Massachusetts.
- Labov, W., 2001. *Principles of Linguistic Change: Social Factors*. Blackwell, Massachusetts.
- Labov, W., Ash, S., 1997. *Understanding Birmingham*. In: Bernstein, C., Nunnally, T., Sabino, R. (Eds.), *Language Variety in the South Revisited*. University of Alabama Press, Tuscaloosa, Alabama, pp. 508–573.
- Labov, W., Malcah, Y., Richard, S., 1972. *A Quantitative Study of Sound Change in Progress*. U.S. Regional Survey, Philadelphia.
- Labov, W., Ash, S., Boberg, C., 2006. *The Atlas of North American English: Phonetics, Phonology and Sound Change*. De Gruyter, Berlin.
- Luthin, H., 1987. The story of California (ow): the coming-of-age of English in California. In: Denning, K.M., Inkels, S., McNair-Knox, F., Rickford, J. (Eds.), *Variation in Language: NWAV-XV at Stanford (Proceedings of the Fifteenth Annual New Ways of Analyzing Variation Conference)*. Department of Linguistics, Stanford University, Stanford, pp. 312–324.
- Macaulay, R., 1991. *Locating Dialect in Discourse: The Language of Honest Men and Bonnie Lassies in Ayr*. Oxford University Press, New York/Oxford.
- Newman, R., 2003. Using links between speech perception and speech production to evaluate different acoustic metrics: a preliminary report. *Journal of the Acoustical Society of America* 113, 2850–2860.
- Niedzielski, N., 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18, 62–85.
- Niedzielski, N., Preston, D., 2003. *Folk Linguistics*. Mouton De Gruyter, Berlin.
- Nygaard, L., Sommers, M., Pisoni, D., 1994. Speech perception as a talker-contingent process. *Psychological Science* 5, 42–46.
- Pierrehumbert, J., 2001. Exemplar dynamics: word frequency, lenition and contrast. In: Bybee, J., Hopper, P. (Eds.), *Frequency and the Emergence of Linguistic Structure*. John Benjamins, Amsterdam, pp. 137–157.
- Pierrehumbert, Janet B., 2006. The next toolkit. *Journal of Phonetics* 34, 516–530.
- Plichta, B., Preston D., 2005. The /ay/s Have It: The perception of /ay/as a North-South stereotype in US English. In: Kristiansen, T., Coupland, N., Garrett, P. (Eds.), *Acta Linguistica Hafniensia* 37, special issue on subjective processes in language variation and change. 243–285.
- Preston, D., 1989. *Perceptual Dialectology*. Foris, Dordrecht.
- Preston, D., 1993. *Folk dialectology*. In: Preston, D. (Ed.), *American Dialect Research*. Prentice Hall, Englewood Cliffs, NJ, pp. 333–378.
- Purnell, T., Idsardi, W., Baugh, J., 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18, 10–30.
- R Development Core Team, 2010. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. , URL: <http://www.R-project.org/>.
- Strand, E., 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18, 86–99.
- Strand, E., Johnson, K., 1996. Gradient and visual speaker normalization in the perception of fricatives. In: Gibbon, D. (Ed.), *Natural Language Processing and Speech Technology*. Mouton, Berlin, pp. 14–26.
- Strange, W., 1995. Cross-language studies of speech perception: a historical review. In: Strange, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. York Press, Baltimore, pp. 3–48.
- Sumner, M., Samuel, A., 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60, 487–501.
- Thomas, E., 2001. *An Acoustic Analysis of Vowel Variation in New World English*. Publication of the American Dialect Society 85. Duke University, Durham, NC.
- Thomas, E., 2002. Sociophonetic Applications of Speech Perception Experiments. *American Speech* 77, 115–147.
- Thomas, E., 2011. *Sociophonetics: An Introduction*. Palgrave Macmillan, New York/Basingstoke, Hampshire.
- Watt, D., Fabricius, A., Kendall, T., 2011. More on vowels: plotting and normalization. In: Di Paolo, M., Yaeger-Dror, M. (Eds.), *Sociophonetics: A Student's Guide*. Routledge, London/New York, pp. 107–118.
- Weinreich, U., William, L., Marvin, H., 1968. Empirical foundations for a theory of language change. In: Lehmann, W., Malkiel, Y. (Eds.), *Directions for Historical Linguistics*. University of Texas Press, Austin, pp. 95–195.
- Willis, C., 1972. Perception of vowel phonemes in Fort Erie, Ontario, Canada, and Buffalo, New York: an application of synthetic vowel categorization tests to dialectology. *Journal of Speech and Hearing Research* 15, 246–255.
- Wolfram, W., Kirk, H., Natalie, S.-E., 1999. *Dialect Change and Maintenance on the Outer Banks*. Publication of the American Dialect Society 81. Duke University Press, Durham, NC.