

University of Pennsylvania Working Papers in Linguistics

Volume 14, Issue 2

2008

Article 12

SELECTED PAPERS FROM NWA 36

The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas

Christian Koops*

Elizabeth Gentry†

Andrew Pantos‡

*Rice University

†Rice University

‡Rice University

Copyright ©2008 by the authors.
<http://repository.upenn.edu/pwpl>

The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas*

Christian Koops, Elizabeth Gentry, and Andrew Pantos

Abstract

In a study of native listeners' implicit knowledge of the social distribution of a phonological variable we found that the perceived distribution does not completely match the findings of production surveys of the same variable in the same community. This presents an interesting challenge to models of linguistic variation based exclusively on production data, and questions how the distribution of sociolinguistic variables is best defined.

The linguistic phenomenon we are investigating is the 'unmerger' of pre-nasal /i/ and /e/ among Anglo Houstonians. This long-standing feature of Southern American English is increasingly losing ground in large metropolitan centers of the South (e.g. Tillery and Bailey 2004).

We conducted a speech perception experiment to measure the degree to which Houstonians expect Anglo speakers of three different age groups to participate in the merger. The experiment was conducted using a head-mounted eye-tracker, which monitored the participants' eye movements as they selected on a computer screen the lexical items which they heard (Tanenhaus et al. 2000).

Of particular interest were participants' eye fixations on /eN/ and /iN/ competitors, i.e. words which are temporarily ambiguous with the target word in a merged production system. We interpret a greater amount of looks to the competitor as an indicator of the listener's readiness to assume that the speaker may have a merged system.

The results show that, as predicted, listeners are more likely to assume a merged system when listening to an "old" speaker than when listening to a "middle-aged" speaker. However, we find no significant difference between the perception of a "middle-aged" and a "young" speaker.

These results generally corroborate our recent production surveys in Houston (Gentry 2006; Pantos 2006), which show the merger to be positively correlated with age. However, there is a discrepancy between our participants' perception and the actual production of Houstonians in the "middle-aged" group. Although, in production terms, native Houstonians in this group pattern

*We would like to thank Katherine Crosswhite and Nancy Niedzielski for their comments on an earlier version of this paper. Special thanks go to Katherine Crosswhite for her eye-tracking and statistical expertise.

with "old" speakers in participating in the merger, listeners expect a "middle-aged" speaker to pattern with a "young" speaker in being less merged.

We suggest that this mismatch is due to recent demographic changes in Houston, with large numbers of non-Southern Anglo speakers moving into the Houston metropolitan area in the course of the Sunbelt migration (Thomas 1997; Klineberg 2006). Native listeners who are exposed to this linguistically mixed population apparently associate merged vowels specifically only with the oldest speakers, even though some younger speakers are also merged.

Our study adds to previous research on the perception of vowel merger, including near-merger (Di Paolo and Faber 1990, Labov et al. 1991, *inter alia*) by providing additional evidence for the role of perceived speaker dialect as a potentially disambiguating factor in speech perception (Hay et al. 2006). Our findings have broad implications for the study of language variation as they challenge sociolinguists to consider whether the variation we seek to describe resides in the "objective" production of speakers native to a community, or rather in the "subjective" experience of its native listeners.

The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas

Christian Koops, Elizabeth Gentry and Andrew Pantos*

1 Introduction

In this paper we concern ourselves with the perceptual correlates of sociophonetic variation. By this, we mean biases affecting speech perception that result from and systematically reflect production variation. Based on previous research, we can distinguish two types of such biases.

1.1 Effects of Listener Dialect

One well-known type of perceptual bias is the effect of a listener's production dialect on the way he or she perceptually categorizes acoustic phonetic information. It has been demonstrated that perceptual categories are partially a function of production categories, such that, for example, speakers of a dialect characterized by a vowel shift relative to another dialect tend to perceive vowel tokens in terms of their own "shifted" categories (Willis 1972). Perception, then, is partially informed by the latent assumption on the part of listeners that other speakers will tend to sound like they themselves do. Listener dialect effects are at the heart of much recent and ongoing research on cross-dialect perception and misperception (e.g. Labov and Ash 1997; Preston 2005).

1.2 Effects of Perceived Speaker Dialect

A second type of perceptual bias is what we will refer to as *perceived speaker dialect*. There has been a steady increase over the past ten years in sociophonetic studies demonstrating that speech perception is also partially informed by listeners' latent assumptions (often referred to as "stereotypes") about what different speakers sound like. When provided with explicit or implicit information about a speaker's social category membership (and, by extension, about their dialect), listeners adjust their perceptual categories to match those assumed to exist in the speaker. One method that has been used to reveal these subtle perceptual adjustments is the matched guise technique, borrowed from social psychology (see Ball and Giles 1988 for an overview). Participants in a speech perception experiment may, for example, be given a visual cue to the speaker's social category in the form of a picture or video clip as they perform a phonetic matching, identification, or discrimination task. Perceptual shifts are then induced by systematically varying the visual cue while holding the auditory stimulus constant, and taking account of any differences in participants' decisions on the task. Studies of this kind have shown that listeners have implicit awareness of sociophonetic variation along axes of variation traditionally studied in production surveys, including gender (Strand 1999; Johnson et al. 1999), national dialect (Niedzielski 1999), age (Drager 2006), and even social class (Hay, Warren and Drager 2006).

We believe that effects of perceived speaker dialect are of particular interest for sociolinguists, perhaps more so than effects of listener dialect. Variationist sociolinguists have always been concerned with the perceptual reality of correlations between linguistic and social variation, going back to Labov's (1966) "subjective reaction tests". Experimental studies of perceived speaker dialect like those reviewed above speak precisely to those aspects of variation which language users "pay attention to", whether they do so explicitly or implicitly. The results of such studies provide a window on the way in which members of particular speech communities themselves process, store and remember the variation which they experience.

From this perspective, the findings of Hay et al. (2006) are particularly interesting. The authors replicate in perception the complex trajectory of a change in progress in a particular community. Hay et al. found that New Zealand listeners' ability to distinguish between words

*We would like to thank Katherine Crosswhite and Nancy Niedzielski for their comments on an earlier version of this paper. Special thanks go to Katherine Crosswhite for her eye-tracking and statistical expertise.

containing two vowels involved in an ongoing merger, the diphthongs in the words *here* and *hair*, closely echoes the complex distribution of this merger in New Zealand society. For example, because young speakers lead the merger, and are thereby less likely to produce a reliable distinction, listeners in perception experiments have less “faith” in pronunciations attributed to younger speakers, as seen in greater error rates on discrimination tasks involving minimal pairs.

The experiment which we report here was designed to develop this line of inquiry. Our specific goal was to determine whether effects of perceived speaker dialect can also be found at a *local* level of variation. Note that the merger which Hay et al. were concerned with is a nationwide change in New Zealand English. Our experiment therefore tested the effect of perceived speaker age on speech perception in the context of a merger configuration unique to a city.

2 The PIN/PEN Merger in Houston

In dialects of English characterized by the PIN/PEN merger the contrast between the high and mid front lax vowels /i/ and /e/ is neutralized before the nasals /m/ and /n/ (the effect of the velar nasal is not completely predictable: see Thomas 2004 for discussion).¹ We will refer to the vowels affected by this merger as PIN and PEN vowels and to the words containing them as PIN and PEN words. The merged vowel quality is often said to be that of the word *pin* in a non-merged system (e.g. Wolfram and Schilling-Estes 1998:70).

The PIN/PEN merger is a hallmark of Anglo dialects in the Southern US. Labov, Ash, and Boberg (2006) show the area in which the merger is continuously present to be largely co-extensive with the area characterized by (some degree of) /ay/-monophthongization, which the authors use as the definitional criterion of the South as a linguistic area. They also found that, as a nationwide trend, the merger is in the process of expanding.

At the same time, the contrast is apparently being regained in some urban centers of the South. Thomas (2004:316) notes that “some Southerners, largely under the influence of schools, have begun to distinguish PIN and PEN,” and Tillery and Bailey (2004) report that while the merger “is still thriving throughout the rural South...in the largest Southern metropolises (areas such as Dallas and Atlanta) it is disappearing.” (331) That the “unmerging” of PIN and PEN should happen specifically in urban locales is not entirely unexpected. It is in line with a general loss of traditional Southern phonology in large cities, which has led, for example, to a dramatic rural-metropolitan split in the Anglo dialects of Texas (Thomas 1997). According to Thomas, what triggered this split was the massive post-war demographic shift known as the Sunbelt migration, which has brought large numbers non-Southerners into the Southern and Southwestern US in response to sustained economic growth in these areas of the country. The Sunbelt immigrants moved specifically into the suburbs of metro areas, and as their children failed to adopt traditional Southern vowel variants, the large cities of Texas, such as Dallas, Fort Worth, Houston and Austin, became “dialect islands” where a new, metropolitan dialect is spoken which contrasts markedly with the speech found in surrounding rural areas.

Our recent fieldwork in Houston clearly confirms this general trend (Gentry 2006; Pantos 2006). To the extent that traditional Southern features are present in Houston Anglos, they are found in older speakers, and even the speech of older Anglo Houstonians often shares more features with Western or Midland dialects than with the rural South. We also find evidence for the decline of the PIN/PEN merger, but this appears to be a more recent and yet incomplete development. While all but one of Gentry’s 18 interviewees above the age of 40 merge PIN and PEN, the majority (7 out of 12) of Pantos’ teenage respondents aged 15 to 19 did not show the merger. This age effect was carried largely by the female teenagers, suggesting that the “unmerger” is led by young females.

We also found an interesting phonetic effect in our acoustic analysis of the quality of the merged PIN and PEN tokens after pooling all tokens produced by female Houstonians who merge the vowels. Instead of the generally reported convergence on or near the vowel /i/, we found as much evidence for raising of PEN as for lowering of PIN. In F1/F2 space the merged tokens show an equal amount of overlap with /i/ and /e/, and their central tendency is exactly intermediate in

¹Throughout this paper we use Labov’s (1994) phoneme labels rather than IPA.

the F1 dimension, with a statistically significant difference relative to both /i/ and /e/. This confirms our auditory impression that the merged quality is often intermediate, and that both vowels seem about equally affected. Only in the F2 dimension are the merged tokens closer to /i/. There is no statistically significant difference relative to /i/, but significant fronting relative to /e/.²

Given these social and phonetic characteristics of the PIN/PEN merger in Houston, we designed an experiment to test native Houstonians' implicit knowledge of the social distribution and phonetic outcome of the merger.

Our main prediction was an effect of perceived speaker age, analogous but opposite to that found for the merger of the *here* and *hair* vowels in New Zealand. Houston listeners should display a neutralization bias, i.e., be more inclined to assume a merged system, when listening to older speakers. On the other hand, they should display a discrimination bias, i.e., rely more on the available acoustic information, when listening to younger speakers. We restricted the experiment to female speech because our production data showed a stronger age contrast for females.

Second, we predicted an effect of listener dialect. To the extent that listeners themselves merge the two vowels in production they should show a general neutralization bias in perception.

Third, we tested for an effect of vowel quality, i.e. whether listeners do in fact expect the vowels to converge closer to /i/. As discussed above, the evidence for such an asymmetry is limited in Houston. But, to the extent that the existing, weak asymmetry serves as a cue to listeners, they should display a neutralization bias when categorizing the (potentially ambiguous) PIN vowel and a discrimination bias for the PEN vowel.

3 Methodology

Participants performed a four-alternative forced choice word identification task while seated in front of a computer screen in a soundproof booth. At the beginning and throughout the duration of each trial a picture of the speaker (discussed below) was displayed in the center of the screen. After 1.5 seconds, four words appeared at the corners of the picture (see Figure 1).

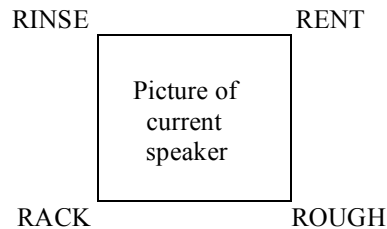


Figure 1: Visual display in the word identification task

Participants had 3.5 seconds time to familiarize themselves with the four words and their position. They then heard a cue to click into the center of the picture, which caused the sound file of the target word (e.g. “rinse”) to be played over a loudspeaker. The target word was spoken with a non-merged pronunciation. Participants were instructed to identify the word they heard by clicking on it as quickly as possible. If no click occurred within 2 seconds after the cue, the trial was aborted and the next trial started. There was a 1-second pause between trials.

The four alternatives in each trial included the target (e.g. *rinse*), a PIN or PEN competitor (e.g. *rent*) and two distractors (e.g. *rack* and *rough*). The target and the competitor were words which, in a merged production system, are phonologically identical up to and including the nasal consonant, and therefore temporarily ambiguous. The distractors were chosen to be phonologically identical to the target up to and excluding the PIN or PEN vowel. Their function was to divert

²We do not want to make too much of this finding in the absence of a complete set of the relevant acoustic measurements (such as F0 and duration), which are the subject of ongoing work. But it is tempting to interpret the lack of strong raising in PEN as part of the generally only mildly Southern dialect of Houston. It may be what Feagin (2003) describes as “not an assimilation to northern patterns but the development of a modified southern speech, a koiné of sorts, with some of the distinctiveness filtered out.” (139)

attention from the fact that words with pre-nasal /e/ and /i/ were at issue. In the following, we will refer to the time period from the onset of the PIN or PEN vowel to the end of the nasal consonant as the *ambiguity phase*.

The point of interest in this design was whether listeners considered the PIN or PEN competitor as a lexical candidate during the ambiguity phase. To the extent that their perception was affected by a discrimination bias (e.g. because the speaker was perceived as relatively young, and therefore less likely to merge PIN and PEN) listeners should discard the competitor from consideration as soon as the vowel was heard. To the extent that their perception was subject to a neutralization bias (e.g. because the speaker was perceived as older, and therefore more likely to be merged), they should continue to consider the competitor throughout the ambiguity phase.

To measure listeners' hypotheses during the ambiguity phase we used an eye-tracking paradigm. As the participants performed the word identification task, they wore a head-mounted EyeLink II eye-tracker which recorded their eye fixations to different screen coordinates at a sampling rate of 250 frames per second. Eye-tracking is today widely used as a measure of lexical activation during the early stages of word recognition. The underlying assumption of the eye-tracking paradigm is a linking hypothesis between the amount of visual attention directed at objects and the activation level of the corresponding concepts during speech perception and speech production (Tanenhaus et al. 2000). Thus, we used the average amount of time which listeners spent looking at the PIN or PEN competitor during the ambiguity phase as a measure of their degree of certainty that it might be the word they were hearing.

The experiment included 36 PIN/PEN trials, as described above, and 40 filler trials. Half of the 36 PIN/PEN trials had a PIN target and the other half had a PEN target. The order of trials and the screen positions of target, competitor, and distractors were randomized.

3.1 PIN/PEN Stimulus Words

Our target-competitor pairs were obtained semi-automatically by searching and matching phonological word forms in the CELEX 2 database (Kerkman et al. 1995). Each word pair had the same stress pattern, the same number of syllables, and, in most cases, also the same number of phonemes and the same number of letters in their orthographic form. We excluded words in which the PIN and PEN vowels were not spelled with the letters *i* and *e* (e.g. *symphony*, *friend*) and words with different initial consonant spellings (as in the pair *Cindy-sender*). Other examples of stimulus pairs are: *dinner-dentist*, *ginger-gender*, *limit-lemon*, *minute-mental*.

3.2 Stimulus Photos

To manipulate perceived speaker age we adopted Hay et al.'s (2006) matched guise design, but with an expanded age spectrum. The auditory stimuli were presented in the guise of three Anglo females of different ages by displaying to the participants one of three photographs (Figure 2).



Figure 2: The three different Anglo females varying by age paired with the auditory stimuli

In the following, we will refer to these pictures as the “young”, “middle-aged” and “old” picture. The pictures were chosen to cover approximately evenly spaced points on the age spectrum from late teen to senior citizen. In our selection of the photographs we attempted to avoid as much as possible visual cues to the speaker's social group membership besides age, especially indicators of social class. To this end, we created tight headshots with a minimal amount of visible background and clothing. We also chose pictures with similar facial expressions.

3.3 Stimulus Voices

Because no single voice can believably cover the entire age spectrum defined by our three stimulus photos, we used two voices, one “younger” and one “older” female voice. In other words, we divided the matched guise design into two separate comparisons, as described below.

Comparison 1		Comparison 2	
‘younger’ voice		‘older’ voice’	
“young” picture	“middle-aged” picture	“middle-aged” picture	“old” picture

Table 1: Separate comparisons in the matched guise design

We obtained word list readings of all stimuli and fillers from five female speakers of appropriate ages, and then identified those two who, in our judgment, were most likely to pass for speakers of either the two younger or two older age categories. Our “younger” speaker was a 32-year old native of Houston. Our “older” speaker was a 49-year old native of Northeastern Missouri. As both speakers are Southerners themselves, we took extra care to ensure non-merged yet natural pronunciations. For the target and competitor words each speaker provided several readings of word lists with separate columns labeled “PIN words” and “PEN words”, thus implicitly drawing their attention to the phonological similarity, but without specifically instructing them to produce distinct vowels. Each author then independently listened to and rated each recorded token. Only tokens which all three of us judged to be non-merged were included in the experiment.

3.4 Experimental Groups

Participants were randomly assigned to one of two voice-picture pairings (Table 2).

Group 1	Group 2
“younger” voice + “young” picture	“younger” voice + “middle-aged” picture
“older” voice + “middle-aged” picture	“older” voice + “old” picture

Table 2: Voice-picture pairings in the two experimental groups

As can be seen in the table above, participants in both groups heard the same voices, but each in different age guises. In addition, participants in each group were run on two separate lists, and the results of both lists were pooled in the statistical analysis. The two lists each contained half of the 36 PIN/PEN stimuli read by each speaker. The separation into two lists is standard practice in psycholinguistic experiments in order to avoid participants seeing the same visual display twice, which inevitably biases their visual attention. For example, their eye gaze may be led to the item that was chosen as the correct response on the first exposure, or to an item they did not choose if they assume that there is a balanced design.

3.5 Listener Production

Following the eye-tracking part of the experiment, and while the participants were still seated in the sound-proof booth, each participant performed three short production tasks and a self-perception task. Participants read a short reading passage with embedded PIN and PEN words, a word list containing PIN and PEN words and an equal number of fillers, and a series of minimal pairs, e.g. *tin* and *ten*, *pen* and *pin*, etc. They were then asked to specify for each of the minimal pairs whether they would, under normal circumstances, pronounce the pair of words “the same”, “close”, or “different”. Their speech during each of these tasks was recorded on a Marantz PMD670 digital recorder using a Shure SM10A head-mounted microphone.

Based on the results of these four tasks we calculated mergedness scores for each participant. For the passage, word list, and minimal pair list data, we each independently made binary auditory judgments of each produced PIN or PEN word token as either ‘merged’ or ‘non-merged’. From the pooled results we calculated the ratio of merged to non-merged tokens, yielding three

scores per participant. The self-perception data were quantified by assigning two points for each judgment of a pair as “the same”, one point for each judgment as “close”, and no point for judgments as “different”, and again calculating a ratio of assigned points out of possible points.

Participants’ average mergedness scores on the self-perception task were larger and more varied than the scores on any of the three reading tasks (mean 0.41, s.d. 0.38).

3.6 Participants

There were a total of 29 participants. The majority of the participants were undergraduate students at Rice University who took part in the experiment for class credit. We later solicited additional participants through class and email announcements and through advertisements on the Rice campus. These later participants were paid \$8 in cash for participation. They included more Rice undergraduates and several Rice administrative staff members. We collected demographic information including residential histories from all participants and analyzed only the data from those who had grown up in the Houston metropolitan area.

Group 1 (n = 15)	Group 2 (n = 14)
Mean age: 23.3 years (median 20) Gender: 11 female, 4 male Ethnicity: 7 Anglo, 3 South Asian, 2 Hispanic, 2 African American, 1 Asian	Mean age: 23.5 years (median 20) Gender: 9 female, 5 male Ethnicity: 7 Anglo, 2 Hispanic, 2 Asian, 1 South Asian, 1 African American, 1 Native American

Table 3: Demographics (age, gender, ethnicity, and size) for the two experimental groups

4 Results

We performed two separate analyses of covariance, one for each of the two comparisons. Our dependent variable was the proportion of time spent fixating the competitor word during the ambiguity phase, averaged over all completed trials per participant. We tested the effect of two independent variables: seeing an older picture and the participants’ mergedness score. Only the mergedness scores based on the self-perception task yielded significant results. Our first comparison (the “younger” voice paired with either the “young” or the “middle-aged” picture) showed no significant effect of perceived age ($F(1,26)=1.01, p=0.325$), but a significant effect of mergedness ($F(1,26)=7.47, p=0.011$). Participants with higher mergedness scores spent more time looking at the competitor. Our second comparison (the ‘older’ voice paired with either the “middle-aged” or the “old” picture) showed a significant effect of perceived age ($F(1,26)=7.07, p=0.013$). Participants spent more time looking at the competitor when they saw the “old” picture. We also found a significant effect of participant mergedness ($F(1,26)=5.74, p=0.024$), in the same direction as for the first comparison.

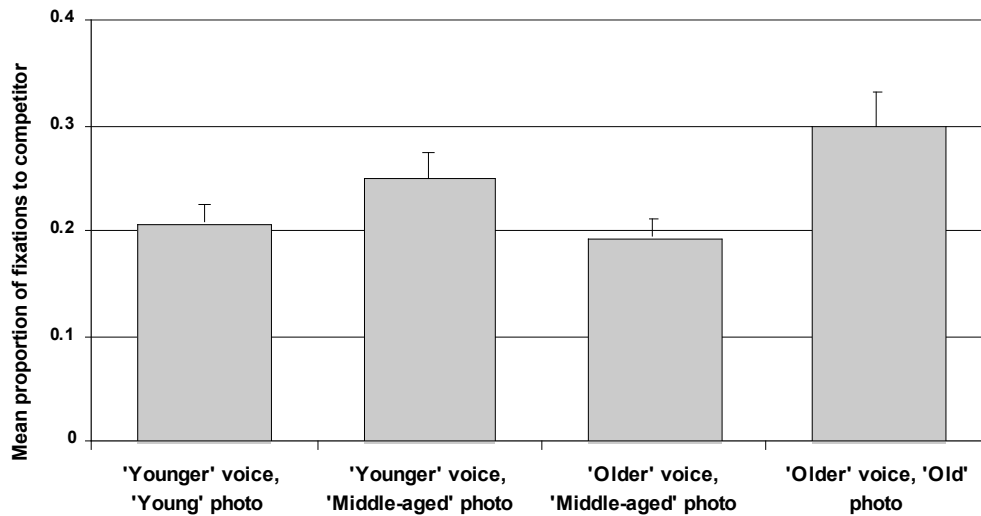


Figure 3: Proportion of fixations to the competitor word for each voice-picture pairing

We then compared the proportion of time spent fixating the competitor word depending on whether the target was a PIN word or a PEN word for all completed trials by all speakers. Although the average proportion was slightly higher for PIN trials than for PEN trials, the difference is not significant (one-tailed t -test, $p=0.143$, see Figure 3).

5 Discussion

Starting with the consistent effect of participant mergedness, it is not surprising that the presence of the merger in the participants' own production should lead to a neutralization bias, given the known effects of listener dialect on perception (see also Conrey, Potts and Niedzielski 2005). The effect is also consistent with Hay et al.'s (2006) findings on the perception of the *here* and *hair* diphthongs in New Zealand, which included an effect of listener mergedness in addition to various perceived speaker effects. What is more interesting is that, again as Hay et al.'s experiment, even though our participants showed the merger to varying degrees themselves, this did not lead to a complete collapse of their discriminatory ability, as seen in the fact that we did find an effect of perceived speaker age, which presupposes the selective activation of the PIN and PEN vowels.

Next, the lack of an effect of vowel quality shows that our participants did not expect the merged pronunciation to approximate /i/ more than /e/. We take the fact that vowel quality does not serve as a cue to listeners as perceptual evidence for our tentative claim in Section 2 that the phonetic result of the PIN/PEN merger in Houston is not that of unilateral raising of PEN, but that both vowels are about equally affected. It would be interesting to test whether this pattern is restricted to Houston (or, perhaps, large Southern cities in which a metropolitan Southern dialect is spoken). As far as we are aware, there have been no previous proposals regarding regional variation in the phonetic details of the PIN/PEN merger.

Finally, the effect of perceived speaker age is more complex, but nonetheless consistent with our overall hypothesis. The age effect we found points in the predicted direction. Seeing an older picture induced a neutralization bias in our participants. However, we found this effect only in the comparison of the "middle-aged" and "old" speaker guise. This does not completely reflect our earlier production results (cf. Section 2). Recall that our production surveys showed native Houstonians above the age of 40 to be merged almost without exception. Thus, in production, the middle-aged Houstonians pattern with the older Houstonians rather than with the younger ones. Our participants, however, did not react differently to the "young" and "middle-aged" guises, and instead displayed the predicted effect with regard to the "middle-aged" and the "old" guise. How do we explain this production-perception mismatch?

We suggest that the age effect we found is indeed based in reality, i.e. in the accumulated experience of our participants with the speech of Anglo females in Houston. What is at odds with reality, rather, is the assumption that the speech which our participants are exposed to in the Houston metropolitan area is, in all relevant respects, that of native Houstonians. It is not unlikely that the group of middle-aged Anglo females in Houston contains a large enough proportion of speakers who moved to the city from non-merging regions to match the rate of younger non-merging females. That would explain why our participants did not treat these two groups differently. It is known that Houston has in the past attracted large numbers of Anglos (as well as other ethnicities) who moved to the city, for example, to take advantage of job opportunities in the oil and gas industry, and that the city continues to attract large numbers of immigrants. As discussed in Section 2, it was the Sunbelt immigrants who, according to Thomas (1997), caused the decline of Southern speech not only in Houston, but in large Texas cities overall. It is not implausible, then, that as a result of this demographic trend, the concentration of native Houstonians in the working age group is low enough to have caused the effect we observed. Moreover, because the Sunbelt migration is a relatively recent event (Tillery and Bailey 2004 date the onset of it to the 1970s), the proportion of non-Houstonians is likely to be still lower in the oldest age groups. This would explain why our participants associated the merged pronunciation only with the oldest age guise. Although we do not have sufficient demographic information to clarify this issue at present, we suggest that the demographic makeup of Houston is the most likely cause of the production-perception mismatch.

What may have further contributed to nature of the observed age effect is a potential problem with our experimental design, which we recognized in retrospect. Our participants were not explicitly told that the speakers they were listening to were Houstonians. Many of the participants may have inferred that the dialect of Houston was at issue from the fact that they were recruited to take part in the experiment specifically because they were from Houston. But the label “Houston” was not mentioned during the experiment. In effect, then, we tested a stronger hypothesis than intended, viz. that Houston listeners would show the predicted biases when listening to any Anglo female. As a result of not having been primed with the concept ‘Houston’, the participants may have perceptually averaged over the experienced speech of a larger set of women than they would have otherwise, including a higher number of non-merging speakers. For example, memories of middle-aged speakers activated by our pictures may have included the speech of recently heard female professors at Rice, a group including a negligible number of native Houstonians.

6 Conclusions

In conclusion, we have found evidence for both types of perceptual correlates of production variation discussed at the outset of this paper. Specifically, we have shown that effects of perceived speaker dialect are also present at a local level of variation, such as the dialect specific to a city. At the same time, we found that socio-perceptual experiments aiming at micro-level variation can show unexpected results, such as apparent production-perception mismatches, in situations where it is not completely clear what assumptions listeners bring to the task. Listeners can only be expected to make detailed, sociophonetically informed judgments to the extent that they possess detailed expectations about what the speech they hear represents in social terms.

References

- Ball, Peter, and Howard Giles. 1988. Speech style and employment selection: the Matched Guise Technique. In *Doing social psychology. Laboratory and field exercises*, ed. G. M. Breakwell, H. Foot and R. Gilmour, 121–149. Cambridge: Cambridge University Press.
- Conrey, Brianna, Geoffrey Potts, and Nancy Niedzielski. 2005. Effects of dialect on merger perception: ERP and behavioral correlates. *Brain and Language* 95:435–449.
- Drager, Katie. 2006. From bad to bed: the relationship between perceived age and vowel perception in New Zealand English. *Te Reo* 48:55–68.
- Feagin, Crawford. 2003. Vowel shifting in the southern states. In *English in the Southern United States*, ed. S. Nagle and S. Sanders, 126–140. Cambridge: Cambridge University Press.

- Gentry, Elizabeth. 2006. Hovering between the South and the West: Houston's merged dialect. Paper presented at NAWAV 35.
- Hay, Jennifer, Paul Warren, and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34:458–484.
- Johnson, Keith, Elizabeth Strand, and Mariapaola D'Imperio. 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27:359–384.
- Kerkman, H., R. Piepenbrock, R. H. Baayen and H. van Rijn. 1995. *The CELEX lexical database*. Philadelphia, PA: Linguistic Data Consortium.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1994. *Principles of linguistic change*. Vol. 1: *Internal factors*. Oxford: Blackwell.
- Labov, William, and Sharon Ash. 1997. Understanding Birmingham. In *Language variety in the South revisited*, ed. C. Bernstein, T. Nunnally and R. Sabino, 508–573. Tuscaloosa: University of Alabama Press.
- Labov, William, Sharon Ash, and Charles Boberg. 2006. *The atlas of North American English*. Berlin, New York: Mouton de Gruyter.
- Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18:62–85.
- Pantos, Andrew. 2006. Redefining the South: Teenage Houstonians and the Southern shift. Paper presented at NAWAV 35.
- Preston, Dennis. 2005. Belle's body just caught the fit gnat: The perception of Northern Cities shifted vowels by local speakers. *University of Pennsylvania Working Papers in Linguistics* 22:133–146.
- Strand, Elizabeth. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18:86–99.
- Tanenhaus, Michael, James Magnuson, Delphine Dahan, and Craig Chambers. 2000. Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research* 29:557–580.
- Thomas, Erik. 1997. A rural/metropolitan split in the speech of Texas Anglos. *Language Variation and Change* 9:309–332.
- Thomas, Erik. 2004. Rural Southern white accents. In *A handbook of varieties of English*. Vol. 1: *Phonology*, ed. E. Schneider, K. Burridge, B. Kortmann, R. Mesthrie and C. Upton, 300–324. Berlin, New York: Mouton de Gruyter.
- Tillery, Jan, and Guy Bailey. 2004. The urban South: phonology. In *A handbook of varieties of English*. Vol. 1: *Phonology*, ed. E. Schneider, K. Burridge, B. Kortmann, R. Mesthrie and C. Upton, 325–337. Berlin, New York: Mouton de Gruyter.
- Willis, Clodius. 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, Walt, and Natalie Schilling-Estes. 1998. *American English: Dialects and variation*. Malden, MA: Blackwell.

Department of Linguistics
 Rice University
 PO Box 1892
 Houston, TX 77251-1892
ckoops@rice.edu
eagentry@rice.edu
ajpantos@rice.edu