University of Pennsylvania Working Papers in Linguistics

Volume 16
Issue 2 *Selected Papers from NWAV 38*

Article 19

1-1-2010

Using Acoustic Trajectory Information in Studies of Merger

Michael Scanlon
University of Washington, ms27@uw.edu

Alicia Beckford Wassink
University of Washington, wassink@uw.edu

Using Acoustic Trajectory Information in Studies of Merger

Abstract
This study investigates the utility of examining acoustic trajectory information indicative of gliding in the case of mergers or near-mergers. It presents a sociophonetic analysis of conversational speech from one African American Seattle native, who perceives the *pin* and *pen* classes as merged. The study finds no difference ("merger") between the speaker's *pin* and *pen* classes by F1 or F2 at vowel midpoint. However, phonemic vowel distinctions are preserved in Euclidean distance and duration, and the vowel classes are more distinct pre-nasally than in non-pre-nasal contexts. A regression of the researcher's perception of distance on vowel class corroborates this pattern. Lastly, multidimensional calculation of overlap using SOAM (Wassink 2006) for a small sample of data from 12 Seattle speakers suggests Seattle African Americans differentiate *pin* from *pen* somewhat by the amount of glide, while Seattle Whites do not.

Using Acoustic Trajectory Information in Studies of Merger

Michael Scanlon and Alicia Beckford Wassink*

1 Trajectory and Merger

Merged *pin-pen*, where the merger of /ɪ/ and /ɛ/ before nasals results in homophonous productions of *pin*/*pen*, *tin*/*ten*, and *since*/*sense*, is described as a relatively uniform feature of African American English throughout the U.S. (Labov et al. 2006, Thomas 2007). The status of merger is usually examined impressionistically (Brown 1990, Edwards 1997, Gordon 2000) or through spectral qualities of the *pin* and *pen* classes at a single measurement point (Thomas 2001).

However, vowel classes in English that appear spectrally merged may be distinguished by the trajectory of the vowel over the course of its duration. Milroy and Harris (1980) investigated the apparent merger of the vowel class of *meat* with that of *mate* in vernacular Belfast English, and argued that the two vowels show overlapping but different distributions both in vowel height and in the incidence of an inglide. Labov et al. (1991) describes a near-merger in the *too* and *toe* classes for Norwich speakers, where the nuclei of the two vowel classes show overlap along F1 and F2, but the vowel classes differ in their trajectory. Even for nominally monophthongal vowels, spectral change can be an important factor in vowel identification (see Morrison 2008). Acoustic trajectory information has not yet been utilized in studies of the *pin-pen* merger. In the present study, our initial auditory impressions were that some Seattle African Americans may differentiate the *pin* and *pen* classes by their degree of gliding.

2 Case Study: J

2.1 Overview

The data for this paper are drawn from an ongoing study of sociolinguistic variation in Seattle, Washington. The data are from a set of 18 interviews conducted by J, a 65-year-old African American raised in Yesler Terrace (YT), a multi-ethnic public housing community in Seattle. J moved to the Seattle metropolitan area from Louisiana when she was 1 year old, and moved into YT when she was 7. She resided in YT through high school, and had family ties in YT until she was 27. J is an anthropologist by profession, and self-identifies as middle-class today. Between 2006 and 2009, J interviewed subjects who spent their formative years in YT about their experience growing up in this public housing community. The data for this case study are drawn from J's speech in these interviews. J's status as interviewer allowed us to collect a large number of *pin/pen* tokens and conduct statistical analyses for the data.

J was also a participant in a separate but related research project, which fortuitously afforded an opportunity for her to provide metalinguistic commentary. (1) shows a metalinguistic comment J made about the status of the *pin/pen* merger in Seattle.

(1) J: But another thing that we do is, we don't, is that we say, we don't make a difference between [pɪn] and [pɛn] up here—it's all [pɪn].
Interviewer: For the whole—for the whole Northwest?
J: I *know* it's for this area, around here.

In e-mails exchanged after the data collection phase of this study, J clarified that she believed the two vowels were merged in the region, though the actual vowel sound may vary by ethnicity. For this reason we wanted to account for any interlocutor ethnicity effects in the data, as well as describe the trajectory of the two vowel classes.

*This presentation is made possible through NSF grant #BCS-0643374. Thanks to the English in the Pacific Northwest team, to the UW Sociolinguistics Brownbag, the UW Phonetics Lab, and to Jean Harris.

2.2 Methods

J interviewed 18 subjects (14 African American, 4 White) who were raised in Yesler Terrace, and were born between 1935 and 1955. All were either born in the Seattle area or arrived in Yesler Terrace between birth and 13 years of age, and were raised in Yesler Terrace throughout adolescence. The interviews were centered on the topic of growing up in Yesler Terrace, but were casually conducted, and the subject was allowed to stray off-topic. 128 total tokens of pre-alveolar-nasal /ɛ/ and /ɪ/ in mono- and bisyllabic words were extracted from the interviews (91 *pen*, 37 *pin*). No more than 3 tokens per word were extracted for each interview. Tokens were coded by preceding place of articulation (bilabial, labiodental, dental, alveolar, post-alveolar, velar, glottal).[1] Tokens were coded for interlocutor ethnicity (African American, Caucasian). Tokens were auditorily coded as reduced or fully diphthongal. A token was considered reduced if vowel quality change other than what would be expected due to coarticulatory effects could be discerned. Vowels were measured for duration, and for the first two formant frequencies (F1 and F2) at vowel onset, 20%, 50%, 80%, and offset. The formant measurements, independently, reflect only static features of the vowel. To reflect the trajectory, or change in vowel quality, over the duration of the vowel, measures which could reflect vowel quality change were also taken. Change in F1 (ΔF1 hereafter) and change in F2 (ΔF2) were calculated by subtracting formant values at the 20% point from formant values at the 80% point. Euclidean distance (hereafter, *distance*) was calculated using the formula in (2) (Fabricius 2007):

$$(2) \quad \sqrt{(\Delta F1)^2 + (\Delta F2)^2}$$

Since distance is a composite statistic reflecting distance between vowel measurements at two points in time, it is interpreted here as a rough measure of the degree of gliding for a vowel. Distance is thus a useful statistic in cases where it is not clear whether ΔF1 or ΔF2 differentiate a pair of vowel categories, or when it is likely that both ΔF1 or ΔF2 serve to differentiate a pair. Reduced vowels will show smaller distances, while distances will be larger for vowels that begin in one part of the acoustic vowel space and end in another.

2.3 Results

Figure 1 shows an F1-by-F2 plot of J's averaged *pin* and *pen* class data, broken down by interlocutor ethnicity. (The *pin* class is graphically represented with "in," and the *pen* class with "en.") The mean 20% point for each vowel class is represented with a symbol, and arrowheads reflect the mean 80% point for each vowel class. The mean distance for each vowel class, then, is reflected in the length and direction of the arrow. To provide a sense of the location of these vowels in the wider system, means for /ɪ/ and /ɛ/ in non-prenasal contexts and for /i/, /e/, /u/, and /a/ are also displayed for comparison.

A few observations about J's vowel plot can be made. First, the *pin* and *pen* classes both have a downward and centralizing glide. The prenasal classes are proximal to [ɛ] with African American interlocutors (symbols with open squares), but with Caucasian interlocutors (filled circles), the prenasal classes appear to be in a position between [ɪ] and [ɛ]. The *pen* class with Caucasian interlocutors appears to start in a higher and more fronted position than the other three prenasal classes. Lastly, there appears to be a difference between the /ɛ/ and /ɪ/ classes in degree of gliding; /ɛ/ shows a longer glide than /ɪ/. This difference holds prenasally as well.

[1]Tokens were also coded for number of syllables and voicing of the preceding environment (after Brown 1990), but these factors were not found to affect *pin-pen* production in the statistical analyses.

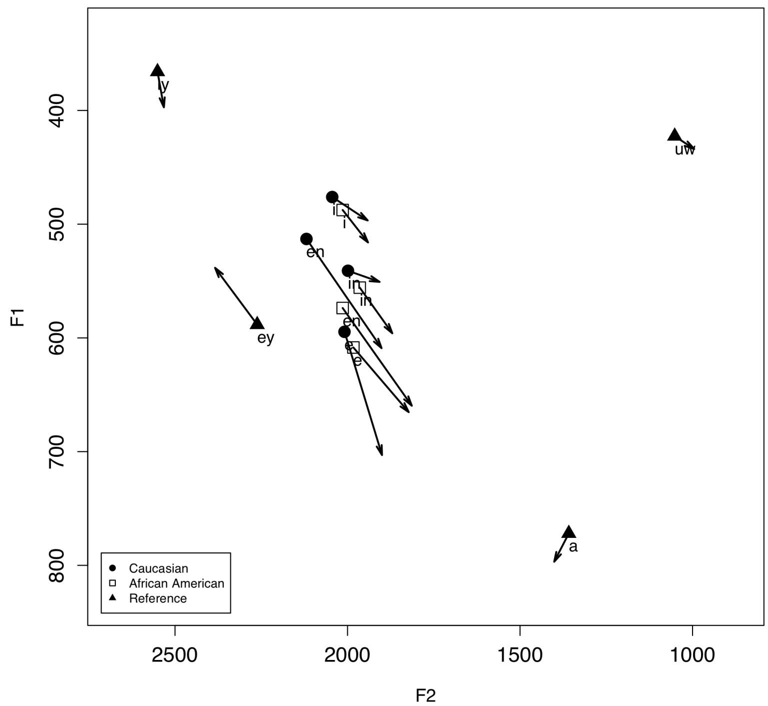

Figure 1: Plot of J's averaged *pin-pen* data, by interlocutor ethnicity.

Conventional and alternative statistical tests of merger were conducted. As a conventional test, paired t-tests were used to evaluate if there was a difference between *pin* and *pen* F1 and F2 values at midpoint (after Labov et al. 1991). Three alternative tests of merger were run. First, regressions of *pin* and *pen* F1 and F2 values at midpoint were conducted; this method is comparable to the conventional method, but controls for any environmental effects or interlocutor characteristics. Second, regressions were run using alternative acoustic measures. A regression of distance on vowel class was run, to determine if *pin* and *pen* differ in degree of gliding, controlling for environment and interlocutor ethnicity. A regression of duration on vowel class was run, to determine if *pin* and *pen* differ in duration, controlling again for environment and interlocutor ethnicity. Lastly, a regression of auditory impressions was run, to determine if there were differences in the way different vowel classes were perceived by the investigator (MS).

T-tests of F1 and F2 for *pin* and *pen* at midpoint did not show *pin* and *pen* to be significantly different, although *pen* has a marginally lower height than *pin*, at p=0.061.

For regressions of F1 and F2 at midpoint, ethnicity was a significant factor; F1 was 72 Hz lower with Caucasian interlocutors than with African American ones (p=0.012). Vowel class was not found to be significant in the midpoint regression analyses (p=0.756).[2]

The regressions of distance and duration did show vowel class as a significant factor. The *pin* class had, on average, a glide 95 Hz shorter than the *pen* class (p=0.001), and *pin* was, on average, 30 ms shorter than *pen* (p=0.001), holding other factors constant. Distance and duration appear closely related as they show a strong correlation (0.762, p=0.000). Also, there was an interactive effect on distance, as shown in Figure 2a. The solid line represents the /ɪ/ class, and the dotted line represents the /ɛ/ class. The y-axis displays the mean distance for each vowel class. Thus, the distance between the two lines represents the mean difference in distance between the two vowel classes. The x-axis serves to separate the two conditions, non-prenasal and prenasal. *Pin* and *pen* distances (on the right side of the graph) were more distinct than /ɪ/ and /ɛ/ in non-prenasal contexts (on the left side), holding other factors constant (p=0.044). The same interaction pattern

[2]Preceding place of articulation was found to be significant in the regression analyses, but since it was included only as a controlling factor, the results are not described here.

emerged in the regression of auditory impression on vowel class, as shown in Figure 2b; here, the y-axis displays the percentage of tokens that were coded as diphthongal. Nasals were more likely to be coded as diphthongal (p=0.003), and *pen* was more likely to be coded as diphthongal than *pin*, holding other factors constant (p=0.027). Interlocutor ethnicity was not found to be a significant factor in degree of gliding, duration, or auditory impression.

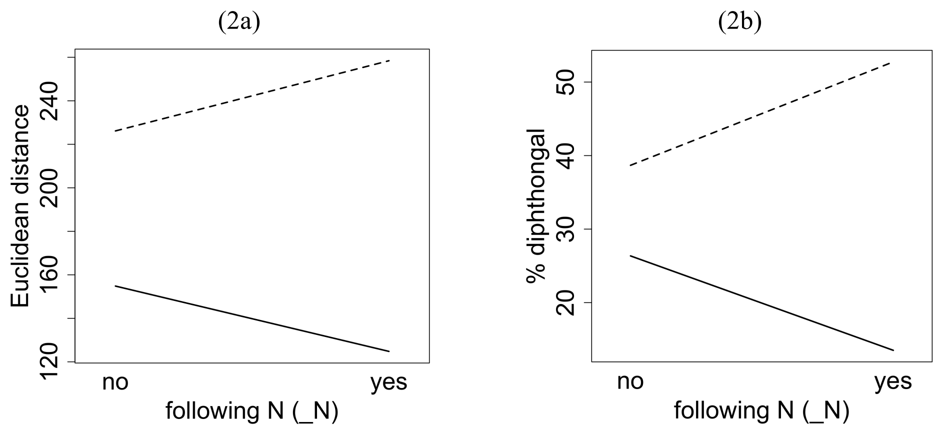

Figure 2: Interaction plot of J's /ɪ/ and /ɛ/ vowel classes.

To summarize the results of the *pin-pen* analysis: J shows higher *pin* and *pen* vowels with Caucasians than with African American interlocutors; the prenasal and non-prenasal /ɛ/ classes have a greater degree of gliding and longer duration than the prenasal and non-prenasal /ɪ/ classes; the *pin* and *pen* classes were more distinct than /ɪ/ and /ɛ/ non-prenasally, and the auditory analysis also shows greater differentiation in the prenasal condition.

2.3 Discussion

We find no difference (i.e., "merger") between *pin* and *pen* by F1 or F2 at midpoint. This finding holds even when accounting for environmental and context factors. However, spectral change can differentiate *pin* and *pen* where one-point spectral measures do not. There is strong evidence that J differentiates the *pin* and *pen* classes by degree of gliding, and that differences in spectral movement of the /ɪ/ and /ɛ/ vowels are greatest when the difference in height between the vowels is the smallest (i.e., before [n]). The acoustic difference between the *pin* and *pen* classes is robust; distance and duration are highly correlated, and both acoustic cues tend to be present simultaneously.

This study has focused on production, and we cannot make any definitive statements about J's (or other listeners') perception of the vowels in question. However, the data suggests a near-merger in production (see Labov 1994), and a commutation test or similar perception data would help test whether J in fact has a near-merger for *pin-pen*.

3 A Local African-American Pattern?

At this point, it is not clear whether the results so far are idiosyncratic to J, or whether they reflect local dialectal patterns. While it was not possible to run parametric statistical tests for each of J's interlocutors due to small token counts, it is possible to estimate the degree of overlap between the *pin* and *pen* classes for some of her interlocutors using Wassink's (2006) Spectral Overlap Assessment Metric (SOAM). SOAM estimates best-fit ellipses for vowel distributions from raw token data and calculates the proportion of multidimensional overlap between two vowel distributions based on those ellipses. (See Wassink 2006 for details on the SOAM calculations.) Figure 3a shows two-dimensional overlap between J's *pin* and *pen* classes (in green and red, respectively), based on F1 and F2 values at midpoint. J shows 94% overlap between the *pin* and *pen* classes in two dimensions. Figure 3b shows overlap between the two classes based on three dimensions: F1 at midpoint, F2 at midpoint, and distance. The overlap based on three dimensions is somewhat smaller, at 76%. The difference between any speaker's 3D overlap percentage and their 2D overlap percentage (hereafter, *change in overlap*) yields a rough measure of the contribution of spectral change to the *pin-pen* distinction for that speaker.

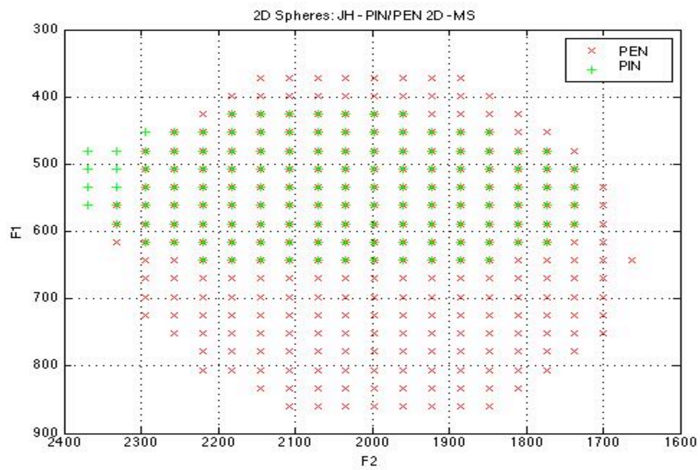

Figure 3a: F1-by-F2 plot of overlap between J's *pin* and *pen* distributions.

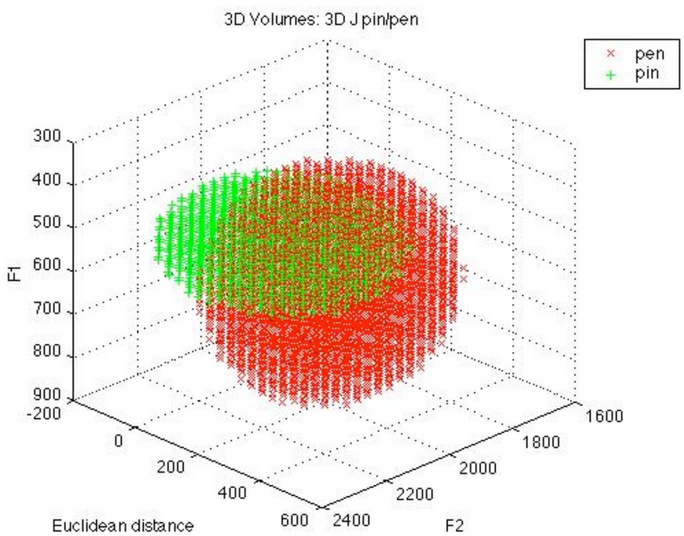

Figure 3b: F1-by-F2-by-distance plot of overlap between J's *pin* and *pen* distributions.

Figure 4 shows the change in overlap for 12 speakers from Yesler Terrace. A positive change in overlap means less overlap with the inclusion of the distance variable. The change in overlap ranges from -3 to 21 percent. The majority of African American speakers (7 out of 9, including J) show a sizable change in overlap, while the 3 Caucasian speakers show change in overlap comparable to zero. This suggests that African Americans in the community differentiate *pin* and *pen* by their trajectories, while Caucasians in the community do not. While this analysis uses a small sample, it suggests that J's pattern of *pin-pen* production may be representative of a local and distinct African American pattern.

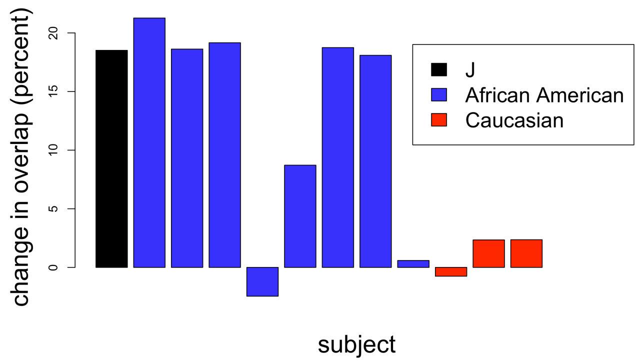

Figure 4: Change in overlap for 12 Seattle speakers.

4 Conclusions

This study has demonstrated that acoustic trajectory information can be used to distinguish the *pin* and *pen* classes, and that, at least from the perspective of the researchers, this difference is perceptible. This study's findings emphasize the utility of examining acoustic trajectory information indicative of gliding in the case of mergers or near-mergers, and question previous reports of fully merged *pin-pen*. The study also gives a preliminary account of *pin-pen* patterns for Seattle African American speakers.

References

Brown, Vivian R. 1990. Phonetic constraints on the merger of /ɪ/ and /ɛ/ before nasals in North Carolina and Tennessee. *The SECOL Review* 14:87–100.

Edwards, Walter F. 1997. The variable persistence of Southern vernacular sounds in the speech of inner-city Black Detroiters. In *Language Variety in the South Revisited*, ed. C. Bernstein, T. Nunnally, and R. Sabino, 76–86. Tuscaloosa, AL: University of Alabama Press.

Gordon, Matthew J. 2000. Phonological correlates of ethnic identity: Evidence of divergence? *American Speech* 75:115–136.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The Atlas of North American English: Phonetics, Phonology, and Sound Change.* New York: Mouton de Gruyter.

Labov, William, Mark Keren, and Corey Miller. 1991. Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3:33–74.

Milroy, James, and John Harris. 1980. When is a merger not a merger? The MEAT/MATE problem in a present-day English vernacular. *English World-Wide* 1:199–210.

Morrison, Geoffrey Stewart. 2008. Theories of vowel inherent spectral change: A review. Ms., University of Alberta, Edmonton.

Thomas, Erik R. 2007. Phonological and phonetic characteristics of African American Vernacular English. *Language and Linguistics Compass* 1:450–475.

Thomas, Erik R. 2001. *An Acoustic Analysis of Vowel Variation in New World English.* Publication of the American Dialect Society 85. Durham, NC: Duke University Press.

Wassink, Alicia Beckford. 2006. A geometric representation of spectral and temporal vowel features: Quantification of vowel overlap in three linguistic varieties. *Journal of the Acoustical Society of America* 119:2334–2350.

Department of Linguistics
University of Washington
Seattle, WA 98115
ms27@uw.edu
wassink@uw.edu