

Article

Production and perception of the PIN-PEN merger

Martha Austen

Department of Linguistics, The Ohio State University, 300 Oxley Hall, 1712 Neil Avenue, Columbus, OH 43210

Abstract

This study presents the first US-wide survey of the PIN-PEN merger since Labov, Ash & Boberg (2006). Production and perception data were collected from 277 speakers from across the country, with perception-only data from an additional 94 speakers; these data largely replicated previous findings about the social and geographic distribution of the merger. An examination of production and perception data together showed that near merger—in which speakers cannot *hear* the difference between PIN and PEN words, yet pronounce them differently—was relatively common, although this phenomenon has received little attention in the literature on the merger. Additionally, an investigation of how merged speakers phonetically realized their merged PIN-PEN vowel revealed that, in contrast to previous findings, speakers were equally as likely to merge to [ɛ] (*tw[ɛ]n* for *twin*) as they were to [ɪ] (*h[ɪ]n* for *hen*). However, there was no apparent social or geographic patterning to this phonetic realization.

Keywords: PIN-PEN merger; near merger; Southern English; African American English; phonetic realization

(Received 20 September 2018; revised 15 May 2019; accepted 17 September 2019)

1. Introduction

Despite being one of the most widely studied features of Southern English, the PIN-PEN merger—the merger of /ɪ/ and /ɛ/ before nasals—has rarely been examined or mapped on a broad geographical scale; most studies have focused on the merger within particular communities or particular regions. A notable exception to this is the Atlas of North American English (Labov, Ash & Boberg, 2006), which reports that although the merger predominantly occurs in the US South, it is also present in other parts of the country (as has also been reported by a number of previous studies), and that the merger appears to be expanding in apparent time. This study updates and expands on the results of the ANAE by closely examining the phonetic realization of the merger. Additionally, like the ANAE, this study examines production and perception together, focusing particularly on the incidence of near merger.

2. Previous research

2.1 Regional and social distribution

The PIN-PEN merger has been widely documented in the dialectological and sociolinguistic literature, though as noted above most studies have focused only on a particular community or region. Generally, the merger is associated with the US South and with African American speakers, and to a lesser extent with rural areas and lower class/less educated speakers. The merger is widespread throughout the South, with the exception of New Orleans (Labov et al., 2006). Outside of the South, merged white speakers have been reported in Detroit (Edwards, 1997), southern Illinois (Bigham, 2005), southern Indiana (Labov et al., 2006), Kansas (Labov et al., 2006), Oklahoma

(Bakos, 2013), and parts of California (Geenberg, 2014; Labov et al., 2006; Podesva et al., 2015; Warren & Fulop, 2014).

Labov et al. (2006) report that Black speakers are more consistently merged than white speakers both in Southern cities and outside of the South. Likewise, in the Linguistic Atlas of the Gulf States (Pederson, McDaniel & Adams, 1986), Black speakers are more likely to be merged than white speakers.¹ However, Black speakers are not universally merged, and are less likely to be merged in some cities than in others: Coggshall and Becker (2009) report a lack of merger in New York City (but see Labov (1968)), whereas in Pittsburgh (Eberhardt, 2009) and Detroit (Edwards, 1997) African American speakers tend to be merged. It is unclear whether there are larger regional patterns to be observed within African American English—for instance, if Black speakers are more likely to be merged in the Midwest than in the Northeast—or whether this variation is city-specific.

In some communities, the merger correlates with lower social status and/or lower levels of education. At a national level, less educated speakers are more likely to be merged (Labov et al., 2006), and in some individual communities like Charleston, South Carolina, lower class speakers are more likely to be merged (Baranowski, 2013). In other places, there is no apparent social stratification and speakers are equally likely to be merged across different levels of education, such as in West Virginia (Hazen, 2005). Within the same city, social stratification of the merger may differ across racial/ethnic groups: Edwards (1997) reports that in Detroit, both middle class and working class Black speakers are merged, but only working class white speakers are merged. In contrast to the typical association between the merger and lower social status or education, in the Linguistic Atlas of the Gulf States (Pederson et al., 1986), *more* educated speakers are more likely to be merged, and middle and upper-middle class speakers are more likely to be merged than lower class speakers (though upper class speakers are the least likely to be merged). This contrast to the pattern reported elsewhere could be attributable to the older nature of the Atlas data. The merger may have once been a prestige feature in the South (cf. Brown, 1991), but has subsequently lost its prestige.

Author for correspondence: Martha Austen, Email: austen.14@osu.edu

Cite this article: Austen M. (2020) Production and perception of the PIN-PEN merger. *Journal of Linguistic Geography* 8: 115–126, <https://doi.org/10.1017/jlg.2020.9>

© Cambridge University Press 2021.

The relationship between the PIN-PEN merger and gender and age is unclear. For gender, some studies find that men are more likely to be merged than women (across the United States: Labov et al., 2006), others that women are more likely to be merged than men (in West Virginia: Hazen, 2005; in the South generally: Pederson et al., 1986), and others no gender differences at all (among white speakers in Charleston, South Carolina: Baranowski, 2013; among African Americans in Detroit: Edwards, 1997).

Whether the PIN-PEN merger represents a change in progress is similarly inconsistent across communities. At a national level, Labov et al. (2006) find that the merger is expanding in apparent time; this also seems to be true in West Virginia (Hazen, 2005) and among white speakers in Charleston, South Carolina (Baranowski, 2013). In contrast, in large Southern cities such as Atlanta, Dallas, and Houston, the merger is reported to be reversing itself, with younger speakers less likely to be merged (Koops, Gentry & Pantos 2008; Tillery & Bailey, 2004). Taken in combination, these results suggest that the merger may be expanding geographically to areas that were historically non-merged—such as West Virginia and Charleston, South Carolina—and may especially be expanding in rural areas, potentially moving towards becoming a marker of rural versus urban speech rather than Southern versus non-Southern speech.

2.2 Phonetic realization

Where phonetic descriptions are given, the PIN-PEN merger is typically described as a merger towards [ɪ], such that *pen* sounds like *pin* to non-merged listeners (e.g., Bailey & Thomas, 1998:101; Bakos, 2013; Brown, 1991; Edwards, 1997; Hazen & Fluharty, 2004; Wise, 1933; Wolfram & Schilling, 1998). However, some descriptions also mention speakers who merge towards [ɛ], such that *pin* sounds like *pen* to non-merged listeners (Bigham, 2005; Ito & Campbell-Kibler, 2011; Labov, 1968; Pederson, 1965:41–42; Thomas, 2004). Still, other studies describe a merged vowel that is intermediate between [ɪ] and [ɛ] (Koops et al., 2008; Baranowski, 2013). From these studies, it is clear that a variety of realizations of the merger exist (see Bigham (2005) for a thorough review of the literature). However, the relative frequency of these different pronunciations is unclear: Thomas (2004), discussing rural white Southern speech, states that a majority of speakers merge towards [ɪ], with fewer speakers merging towards [ɛ], whereas in southern Illinois, Bigham (2005) finds that a majority of his participants merge towards [ɛ] rather than towards [ɪ]. In the Linguistic Atlas of the Gulf States (Pederson et al., 1986), 36% of participants merge towards [ɪ] (across seven PEN words), whereas only 17% merge towards [ɛ] (in the item *pin*).

It is unclear whether these different descriptions of the merger's phonetic realization can be attributed to the different populations studied: to date, no study has examined the extent to which this variation in the merger's phonetic realization is constrained by geographic or social factors. Comparing across studies, there is no obvious geographic, social, or temporal pattern as to which studies report [ɪ] versus [ɛ]. For studies which report multiple realizations within a particular community, there are no clear social differences between speakers who merge towards [ɪ] versus those who merge towards [ɛ].

2.3 Production versus perception

Most previous work on the PIN-PEN merger has focused on the merger in production, i.e., whether speakers pronounce PIN and PEN words differently. However, merger may also be assessed in terms of perception, i.e., whether speakers can hear the

difference between PIN and PEN words. Merger in perception need not always go hand in hand with merger in production: in a “near merger,” speakers cannot *hear* the difference between two sounds, yet still pronounce them at least subtly differently (Labov, 1994:ch.12). Near mergers have been attested for a number of mergers (e.g., Bowie, 2000; Di Paolo, 1992; Labov, 1994). The converse situation, a merger in production but not in perception—i.e., producing two sounds the same, but being able to perceive the difference between them—seems to be less common or nonexistent. Labov et al. (2006:29) report some speakers who appear to be merged in production but not perception, but comment that this is likely the result of participants thinking that words are pronounced differently because they are spelled differently. Other studies have found speakers who are merged in production, but who can distinguish between the two sounds at higher-than-chance levels when produced by non-merged speakers (Hay, Drager & Thomas, 2013; Thomas & Hay, 2005).

Despite the widespread attestation of near merger for other mergers, little work on the PIN-PEN merger has examined production and perception together. An exception is Labov et al. (2006), which collected both production (elicitation of *pin* and *pen*) and perception (asking participants whether they judged *pin* and *pen* to be the same or different) data. Labov et al. found that fifty participants judged *pin* and *pen* to sound the same, but produced them differently—i.e., had a near merger—as compared to 183 who were merged in both production and perception, and only eight who judged *pin* and *pen* to sound different but pronounced them the same. Labov's earlier work on African American speakers in New York City (Labov, Cohen, Robins & Lewis, 1968:119–120) also found a similar pattern, where “quite a few” participants produced PIN and PEN differently but perceived them as being the same. In both of these works, what demographic characteristics might make a speaker more likely to be near merged are not examined. In looking at the map in the atlas, there is no obvious geographic pattern as to where these speakers are located—they are found in roughly equal numbers both inside and outside of the isogloss for the merger and are not confined to any particular region. Thus, while this work demonstrates that near merger for PIN-PEN exists, it is still unclear whether there is any geographical or social pattern as to which speakers are more likely to be near merged.

The results of other work examining production and perception together have been contradictory: in Charleston, South Carolina, Baranowski (2013) found that speakers were more likely to be merged in production than in perception, where perception was assessed by asking speakers to judge whether pairs of words sounded the same during a minimal pair wordlist reading task. Baranowski speculates that these results are due to the merger being above the level of awareness in Charleston: speakers might be more likely to judge two words as different because they know they are “supposed” to be different (under a standard language ideology), regardless of whether they can actually hear the difference between the two words. Thus this task may not be measuring speech perception per se but rather speakers' ideologies about the merger and about their own speech. Potentially the inverse ideological stance can be found in southern Maryland (Bowie, 2000), where some participants reported that they merged PIN and PEN but in fact did not. These studies, then, suggest that speakers' reports of their own production and perception may not always reflect their actual linguistic behavior.

3. Research questions

As outlined above, no study other than Labov et al. (2006) has examined the PIN-PEN merger on a US-wide scale, and no study has examined the merger's phonetic realization on a US-wide scale, despite the differing descriptions of the realization across (and sometimes within) studies. Additionally, while previous studies have shown that some speakers have mismatches between production and perception of the merger, whether there is any geographical or social patterning to these mismatches remains unclear. Thus, this study asks four research questions:

1. What is the geographic and social distribution of the merger *in production*?
2. What is the geographic and social distribution of the merger *in perception*?
3. What is the relationship between production and perception of the merger, and is this relationship geographically/socially conditioned?
4. Among speakers who are merged in production, is there any geographic or social conditioning to the phonetic realization of the merged vowel (towards [ɪ] or towards [ɛ])?

4. Methods

4.1 Participants

Participants from across the United States were recruited via Amazon Mechanical Turk, an online platform where people do tasks for pay over the Internet. In the past several years, Mechanical Turk and similar platforms have been used for a variety of linguistic research (e.g., Degen & Goodman, 2014; Kanwal, Smith, Culbertson & Kirby, 2017; Kim, Wyszogrod, Reddy & Stanford, 2016; Tamminga, 2017), and have been shown to produce results similar to those of laboratory studies (Berinsky, Huber & Lenz, 2012; Enochson & Culbertson, 2015; Tamminga, 2017; Wang, Huang, Yao & Chan, 2015). Participation was limited to users in the United States, and data from participants who reported having grown up outside of the United States or not speaking English as a native language were discarded ($n=12$). Data from an additional 9 participants were discarded for failing attention checks in the perception task (described below). Participants were paid \$2.50 each.

Production and perception data were analyzed from 277 participants, with perception-only data from an additional 94 participants from whom production data could not be collected due to technical difficulties. Participant hometowns were fairly evenly scattered across the US, as shown in Figure 1a. Of the 277 participants with both production and perception data, 48% were female and 52% were male. The median age of these participants was 34 (Figure 1b). As compared to the US population as a whole, white, Asian, and mixed race speakers were overrepresented (77% white participants compared to 64% of the US population; 6% Asian as compared to US 5%; 4% mixed race as compared to US 3%), and Black and Latino speakers were underrepresented (8% Black as compared to US 12%; 5% Latino as compared to US 16%) (Figure 1c). On average, participants were more educated than the general US population (47% of participants had achieved a Bachelor's degree or higher, as compared to 31% of the US population) (Figure 1d). Demographics for the perception-only participants were similar (median age 36; 73% white, 7% Black, 10% Latino, 7% Asian, 2%

Table 1. Words read by participants in production task.

PEN	PIN	BET	BIT	fillers	
pen	pin	pet	pit	cat	pal
meant	mint	dead	did	bat	robin
Ben	bin	bet	bit	dog	cow
ten	tin	head	hid	talking	frog
cents	since	set	sit	task	fog
Jen	gin	bed	bid	water	had
Ken	kin	Keds	kit	monster	carrot
				ham	

mixed race; 46% with a Bachelor's degree or higher), though more heavily male (37% female, 63% male).

4.2 Procedure

At the beginning of the experiment, participants completed a matched guise task involving PIN-PEN words realized both with [ɪ] and [ɛ]. The structure of this task and its results will be described in future work by this author (see also Austen, 2017) and will not be discussed here, but are mentioned because the PIN-PEN words in the task may have prompted participants to be more aware of PIN-PEN words in the subsequent tasks described below. In the task, participants heard talkers read sentences aloud and were then asked to rate the talkers along a series of Likert scales (e.g., How educated does this person sound?). Each target sentence contained one PIN-PEN word. Participants heard an equal number of merged and non-merged pronunciations. Of the merged pronunciations, participants heard an equal number of mergers towards [ɪ] and towards [ɛ].

4.2.1 Production: wordlist

Following the matched guise task, participants recorded themselves reading a wordlist presented in random order, using either their computer's microphone or an external microphone. The wordlist contained 7 pairs of PIN-PEN words and 15 fillers (listed in Table 1); in addition, there were 7 pairs of BIT-BET words—non-pre-nasal /ɪ/ and /ɛ/—which should be unaffected by the merger. For merged speakers, these BIT-BET productions were used to determine which vowel the speaker merged towards. Due to incompatibilities between the recording script and certain browsers, as well as some participants lacking microphones, production data were available for only 277 of the participants (74%).

It should be noted that participants' use of their own microphones is a limitation of this study. As an anonymous reviewer points out, different recording devices can yield different acoustic measurements (De Decker, 2015; De Decker & Nycz, 2011), which could lead to participants using certain microphones seeming more or less merged than they would have with another microphone. Unfortunately, participants were not asked what type of microphone they were using, meaning that microphone type cannot be included as a factor in the analysis. However, as we shall see in the results section, the findings about merger in production largely mirror those of previous studies. This suggests that, even with the noise introduced by different recording devices, the data collected in this study are still reliable.

4.2.2 Perception: word identification

To test for merger in perception, participants completed a forced-choice word identification task: they heard a word (e.g., "pen") and

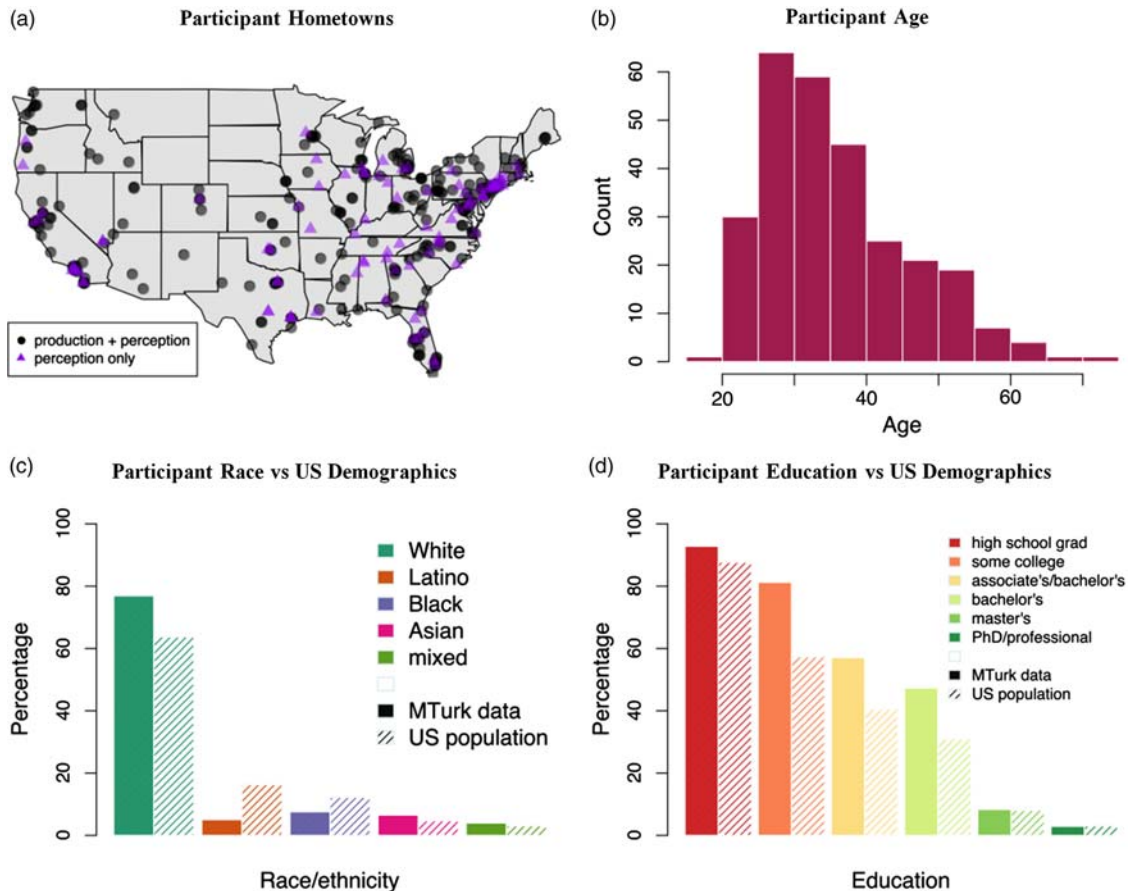


Figure 1. Demographic characteristics of participants.

were asked to click on the word they heard (e.g., “pen” or “pin”). Stimuli were produced by a standard-sounding talker in her mid-20s. Listeners heard two productions each of four pairs of PIN-PEN words (*pin/pen*, *bin/Ben*, *tin/ten*, and *mint/meant*), plus two productions each of three pairs of fillers (*dog/dock*, *cat/bat*, and *bit/bet*). Stimuli were presented in random order. To ensure that participants were paying attention and could distinguish between [i] and [ɛ] in non-pre-nasal contexts, data from participants who misperceived two or more of the *bit/bet* or *cat/bat* filler stimuli were discarded (n=9).²

4.2.3 Rhyming

An additional manner in which speakers may be merged is in rhyming: merged participants should say that pairs such as “Jen” and “kin” rhyme, whereas non-merged participants should not. Participants were shown pairs of words and asked whether they rhymed. Participants gave judgments on four pairs of PIN-PEN words, along with eleven fillers. To check that participants were not unduly influenced by orthography, the fillers included pairs that rhymed but differed orthographically (e.g., “weigh” and “pay”), or matched orthographically but did not rhyme (e.g., “cow” and “tow”).

4.3 Demographic questionnaire

At the end of the experiment, participants were asked where they grew up, languages they spoke by the age of five, gender, race/ethnicity, year of birth, occupation, highest level of education

completed, current zip code, and what they thought the experiment was about.

Based on responses to this questionnaire, participants’ race/ethnicities were coded as “white,” “Black” “Latino,” “Asian,” or “mixed race.” Participants checking the “white” box in combination with the “Hispanic/Latino/a” box were counted as Latina/o; any other combination of boxes was counted as “mixed race”. Three participants declined to provide their race; they were coded as “white” because this was the most frequent category in the data.

For participants who stated that they had grown up in multiple places (n=27), hometown was coded based on the location they reported having lived at the longest between the ages of 5 and 15. Education level was coded along a 9 point scale (1=some high school, 2=GED, 3=high school graduate, 4=some college, 5=associate’s degree, 6=bachelor’s degree, 7=some graduate school, 8=master’s degree, 9=PhD or professional degree).

4.4 Acoustic analysis

The vowel in each PIN-PEN and BIT-BET word was measured at the midpoint in Praat (Boersma & Weenink, 2002), with the onset of each vowel defined at the beginning of the second formant and the offset either by the point at which the spectrogram abruptly lightened, indicating the onset of the following nasal consonant (for PIN-PEN words), or at the end of the second formant (for BIT-BET words). Vowels were measured automatically using Praat’s formant tracker, but checked by hand to ensure the formant tracker’s accuracy.

5. Results

5.1 Merger in production

Here I first examine what demographic factors are associated with the PIN-PEN merger in production. The degree to which each participant was merged was calculated as a Pillai score, the output of a MANOVA which measures the distinctiveness of two distributions (here, the distributions of pre-nasal /i/ and /ɛ/) (Hall-Lew, 2010; Hay, Warren & Drager, 2006). The higher the Pillai score, the more distinct the two vowels. Participants' Pillai scores ranged from 0.005 to 0.994, of a maximum possible range of 0 to 1.

A reviewer notes that Pillai scores have been criticized (Johnson, 2015) and may be unreliable for small token counts. Because of this concern, merger was additionally measured using Bhattacharyya's affinity (Bhattacharyya, 1943). Calculating Bhattacharyya's affinity requires at least five tokens per vowel class; 42 (15%) of the participants did not meet this criterion due to background noises or recording errors that made fewer than five tokens of at least one vowel class measurable. Therefore, the dataset with Bhattacharyya's affinity is slightly smaller than that with Pillai score. Analyses with Bhattacharyya's affinity produced qualitatively similar results to Pillai score, so in the sections of this paper that follow I report primarily on Pillai score (since Pillai score allows for analysis of the full production dataset), but describe any differences observed in the analysis with Bhattacharyya's affinity.

To determine which demographic categories of speakers were most likely to be merged, social factors affecting Pillai score (or Bhattacharyya's affinity) were modeled using a Generalized Additive Model (GAM). Since previous research has shown that region correlates with the PIN-PEN merger, a GAM offers a better model than linear regression: while traditional regression can only model predictors as straight lines, GAMs can model arbitrarily curvy lines (i.e., functions with any number of exponents). As regional differences do not usually occur along a straight line, this makes the GAM a superior tool for modeling regional variation. Here regional variation was modeled as a smooth curve of latitude and longitude of the participant's hometown. Predictors tested for the model were race (treated as a sum contrast), age, gender, education level, and logarithm of hometown population according to the 2010 US Census. Where the participant had left a question blank, it was replaced with the overall participant mean. The model was built using a step-up approach, using log-likelihood comparisons with an alpha of 0.05; when p-values were identical, AIC was used as the selection criterion.

The final model consisted of the latitude/longitude smooth curve, education level, race/ethnicity, hometown population, and year of birth; only gender was not significant (Table 2). (With Bhattacharyya's affinity, educational level and hometown population were not significant, but still trended in the same direction.) Model predictions by latitude and longitude are shown in Map 1: speakers in the South are predicted to be more merged (have lower Pillai scores) than elsewhere, replicating previous research. Less educated speakers ($t(260)=2.098$, $p=0.037$), Black speakers ($t(260)=-3.59$, $p<0.001$), and speakers from smaller towns ($t(260)=2.468$, $p=0.016$) are more likely to be merged, also matching previous research. Additionally, older speakers are more likely to be merged than younger speakers ($t(260)=2.468$, $p=0.014$), suggesting that the merger may be reversing itself, at least in some communities.

Table 2. Production Model.

Term	Estimate	Std Error	t value	p value
(Intercept)	-7.92	3.341	-2.371	0.018 *
education level	0.02	0.01	2.098	0.037 *
race - Asian	0.035	0.06	0.583	0.56
race - Black	-0.194	0.054	-3.59	< 0.001 ***
race - Latino	0.086	0.064	1.346	0.18
race - mixed	0.057	0.072	0.787	0.432
hometown population	0.021	0.009	2.432	0.016 *
year of birth	0.004	0.002	2.468	0.014 *
Smooth Terms	edf	Ref df	F	p value
te(longitude, latitude)	8.88	10.784	2.501	0.006 **
Significantly differs from simpler models without ...				
... education level: $p=0.037$ *				
... race: $p=0.010$ **				
... hometown population: $p=0.013$ *				
... year of birth: $p=0.015$ *				
... longitude/latitude: $p<0.001$ ***				

Generalized additive model predicting Pillai scores (smaller = more merged). Race is treated as a sum contrast, so the coefficient for white speakers is -1 times the sum of the coefficients for other races (-0.016).

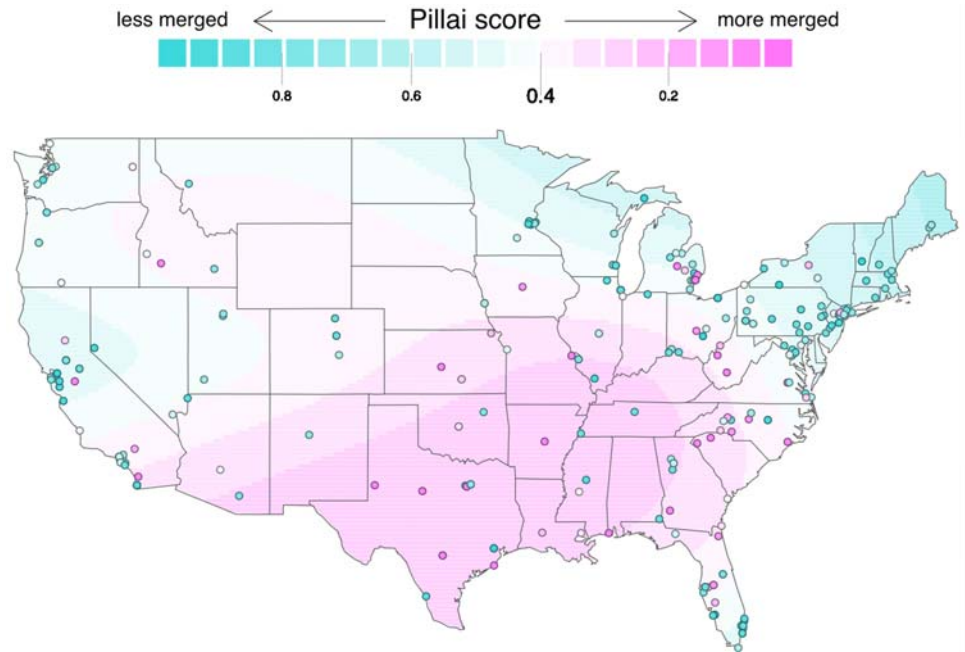
5.2 Phonetic realization of the merger

How did merged participants vary in the quality of their merged vowel, and were there demographic differences in who merged towards [ɛ] versus [ɪ]? Merged participants were defined as those with a Pillai score of 0.4 or lower ($n=64$). This cutoff was chosen because above 0.4 most MANOVAs used to calculate the Pillai score indicated that the distributions of PIN and PEN were significantly distinct (Figure 2)³.

For each merged participant, the vowel in each BIT and BET word (i.e., non-pre-nasal /i/ and /ɛ/) was measured at the midpoint. Then, two Pillai scores were calculated: one testing for distinction between PIN-PEN words and BIT words, and one between PIN-PEN words and BET words. A low PIN-PEN:BIT Pillai score indicates that the merged vowel is similar to [ɪ], whereas a low PIN-PEN:BET score indicates that the merged vowel is similar to [ɛ]. (A low score for both would indicate that all three vowel classes are similar, i.e., that PIN-PEN, BIT, and BET have highly overlapping distributions.) From these two Pillai scores, a directionality score was calculated as the natural logarithm⁴ of the ratio of the PIN-PEN:BIT score to the PIN-PEN:BET score:

$$\text{Direction Score} = \ln(\text{Pillai}(\text{PIN-PEN:BIT})/\text{Pillai}(\text{PIN-PEN:BET}))$$

A positive score indicates that the participant's merged vowel is closer to [ɛ]; a negative score that the vowel is closer to [ɪ]; and a score of 0 that the vowel is equidistant between [ɪ] and [ɛ]. Sample vowel plots of participants with positive, negative, and near-zero directionality scores are shown in Figure 3. Scores ranged from -7.32 to 5.08, with a median score of 0.22 (i.e., a vowel somewhat in the middle, but slightly closer to [ɛ].) Figure 4 plots the scores for each participant. These results indicate that a merger towards [ɛ] is more common than the literature suggests: 40 of the participants (63%) have a positive directionality score, i.e., a merged vowel closer to [ɛ]. With Bhattacharyya's affinity, defining "merged" participants as those with affinity scores greater than 0.75 ($n=48$), the median directionality score is -0.02 (i.e., an intermediate vowel, slightly closer to [ɪ]), and 48% of participants have a positive directionality score.



Map 1. Map of Predicted vs Actual Pillai Scores. Predicted (background colors) and actual (foreground dots) Pillai scores. Model predictions are for a 35-year-old white speaker who has attended some high school, from a hometown of population 400.

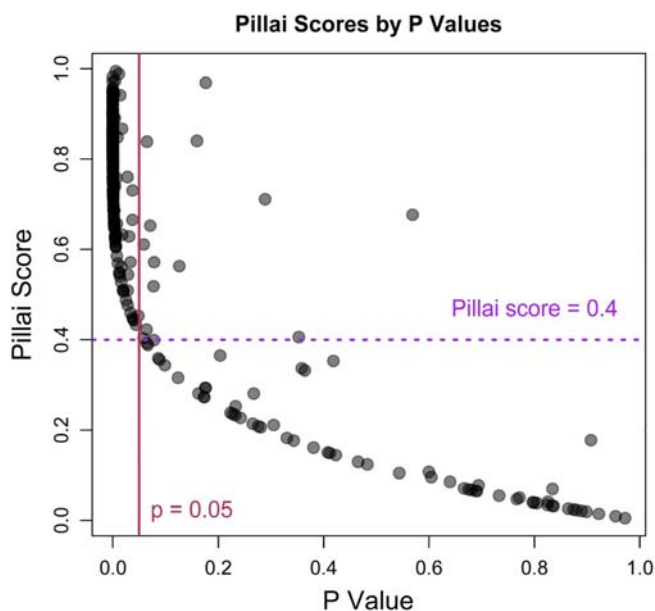


Figure 2. Pillai Scores by P-Values. Pillai scores versus MANOVA p-values. Most participants with Pillai scores greater than 0.4 have p-values of 0.05 or smaller, indicating that their [ɪn] and [ɛn] are distinct; thus, a Pillai of 0.4 is used as the cutoff between merged and distinct participants.

To test for the influence of social factors (age, education level, gender, hometown population, race, and latitude/longitude) on the realization of the merger, a model with direction score as the dependent variable was constructed in the same manner as above. No factor, including latitude/longitude, was significant with Pillai score.

Whether the lack of effect for any other social factor means that these factors do *not* condition the variation between [ɪ] and [ɛ] remains unknown: the number of merged participants might be too small to yield a significant effect, or the variation might be socially stratified differently in different communities across the

United States (e.g., [ɛ] being produced by more educated speakers in one community, but by less educated speakers in another). The first of these possibilities, that the lack of significant results stems from an insufficient number of participants, seems unlikely, as none of the social factors show any trends (see Figure 5), with the exception of a clustering of speakers who merge to [ɪ] in California (Figure 5f). (With Bhattacharyya's affinity, longitude—but not latitude—is a significant ($p=0.03$) predictor of directionality score. This model predicts that speakers in the West are more likely to merge towards [ɪ] than in the Midwest or the South.) Thus, these results seem to suggest that the variation between [ɪ] and [ɛ] is largely random, and not conditioned by any of the social factors examined in this study.

5.3 Merger in perception

Merged listeners were expected to perform at chance on the word identification task. However, nearly all participants scored better than chance: of the 372 participants, only 11 (3%) were at or below 50% accuracy (see Figure 6). (Similar results, where merged participants can identify words with higher than chance accuracy, have also been found with other mergers, e.g., Thomas and Hay (2005) and Arnold (2015).)

Because most participants performed better than chance, merger in perception was defined as whether a listener could *consistently*—i.e., with 100% accuracy—distinguish between PIN and PEN. Participants who identified all sixteen of the PIN-PEN tokens correctly were classified as non-merged ($n=185$, 50%), and participants who misidentified two or more were classified as merged ($n=123$, 33%). Those who misidentified only one token ($n=64$, 17%) were excluded from analysis, as this group of participants was likely to include both truly merged participants who performed exceptionally well by chance and non-merged participants who accidentally clicked on the wrong word despite accurately perceiving it.

To examine social factors affecting merger in perception, a model was constructed in the same manner as for production, with the participant's classification as "merged" or "non-merged" as the dependent variable. The final model included a smooth for latitude

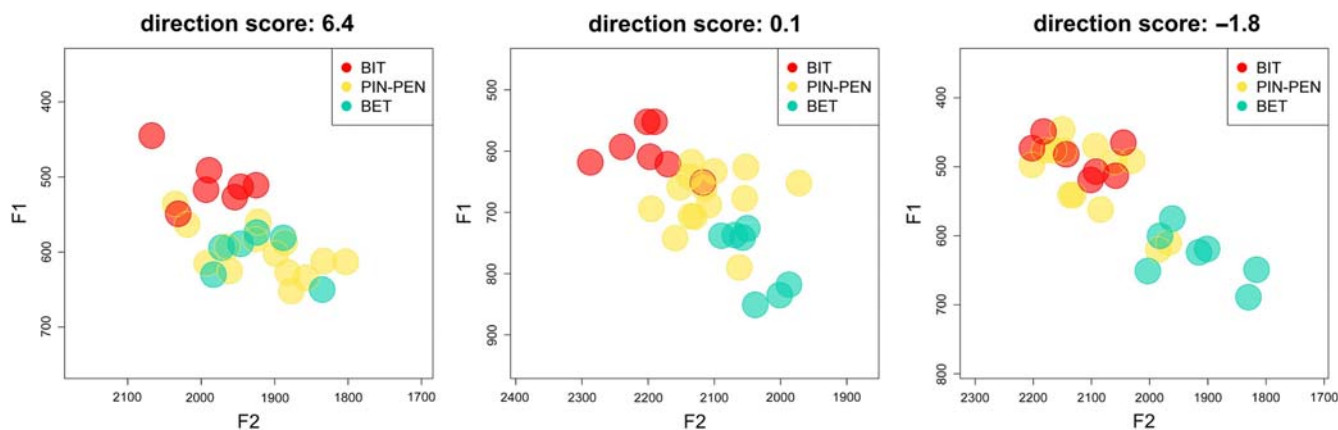


Figure 3. Sample Direction Scores. Vowel plots from a participant with a positive directionality score (6.4, on the left), indicating a merger towards [ε]; a directionality score close to 0 (0.1, in the middle), indicating an intermediate vowel; and a negative directionality score (-1.8, on the right), indicating a merger towards [i].

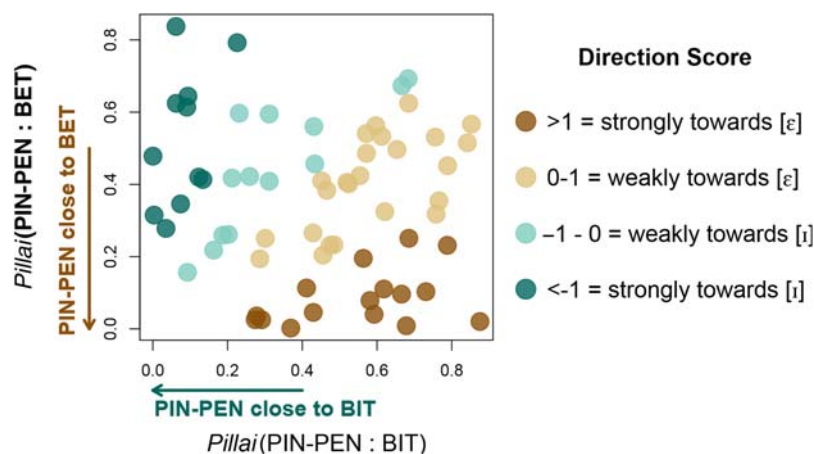


Figure 4. Direction Scores. PIN-PEN:BIT and PIN-PEN:BET Pillai scores for each merged participant, color-coded by directionality score. Speakers in the top left quadrant can be interpreted as merging towards [i], in the top right as having an intermediate vowel midway between [i] and [ε], in the bottom right as merging towards [ε], and in the bottom left as having minimal distinction between [i] and [ε] even in non-pre-nasal contexts.

and longitude, education level, and race (Table 3). As with production, speakers in the South were most likely to be merged (Map 2). The lower the level of education, the more likely that a participant would be merged ($t(290)=-2.56, p=0.01$). Black speakers were significantly more likely to be merged than speakers of other races/ethnicities (i.e., had a higher probability of being merged than the mean probability across speakers of all races/ethnicities, $t(290)=2.809, p=0.005$), replicating previous research.

5.4 Rhyming

In the rhyming task, it was expected that individual participants would either say that *all* of the PIN-PEN words rhymed, or that none of them did. However, this was not the case: 86 participants (23%) said that only *some* of these words rhymed (Figure 7); most of these participants (84%) were merged in either production or perception. As the instructions for the task simply asked “Do these words rhyme?”, participants may have been confused about for whom the words were supposed to rhyme, perhaps being aware that some speakers pronounce PIN and PEN differently. One participant commented that she didn’t hear any sound playing during this task, suggesting that she was expecting to listen to someone else’s productions of the words and then judge whether they rhymed. Participants may have also been unduly influenced by orthography, saying that some PIN-PEN words did not rhyme because they

were spelled differently. Of the control rhyming pairs, 21% of participants incorrectly said that *cow* and *tow* rhyme, suggesting that orthography did play a large role in responses.

Since participants did not respond to the task as expected, no further analysis of rhyming scores was undertaken.

5.5 Production versus perception

As noted earlier, production does not always equal perception. Previous research shows that speakers may be merged in perception but not in production (“near merger”). The converse phenomenon, merger in production but not in perception, is reportedly rare or nonexistent, and in sound change merger in perception is said to come before merger in production (Labov, 2011:334). The present study’s results support this general finding: while 33 participants were merged in perception but not in production (i.e., near merged), only 6 were merged in production but not in perception.⁵ (51 were merged in both production and perception. With Bhattacharyya’s affinity, 42 participants were merged in both production and perception, 30 were near merged, and 5 were merged in production but not perception.)

To examine which social factors predict near merger, two separate models were constructed: one comparing near merged participants to fully merged participants (excluding fully non-merged participants), and one comparing near merged participants to fully

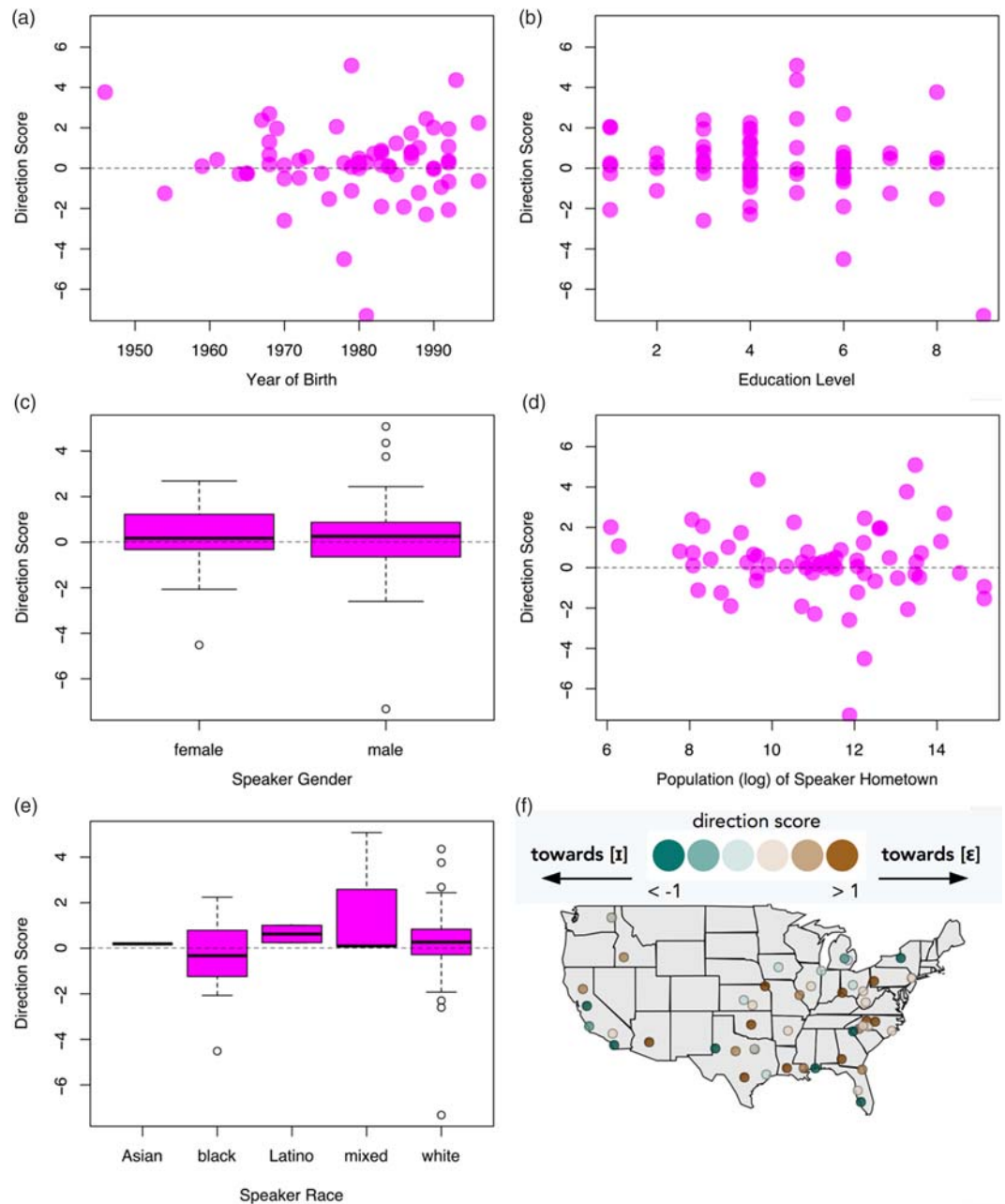


Figure 5. Direction Score by Social Factors. Direction score by various social factors, showing a lack not just of significant effects but also of any noticeable trends.

non-merged participants (excluding fully merged participants). As before, models were built using a step-up approach with latitude/longitude, age, gender, education level, and race/ethnicity as potential predictors. A smooth for latitude/longitude failed to improve upon either null model, so the models were built as generalized linear models rather than generalized additive models.

The model comparing near merged speakers to fully merged ones is given in Table 4, containing a significant predictor of gender. Men were more likely to be near merged than women (Wald $z=-2.248$, $p=0.025$) (Figure 8). (With Bhattacharyya's affinity, gender was not significant, though still trending ($p=0.090$) in the same direction.)

In comparing near merged speakers to fully non-merged ones, no factor significantly improved upon the null model based on Pillai score. With Bhattacharyya's affinity, less educated speakers were more likely to be near merged.

6. Discussion and conclusion

This study's findings with regards to production and perception of the PIN-PEN merger largely replicate previous results: in production, less educated speakers, Black speakers, older speakers, and speakers from smaller towns were most likely to be merged; in perception, less educated speakers and Black speakers were most likely to be merged. Across both production and perception, merged speakers were most highly concentrated in the South, but were also found in sizeable numbers in parts of the Midwest (Illinois, Indiana, Ohio, and especially southeastern Michigan) and in California. The presence of the merger in these regions is likely due to the influx of Southerners in the early to mid-20th century. Both Black and white Southerners (especially Appalachians) migrated in large numbers to California and Midwestern cities

Table 3. Perception Model.

Term	Estimate	Std Error	t value	p value
(Intercept)	0.474	0.467	1.015	0.31
education level	-0.205	0.08	-2.56	0.01 *
race - Asian	-0.694	0.588	-1.179	0.238
race - Black	1.292	0.46	2.809	0.005 **
race - Latino	0.569	0.503	1.13	0.258
race - mixed	-1.008	0.564	-1.788	0.074
Smooth Terms	edf	Ref df	F	p value
te(longitude, latitude)	11.735	13.635	46.342	< 0.001 ***
Significantly differs from simpler models without ...				
... education: p=0.010 **				
... race: p=0.007 **				
... longitude/latitude: p<0.001 ***				

Binomial GAM predicting whether a speaker will be merged in perception (higher coefficient = more likely to be merged). Race is treated as a sum contrast, so the coefficient for white speakers is -1 times the sum of the coefficients for other races (0.159).

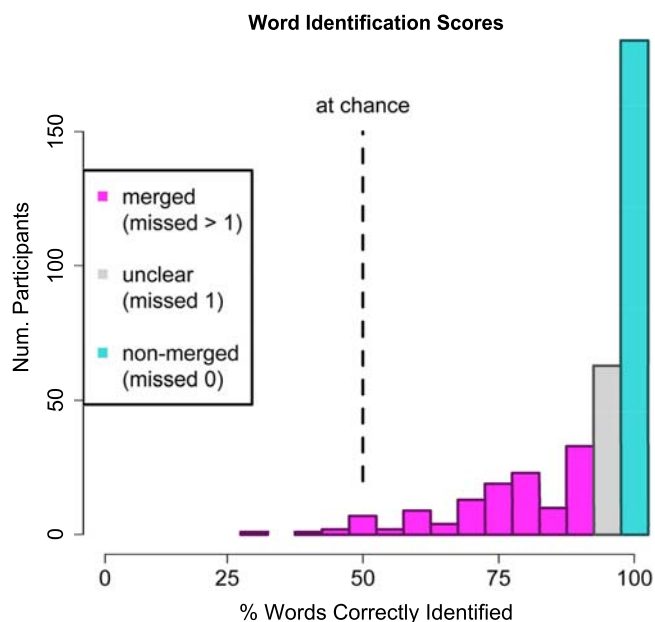


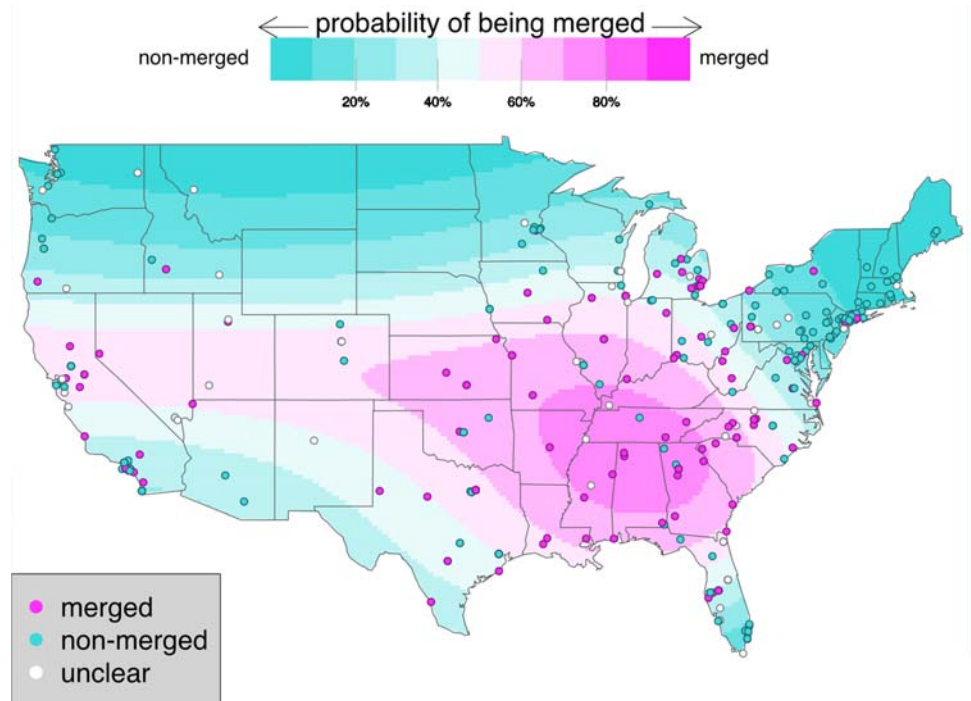
Figure 6. Word Identification Scores. Histogram of scores on the word identification task, color coded to show how scores were classified as “merged”, “unclear”, or “non-merged”.

such as Chicago, Detroit, Indianapolis, and Columbus, Ohio (Alexander, 2006; Franklin, 2002; Johnson, 2006:33-35; Kirby, 1983; Tolnay, 2003). (This period of migration also encompasses the well-known movement of the “Okies”—who were not just from Oklahoma, but also commonly from Texas, Arkansas, or Missouri—to California during the Dust Bowl in the 1930s (Kirby, 1983).) These migrants likely brought the PIN-PEN merger with them to their new homes; the merger may have persisted over time due to the continued isolation of migrants from the local white speakers. (Like African Americans and “Okies,” working-class white Southerners faced substantial discrimination, often living in so-called “hillbilly ghettos” (Alexander, 2006; Anderson, 2008; Harkins, 2001; Johnson, 2006:16).) The merger may have become a marker of Southern identity, as Anderson (2008) shows for /ay/ monophthongization among Black and white migrants to

Detroit. In other places, the merger may have over time become associated with other locally relevant identity categories; for example, Geenberg (2014), finds that in Trinity County, California, use of the merger is best correlated not with Southern ancestry, but rather with speakers who partake in a survivalist, living-off-the-land lifestyle.

When examining the relationship between production and perception, this study found that near merger—merger in perception but not production—was roughly half as common as full merger (33 near merged versus 51 fully merged participants, 39% of all participants merged in perception). This is nearly double the proportion of near-merged speakers reported in Labov et al. 2006, where only 21% of speakers who were merged in perception were near-merged. One possible reason for this discrepancy is the difference in tasks used to classify merger in perception. In Labov et al. (2006), participants were asked to judge whether the lexical items *pin* and *pen* sounded the same, with judgments about a further pair *him* and *hem* elicited only for participants who the researcher judged to produce *pin* and *pen* “close or the same”. In contrast, in this study participants were asked to classify a given token of *pin/pen*, *bin/Ben*, *tin/ten*, or *mint/meant* as a PIN or PEN word. Participants in the same-different judgment task used in Labov et al. (2006) may have been more influenced by explicit knowledge about which words are “supposed” to be pronounced differently, leading them to judge words as different even if they would not be able to perform 100% accurately in this study’s word identification task. The particular words *pin* and *pen* might be especially affected by conscious language ideologies, as the merged pronunciation of these words is often commented upon. It may be that speakers who are more likely to be near merged are also more likely to be aware of standard language ideologies around the PIN-PEN merger, and thus more likely to judge PIN-PEN as sounding differently, leading to a smaller percentage of apparent “near-merged” speakers. Alternatively, differences in the samples of the two studies may have led to these differences: Labov et al. (2006) attempted to survey as random a sample as possible, whereas the Mechanical-Turk-recruited participants in this study were on average more educated than the US population as a whole.

This study also found that men were more likely to be near-merged (as compared to being fully merged) than women, a finding that to the best of my knowledge has not been found for any other near merger, although most studies of near merger have not looked for such an effect. Given that women tend to lead changes in progress (Labov, 2011), and that this study suggests that the merger is reversing itself (as younger speakers are less likely to be merged in production), this is surprising: if near-merger is a transitional stage between having full merger and full distinction, then we would expect more women to be near-merged than fully merged. However, it is not clear that near merger is in fact a transitional stage in the reversal of a merger: while perception tends to precede production in mergers in progress, leading to near merger, more research is needed to determine whether near-merger is also an intermediate stage in the reversal of a merger. If the merger is spreading in some places and reversing itself in others—as suggested by the contradictory findings in previous research about whether the merger is a change in progress—then this effect of men being more likely to be near-merged may be a reflection of men being slower to fully adopt a merger in those places where the merger is spreading. More data would be needed to test the prediction, but future studies focusing on individual communities may be able to explore this potential interaction between gender, changes in progress, and production versus perception.



Map 2. Perception Data and Model Predictions. Background colors: model predictions of the likelihood of being merged in perception for a white speaker who has attended some college. Foreground dots: merger status in perception of actual participants.

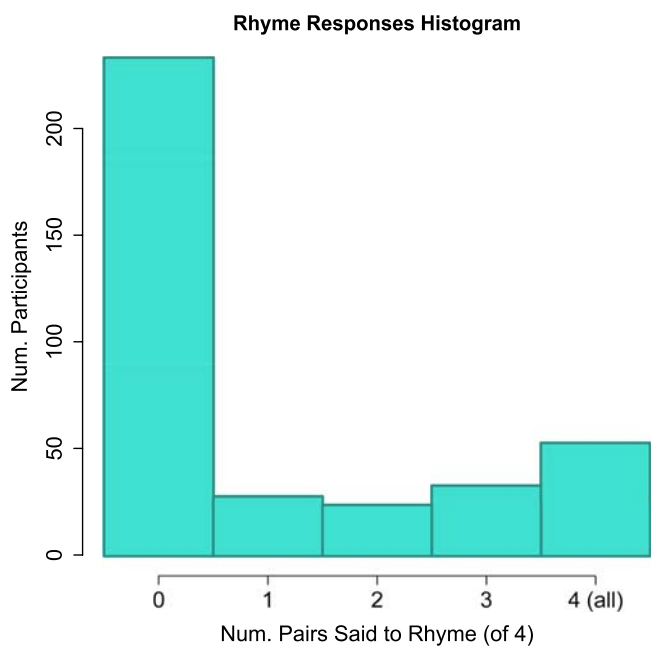


Figure 7. Responses to Rhyming Task. Histogram of the number of PIN-PEN pairs that a participant said rhymed.

An additional finding was that, contrary to expectations, almost all speakers who were merged in production were able to identify PIN and PEN words with higher-than-chance, but less than 100%, accuracy in the perception task. Similar results have been reported for other conditioned mergers (Thomas & Hay, 2005; Arnold, 2015), and, as Wade (2017) points out, in most perception tasks the cut-off for being counted as “merged in perception” is anything less than 100% accuracy, meaning that this phenomenon may be more widespread than commonly believed. These results suggest that

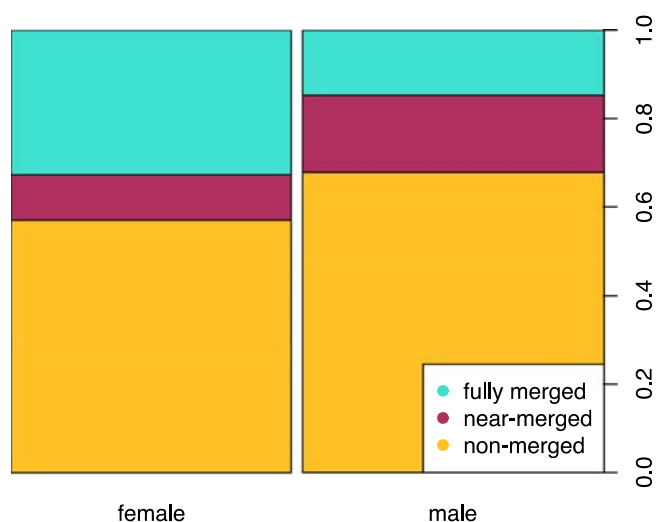
merged speakers have some probabilistic knowledge about the way PIN versus PEN words are produced by non-merged speakers. A subject for future investigation is whether this knowledge comes from merged listeners storing non-merged speakers’ productions on a word-specific level, or from the merged speakers’ knowledge of their own productions of BIT-BET words. That is, do listeners perform better than chance because they probabilistically know how non-merged speakers pronounce these words, or because they know how they (the merged speakers) produce /e/ and /i/ in non-pre-nasal contexts and are able to judge whether a given token is closer to their non-pre-nasal /e/ or /i/? An additional question for future research is to what extent orthography and conscious awareness of the merger affect performance on this perception task.

Finally, this study found that merger towards [e] (*tw[e]n* for *twin*) was equally as common as merger towards [i] (*h[i]n* for *hen*), contradicting the literature’s tendency to describe the merger as being towards [i]. There are several potential reasons for this contradiction: first, this study used acoustic measurements to determine merger directionality, whereas many previous studies relied on the ears of the researcher. It may be that merged tokens that are acoustically closer to a speaker’s BET vowel than to their BIT vowel are still perceptually closer to /i/: that is, just because a token is acoustically similar to [e] does not necessarily mean it is heard as [e]. Secondly, as Bigam (2005) posits, researchers may perceive merged tokens as [i]-like because of their beliefs about the PIN-PEN merger: since early research tended to describe the merger as being towards [i], this may be taken as a given by subsequent researchers and have affected their perception. This may be especially true for more intermediate tokens: if these tokens can be perceived equally well as [i] or [e], believing that they should sound like [i] may lead researchers to perceive them as such. Lastly, the high proportion of merger towards [e] could be a reflection of style-shifting. While merged speakers are likely to remain merged across styles, the quality of their merged vowel might shift according to context. In Austen (2017), I found that merged

Table 4. Near Merger Model.

Term	Estimate	Std Error	t value	p value
(Intercept)	0.932	0.327	2.845	0.004 **
gender – male	-1.037	0.461	-2.248	0.025 *

Logistic model predicting the likelihood of being fully merged (a positive value) versus near-merged (a negative value).

**Figure 8.** Merger Status by Gender. Proportion merger status by gender.

listeners perceived the merger towards [ɛ] as prestigious, such that they heard *tw[ɛ]n* as more prestigious than *tw[i]n*. Thus, it could be that merged speakers shift to [ɛ] in more formal contexts, such as the wordlist-reading task used in this experiment. In less formal contexts, like conversational speech upon which many previous researchers based their observations, these speakers could shift to more [i]-like pronunciations.

Overall, this study suggests that dialectology would do well to consider the relationship between production and perception. Examining phenomena such as near mergers may help us to understand how sound changes like mergers spread across social and geographic space. The PIN-PEN merger appears to be a particularly useful merger to study: if, as this study taken in combination with previous research suggests, it is expanding in some areas and receding in others, then we can simultaneously observe the spread and reversal of a merger. Further study of this merger can offer new insights into the interaction between dialect regions, sound change, and the production-perception interface.

Endnotes

1. Statements about the Linguistic Atlas of the Gulf States in this paper are based on averages calculated across seven PEN items (*ten*, *twenty*, *pen*, *ten cents*, *Tennessee*, *Memphis*, and *Wednesday*) reported on in the Atlas's "Social Matrix" (Volume 6).
2. Misperceptions of the *dog/dock* stimuli did not disqualify participants because I found the talker's productions of these to be ambiguous, as apparently did many of the participants, who misidentified them 12% of the time.
3. An alternate possibility would have been to define merged participants as those with a p-value of 0.05 or greater in the MANOVA, i.e., PIN and PEN vowels that were not significantly distinct from one another. However, because participants read a fairly small number of words and sometimes not all of these words were measurable (e.g., a dog barking during the vowel), there is a problem of statistical power: participants may in fact have distinct PIN and PEN vowels,

but the difference may not reach significance because of the small number of tokens. Thus, here Pillai score offers a way to avoid inadvertently defining non-merged participants as merged; these participants will still have a Pillai score greater than 0.4 even if they do not have a p-value less than 0.05.

4. The natural logarithm of the ratio was taken in order to make the scores more interpretable, as the natural logarithm converts ratios less than 1 to negative numbers, and ratios greater than 1 to positive numbers.

5. Although the number of participants who appear to be merged in production but not perception is 6, and not zero, this does not necessarily indicate that merger in production but not perception actually exists. Listeners only classified 16 tokens each, so these 6 speakers could well be merged in perception, but happened to have guessed correctly for all 16 tokens. Statistically, this is not unlikely: if merged listeners have an 80% chance of classifying a given token correctly, then they have a 3% chance of correctly identifying all 16 tokens.

References

- Alexander, J. Trent. 2006. Defining the diaspora: Appalachians in the Great Migration. *Journal of Interdisciplinary History* 37(2). 219–247.
- Anderson, Bridget L. 2008. *Migration, accommodation, and language change: Language at the intersection of regional and ethnic identity*. Basingstoke, Hampshire: Palgrave Macmillan.
- Arnold, Lacey. 2015. Multiple mergers: Production and perception of three pre-/l/ mergers in Youngstown, Ohio. *University of Pennsylvania Working Papers in Linguistics* 21(2). 2.
- Austen, Martha. 2017. Pinning down social meaning: How listener phonology shapes social perception. Poster presented at New Ways of Analyzing Variation (NWAV) 46, Madison, WI, November 2–5. https://www.asc.ohio-state.edu/austen.14/materials/nwav2017_qp2_poster.pdf (7 October, 2020).
- Bakos, Jon. 2013. *A comparison of the speech patterns and dialect attitudes of Oklahoma*. Stillwater, OK: Oklahoma State University dissertation.
- Bailey, Guy & Erik Thomas. 1998. Some aspects of African-American Vernacular English phonology. In Mufwene Salikoko, John R. Rickford, Guy Bailey & John Baugh (eds.), *African-American English: Structure, History, and Use*, 85–109. New York: Routledge.
- Baranowski, Maciej. 2013. On the role of social factors in the loss of phonemic distinctions. *English Language and Linguistics* 17(02). 271–295.
- Berinsky, Adam J., Gregory A. Huber & Gabriel S. Lenz. 2012. Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis* 20(3). 351–368.
- Bhattacharyya, Anil. 1943. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society* 35. 99–109.
- Bigham, Douglas S. 2005. *The movement of front vowel allophones before nasals in Southern Illinois White Vernacular English (the PIN PEN merger)*. Austin, TX: The University of Texas at Austin dissertation.
- Boersma, Paul & Weenink, David. 2002. *Praat: Doing phonetics by computer*. <http://www.praat.org/>.
- Bowie, David. 2000. *The effect of geographic mobility of the retention of a local dialect*. Philadelphia, PA: University of Pennsylvania dissertation.
- Brown, Vivian R. 1991. Evolution of the merger of /ɪ/ and /ɛ/ before nasals in Tennessee. *American Speech* 66(3). 303–315.
- Cogshall, Elizabeth L. & Kara Becker. 2009. The vowel phonologies of African American and white New York City residents. *Publication of the American Dialect Society* 94(1). 101–128.
- De Decker, Paul. 2015. The recorder's paradox: Balancing high-quality recordings with spontaneous speech in noisy recording environments. *Journal of the Acoustical Society of America* 137. 2304–2304.
- De Decker, Paul & Nycz, Jennifer. 2011. For the record: Which digital media can be used for sociophonetic analysis? *University of Pennsylvania Working Papers in Linguistics* 17(2). 51–59.
- Degen, Judith & Noah Goodman. 2014. Lost your marbles? The puzzle of dependent measures in experimental pragmatics. *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 36.
- Di Paolo, Marianna. 1992. Hypercorrection in response to the apparent merger of (ɔ) and (α) in Utah English. *Language & Communication* 12(3/4). 267–292.

- Eberhardt, Maev. 2009. African American and white vowel systems in Pittsburgh. *Publication of the American Dialect Society* 94(1). 129–157.
- Edwards, Walter F. 1997. The variable persistence of Southern Vernacular Sounds in the speech of inner-city Black Detroiters. In Cynthia Bernstein, Thomas Nunnally & Robin Sabino (eds.), *Language Variety in the South Revisited*, 76–86. Tuscaloosa: University of Alabama Press.
- Enochson, Kelly & Jennifer Culbertson. 2015. Collecting psycholinguistic response time data using Amazon Mechanical Turk. *PLOS ONE* 10(3).
- Franklin, V.P. 2002. “Location, location, location”: The cultural geography of African Americans: Introduction to a journey. *The Journal of African American History* 87. 1–11.
- Geenberg, Katherine. 2014. *The other California: Marginalization and sociolinguistic variation in Trinity County*. Palo Alto, California: Stanford University dissertation.
- Hall-Lew, Lauren. 2010. Improved representation of variance in measures of vowel merger. *Proceedings of Meetings on Acoustics* 9(1). 1–10.
- Harkins, Anthony. 2001. The hillbilly in the living room: Television representations of southern mountaineers in situation comedies, 1952–1971. *Appalachian Journal* 29(1/2). 98–126.
- Hay, Jennifer, Katie Drager & Brynmor Thomas. 2013. Using nonsense words to investigate vowel merger. *English Language and Linguistics* 17(02). 241–269.
- Hay, Jennifer, Paul Warren & Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics (Modelling Sociophonetic Variation)* 34(4). 458–484.
- Hazen, Kirk. 2005. Mergers in the mountains: West Virginia division and unification. *English World-Wide* 26(2). 199–221.
- Hazen, Kirk & Ellen Fluharty. 2004. Defining Appalachian English. In Margaret Bender (ed.), *Linguistic diversity in the South: Changing codes, practices, and ideology* (Southern Anthropological Society Proceedings 37), 50–65. Athens, GA: University of Georgia Press.
- Ito, Kiwako & Kathryn Campbell-Kibler. 2011. Speaker-adaptation to /I/-/E/ Merger: An Eye-tracking Study. *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*. Hong Kong.
- Johnson, Daniel Ezra. 2015. Quantifying overlap with Bhattacharyya’s affinity and other measures. Paper presented at New Ways of Analyzing Variation 44. Toronto, Ontario. https://danielezrajohnson.shinyapps.io/nwav_44/
- Johnson, Susan Allyn. 2006. *Industrial voyagers: A case study of Appalachian migration to Akron, Ohio, 1900–1940*. Columbus, OH: The Ohio State University dissertation.
- Kanwal, Jasmeen, Kenny Smith, Jennifer Culbertson & Simon Kirby. 2017. Zipf’s Law of Abbreviation and the Principle of Least Effort: Language users optimise a miniature lexicon for efficient communication. *Cognition* 165. 45–52.
- Kim, Chaeyoon, Ezra Wyszogrod, Sravana Reddy & James Stanford. 2016. A large-scale online study of dialect variation in the US Northeast: Crowdsourcing with Amazon Mechanical Turk. Paper presented at New Ways of Analyzing Variation (NWAV) 44, Vancouver.
- Kirby, Jack Temple. 1983. The southern exodus, 1910–1960: A primer for historians. *The Journal of Southern History* 49(4). 585–600.
- Koops, Christian, Elizabeth Gentry & Andrew Pantos. 2008. The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics* 14(2). 12.
- Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(02). 205–254.
- Labov, William. 1994. *Principles of Linguistic Change*. Vol. 1: Internal Factors. Oxford: Blackwell.
- Labov, William. 2011. *Principles of Linguistic Change*. Vol. 3: Cognitive and Cultural Factors. John Wiley & Sons.
- Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Walter de Gruyter.
- Labov, William, Paul Cohen, Clarence Robins & John Lewis. 1968. *A study of the non-standard English of Negro and Puerto Rican speakers in New York City*. Vol. 1: Phonological and grammatical analysis. New York: Columbia University.
- Pederson, Lee. 1965. *The pronunciation of English in metropolitan Chicago: Vowels and consonants*. Publication of the American Dialect Society 44.
- Pederson, Lee, Susan McDaniel & Carol Adams (eds.). 1986. *Linguistic atlas of the Gulf States*. 7 vols. Athens, GA: University of Georgia Press.
- Podesva, Robert J., Annette D’Onofrio, Janneke Van Hofwegen & Seung Kyung Kim. 2015. Country ideology and the California Vowel Shift. *Language Variation and Change* 27(02). 157–186.
- Tammimga, Meredith. 2017. Matched guise effects can be robust to speech style. *The Journal of the Acoustical Society of America* 142(1). EL18–EL23.
- Thomas, Brynmor & Jennifer Hay. 2005. A pleasant malady: The Ellen/Allan merger in New Zealand English. *Te Reo* 48. 69.
- Thomas, Erik. 2004. Rural Southern white accents. In Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton (eds.), *A handbook of the varieties of English. Volume 1: Phonology*, 300–324. Berlin: Mouton de Gruyter.
- Tillery, Jan & Guy Bailey. 2004. The urban South: phonology. In Edgar W. Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie & Clive Upton (eds.), *A handbook of varieties of English. Volume 1: Phonology*, 325–337. Berlin: Mouton de Gruyter.
- Tolnay, Stewart E. 2003. The African American “Great Migration” and Beyond. *Annual Review of Sociology* 29(1). 209–232.
- Wade, Lacey. 2017. The role of duration in the perception of vowel merger. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1):30. 1–34.
- Wang, Shichang, Chu-Ren Huang, Yao Yao & Angel Chan. 2015. Mechanical Turk-based experiment vs laboratory-based Experiment: A case study on the comparison of semantic transparency rating data. *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation*, 53–62.
- Warren, Ron & Sean A. Fulop. 2014. The merged vowel of PIN and PEN as realized in Bakersfield, California. *The Journal of the Acoustical Society of America* 135(4). 2292.
- Wise, C. M. 1933. Southern American dialect. *American Speech* 8(2). 37–43.
- Wolfram, Walt & Natalie Schilling. 1998. *American English: Dialects and Variation*. Wiley.