Journal of Phonetics 40 (2012) 289–306

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Variation in perception and production of mid front vowels in the U.S. Southern Vowel Shift

Tyler Kendall [a,*], Valerie Fridland [b]

[a] Department of Linguistics, 1290 University of Oregon, Eugene, OR 97403, United States
[b] Department of English/MS0098, University of Nevada, Reno, NV 89557, United States

ARTICLE INFO

Article history:
Received 15 March 2010
Received in revised form
9 December 2011
Accepted 19 December 2011
Available online 25 January 2012

ABSTRACT

Looking at speech perception from a sociolinguistic perspective, the paper first explores how speakers from three different regions in the U.S. perform on a vowel identification task for a continuum between /e/ and /ɛ/. Following the general analysis of cross-regional perception, we turn our focus to a subsample of Southern participants who also provided speech data, investigating the nature of the link between their speech production and perception for these vowels. In particular, we are interested in the extent to which participation in a series of shifts affecting the Southern speech region in production (the Southern Vowel Shift or SVS) affects perception in that region. The data includes a set of seven siblings and we also examine whether sibling status affects perceptual variability. Our results suggest that region does play a significant role in mediating perception, particularly in the South, and that SVS participation in production is related to differences in perception within that region, suggesting that both individual and community based norms are crucial in speech processing. Finally, identifying a large amount of familial variability in both perception and production, we find that siblinghood does not seem to play a greater role in speech perception similarity than shift participation.

1. Introduction

A great deal of work in sociolinguistics has focused on differences in production norms for speakers from different social and ideological backgrounds. That a relationship exists between external social factors and linguistic output is undeniable (see Labov (1994, 2001) for fuller discussion), with gender, ethnicity, social class and geography established as playing foundational roles in the formation of individuals' linguistic repertoires (Clarke, Elms, & Amani, 1995; Clopper, Pisoni, & De Jong, 2005; Eckert, 2000; Labov, 1991, 1994, 2001; Labov, Ash, & Boberg, 2006; Labov, Yeager, & Steiner, 1972; Milroy, 1980; Wolfram, 1991). Such socially-based production differences are found across all linguistic levels, though much research has centered on variation in syntax/morphology and phonology. Morphosyntactic variation across socially-based dialects, such as multiple negation or copula absence, is often categorical (present or absent) and, as it can be easily recognized by listeners, often stigmatized (Fasold, 1972; Rickford, 1999; Wolfram, 1969). Thus, listeners can be fairly easily tapped for their knowledge of and familiarity with such structures and their perception of syntactic variation is generally influenced by overt norms and commentary on the use of these features.

Phonological variation, on the other hand, has been embraced by sociolinguists (particularly sociophoneticians) as an object of study in part because of its more covert nature. While certainly a number of socially-based phonetic differences (e.g., *tied*: /tayd/ vs. /taːd/) are clearly noticed by listeners, many (e.g., vowel shifts) are not, allowing researchers to collect naturally occurring data from linguistically naïve speakers, with little or no attention paid to altering the specific phonetic target. But, by its very nature as allegedly subconscious behavior, an enduring question is how or to what degree such phonetic variation is really meaningful for speakers (Campbell-Kibler, 2007; Eckert, 2000; Hay & Drager, 2007; Labov, 1991, 1994, 2001). Put differently, sociolinguists often discuss differences in production as if they are equivalent to differences in perception in that we correlate socially-stratified variation with speaker awareness of such variation at some level. Since different groups often use different variants, even when sharing the same basic phonological system and, in many cases, a great deal of social contact, it seems logical to assume that these differences are perceptually relevant to listeners. But, do all phonetic variations signal some kind of sociolinguistic identity or can some dialect-based variants simply be the result of the natural development of different internal pressures? Certainly, while not all variations will turn out to be cognitively relevant (Labov et al., 2006; Pierrehumbert, 2006), increasing evidence suggests that social variation, in many cases, provides more than simply clues to social identity, playing a pivotal role in the actual

* Corresponding author. Tel.: +1 541 346 3199.
E-mail addresses: tsk@uoregon.edu (T. Kendall), fridland@unr.edu (V. Fridland).

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.12.002

processing and decoding of the speech signal itself (Evans & Iverson, 2004, 2007; Foulkes & Docherty, 2006; Hay & Drager, 2007, 2010; Hay, Nolan, & Drager, 2006; Hay, Warren, & Drager, 2006; Jannedy & Hay, 2006; Niedzielski, 1999; Strand, 1999; Strand & Johnson, 1996). So, while there clearly exists a link between perception and production, exploring how this link is mediated and to what degree a speaker's own productive system mirrors or interacts with his/her perceptual system turns out to be quite complex, particularly in light of the influence of extra-linguistic social factors.

In this paper, we consider the nature of the relationship between speech production and perception. The data under analysis here were collected as part of a larger project investigating perception across regional U.S. dialects. The current paper is particularly interested in determining the effect of productive participation in regional vowel shifts on vowel identification. In order to focus on this aspect in more detail, we center most of our discussion on the linguistic systems of a subset of subjects from Western Tennessee, a region of the U.S. affected by the Southern Vowel Shift. The individuals examined here show different degrees of engagement with local vowel variants in production and, as we attempt to illustrate, perception. Further, many of these subjects are siblings in the same, large family. This fact both complicates and enriches the consideration of individual variation in production and perception, something we discuss further throughout this paper. Focusing on the Southerners while providing an overview of our more general results allows us to establish these Southerners as a sociolinguistically relevant speech community within the larger U.S., but our primary interest here will be the closer analysis of variation in production and perception within our Southern sample.

We proceed by reviewing in the next section the state of knowledge about the relationship between social factors and speech perception, and the relationship between production and perception at the individual level. We also further situate our present study within recent interests in vowel shifts occurring in regional varieties of U.S. English. In Section 3, we begin our empirical inquiry by examining variation in the results of a vowel identification task where listeners in three regions of the U.S. heard synthesized continua between /e/ and /ɛ/ (*bait*∼*bet* and *date*∼*debt*) and had to determine which they heard. To anticipate our results somewhat, this analysis shows that Southerners have, in general, backed /e/ perceptions in comparison to Northern and Western subjects. This perceptual pattern correlates nicely with the backed /e/ production of the Southern Vowel Shift and is taken as an indication that regional productive norms are reflected – at least in this case – in listeners' perceptions. To examine this more closely, we turn in Sections 4 and 5 to vowel production data in addition to and in concert with the vowel perception data for a subset of these Southern subjects. Here we show that, although the connection is quite complex and somewhat noisy, individual variability in the production of the mid front vowels does appear to relate to variability in the perception of those vowel classes. In Section 6, we examine the range of variability found across individual subjects from each region and, in particular, focus our attention on the sibling status of many of our Southern subset subjects to address what else we learn about the production-perception link from considering members of the same family. We close, in Section 7, by considering the potential social and linguistic theoretical ramifications of our results.

2. Background

Researchers have investigated different facets of the potential relationship between social factors and speech perception, ranging from general speech processing studies to those exploring the co-processing of speech and social information. Perception studies in speech sciences have often been oriented to providing empirical support for theories that attempt to explain human perceptual competence and its relationship to production more generally (in terms of our underlying abilities or cross-linguistically) rather than investigating how such abilities might be mediated by our orientation and daily exposure to socially-relevant variation at the level of the individual (see review of such work in Casserly and Pisoni (2010) or a range of relevant papers in Pisoni and Remez (2005)). A number of more recent studies have moved toward incorporating more top-down information into the speech processing equation, looking at how information such as lexical status and talker familiarity influence perception (e.g., Dahan, Drucker, & Scarborough, 2008; Elman & McClelland, 1988; Ganong, 1980; Johnson, 1997; Kazanina, Phillips, & Idsardi, 2006; McClelland & Elman, 1986; Pierrehumbert, 2001, 2002; Pisoni, Nusbaum, Luce, & Slowiaczek, 1985).

Beyond these more general perception studies, some studies have more directly examined the link between individuals' speech production and perception, typically looking at the relationship between listeners' judgments on category goodness and their own production of a vowel target (Frieda, Walley, Flege, & Sloane, 2000; Johnson, Flemming, & Wright, 1993; Newman, 2003). In general, the findings from such research have suggested a relationship between production and perception in that preferred perceptual exemplars were produced with more extreme features (e.g., higher vowels higher, lower vowels lower) than those typically produced (the 'hyperspace' effect). Using a different approach to assessing the production-perception link, Bell-Berti, Raphael, Pisoni, and Sawusch (1979) found evidence from EMG studies of two different production strategies for the vowel pairs /i/–/ɪ/ and /e/–/ɛ/ for American English speakers. For some speakers, tongue height differentiated the vowels in the pairs, while for others it was tongue tension. Formant patterns, however, did not reveal a spectral distinction. To determine if these differences in production were related to differences in perception, vowel identification tasks were administered to the EMG subjects who were asked to label each step of a seven-step continuum either /i/ or /ɪ/ in both an equal-probability (control) format and a format where /i/ values were more likely (the 'anchored' format). The researchers found that most listeners' /i/∼/ɪ/ boundary shifted toward /i/ in the anchored formant. However, this effect was much larger for those speakers whose EMG data showed tongue height to be the primary means of differentiating the tense/lax vowel pairs compared to those relying on tongue tension. Thus, this difference in production reduced this latter group's susceptibility to anchoring effects. Bell-Berti et al. suggest these findings point to a link between speech production and speech perception.

Research such as this, though indicative of a relationship between production and perception, does not delve into how speech processing might also be mediated by our orientation and daily exposure to socially-relevant variation. Examining if and how perception is tied to dialect variation productively is important both on a theoretical level (in terms of theories of speech processing and language change) but also on an applied level: What if everyone else (in school, at work, in the courtroom) does not hear what you hear? Cross-dialect comprehension is a key component in the development of educational materials and tests as well as voice recognition technologies, not to mention in our daily interactions. Thus, establishing how individuals vary in terms of perception is a first step toward ensuring such dialect-based perceptual differences (and related production differences) are not misinterpreted as disordered or insufficient (in terms of educationally-based expectations, for example).

In cross-language perception studies, unfamiliarity with foreign accented speech has been found to affect speech processing (see, e.g., Bohn & Munro (2007) or Strange (1995)) while prior exposure to foreign accented speech has been found to increase adaptation (Bradlow & Bent, 2003; Clarke & Garrett, 2004; Gass & Varonis, 1984). The effects on perception of dialect differences have not been as thoroughly researched, but several studies have shown that unfamiliar dialects also cause processing delays (Adank & McQueen, 2007; Floccia, Groslin, Girard, & Konopczynski, 2006) and affect speech intelligibility, especially in noise (Clopper & Bradlow, 2008; Labov & Ash, 1997; Mason, 1946). Clearly, as we learn more about the role of sociolinguistic variation in the processing of speech, understanding the effect of dialectal variation is increasingly important.

A number of more sociolinguistically-oriented perception studies have shown that expectations based on socially-meaningful linguistic experience indeed influence participants' perceptual behavior (see, e.g., Thomas (2002) or Hay and Drager (2010) for reviews), though most sociolinguistic research has focused either on production (e.g., Labov, 1994, 2001) or perception without much treatment of a relationship between them. A few recent studies, however, have begun to bridge this divide and have suggested that individual variation in production influences how individual listeners interpret speech stimuli, indicating that how one speaks, not just the productions surrounding him/her, has influences on perception.

In one such study, Sumner and Samuel (2009) examined how dialect variation affected processing for *–er* final words for listener groups whose own production norms included only rhotic (/r/-ful) forms (General American, GA, and rhotic New York, NY, speech) or non-rhotic (/r/-less) forms (non-rhotic NY speech). Their research attempted to determine whether dialect differences were observable both in terms of short-term processing and in terms of representation (long-term processing). In an immediate form-priming task, they found that non-rhotic primes aided form recognition for both rhotic and non-rhotic NY speakers while GA speakers unfamiliar with non-rhoticity did not have the same priming effects when provided non-rhotic primes. The researchers suggest these results show that regular exposure to a dialect, whether used productively or not, makes listeners more "flexible in form processing" (2009, p. 493), as the variable forms are treated as perceptual equivalents (even if not used productively). Importantly, they also found that the boost provided by regular exposure to both rhotic and non-rhotic speech only assisted recognition in terms of immediate form activation. Introduction of a time lag between the prime and the target form resulted in significant differences in reaction times and error rates between the non-rhotic NY speakers and both the rhotic NY and the GA speakers, with only the GA form resulting in long-term representational effects for the rhotic speakers.

So, in other words, speakers' representation of the target words matches primarily that which they produce, though they may still be able to recognize that familiar variants (such as non-rhotic variants) are related to those representations, as revealed in the immediate priming tasks. Sumner and Samuel describe this ability as 'fluent listening,' which involves the ability to recognize a number of different surface forms as related to one that is produced, but does not affect long-term storage and representation of that item as yet a single underlying (rhotic) form. In contrast, speakers exhibiting non-rhotic speech showed no long-term advantage for the General American rhotic form as both forms were equally able to assist recognition in long-term tasks, suggesting that they have multiple surface *and* underlying representations for that item. Thus, while not a straightforward relationship, these results suggest that dialect experience plays a large role in perceptual processing, but that a speaker's own production norms, not only broader community norms, play a crucial role in whether and how variable forms are stored.

In another study examining sociolinguistic perception and its relationship to production, Evans and Iverson (2004) asked subjects to select best exemplars of vowels embedded in carrier phrases that were produced with either a Northern British or (Standard) Southern British accent. Subjects with Northern and Southern English backgrounds living in London (with presumed multidialectal exposure) were tested as well as Northern English speakers remaining in the North. Their results showed that different formant frequencies were selected as best exemplars for *bud* and *cud* dependant on their embedding in accent-distinguished carrier phrases and that subject background played a significant role. Listeners from the same dialect background as used in the carrier phrase were more accurate in their adjustment. Listeners who remained in Northern England showed little evidence of accent normalization in any context, instead choosing traditional Northern English vowels in both carrier sentence types. In a subsequent study, Evans and Iverson (2007) found that young adults were able to adjust their accent over time, though there was no measurable change in their perception of these categories across the same span of time. However, they did find that individual differences in production were also reflected in differences in perception, suggesting that a speakers' own production, in addition to any adherence to community norms, interacts with speech processing.

Addressing the importance of listener- *and* speaker-specific characteristics in speech processing, Hay, Warren, and Drager (2006) found that perception of words involved in the expanding *near-square* merger in New Zealand English (NZE) was dependent both on listeners' own production of the merger and on social attributes associated with the speaker. In the study, participants who were measured for and also self-reported their own degree of merger were played words produced by 4 speakers maintaining the distinction and shown a corresponding picture that depicted varying age and class attributes (in younger/older, lower class dress/higher class dress, no photo conditions). Though the recording was identical and only the photograph actually varied across listens, both the social attributes of the speakers and the merger status of the hearer affected the results. Those with the merger were more likely to inaccurately identify the distinct vowels involved in word pairs compared to those who maintained a distinction, not only overall, but also particularly on word pairs they identified as the same in their own speech. In addition, perceived age and social class of talker also had an effect on accuracy, paralleling the fact that the merger is more often identified with younger and less socially prestigious speakers. Merged participants did not show differences in error rates based on the age manipulation. So, in other words, they were not sensitive to differences in merger by age. However, those who reported having non-merged vowels in a specific word-pair showed significantly higher error-rates in the younger photo condition. However, they were much more accurate in the older-age condition for that same word-pair, suggesting their expectations of merger for younger vs. older talkers were different. Social class manipulation results were less straightforward and greatly depended on how distinct the word pairs were for the sample talkers, but merged participants were most sensitive to the class manipulation and error rates increased as the social class depicted in the photo decreased.

It is clear from such studies that both social information about a talker *and* a listener's own production norms play key roles in speech perception. Hoping to add to these recent attempts to account for the role of production in speech processing, the current study examines the link between U.S. English speakers' own production of regionally variable vowel categories and their perception of these same categories. Regional differences across American dialects are widely recognized and mark historically and enduringly significant divides due to settlement, migration, and ethnicity, among other factors. Despite increasing mobility and contact among regions, most recent studies have shown that speech in the U.S. is being affected by a number of shift patterns that lead to greater distinctiveness in how vowels are realized phonetically in the (Inland) North and South (Eckert, 1988, 2000; Evans, 2001; Feagin, 1986; Fridland, 2000, 2001, 2003a, 2003b, 2004; Fridland & Bartlett, 2006; Gordon, 1997; Labov, 1991, 1994, 2001; Labov et al., 2006; Thomas, 1989, 1997a, 1997b, 2001). Northern speech is affected by what is known as the Northern Cities Shift (NCS), with backed and/or lowered /ɛ/ and /ɪ/ as part of a series of changes involving the front and low back system. As a result, many Northerners show a widely separated system between the tense and lax front vowels, with the /i/ and /e/ classes much more peripheral acoustically than /ɪ/ and /ɛ/. In contrast, affected by the Southern Vowel Shift (SVS), contemporary Southern speech typically shows centralized /e/ (and sometimes /i/) classes and peripheralized /ɛ/ (and sometimes /ɪ/) classes. The result is often a reversed or overlapping position acoustically of the mid (and, less commonly, high) front vowels for many Southern speakers (see Labov (1994) or Gordon (2002) for greater detail and schematizations of these shifts).

Owing to these different shift patterns, the Northern and Southern front vowel systems show strikingly different relationships between the tense and lax vowels. In the West, Luthin (1987) and Thomas (2001) suggest that some areas show evidence of the Canadian Vowel Shift (CVS) as described by Clarke et al. (1995). In the CVS, also referred to as the California Vowel Shift,[1] the lax front vowels are primarily falling (toward /æ/), rather than backing as in the NCS. In sum, these various shift patterns have lead to greater differentiation in the vowel systems across regions. As such, it would not be surprising that regional differences in perception might accompany these differences in production, as has been found previously (e.g., Janson, 1986; Labov & Ash, 1997; Plichta & Preston, 2005; Willis, 1972). Indeed, our work, as will be discussed, suggests that perception does vary across regions in line with these differences in production, particularly between the South and the other regions.

But, in addition to general differences in regional perception, we suspect, as Sumner and Samuel's (2009) research suggests, that speakers are most affected by speech norms they use themselves, not simply norms used more generally within their community, as it is these features that most directly constitute an individual's *relevant* dialect experience. While it is certainly true that speakers who share ideological, geographical, and social proximity will be more likely to also share similar linguistic systems, we cannot necessarily take this as evidence that they are similarly perceptually affected by surrounding (local) variation that is not productively relevant to them as individuals. In other words, we hear much more than we say, but somehow we must organize that perceptual exposure according to a number of different social references which those variants index in ways that are meaningful both socially and linguistically. Certainly, the variants one uses and the identity they index should play an influential role in guiding interpretation of surrounding speech. One filters speech through the lens of individual, not exclusively collective, experience.

Thus, building on preliminary work that pointed toward differences within our Southern sample (Kendall & Fridland, 2010), our current research primarily explores in greater detail how variation in the degree of individual involvement in the Southern Vowel Shift affects how Southern participants perform perceptually. Due to our focus here on intra-regional distinctions, we address cross-regional variation more comprehensively in a separate paper (Fridland & Kendall, in press).

3. Vowel identification test

Here we turn our focus to the results of a forced choice, vowel identification task designed to determine if regional dialects have different perceptual tendencies. We begin by explaining our methods and experiment design and then discuss its results.

3.1. Method

The vowel identification test was developed using a forced choice design, where synthesized vowel tokens based on the same talker's realizations were randomly played for listeners who were then asked to determine the token they just heard from two choices. For example, after hearing a synthesized token drawn from a *bait*∼*bet* continuum, listeners had to decide whether they heard *bait* or *bet*. In order to test a large number of regionally diverse subjects in a widely accessible setting, the vowel identification test was developed and administered through a website that participants could access with a password from their own computers.

3.2. Participants

For the perception study, participants from four field sites were recruited, one in the West, one in the Inland North, and two in the South (regions are labeled according to the regions of the *Atlas of North American English*, Labov et al., 2006). Western participants were students at the University of Nevada, Reno and were natives of California and Nevada. Northern participants were students at SUNY-Oswego and were natives of New York state and surrounding areas in the Inland North. Southern participants were recruited from Tennessee (TN) attending the University of Memphis and from Virginia (VA) attending Virginia Tech. Thirteen participants from the Tennessee site also provided speech samples and we examine this subsample of the subjects at length in later sections of the paper.

In total, 420 participants from these different regions in the U.S. took the online study. From these total participants, we selected those participants for analysis who were native English-speaking European Americans who were raised (from age 4) in their respective region and who reported normal hearing. 217 participants are analyzed here for the vowel identification test, 91 West, 58 North, and 68 South (20 TN, 48 VA).

3.3. The vowel identification test

Natural speech data provided by a single middle-aged male speaker with no noticeable dialect features (following Clopper & Pisoni, 2004) were used as the source stimuli for the continua synthesis. He was recorded producing monosyllabic word pairs (minimal pairs) selected to represent vowel category endpoints (e.g., *bead*∼*bid*, *bait*∼*bet*) in two consonant onset contexts (/b/ and /d/; e.g., *bait*∼*bet* and *date*∼*debt*). The vowel test included 5 such vowel continua for each context, though the current paper focuses exclusively on the results of the mid front vowels, /e/∼/ɛ/.

A 24-bit, 48,000 Hz digital recording of the sample speaker was made. Audio files containing individual words were then reduced to 10,000 Hz. No other digital signal processing techniques, such as low-pass filtering, were applied. The target vowels

[1] See http://www.stanford.edu/∼eckert/vowels.html for more detail on the California Vowel Shift.

were analyzed to determine their static and dynamic properties and to ensure that they fell within the normal range for each vowel class based on a summary of vowel frequency ranges presented in Kent and Read (2002). For each vowel pair, a seven-step continuum range was determined based on the sample speaker's production values for each of the two selected endpoint vowel categories. Thus, in synthesizing the /e/~/ɛ/ continua, the speaker's own /e/ and /ɛ/ class values were used as endpoints. Based on these endpoints, the stimuli were created through a vowel synthesis program designed by Bartek Plichta (available at http://bartus.org). The program involves an algorithm that estimated subsequent steps in a formant trajectory change over time. The algorithm interpolated missing formant, bandwidth, and amplitude values to create the stimuli trajectories, and dynamic information for diphthongal vowels was altered for each step by estimating formant trajectory and formant transitions.[2] To hold duration constant across each vowel pair, an intermediate duration based on the original target vowels' durations was adopted by means a PSOLA-based computational method. (See Appendix A for formant information at each step.)

The vowel identification test followed the one-interval two-alternative forced choice task design often used in the speech sciences (Hillenbrand, Getty, Clark, & Wheeler, 1995; Strange, 1995; Thomas, 2002), except that it was accessed online. Each trial presented a single vowel continuum step (played once) and participants were asked to indicate the token they just heard as one of two choices drawn from the relevant vowel categories (e.g., *bait* or *bet*). Participants made a decision by clicking the box before their selection and then they were asked to submit their answer by pressing the 'continue' button. Once submitted, answers could not be altered. Each step in each vowel continuum had four iterations, i.e., was played 4 times, randomized over the course of the study. The test included 560 trials. The study was also randomized by trial. Before beginning the study, participants were asked their hearing history and a series of demographic questions and then they were given a 4-trial practice test using a non-tested vowel continuum. Subjects were advised that they were to take the test in a location with minimal noise and to turn off all external noise sources. As the goal of the larger study from which this sample of participants is taken was to be a broad survey of perception across U.S. regions, it was designed to be taken in any quiet setting with a computer. We were unable to require that headphones were worn though headphones were recommended in the experiment instructions and subjects were asked to indicate whether they wore headphones or used other speakers. Of the participants taking the test, 70% reported using headphones, 15% reported using "good" external speakers, and 15% reported using laptop/internal speakers. It is not ideal that several subjects did not wear headphones. Nonetheless, we include all of these listeners in our data and, in order to assess whether headphone-wearing affected the results, we include this information as a factor in our statistical analyses below.

3.4. Vowel identification results

We now present the perception study results for the 217 participants from the three regions. Given that each of these regions has different productive tendencies for the mid front vowels (as discussed in Section 2), a reasonable hypothesis approaching this identification test would be that subjects from the different regions have different perceptual tendencies. In particular, we would expect that Southerners, for whom the Southern Vowel Shift involves a reversal or spectral overlapping of /e/ and /ɛ/, would exhibit different perceptual tendencies for the mid front vowels compared to the other regions where these vowels are produced with greater spectral distinction (like the West and North). Since the NCS is characterized by more distant /e/ and /ɛ/ vowels than found in other regions, we might also (or alternatively) expect listeners from the North to show different perceptions than listeners from regions (like the South and West) with more proximate mid front vowels. Figs. 1 and 2 show the mean step-by-step results for the three regions across the /e/ to /ɛ/ continuum, in /b/ onset and /d/ onset conditions, respectively. In all such perception plots throughout this paper, the *x*-axis refers to the continuum step with 1 indicating the most /e/ -like synthesized token while step 7 indicates the most /ɛ/ -like token; the *y*-axis displays the percentage of tokens heard as /ɛ/ and not /e/.

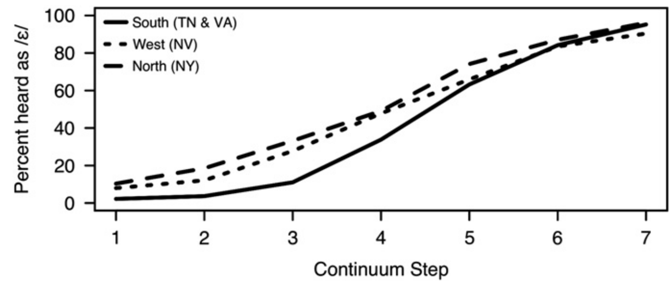

Fig. 1. *bait~bet* perception results for 217 listeners.

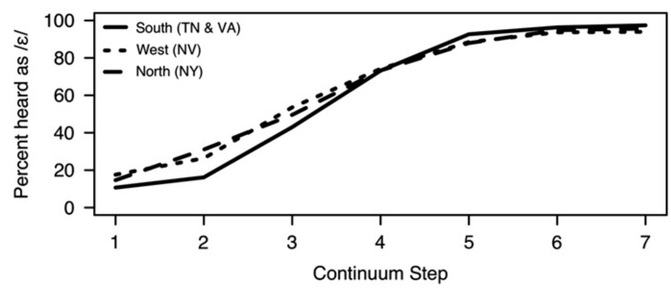

Fig. 2. *date~debt* perception results for 217 listeners.

As both figures indicate, listeners in all three regions generally heard the extreme ends of the continuum categorically or near categorically as /e/ (step 1) and /ɛ/ (step 7) with intervening steps capturing a range of response rates. There are some noticeable differences between the /b/ and /d/ contexts – for instance, listeners reach near categorical perceptions of /ɛ/ earlier in the continuum for /d/ – but both contexts indicate that Southerners perceive the continuum differently than Northern and Western listeners. For the /b/ context, there is also a just visible difference between the North and West, with the North hearing slightly more /ɛ/ than the other regions. Southerners differ from the other regions until about step 5 in the /b/ context (Fig. 1), identifying more tokens as /e/ than /ɛ/ in comparison with the North and West. For /d/ (Fig. 2), Southerners also identify more tokens as /e/ than the North and West, although the difference is less visible overall and disappears at step 4, a step earlier than for /b/.

In order to assess whether these visible patterns are significant, and whether other factors available in our data are significant (such as listener sex and headphone-usage), mixed-effect logistic regression models were fitted to each of these continua data, testing perception as /ɛ/ in each trial as the dependent variable. Modeling proceeded beginning with a minimal, intercept-only model with random intercepts for subject, and built up

[2] Dynamic versions and synthesized static versions of each set of stimuli were tested so that we could compare how categorizational behavior was affected by the removal of dynamic cues. We do not examine static stimuli here as this contrast is not relevant to the current paper's topic.

Table 1
Regression model results for *bait*∼*bet* perception for 217 listeners.

Factor	Log-odds est.	Std. err.	Z-value	p
(Intercept)	−9.615	0.647	−14.870	<0.000001
Continuum step	2.123	0.128	16.657	<0.000001
North vs. South	2.983	0.891	3.348	<0.001
West vs. South	3.583	0.828	4.326	<0.0001
Ext. spkrs vs. headphones	−0.766	0.477	−1.608	=0.11
Int. spkrs vs. headphones	−1.354	0.481	−2.815	<0.01
Step × North vs. South	−0.416	0.179	−2.34	<0.05
Step × West vs. South	−0.540	0.159	−3.389	<0.001

Somers' D_{xy}=0.925, C=0.963.

Table 2
Regression model results for *date*∼*debt* perception for 217 listeners.

Factor	Log-odds est.	Std. err.	Z-value	p
(Intercept)	−6.557	0.520	−12.622	<0.000001
Continuum step	2.048	0.132	15.546	<0.000001
North vs. South	1.715	0.735	2.334	<0.05
West vs. South	1.820	0.666	2.734	<0.01
Step × North vs. South	−0.437	0.184	−2.376	<0.05
Step × West vs. South	−0.474	0.167	−2.837	<0.01

Somers' D_{xy}=0.922, C=0.961.

to more complex models, testing all possible interactions for main effects that were significant in earlier models of either continuum dataset. Our modeling practices closely follow those recommended in Baayen (2008). The final regression models, including subject as a random intercept and continuum step as a random by subject slope, confirm that Southerners are more likely to hear /e/ (i.e., less likely to hear /ɛ/) than the other regions. The results from these regression models are presented in Table 1, for the /b/ context, and Table 2, for the /d/ context.[3] Only factors that were determined to improve the models are included in the tables. Likelihood ratio tests were used for model comparison and measures of model fit, Somers' D_{xy} and C statistics, were calculated to determine how well the models account for the data. The extremely high Somers' D_{xy} and C statistics indicate that the models fit the data quite well.

Both models yield significant main effects for continuum step and region. Listener sex is not significant in any of our perception data. Step is extremely significant, in line with our expectations from Figs. 1 and 2, with the log-odds of hearing /ɛ/ increasing by 2.1 for each increase in the continuum for /b/ and 2.0 for each increase in the continuum for /d/. For both contexts, the South is significantly less likely to hear /ɛ/ than the North and the West. The South is the reference level for both of the presented models. Identical models (not shown) were run with the region factor re-leveled with North as the reference level in order to assess whether the Northerners are significantly different than the West. This was not the case, even for the /b/ context where there is some visible indication of difference in Fig. 1 (West vs. North: p=0.4 for /b/; p=0.9 for /d/). In addition to the main effects of region and step, there are significant region x step interactions in both models, where the South "catches up" (i.e., increases relatively more in the likelihood of hearing /ɛ/) to the North and West a little with each increase in step. In summary, a hypothesis that the South would show different perceptual tendencies aligning with substantive differences in mid front vowel production was confirmed. The North, on the other hand, did not show as much general distinction in perception despite participation in the NCS shift pattern, as they only differed significantly from the South (not the West). Based on these results, it seems that Southern perception is the most altered away from the other regions. The fact that the North and West, which show highly (North) vs. moderately (West) separated mid front vowels due to the NCS and CVS, respectively, are not significantly distinct in perception may simply stem from the fact that they are also not as distinct in production when compared to SVS affected Southerners and that Northern and Western speech differences are also not as salient as regional markers as those that separate Southern speech.[4]

For the /b/ context, but not the /d/ context, there is also a significant main effect for the headphone-usage factor. Listeners who used laptop/internal speakers were significantly less likely to hear /ɛ/ than listeners wearing headphones. The middle category for this factor, "good external speakers" is not significantly different from headphones, but is about halfway between the disfavoring effect of laptop speakers and the baseline of headphones. This factor is not significant for the /d/ context, although preliminary models, which included the headphone factor, indicated that the effect was in the same direction. We tested for interactions between headphone use and region and this interaction was not (even near) significant. Thus, it appears that headphone use, or rather lack of headphone use, and the lower quality of listening resulting from other listening devices, has the same effect across all regions. Listeners appear more likely to hear /e/ with less sound fidelity. This is an interesting finding, but further consideration is outside of this paper's focus of interest and is left for research elsewhere. For our present purposes, we argue that by including headphone use as a factor in our statistical models we adequately control for the potential influence this has on our current investigation.

In terms of category crossover (measuring crossover as when the perception of /ɛ/ becomes greater than 50%), Southerners, in the /b/ context condition, maintained a longer perception of /e/ across the continuum compared to other listeners. An ANOVA testing the crossover point location for the /b/ context yields (F(2, 214)=4.1796) p=0.017; a post hoc Tukey test shows that the North vs. South comparison is significant (p=0.02) and that the West vs. South comparison almost reaches significance (p=0.07). The West vs. North comparison, as expected from the figure and data above, is not significant (p=0.7). The crossover in the /d/ context happens earlier for all three groups and despite the fact that the data for the /d/ context (Fig. 2 and Table 2) indicate similar tendencies to the /b/ context, the regions do not have significantly different crossovers in this condition (ANOVA: F(2, 214)=0.7009, p=0.5).

Nonetheless, the overall identification test results indicate that the Southern listeners hear the /e/∼/ɛ/ boundary differently than the listeners from the North and the West. Coupled with our knowledge of production in the Southern Vowel Shift, which is categorized by a backing of /e/ relative to, and closer to, /ɛ/, this finding is suggestive of a link between production and perception, with the backed /e/ of the Southern Vowel Shift creating a broader

[3] Mixed-effect regressions were run using the lmer() function of the lme4 package in R (Bates & Maechler, 2010) and additional functions from the languageR package in R (Baayen, 2008, 2010).

[4] As an anonymous reviewer pointed out, it could be possible that the shifted /ɛ/ productions common in the Northern and Western varieties were not represented in the stimulus materials and so differences between these groups were missed. However, the sample speaker, though lacking salient dialect markers, spent most of his childhood in Reno, NV where the Western participants were recruited and it would be surprising that his /e/ and /ɛ/ production differed markedly from others in the West. Further, the fact that the vowel identification task obtained the full range of possible percepts (near 0% to near 100%) indicates that the stimuli did capture a wide range of possible /e/ and /ɛ/ productions.

range for tense vowel (i.e., /e/) perception. In other words, as suggested above, these cross-dialect perception differences reflect a difference generally in Southern production and perception behavior compared to other regions.

These results suggest that Southerners form a unique speech community not just in production (as previously noted) but in perception as well. In the next section, we look in greater depth within this region, examining more specifically how perception and production are linked at the level of the individual (beyond simply generalized norms within a community).[5] Do *individuals* who exhibit more shifted SVS vowels (especially mid front vowels) also exhibit perceptual tendencies that mirror these shifted productions when compared to other individuals from the same locale?

4. Subset production data methods and analysis

In the previous section, we saw evidence that listeners in different regions have different perceptual tendencies and that, for the South in particular, these tendencies appear to be in line with expectations based on the productive processes ongoing in the regions. The Southern listeners showed a perceptual identification that corresponded to the Southern Vowel Shift's tendency for /e/ to back and, potentially, overlap with /ɛ/ in productive, spectral space. One of our primary interests in this paper is assessing whether such a relationship between production and perception is maintained when we consider production and perception at the individual level, rather than overall community norms. We now turn to examine this more closely by considering the production data from a subset of our Southern listeners. Then, in Section 5, we compare these data to perception data for the same subjects.

4.1. Method

In addition to participating in the vowel identification tasks, a subset of perception participants from each region were also recorded reading a reading passage and word list following the perception task. Here, we examine thirteen participants from the Tennessee field site (in the South). These subjects were all recruited and then recorded by a local fieldworker using the friend of a friend approach commonly adopted in sociolinguistic research (Milroy, 1987). Seven of these subsample subjects were siblings. All of the subjects who provided speech data were either currently enrolled in college or had completed their degree. All 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 or home (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 item recitation).

Table 3 provides demographic information for these thirteen Southern participants, including their relationships (whether they are siblings in the same (large) family or not).

Table 3
Southern speakers included in production subsample.

Speaker	Raised	Sex	Age	Is sibling?
Abbey503	Memphis	Female	18–25	Sibling
Andrew809	Memphis	Male	18–25	Sibling
Brittany371	Memphis	Female	18–25	–
Erica2248	Union City	Female	18–25	–
Isaac815	Memphis	Male	31–40	Sibling
Kim1111	Memphis	Female	26–30	–
Laura816	Ripley	Female	18–25	–
Matt2526	Memphis	Male	26–30	Sibling
nawill10	Memphis	Male	18–25	Sibling
Nick3218	Memphis	Male	18–25	–
Oscar1610	Memphis	Male	26–30	Sibling
Roj1518	Memphis	Female	18–25	–
Tbone808	Memphis	Male	18–25	Sibling

4.2. Acoustic analysis

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 (these were used to expand our inventory of front vowels). All vowels were measured from monosyllabic words, with the exception of three words in the reading passage, where we measured the stressed, first syllable. The word types are listed in Appendix B. Although all speakers read the same materials, one speaker (nawill10) misread one word (*bead*), which was excluded from his production data.

Praat (Boersma & Weenink, 2009) was used for all acoustic measurements. Vowel data, including f0, F1, F2, and F3 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.[6] In the remaining discussion, all vowel data are manipulated and (as needed) normalized using the Vowels package (Kendall & Thomas, 2009; http://cran.r-project.org/web/packages/vowels/) for R (R Development Core Team, 2008). In the following analysis, we primarily use the measurement taken at the 1/3 temporal point in each vowel's duration as the comparative reference point. We describe this as the vowel "nucleus" below.

As we are interested primarily in the mid front vowels and the extent to which they exhibit SVS shift acoustically in the subsample, a Euclidean distance measure between speakers' mean tense and lax vowel nuclei for the mid front subsystem (i.e., /e/ vs. /ɛ/) was taken. First, all vowel data was normalized using the Lobanov method of vowel normalization (Lobanov, 1971). In addition to the normal advantages of making the vowel data more comparable across speakers – such as by minimizing acoustic differences having to do with differing vocal tract sizes between, say, females and males (cf. Thomas & Kendall, 2007, Watt, Fabricius, & Kendall, 2011) – Lobanov normalization also results in generating similarly scaled F1 and F2 values, making the data more appropriately described by Euclidean distance (ED) measures than many other normalization processes, or unnormalized data. The ED measure allows us to determine the overall acoustic distance between the tense and lax vowel pair, /e/ and /ɛ/, thus providing a quantitative measure of overlap/reversal.[7]

[5] As mentioned in the introduction, production and perception results *across regions* are examined more fully in a separate paper (Fridland & Kendall, in press).

[6] The script is available for public download at http://ncslaap.lib.ncsu.edu/tools/scripts/vowel_capture_aug09.praat.

[7] Recent work (e.g., Hay, Warren, and Drager 2006; Hall-Lew, 2009) has suggested that Pillai scores may be useful for comparing vowel distributions, especially for cases of potential merger. We tested this by generating Pillai scores for each speaker's /e/–/ɛ/ comparison, but in all cases this lead to less significant (but similar) results in our statistical models than for our ED measure, both for raw scores (ED vs. Pillai) and for ranking individuals based on these measures (see below). The fact that ED – a measure of the distance between central tendencies – outperforms Pillai scores – a measure of distribution overlap – may indicate that what matters for explaining the perception data is speakers' central tendencies (i.e., vowel targets) more than the spread of their actual productions.

Table 4
Southern speakers degree of SVS production.

Speaker	Is sibling?	/e/–/ɛ/ ED	/e/–/ɛ/ Rank	Is shifter?
Isaac815	Sibling	0.320	1	Yes
Laura816	–	0.435	2	Yes
nawill10	Sibling	0.444	3	Yes
Nick3218	–	0.702	4	Yes
Tbone808	Sibling	0.770	5	Yes
Oscar1610	Sibling	0.893	6	Yes
Roj1518	–	1.100	7	–
Andrew809	Sibling	1.142	8	–
Erica2248	–	1.180	9	–
Kim1111	–	1.341	10	–
Matt2526	Sibling	1.585	11	–
Brittany371	–	1.792	12	–
Abbey503	Sibling	1.836	13	–

These metrics were checked against an impressionistic analysis of each speaker's system in terms of degree of Southern Shift advancement for the mid front vowels and were in line with our impressionistic assessment of the individuals.

Table 4 presents the Euclidean distance measure for each Southerner in the subsample, the ranking of each individual (ordered from most overlapping to least) by the ED measure, and a categorical designation of whether the individual was an "advanced" Southern Shifter or not. This final category was based on the acoustic ED measure (all and only individuals designated as Shifters have an ED distance < 0.9[8]). As can be clearly seen in the ED scores, the most advanced shifter barely maintains a nucleus distinction between the two classes while the least affected speaker shows quite distinct nuclei, a contrast explored in greater detail below. Table 4 also re-iterates each individual's familial status, to indicate that advancement in the SVS does not correlate with sibling status. Both the most advanced (Isaac815) and least advanced (Abbey503) speakers are siblings and, notably, the siblings are almost evenly split between Shifters and Non-shifters. We consider sibling status further in Section 6.

Figs. 3 and 4 provide general pictures of the vowel spaces for these Southerners. Fig. 3 provides a vowel plot of the means for the Southern subsample divided into two groups, a +SVS and −SVS group, as indicated in Table 4 (with "Yes" indicating +SVS). This binary classification of these speakers is not meant to indicate hard and fast that some speakers sound Southern and that some do not, nor is it meant to indicate that one can readily draw a line between these two sets of speakers in terms of their integration within the Southern speech community (as there are numerous ways that Southern speech and identity can be realized). This binary classification is primarily a convenient way to examine and display the distinctions in the production systems of these individuals in terms of their participation in the regional shift process, as in Fig. 3, and compare them to the perception results for the mid front vowel classes.

As can be seen in Fig. 3, there are clear differences between the +SVS and −SVS speakers acoustically. While the difference between the mid front, /e/ and /ɛ/, vowels is, of course, derived by the method used to bin the speakers, in general, the short front system appears to differ more generally between the two sets of speakers, suggesting the peripheralization of the front lax system

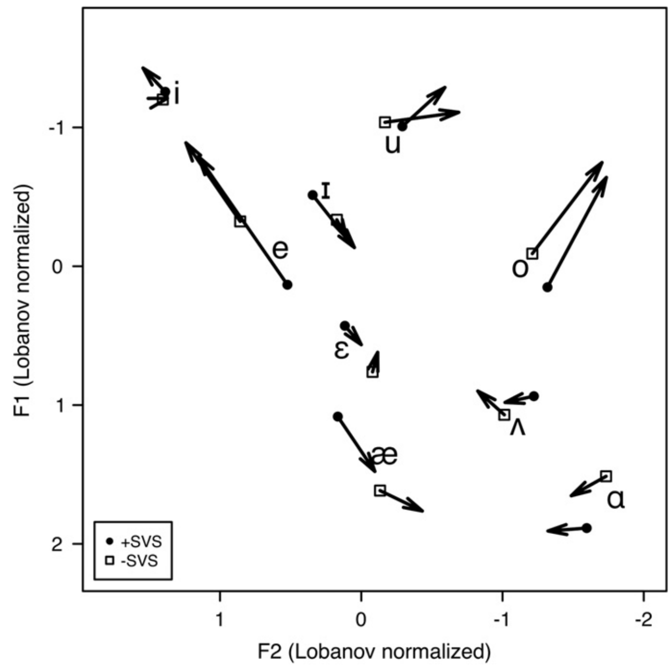

Fig. 3. Southerners mean vowels, in +SVS and −SVS groups.

in line with SVS patterns. Of particular interest here, the mid front vowels are much closer for the +SVS speakers, with *both* the tense and lax mid front vowel quite distinct from the Non-shifters' acoustic position.

To investigate the production data more closely in terms of individual speakers' systems, Fig. 4 presents nuclei measurements for the individual vowel tokens for the front vowels (/i/, /ɪ/, /e/, /ɛ/, and /æ/) for four individuals, Isaac815 and Laura816, the two most advanced SVS speakers by the /e/–/ɛ/ ED metric, on the left, and Matt2526 and Abbey503, two of the least advanced speakers, on the right.

Recall that in the SVS, mid front vowels essentially move toward less phonetic distinction, with /e/ and /ɛ/ often overlapping in spectral space. This kind of acoustic relationship is strongly exhibited for the two speakers on the left of Fig. 4. As is evident, Isaac815 and Laura816 show extensive overlap in their mid front vowel system. Clearly, these speakers participate substantially in the SVS as it affects the mid system. In contrast, looking at Matt2526 and Abbey503, it is evident they maintain very distinct mid front classes, showing little indication of shift at all. In short, some of these speakers are strong participants in the Southern Vowel Shift while some are clearly not despite the similarity in their community exposure. The high front system is not as distinct across speakers, with some speakers showing closer /i/ and /ɪ/ classes than others but no overlap acoustically.

Unlike cross-regional differences, where it is not a leap to assume that speakers from different regions do not come in contact regularly leading to the development of vowel differences, these differences in production among Southern participants are much harder to reconcile, particularly since many of these participants are siblings from the same large family (including Isaac815, Abbey503, and Matt2526 above). Clearly, it is not dialect exposure alone that establishes different productive targets, as we would assume these siblings had much overlap in terms of their exposure to vowel variants as small children. And, if speakers from the same family show this much variability in terms of vowel production, we would have to assume that unrelated listeners receive even more diverse input in childhood and in their daily interactions with other members of their same

[8] A Euclidean distance value of 0.9 would seem to be an arbitrary cut-off point on its own for designating speakers as Shifters or Non-shifters, but when we compare the Southern subsample to production data we have from Reno, NV (West) and Oswego, NY (North), not otherwise examined in this paper (see Fridland & Kendall, in press), we find that the shortest Euclidean distances among the non-Southerners is > 0.9. This, coupled with our impressionistic assessment, justified our decision to use this value as the cut-off.

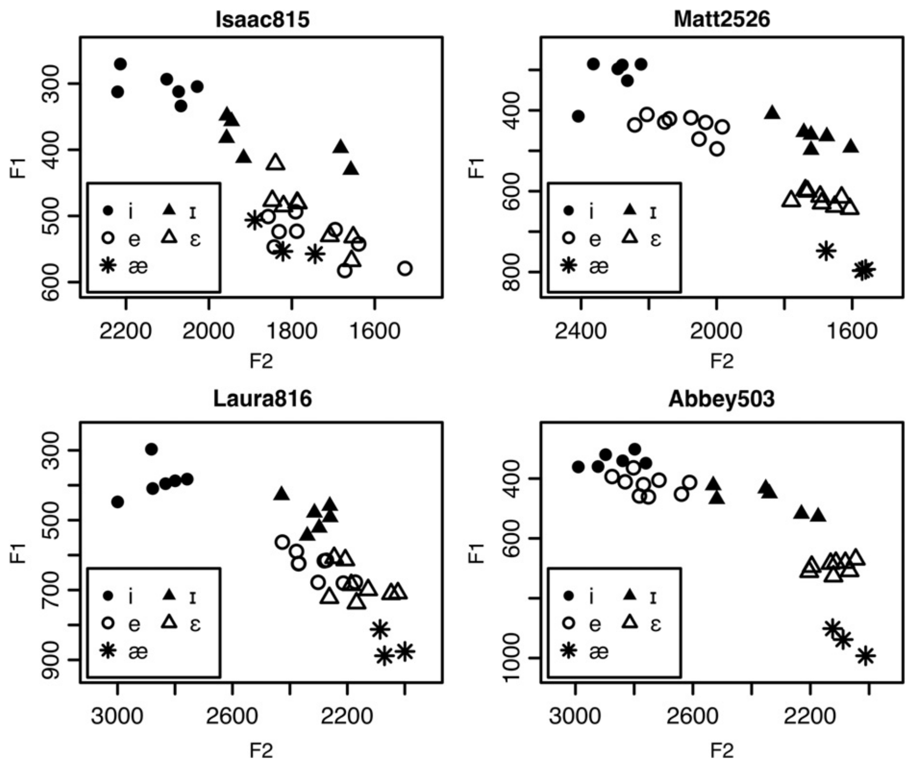

Fig. 4. Four individual Southerners, front vowels, nuclei only.

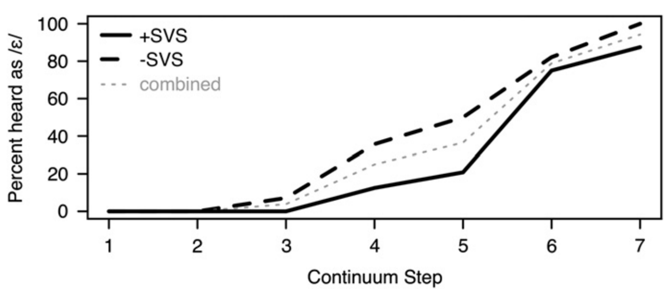

Fig. 5. Subsample *bait*~*bet* perception results.

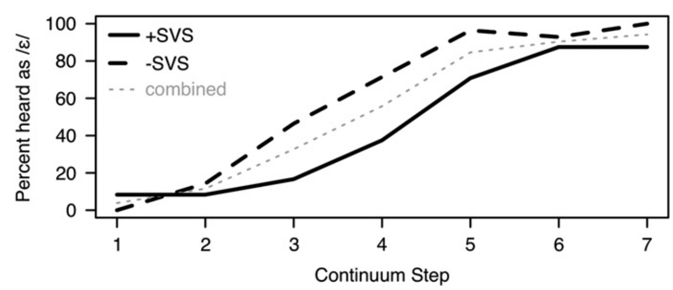

Fig. 6. Subsample *date*~*debt* perception results.

community. (We will address the role of familial status more extensively in Section 6.) Our data suggest that even within very similar groups, individual behavior is remarkably variable. This difference in shift participation is most likely not due to substantially different linguistic exposure, but more a matter of how similar linguistic exposure creates different opportunities for the development of an individual's productive system in line with how they encode and identify with surrounding exemplars (socially and linguistically).

To pursue this line of inquiry in the next section, we will examine how the individual Southern speakers' production relates to how they perceive the mid front vowel continuum. In other words, might differences in the degree of identification with locally-based linguistic variation (as evidenced by shift participation) play a role in shaping how vowel tokens are categorized?

5. Relating perception and production

Turning back now to the perception data for these subsampled subjects, we consider the perceptual variability across subjects in the vowel identification test and whether the (mid front vowel) production data for these subjects aids in understanding the variability in the perception data.

5.1. Perception study results for subsample

Figs. 5 and 6 show the aggregate identification functions for the Southern subset participants for the *bait*~*bet* and *date*~*debt* continua, respectively. (Individual identification functions for each talker-listener can be found in Appendix C.) Each figure presents the perception data separated by the $\pm$SVS distinction from the production data, with the dashed line indicating the Non-shifters and the solid line indicating the Shifters. The fainter, dotted line indicates the overall mean across all 13 subjects.

Looking at these aggregate Shifter versus Non-shifter plots, we see that on average the +SVS Southerners have a later shift in perception to /ɛ/ as they move through both continua. As in Section 3.4, we use mixed-effect regression modeling, with random intercepts for subject and random by subject slopes for continuum step, to gain more insight into this pattern. For the /b/ context, we find a near significant ($p=0.0536$) relationship between +SVS and −SVS status. The model statistics and fixed-effect terms (only) for this model are shown in Table 5. As before (i.e., the models in Section 3.4), we obtain quite good fitting

Table 5
Regression model results for *bait*∼*bet* perception for Southern subsample, testing categorical Shifter status.

Factor	Log-odds est.	Std. err.	Z-value	p
(Intercept)	−11.307	1.802	−6.274	< 0.000001
Continuum step	2.303	0.288	7.997	< 0.000001
+SVS vs. −SVS	−1.822	0.944	−1.930	=0.0536

Somers' D_{xy}=0.948, C=0.974.

models for the data with extremely high Somers' D_{xy} and C statistics. For the /d/ context, the +SVS vs. −SVS category is further from significance (with p=0.2) and we do not show the model results. Yet, as Fig. 6 indicates, there is nonetheless a trend in the same direction as for the /b/ context.[9]

Models that test the actual /e/–/ɛ/ Euclidean distance come close to yielding significant results for both contexts (/b/ context: /e/–/ɛ/ ED obtains p=0.087; /d/: /e/–/ɛ/ ED, p=0.096) and models that test the rank order of the Euclidean distances (the order of the individuals in Table 4 and Appendix C) obtain better – though still not quite significant – results (/b/: /e/–/ɛ/ rank, p=0.075; /d/: /e/–/ɛ/ rank, p=0.062). Overall, although the effects do not reach significance, the shifted Southerners appear less likely than the non-shifted Southerners to hear /ɛ/. More data than is available for these 13 subjects alone would help give resolution to these patterns, but we take this as suggestive of a pattern of perception that relates to differences in degree of shift participation by our Southern subjects.[10]

As in the cross-regional data which indicated that Southerners had significantly later crossover to /ɛ/ compared to other regions, it also appears that those Southerners most affected by Southern Vowel Shift processes in the mid front subsystem have the latest shift in perception. Such a relationship mirrors production as the Southern shifters showed much more centralized /e/ tokens in production which would, in effect, result in a broader distribution of tokens perceptually that might be categorized as /e/ for such speakers. We see here some suggestion that perceptual differences are indeed patterned similarly to differences in production, with those more greatly orienting toward SVS norms doing so both in production and perception. So, although the evidence is not crystal clear, there are indications here that exposure to community norms is filtered differently by speakers despite similar demographic and experiential background. Our results suggest, however, that these more general community patterns do play the most influential role, as seen by our finding of more significant differences cross-regionally as discussed in Section 3.

6. Variance across individual subjects, the South, and siblinghood

Turning our focus to the question of how much of a role similar input plays in shaping the perceptual system, we now look more closely at these Southern subjects to ask to what extent sibling status (in the same family) impacts individual variability. We also examine, more broadly, the range of individual variability we find in the perception data across the three regions to ask to what extent individuals in the South are more or less variable (across individuals) in their perception when compared to non-Southerner listener groups. This part of the investigation is, in large part, motivated by the fact that seven of the 13 participants in our Southern production-perception subsample are siblings. Such familial relationship among this many subjects raises some additional questions of interest, which we believe are worth further consideration, although we begin this section with a quick note of hesitancy. The goal of this experiment was not to examine perception among siblings and the sibling status of these participants was, ultimately, somewhat of an accident – a result of a fieldworker using her family as a convenient source of local subjects. Other than the basic facts of their sibling status and the demographic information supplied by each of these (and all) participants in the listening experiment, we do not have extensive information about their background. Nonetheless, having production and perception data from *seven* siblings seems like a rare opportunity to consider the extent to which, or whether at all, sibling status has an effect on perception.

We might expect – in purely social terms (and abstracting away from the on-going debate in numerous fields over "nature vs. nurture") – that being raised in similar environments would lead to shared experiences, which in turn would impact linguistic representations and perceptual tendencies. In other words, unlike a randomly matched subset of seven unrelated participants from the same area who receive broadly similar community input, we might expect that, though still likely variable, the range of linguistic and social variability within a family unit might be more constrained by the similarity of linguistic, social, and ideological background to which siblings are exposed in comparison to unrelated but demographically similar subjects. We have seen above that our sample of siblings show surprising variability in terms of their linguistic output (recall that the largest and smallest /e/–/ɛ/ Euclidean distances are for two of the siblings). However, the vernacular reorganization that occurs in childhood and adolescence (Kerswill, 1996; Kerswill & Williams, 2000; Tagliamonte & D'Arcy, 2009) might explain this variability in production since peer group influence on language use appears to be quite extensive through adolescence, giving rise to phonetically variable output and increasing distance from parental norms. Yet, initial exposure to the same linguistic input during the period in which phonemic representation is established would, one might hypothesize, make perception more similar across siblings even if later exposure creates the opportunity for fine tuning within those categories (and, not surprisingly, the speech of very young children has been found to be most closely approximate to that of their primary caregiver but diverges as children age; Kerswill, 1996). With such a large group of siblings available in our dataset, we ask here whether there is any evidence for effects of sibling status on perception among these speakers?

To test this, we first include sibling status as a factor in the mixed-effect regression for the TN subjects. Here, we include all twenty perception subjects from Tennessee, not just the thirteen for whom we have production data, since this comparison does not require insight from the production data. As before, individual subject is included as a random intercept in the model and continuum step is included as a by subject random slope. As the model statistics and fixed-effect terms in Table 6 indicate, sibling status arises as a significant predictor for the /b/ context. Sibling status does not yield significance for the /d/ context, mirroring the earlier findings that the variability in the /d/ context perception data is less patterned (at least in terms of the available

[9] In general, the /d/ context data are noisier than the /b/ context data. Other explorations of these data, such as a binary comparison of subjects with ED above or below 0.75 (instead of 0.9), not shown here, do yield significance for the productive "Shifter" vs. "Non-shifter" category in mixed-effect regressions for the /d/ context. We are unable, at this point, to determine whether the differences between the /b/ and /d/ context data are a linguistic phenomenon (i.e., having to do with different perceptual characters of the vowels following /b/ and /d/) or more simply differences in our synthesized stimuli.

[10] Fridland and Kendall (in press) examine a larger number of subjects from across several regions and find support for a direct Euclidean distance relationship between production and perception for the mid front vowels.

Table 6
Regression model results for *bait*∼*bet* perception for Tennessee subjects, testing sibling status.

Factor	Log-odds est.	Std. err.	Z-value	p
(Intercept)	−9.954	1.392	−7.150	< 0.000001
Continuum step	2.190	0.234	9.353	< 0.000001
Is sibling?	−2.276	0.829	−2.746	< 0.01

Somers' D_{xy}=0.951, C=0.976.

predictors). No other factors, or interactions, arise as significant in any of the models we tested. The model indicates that the seven siblings are less likely to hear /ε/ than the 13 other subjects (in the /b/ context). As before, the high D_{xy} and C values indicate that the model fits the data quite well.

Recall from Table 4 that the siblings are roughly split between Shifters and Non-shifters, so the influence of sibling status is not simply a proxy for +SVS. Since the model here is based on the perception data for all 20 listeners from Western Tennessee and not just the 13 for whom we have production data, this model cannot be directly compared to the model in Table 5 which tested the influence of the ±SVS production measure.[11]

Altogether, the significance of the sibling factor (for the /b/ context and full 20 listeners from TN) is an interesting finding as it may suggest that production is more available to fine-tuning in later childhood than perception given the variability found for these siblings in production, but it is not the only comparison to make to investigate the importance of sibling status. In fact, we are interested more in how similar the siblings are to one another than in whether sibling versus non-sibling status matters in absolute terms, i.e., in terms of how likely the siblings are to hear /ε/ in comparison to non-siblings. We assume it is arbitrary where in the range of possible perceptual tendencies these siblings appear in comparison to others who do not share sibling status with them. At this point, we ask, is it the case that these seven siblings show more similar perceptions to one another than, say, seven randomly selected participants from the same region? To determine whether it might just be coincidental that the sibling set showed differences in perception compared to the other subjects, we must also consider more generally how variable subjects are within each region.

To assess this, we run a Monte Carlo experiment whereby we generate 10,000 randomly selected groups of seven individuals from each of the three regions and examine the variance within the groups' data. For simplicity, we focus on the crossover points rather than the full perception data. This simulation allows us to judge what kind of variability occurs in the perception (crossover point) data. Specifically, what distribution of standard deviations do we find across all of these random groups of seven? (The Southerners could be further separated into Tennesseans and Virginians but all assessments of their /e/∼/ε/ perception data and preliminary Monte Carlo simulations indicate that they are not different for the /e/∼/ε/ continua, and thus this separation is not warranted by the data.)

Examining the standard deviations for these randomly generated groups allows us to estimate the overall variability we should expect across subjects from each region. This in turn allows us to ask whether the seven siblings cluster together better than chance (in comparison to others from their region). The comparison between the siblings' standard deviation and the Monte Carlo based estimation of the Southerners is especially relevant here, although we also generate the same simulated data for the Northern and Western subjects to assess how similar the variability is across regions. Monte Carlo simulations, i.e., iterating over many random subsamples of the data, are useful here as we are interested in comparing the perceptions of individuals in groups of seven (as that is the number of siblings) without presupposing which seven subjects to consider at any given time. A more comprehensive consideration might assess the variability found across *every* possible combination of seven subjects. However, generating a full combinatorial analysis of all sets of seven individuals from each region would be unwieldy and unnecessary – for instance, there are over 300 million groups of seven that can be subsampled from our 58 Northern listeners – and much smaller, randomly determined comparisons (e.g., 10,000) should generate stable patterns.

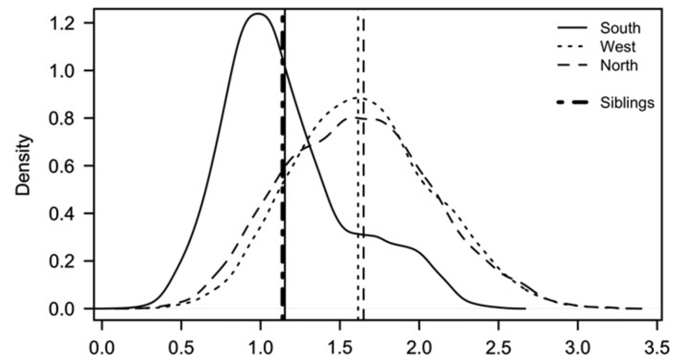

Fig. 7. Distribution of the standard deviations for the /b/ crossover points, from 10,000 random groupings of seven subjects; the vertical lines indicate the standard deviation means for each group, the bold dot-dash vertical line indicates the standard deviation among the seven actual Southern siblings.

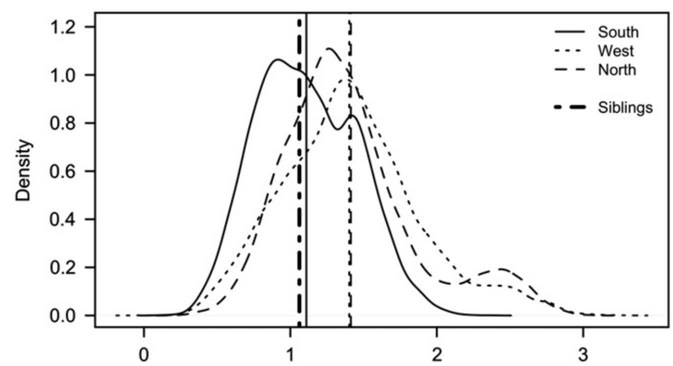

Fig. 8. Distribution of the standard deviations for the /d/ crossover points, from 10,000 random groupings of seven subjects; the vertical lines indicate the standard deviation means for each group, the bold dot-dash vertical line indicates the standard deviation among the seven actual Southern siblings.

The kernel density functions of the standard deviations yielded by this for each regional group are plotted for the /b/ context in Fig. 7 and for the /d/ context in Fig. 8 (these density plots provide smoothed histograms of the data). The Southerners – irrespective of sibling status – have a generally tighter distribution than the other two regions, seen by their much lower standard deviations. Since the distributions are not entirely normal, Kruskal-Wallis rank sum tests were used (in lieu of ANOVAs) and these confirm that the region-based distributions are different (/b/ context: χ^2=6675.6, df=2, $p < 0.000001$; /d/ context: χ^2=2797.9, df=2, $p < 0.000001$).

[11] Testing the sibling factor on the 13 subsampled participants only yields a just under significant p=0.0511. A likelihood ratio test indicates that a model containing the sibling factor is actually slightly worse than the model with the ±SVS factor (Table 5). (A model which includes *both* sibling *and* ±SVS factors obtains p=0.09 for both of these factors and is worse than models which include only one of the two.) In sum, while both models fall just short of reaching significance at the $p < 0.05$ level, the model for the /b/ context shown in Table 5 in the previous section appears to slightly outperform a model on the same data which includes sibling status rather than ±SVS status.

Table A1
/e/–/ɛ/ /b/ context formant values step 1 to step 7.

Time (s)	Step 1		Step 2		Step 3		Step 4		Step 5		Step 6		Step 7	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
0.01	486	1681	486	1681	486	1681	486	1681	486	1681	486	1681	486	1681
0.03	432	1945	432	1945	432	1945	432	1945	432	1945	432	1945	432	1945
0.05	436	2041	436	2041	436	2041	436	2041	436	2041	436	2041	436	2041
0.07	425	2118	441	2077	457	2036	473	1995	489	1954	505	1913	521	1872
0.09	417	2162	436	2115	455	2068	474	2021	493	1974	512	1927	531	1880
0.11	399	2158	419	2112	439	2066	459	2020	479	1974	499	1928	519	1882
0.13	376	2209	398	2167	420	2125	442	2083	464	2041	486	1999	508	1957
0.15	348	2220	359	2173	370	2126	381	2079	392	2032	403	1985	414	1938
0.17	349	2249	371	2201	393	2153	415	2105	437	2057	459	2009	481	1961
0.19	322	2214	355	2161	388	2108	421	2055	454	2002	487	1949	520	1896
0.21	324	2016	358	1980	392	1944	426	1908	460	1872	494	1836	528	1800
0.23	308	1840	337	1822	366	1804	395	1786	424	1768	453	1750	482	1732
0.25	349	1623	395	1646	441	1669	487	1692	533	1715	579	1738	625	1761
0.27	484	1965	554	1964	624	1963	694	1962	764	1961	834	1960	904	1959
0.29	823	1973	840	1975	857	1977	874	1979	891	1981	908	1983	925	1985
0.31	1160	2440	1160	2416	1160	2392	1160	2368	1160	2344	1160	2320	1160	2296
0.33	989	1971	958	1961	927	1951	896	1941	865	1931	834	1921	803	1911
0.35	782	1970	776	1957	770	1944	764	1931	758	1918	752	1905	746	1892
0.37	682	1882	690	1880	698	1878	706	1876	714	1874	722	1872	730	1870
0.39	702	1865	683	1852	664	1839	645	1826	626	1813	607	1800	588	1787
0.41	857	1828	857	1828	857	1828	857	1828	857	1828	857	1828	857	1828

Table A2
/e/–/ɛ/ /d/ context formant values step 1 to step 7.

Time (s)	Step 1		Step 2		Step 3		Step 4		Step 5		Step 6		Step 7	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
0.01	1110	1894	1110	1894	1110	1894	1110	1894	1110	1894	1110	1894	1110	1894
0.03	644	1962	644	1962	644	1962	644	1962	644	1962	644	1962	644	1962
0.05	407	1956	407	1956	407	1956	407	1956	407	1956	407	1956	407	1956
0.07	443	1971	446	1962	449	1953	452	1944	455	1935	458	1926	461	1917
0.09	439	2038	448	2014	457	1990	466	1966	475	1942	484	1918	493	1894
0.11	413	2106	427	2068	441	2030	455	1992	469	1954	483	1916	497	1878
0.13	406	2145	425	2104	444	2063	463	2022	482	1981	501	1940	520	1899
0.15	364	2172	386	2131	408	2090	430	2049	452	2008	474	1967	496	1926
0.17	340	2264	360	2211	380	2158	400	2105	420	2052	440	1999	460	1946
0.19	345	2145	357	2109	369	2073	381	2037	393	2001	405	1965	417	1929
0.21	302	1961	324	1932	346	1903	368	1874	390	1845	412	1816	434	1787
0.23	399	2057	415	2033	431	2009	447	1985	463	1961	479	1937	495	1913
0.25	464	1989	473	1948	482	1907	491	1866	500	1825	509	1784	518	1743
0.27	869	2009	830	1974	791	1939	752	1904	713	1869	674	1834	635	1799
0.29	888	1854	878	1838	868	1822	858	1806	848	1790	838	1774	828	1758
0.31	615	1992	670	1977	725	1962	780	1947	835	1932	890	1917	945	1902
0.33	882	1869	850	1860	818	1851	786	1842	754	1833	722	1824	690	1815
0.35	734	1913	738	1895	742	1877	746	1859	750	1841	754	1823	758	1805
0.37	974	1895	945	1890	916	1885	887	1880	858	1875	829	1870	800	1865
0.39	701	1897	723	1879	745	1861	767	1843	789	1825	811	1807	833	1789
0.41	834	1878	834	1878	834	1878	834	1878	834	1878	834	1878	834	1878

Wilcoxon rank sum tests of the difference between the Southern distribution and the combined Northern and Western distributions[12] confirm further that the lower standard deviations of the South are significantly different (/b/ context: $W=42351364$, $p<0.000001$; /d/ context: $W=62718086$, $p<0.000001$). Simply put, Southerners have less perceptual variability than the other two regions. This is especially striking considering that the Southerners in the simulation are sampled from two different field sites in the South – one in Western Tennessee and one in Virginia – and we might expect that fact to add variability over what we might find in individual field sites elsewhere. Although more research is necessary in order to determine the root of this difference, one putative interpretation may relate to the clear influence of regionally-based community norms in the Southern sample compared to the Northern and Western sample. Recalling the results in Section 3, which indicate that Southern perception patterns differed significantly from Northern and Western perception, Southerners appear to maintain more distinct perceptual norms, which may result in greater perceptual cohesion within the group compared to the other regions.

Further, we see that the standard deviations for the actual seven Southern siblings (/b/: 1.138; /d/: 1.062), indicated by the bold dot-dash lines in the figures, are ultimately similar to the overall distributions of standard deviations for the Southerners. The values are lower than the mean standard deviations for the randomly sampled Southerners for both contexts (/b/: 1.153; /d/: 1.109), but this difference is slight. In fact, the value for the

[12] We combine the North and West distributions for the Wilcoxon rank sum test. They are quite similar and we are primarily interested in confirming the difference between the South and non-South.

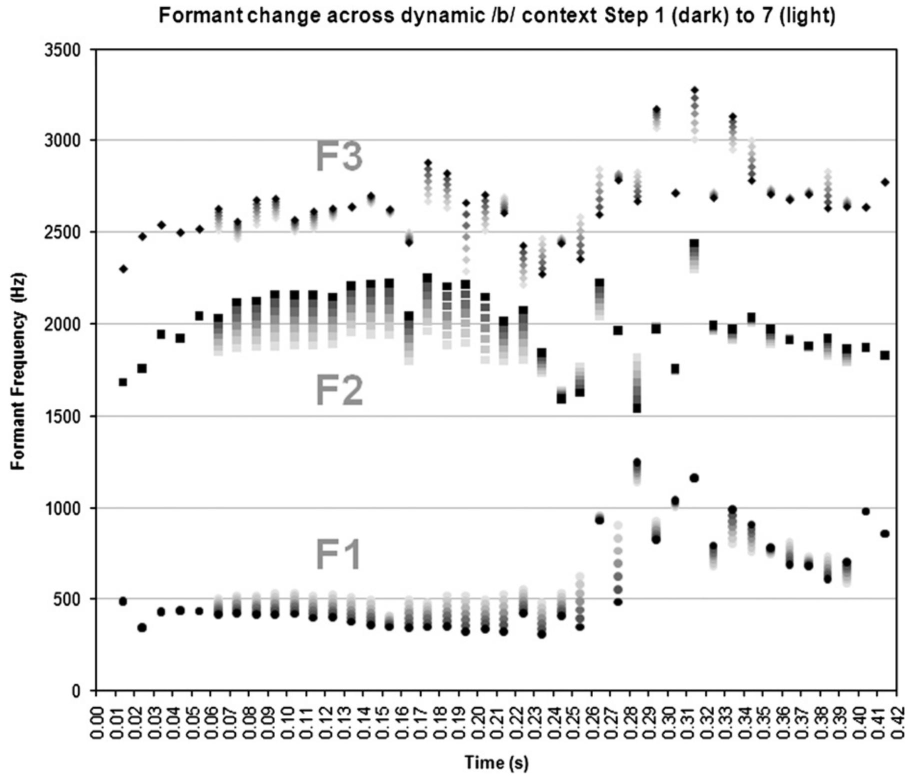

Fig. A1. Formant frequencies for *bait*∼*bet* stimuli.

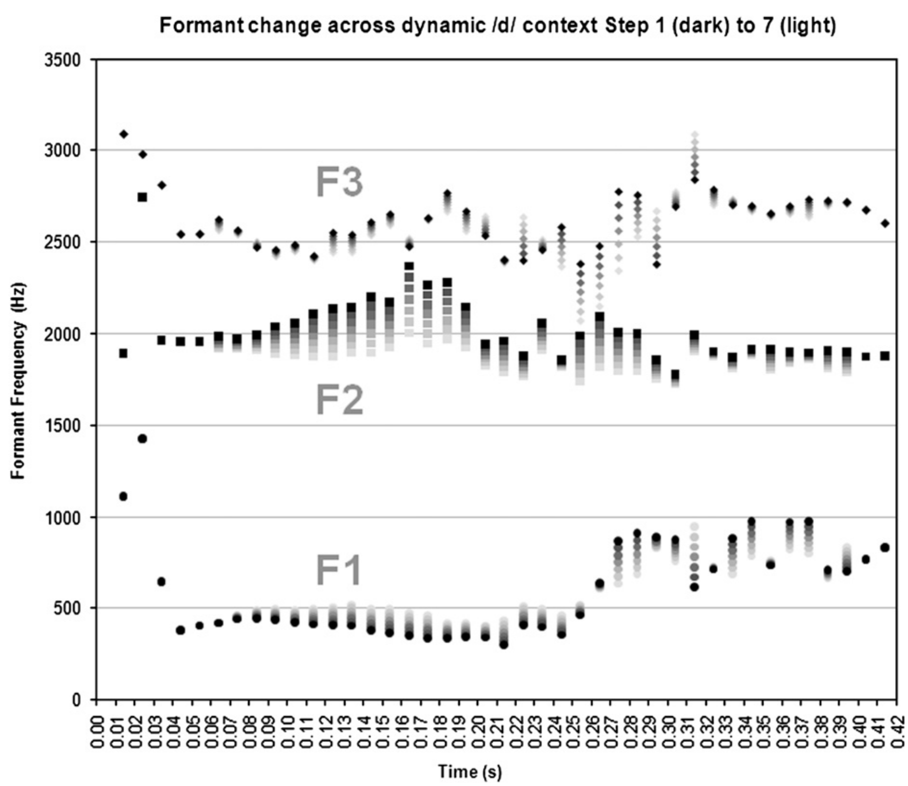

Fig. A2. Formant frequencies for *date*∼*debt* stimuli.

siblings is actually slightly higher than the median standard deviation for the Southerners for the /b/ context (of 1.075; but not the /d/ context: 1.086). Overall, we conclude from this experiment that there is no evidence here that the seven Southern siblings are more similar to one another than seven randomly selected Southerners in our dataset. It may be the case that these siblings are substantially contributing to the overall tendency towards more /e/ perception for Southerners than the other regions, but they do not appear to be doing so in a way that is different from the other Southern subjects who perceive more /e/.

Table B1
Word types included in production study.

Word list	Reading passage
BADE	big
BAIT	cake
BEAD	eat(ing)
BED	head(first)
BEG	Kate
BET	Peg (× 2)
BID	sit
BIT	teas(ing)
BOAT	
BOOED	
BOOT	
BUT	
DAD	
DATE (× 2)	
DEAF	
DEBT	
DEED	
DEEP	
DID (× 2)	
DOCK	
DOES	
DOZE	
DUDE	
DUTCH	
GATE	
GAVE	
PAD	
PEEK	
POD	
POKE	
POT	
SAD	
SOAP	
SOUP	
TAKE	
TUCK	
UP	

7. Conclusions

To close this investigation, it seems clear that speaker-hearers have a variety of linguistic and social resources surrounding and available to them. Those aspects that they attend to perceptually and productively are variable, even within the same community and within the same family, and the utilization of particular aspects of these resources does appear to reflect and be reflected in how perception is shaped. Our results indicate that, at least to some extent, listeners do process speech through the filter of their own production (not just their experience with speech around them). However, the speech around them clearly matters, as is indicated by the cross-regional differences found in our main perception data that indicate the South shows a significant effect on perceptual processing.

Such an analysis suggests that perception, like production, is a dynamic process with individuals attending to different aspects of experience. A similar sentiment was expressed by Evans and Iverson (2007) in their study of Northern and Southern British varieties discussed earlier:

> Given that these individual differences in production are likely due to sociolinguistic influences, it was surprising that they were directly linked to differences in perception. That is, even though spoken accent is a marker of social identity, perception is essentially private and it would be adaptive for individuals to be able to understand as wide a range of speakers as possible. Our subjects likely had very similar perceptual experiences when growing up, but the idiolectal variation in production had influences on both subjective ratings (i.e., best exemplar locations in experiment 2) and speech-in-noise recognition. It thus appears that the sociophonetically driven differences found in production had an impact on perceptual processing, suggesting that there is a strong perception-production link (2007: 3824).

In other words, both their work and ours suggests that how one's experience is categorized (and, crucially, relates to production) is not as much dependent on a listeners' passive exposure, but on their active engagement with the local system.

Similarly, returning to the findings by Hay, Warren, and Drager (2006) regarding the *near-square* merger in New Zealand English (NZE), not only did they find participants' own production affected by whether they were able to identify distinct vowels in word pairs, but listeners also seemed to integrate information about the speakers (such as age and class) in deciding whether pairs were merged or distinct. Perception was based both on what speakers did themselves (as we discovered in our results) and the social association they made with that variant. Crucially, their participants were exposed to both variants perceptually in daily interaction, but social information was processed alongside the acoustic information, predisposing participants towards one interpretation or another. Such a finding dovetails nicely with the suggestion that listeners encode social as well as linguistic representation of variants, as in a probabilistic model of representation along the lines of that proposed by exemplar theories (e.g., Hay, Warren, and Drager 2006; Johnson, 2006; Pierrehumbert, 2001, 2006). Thus, linguistic processing is then affected by these top-down and bottom-up factors, which results in the type of variable perception found above.

The variable use of SVS patterns evidenced above is consistent with recent findings suggesting a young retreat from Southern Vowel Shift patterns quite generally in the South (Fridland, 2000; Labov et al., 2006), particularly in urban areas. The greater advancement of SVS patterns in rural dialects and this retrenchment in use more generally by young urban speakers has been taken as evidence for the rural origins of the shift and the slowing expansion of SVS patterns in urban speakers may be due to this rural association (Irons, 2007). As rural migration into Southern urban centers increases, the shift advancement that came with it appears to have maintained an association with this rural identity. Support for such attitudes toward the shift comes from a "folk linguistic" study by Fridland and Bartlett (2007; see also Niedzielski and Preston, 2003) which asked Southerners to rate SVS shifted and non-shifted front vowel variants on scales of ruralness, education, and pleasantness and demonstrated that Southern shift norms are perceived as more rural than non-shifted variants. Specifically, SVS shifted front tense vowel variants were rated higher on ruralness scales and lower on education and pleasantness scales, suggesting that, in Memphis where the study was administered, they were viewed as non-urban and often relatively negatively compared to non-shifted variants.

Thus, the movement away from these rurally-associated norms among some of the youngest speakers in urban areas like Memphis may reflect a difference in these speakers' orientation to what it means to be Southern and the type of urban identity they want to project. However, as evidenced in the data above, not all younger speakers are retreating from all aspects of the Southern Vowel Shift[13] which would suggest a difference in how speakers with similar backgrounds process and respond to the variants and their associated context within their communities, much like that found in Hay, Warren, and Drager's work (2006) and we likewise

[13] Note that age does not arise as a significant predictor in any of the data presented in this paper. However, this is not surprising given the fact that almost all of our speakers were in younger age groups and so our data was not well distributed to examine age as a factor.

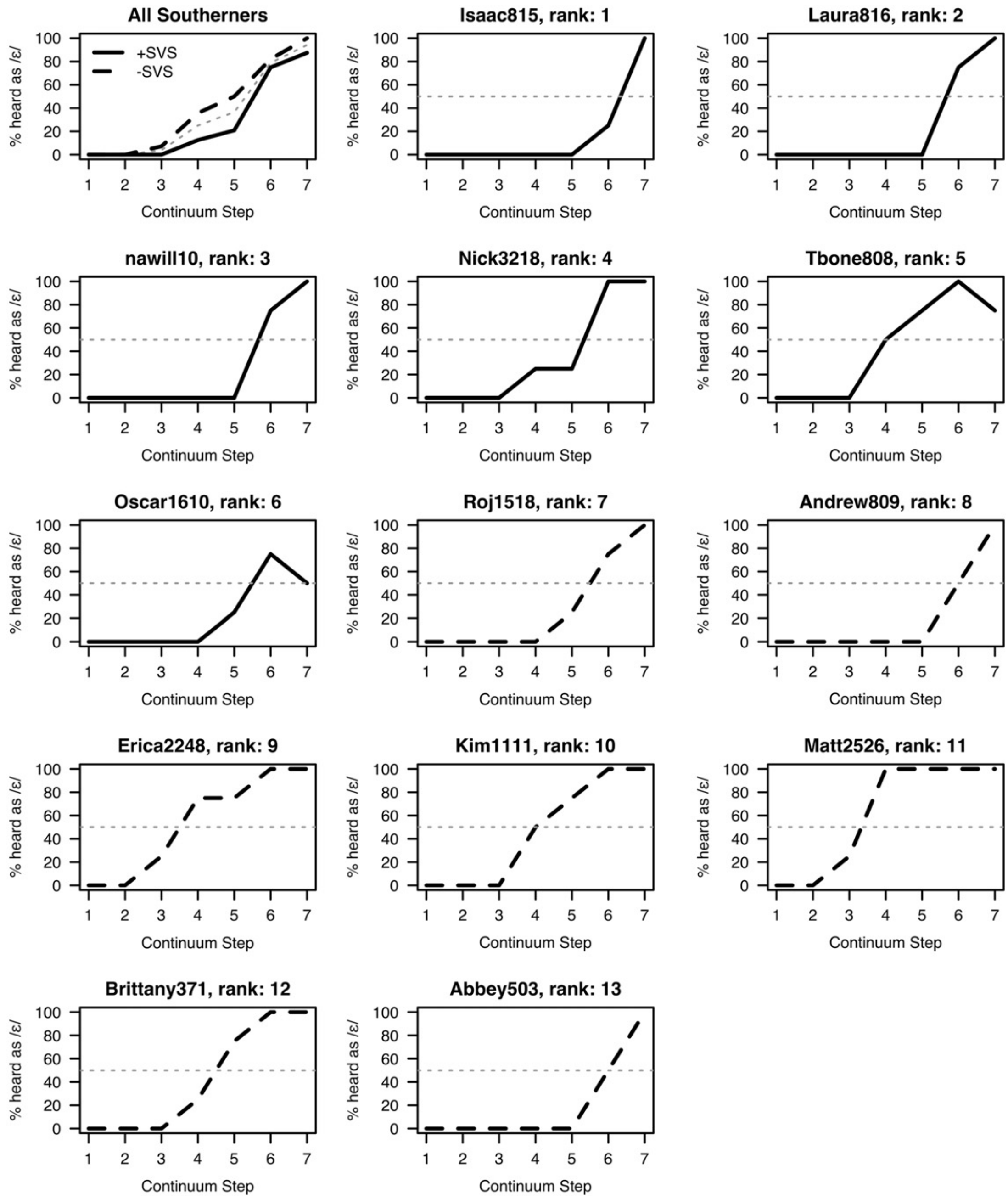

Fig. C1. Subsample *bait*∼*bet* perception results.

suggest our work points toward an exemplar-based model of speech processing.

This orientation toward an experience-based explanation of representation and processing is necessary if sociophonetic research is to move to the level of embracing and explaining the role of the social in phonetic and phonological theory (as discussed in Foulkes & Docherty, 2006) and is motivated by the clear involvement of social factors in speech processing. Traditional accounts of language processing and language change cannot account for the circularity of effects of the social on the cognitive and the cognitive on the social, or, in other words, that social identity fundamentally contributes to both how we talk and how we perceive talk. In order to explain how the social can have such an effect on the language system, some social categories must be internally encoded and interactive with linguistic categories in contrast to the external status currently afforded by traditional phonological theory. The variability of the signal is, after all, a consequence of both linguistic factors (e.g., context) and social factors (e.g., regional orientation) and we seem to be able to decode and interpret both types of information with surprisingly little difficulty. We conclude here by suggesting that our results most strongly point toward the need for more sociolinguistically informed work in speech perception and greater attention to the sociolinguistic variability of individual, not just community-based, norms as a factor in speech processing.

Acknowledgments

An earlier version of some of this work was printed as a working paper in the *Penn Working Papers on Linguistics* (Kendall & Fridland, 2010). Further analysis of these data is addressed in an

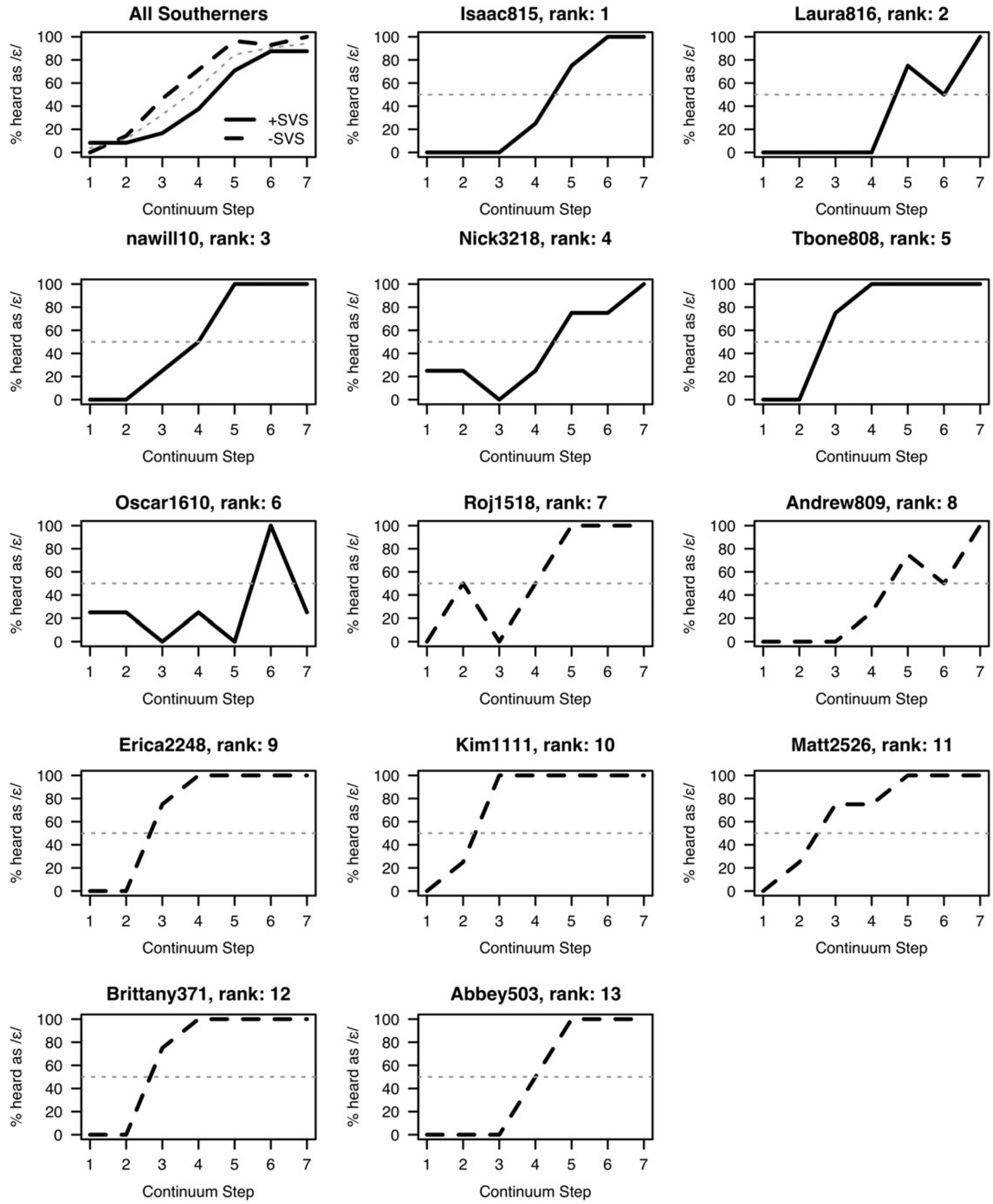

Fig. C2. Subsample *date*∼*debt* perception results.

investigation of cross-regional production-perception patterns in Fridland and Kendall (in press). This research has been supported by a grant from the National Science Foundation Linguistics Program BCS#0518264. We would also like to thank Kristen Link at SUNY-Oswego and Clare Dannenberg at Virginia Tech for their help with data collection and Sohei Okamoto for his help with the database. We thank the anonymous reviewers and editor for helpful comments on earlier versions of the paper.

Appendix A. Formant values and tracks at all steps along synthesized vowel continua from perception study stimuli

See Tables A1 and A2 and Figs. A1 and A2.

Appendix B. Word types included in production study

See Table B1.

Appendix C. Individual perception results for the 13 Southern subsample subjects, ranked by smallest to largest Euclidean distance between /e/ and /ɛ/

See Figs. C1 and C2.

References

Adank, P., & McQueen, J. M. (2007). The effect of an unfamiliar regional accent on spoken-word comprehension. In: J. Trouvain, & W. J. Barry (Eds.), *Proceedings*

of the XVIth international congress of phonetic sciences (pp. 1925–1928). Dudweiler: Pirrot.
Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
Baayen, R. H. (2010). *languageR: Data sets and functions with "Analyzing Linguistic Data"*. R package version 1.0. [Software: 〈http://CRAN.R-project.org/package=languageR〉].
Bates, D., & Maechler, M. (2010). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-37. [Software: 〈http://CRAN.R-project.org/package=lme4〉].
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. (2009). *Praat: Doing phonetics by computer*. [Software: 〈http://www.praat.org/〉].
Bohn, O.-S., & Munro, M. (Eds.). (2007). *Language experience in second language speech learning (in honor of James Flege)*. Amsterdam: John Benjamins.
Bradlow, A. R., & Bent, T. (2003). Listener adaptation to foreign accented English. In: M. J. Sole, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences* (pp. 2881–2884). Barcelona, Spain: Universitat Autonoma de Barcelona.
Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listeners perceptions. *American Speech*, *82*, 32–64.
Casserly, E., & Pisoni, D. (2010). Speech perception and production. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 629–647.
Clarke, S., Elms, F., & Amani, Y. (1995). The third dialect of English: Some Canadian evidence. *Language Variation and Change*, *7*, 209–228.
Clarke, C., & Garrett, M. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, *116*, 3647–3658.
Clopper, C., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech*, *51*, 175–198.
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., & De Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America*, *118*, 1661–1676.
Dahan, D., Drucker, S., & Scarborough, R. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition*, *108*, 710–718.
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*. Malden, MA/Oxford: Blackwell.
Elman, J., & McClelland, J. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically stored phonemes. *Journal of Memory and Language*, *27*, 143–165.
Evans, B. E. (2001). *Dialect contact and the Northern Cities Shift in Ypsilanti, Michigan*. Dissertation. Michigan State University.
Evans, B. G., & 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. G., & 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.
Fasold, R. (1972). *Tense marking in Black English*. Washington, DC: Center for Applied Linguistics.
Feagin, C. (1986). More evidence for vowel change in the South. In: D. Sankoff (Ed.), *Diversity and diachrony* (pp. 83–95). Amsterdam: John Benjamins.
Floccia, C., Groslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing?. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1276–1293.
Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*, 409–438.
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. (2003a). Network strength and the realization of the Southern Vowel Shift among African-Americans in Memphis, TN. *American Speech*, *78*, 3–30.
Fridland, V. (2003b). Tide, tied and tight: The expansion of /ai/ monopthongization in African-American and European-American speech in Memphis, TN. *Journal of Sociolinguistics*, *7*, 279–298.
Fridland, V. (2004). The spread of the cot/caught merger in the speech of Memphians: An ethnolinguistic marker? *Presented at LAVIS III (language and variation in the South III) conference*. Tuscaloosa, AL.
Fridland, V., & Bartlett, K. (2006). The social and linguistic conditioning of back vowel fronting across ethnic groups in Memphis, TN. *English Language and Linguistics*, *10*, 1–22.
Fridland, V., & Bartlett, K. (2007). Southern or rural? The social perception of intraregional vowel distinctions. *Southern Journal of Linguistics*, *31*, 38–62.
Fridland, V. & Kendall, T. Exploring the relationship between production and perception in the mid front vowels of U.S. English. *Lingua*, in press.
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.
Ganong, W. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 110–125.
Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*, 66–85.
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: J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 244–266). Malden, MA/Oxford: Blackwell.
Hall-Lew, L. (2009). *Ethnicity and variation in San Francisco English*. Dissertation. Stanford University.
Hay, J., & Drager, K. (2007). Sociophonetics. *Annual Review of Anthropology*, *36*, 89–103.
Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*, 865–892.
Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, *23*, 351–379.
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., Getty, L., Clark, M., & Wheeler, K. (1995). Acoustic characteristics of English vowels. *Journal of the Acoustical Society of America*, *97*, 3099–3111.
Irons, T. (2007). On the Southern Shift in Appalachian English. *Penn Working Papers in Linguistics*, *13*, 121–134.
Jannedy, S., & Hay, J. (2006). Modeling sociophonetic variation. *Journal of Phonetics (Introduction to a Special Issue)*, *34*, 1–4.
Janson, T. (1986). Sound change in perception: An experiment. In: J. Ohala, & J. Jaeger (Eds.), *Experimental phonology* (pp. 253–260). Orlando, FL: Academic Press.
Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In: K. Johnson, & J. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–166). San Diego: Academic Press.
Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, *34*, 485–499.
Johnson, K., Flemming, E., & Wright, R. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language*, *3*, 505–528.
Kazanina, N., Phillips, C., & Idsardi, W. (2006). The Influence of word meaning on the perception of speech sounds. *Proceedings from the National Academy of Sciences (PNAS)*, *103*, 11391–11396.
Kendall, T., & Fridland, V. (2010). Mapping production and perception in regional vowel shifts. *Penn Working Papers in Linguistics*, *16*, 2.
Kendall, T., & Thomas, E. R. (2009). *Vowels: Vowel manipulation, normalization, and plotting in R*. R package, version 1.0-2. [URL: 〈http:/cran.r-project.org/web/packages/vowels/〉].
Kerswill, P. (1996). Children, adolescents, and language change. *Language Variation and Change*, *8*, 177–202.
Kerswill, P., & Williams, A. (2000). Creating a New Town koine: Children and language change in Milton Keynes. *Language in Society*, *29*, 65–115.
Kent, R., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). Albany, Delmar: Thompson Learning.
Labov, W. (1991). The three dialects of English. In: P. Eckert (Ed.), *New ways of analyzing variation* (pp. 1–44). New York: Academic Press.
Labov, W. (1994). *Principles of linguistic change: Internal factors*. Malden, MA/Oxford: Blackwell.
Labov, W. (2001). *Principles of linguistic change: Social factors*. Malden, MA/Oxford: Blackwell.
Labov, W., & Ash, S. (1997). Understanding Birmingham. In: C. Bernstein, T. Nunnally, & R. Sabino (Eds.), *Language variety in the South revisited* (pp. 508–573). Tuscaloosa: University of Alabama Press.
Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English: Phonetics, phonology and sound change*. Berlin: De Gruyter.
Labov, W., Yeager, M., & Steiner, R. (1972). *A quantitative study of sound change in progress*. Philadelphia: U.S. Regional Survey.
Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, *68*, 1636–1642.
Luthin, H. (1987). The story of California (ow): The coming-of-age of English in California. In: K. M. Denning (Ed.), *Variation in language: NWAV-XV at Stanford (Proceedings of the fifteenth annual New Ways of Analyzing Variation conference)* (pp. 312–324). Stanford: Department of Linguistics, Stanford University.
Mason, H. M. (1946). Understandability of speech in noise as affected by region of origin of speaker and listener. *Speech Monographs*, *13*, 54–68.
McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, l–86.
Milroy, L. (1980). *Language and social networks* (1st ed.). Baltimore, MD: University Park Press.
Milroy, L. (1987). *Observing and analysing natural language*. Malden, MA/Oxford: Blackwell.
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*. Berlin: Mouton De Gruyter.
Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In: J. Bybee, & P. Hopper (Eds.), *Frequency effects and the emergence of lexical structure* (pp. 137–157). Amsterdam: John Benjamins.
Pierrehumbert, J. (2002). Word-specific phonetics. In: C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology*, Vol. 7 (pp. 101–139). Berlin: Mouton de Gruyter.
Pierrehumbert, J. (2006). The next toolkit. *Journal of Phonetics*, *34*, 516–530.

Pisoni, D., Nusbaum, H., Luce, P. A., & Slowiaczek, L. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, *4*, 75–95.

Pisoni, D., & Remez, R. (2005). *Handbook of speech perception*. Malden, MA/Oxford: Blackwell.

Plichta, B., & Preston, D. (2005). The /ay/s have it: The perception of /ay/ as a North–South stereotype in U.S. English. In: T. Kristiansen, N. Coupland, & P. Garrett (Eds.), *Special issue on subjective processes in language variation and change, Acta Linguistica Hafniensia*, *37*, 107–130.

R Development Core Team. (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. [URL: 〈http://www.R-project.org/〉].

Rickford, J. (1999). *African American Vernacular English*. Malden, MA/Oxford: Blackwell.

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: D. Gibbon (Ed.), *Natural language processing and speech technology* (pp. 14–26). Berlin: Mouton de Gruyter.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In: W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3–48). Baltimore: York Press.

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.

Tagliamonte, S., & D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation and language change. *Language*, *85*, 58–108.

Thomas, E. R. (1989). The implications of /o/ fronting in Wilmington, North Carolina. *American Speech*, *64*, 327–333.

Thomas, E. R. (1997a). A rural/metropolitan split in the speech of Texas Anglos. *Language Variation and Change*, *9*, 309–332.

Thomas, E. R. (1997b). *A compendium of vowel plots. A publication of the North Carolina Language and Life Project*. Raleigh, NC: Department of English, North Carolina State University.

Thomas, E. R. (2001). *An acoustic analysis of vowel variation in New World English*. Publication of the American Dialect Society (PADS), Vol. 85. Durham, NC: Duke University Press.

Thomas, E. R. (2002). Sociophonetic applications of speech perception experiments. *American Speech*, *77*, 115–147.

Thomas, E. R., & Kendall, T. (2007). *NORM: The vowel normalization and plotting suite*. [URL: 〈http://ncslaap.lib.ncsu.edu/tools/norm/〉].

Watt, D., Fabricius, A., & Kendall, T. (2011). More on vowels: Plotting and normalization. In: M. Di Paolo, & M. Yaeger-Dror (Eds.), *Sociophonetics: A student's guide* (pp. 107–118). London & New York: Routledge.

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. (1969). *A sociolinguistic description of Detroit Negro Speech*. Washington, DC: Center for Applied Linguistics.

Wolfram, W. (1991). *Dialects and American English*. Englewood Cliffs, NJ: Prentice Hall.