Lexical Neighborhoods and Phonological Confusability in Cross-Dialect Word Recognition in Noise

Cynthia G. Clopper, Ohio State University
Janet B. Pierrehumbert, Northwestern University
Terrin N. Tamati, Indiana University

Abstract
Lexical neighborhood density is a well-known factor affecting phonological categorization in spoken word recognition. The current study examined the interaction between lexical neighborhood density and dialect variation in spoken word recognition in noise. The stimulus materials were real English words in two regional American English dialects. To manipulate lexical neighborhood density, target words in the word-competitor condition were selected so that predicted phonological confusions across dialects resulted in real English words. The target words in the nonword-competitor condition were selected so that predicted phonological confusions did not result in real English words. Word and vowel recognition performance were more accurate in the nonword-competitor condition than the word-competitor condition for both talker dialects. An examination of the responses to specific vowels revealed the role of dialect variation in eliciting this effect. For example, in the word-competitor condition, Northern /ɛ/ was misidentified as /æ/ significantly more often than chance, consistent with the Northern backing and lowering of /ɛ/. Thus, when the predicted phonological confusions were real lexical neighbors, listeners could respond with either the target word or the confusable minimal pair. When the predicted phonological confusions were not real words, the listeners exhibited less lexical competition and responded with the target word.

1. Introduction

Lexical neighborhood density is one of the well-known factors affecting spoken word recognition. Words with many lexical neighbors, differing by one phoneme insertion, deletion, or substitution, are more difficult to recognize than words with few lexical neighbors (Luce and Pisoni 1998). This effect of lexical neighborhood is pervasive in word recognition studies and has been observed for normal-hearing native listeners (Dirks et al. 2001; Sommers et al. 1997; Takayanagi et al. 2002), normal-hearing non-native listeners (Takayanagi et al. 2002), hearing-impaired native listeners (Dirks et al. 2001; Kirk et al. 1997; Sommers et al. 1997; Takayanagi et al. 2002), and hearing-impaired non-native listeners (Takayanagi et al. 2002). In addition, the lexical neighborhood effect is limited to open-set word recognition tasks and closed-set tasks in which many phonologically similar words are presented as the alternatives (Clopper et al. 2006), providing further support for the interpretation of the effect as resulting from competition in lexical access between phonologically similar words (Luce and Pisoni 1998).

In Luce and Pisoni's (1998) Neighborhood Activation Model, phonological neighbors are defined as words that differ by one phoneme insertion, deletion, or substitution. However, not all segments are equally perceptually confusable with one another: more phonetically similar segments tend to be more perceptually confusable than phonetically more different segments. For example, Peterson and Barney's (1952) vowel confusion matrix reveals many confusions between the low back vowels /ɑ/ and /ɔ/, but no confusions between low front /æ/ and high back /u/ (see also Cutler et al. 2004; Hillenbrand et al. 1995). Similarly, Miller and Nicely's (1955) consonant confusion matrices reveal many more confusions between the voiceless stops /p/ and /t/ than between the voiceless stop /p/ and the nasal /m/ at all but the most difficult signal-to-noise ratio (see also Benkí 2003; Cutler et al. 2004; Wang and Bilger 1973). Thus, the lexical neighborhood density effect may be driven by specific error patterns between perceptually confusable segments. In addition, vowels may be perceptually more confusable than consonants in tasks that involve lexical access. For example, Van Ooijen (1996) found that listeners were more likely to produce false alarms involving vowel changes than consonant changes in an auditory word-spotting task. Similarly, listeners are more likely to ignore vowel differences than consonant differences in lexical learning tasks (Bonatti et al. 2005). These results suggest that lexical competition effects may be stronger for phonological neighbors differing by one vowel than for phonological neighbors differing by one consonant.

In American English, vowel differences across dialects involve mostly phonetic shifts in the vowel space that could potentially increase the perceptual confusability between some pairs of vowels. Vowels that belong to different phonological categories in two different dialects, but that are phonetically similar to each other, should be perceptually confusable (see Best 1995). Rakerd and Plichta (2003) provide some evidence for the effect of dialect variation on perceptual vowel categories. They found that listeners shifted their phonetic category boundary for a synthetic continuum from /ɑ/ to /æ/ depending on the dialect of the preceding carrier phrase. For carrier phrases with fronted /ɑ/s, the perceived category boundary was closer to /æ/, whereas for carrier phrases with backed /ɑ/s, the boundary was closer to /ɑ/ (c.f., Kraljic et al. 2008; Stevens et al. 2007, where isolated words did not elicit phonetic category boundary shifts). While this finding is based on the perception of synthetically produced stimulus materials, it suggests that listeners are sensitive to phonetic vowel differences across dialects and that dialect variation can affect phonological category boundaries and, by extension, perceptual vowel confusability. When the category boundary is closer to the /æ/ end of the continuum, the phonetic difference between /æ/ and (fronted) /ɑ/ is smaller than when the category boundary is closer to the /ɑ/ end of the continuum, which may lead to greater perceptual confusability of /ɑ/ and /æ/. Thus, perceptual confusions between specific pairs of segments may differ across dialects, and lexical neighborhood effects may be stronger for some pairs of segments within dialects (in the case of a vowel merger) or across dialects (in the case of a vowel shift) than for other pairs of more phonetically different segments.

Listener experience with dialect variation has also been shown to affect the perception of phonological information across dialects. Clopper and Bradlow (2008) observed significant effects of dialect familiarity on sentence transcription performance in noise. Listeners more accurately transcribed sentences produced in a familiar dialect than in an unfamiliar dialect (see also Mason 1946). Similarly, Adank and McQueen (2007) and Labov and Ash (1997) found that listeners exhibited faster and more accurate word comprehension performance for words produced in a familiar dialect than in an unfamiliar dialect. The results of these cross-dialect speech intelligibility tasks suggest that the phonetic vowel shifts in unfamiliar dialects lead to phonological confusions, which, in turn, lead to slower and less accurate lexical processing. This interpretation is further supported by the results of lexical decision tasks across dialects. Floccia et al. (2006) reported that lexical decision performance was slower and less accurate for unfamiliar dialects than familiar dialects, suggesting that phonological ambiguity due to unfamiliar vowel shifts also leads to slower and less accurate lexical decision judgments (see also Sumner and Samuel 2009). Thus, more perceptual confusions are expected for shifted vowels in less familiar dialects that are phonetically similar to vowels in more familiar dialects than for shifted or merged vowels in more familiar dialects.

The goal of the current study was to examine the interaction between lexical neighborhood effects and naturally-occurring dialect variation in a spoken word recognition task. Based on previous descriptions of regional dialect variation in the United States (e.g., Labov et al. 2006), a set of predicted phonological confusions was established for vowels in two dialects of American English (Midland and North). In one condition of the word recognition task, the target words were selected so that the predicted phonological confusions resulted in real English word neighbors (the word-competitor condition). In the other condition, the target words were selected so that the predicted phonological confusions did not result in real English words (the nonword-competitor condition). The listeners included lifetime residents of the two target dialects, as well as childhood residents of multiple different dialect regions. If dialect variation leads to greater perceptual confusability between some vowel pairs for one of the two dialects, we expect to observe a three-way interaction between lexical condition, talker dialect, and target vowel, such that the effect of lexical condition is stronger for some vowel pairs for one dialect than the other. In addition, if listener familiarity with the talker dialects affects perceptual confusability, we expect to observe a four-way interaction between lexical condition, talker dialect, listener dialect, and target vowel.

2. Methods

2.1. Listeners

Fifty-five Ohio State University undergraduates were recruited to participate in the word recognition experiment. Data from sixteen of the participants were excluded from the analysis: seven participants were bilingual, four participants reported a history of a hearing or speech disorder, two participants were substantially older than the other participants, and three sets of data were not recorded due to a computer error. The remaining 39 listeners were all monolingual native speakers of American English between the ages of 17 and 28 years old with no reported history of hearing or speech disorders. The residential histories of the listeners varied. Twenty of the listeners were lifetime residents of the Midland dialect region, one listener was a lifetime resident of the New England dialect region, nine listeners were lifetime residents of the Northern dialect region, and the remaining nine listeners had lived in more than one dialect region before age 18. The lifetime Midland and New England residents were combined to form the General American listener group[1]. The lifetime Northern residents comprised the Northern listener group, and the listeners who had lived in more than one dialect region before age 18 formed the Mobile listener group. Nineteen of the listeners participated in the word-competitor condition. The other 20 listeners participated in the nonword-competitor condition. All of the listeners received partial course credit in an introductory linguistics course for their participation.

2.2. Predicted phonological confusions

Two regional dialects of American English, Midland and North, were selected for the current study. Both dialects are spoken in the state of Ohio and are, therefore, familiar to the Ohio State University undergraduates who participated as listeners. However, the Northern dialect should be more familiar than the Midland dialect for the Northern listeners (and vice versa for the General American listeners), given their prolonged exposure to their native dialect and more recent exposure to the second dialect. A map depicting the two dialect regions is shown in Figure 1. The state of Ohio is indicated by a star in the figure. Schematics of the vowel systems of the Midland and Northern dialects are shown in Figure 2. The Midland dialect is characterized by fronting of the back vowels /u/ and /ow/ and a merger of the low back vowels /ɑ/ and /ɔ/ (Labov et al. 2006). The Northern dialect is characterized by the Northern Cities Chain Shift, which includes the fronting and raising of /æ/, fronting and lowering of /ɑ/ and /ɔ/, backing and/or lowering of /ɛ/ and /ɪ/, and backing of /ʌ/ (Labov et al. 2006). An initial acoustic analysis of the stimulus materials also suggested lowering of /ʊ/ in the Northern dialect (see also Labov et al. 2006).

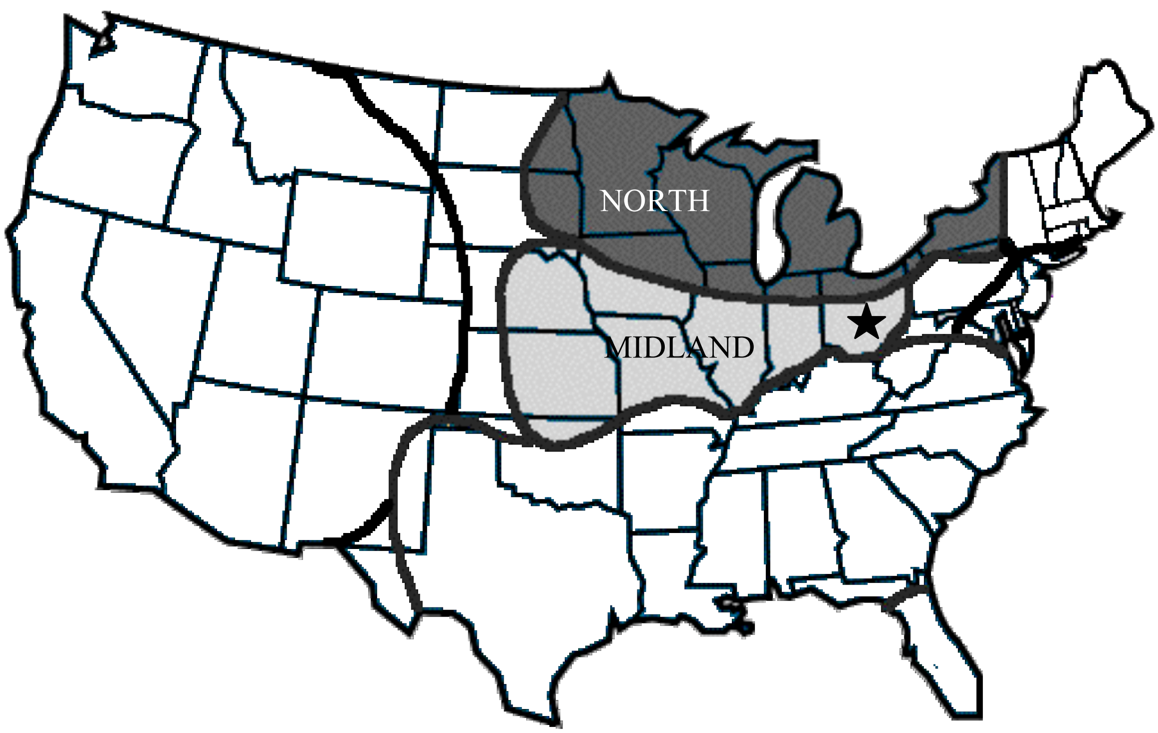

Figure 1. Map of the Northern (dark gray) and Midland (light gray) dialect regions in the United States. The state of Ohio is indicated by the black star.

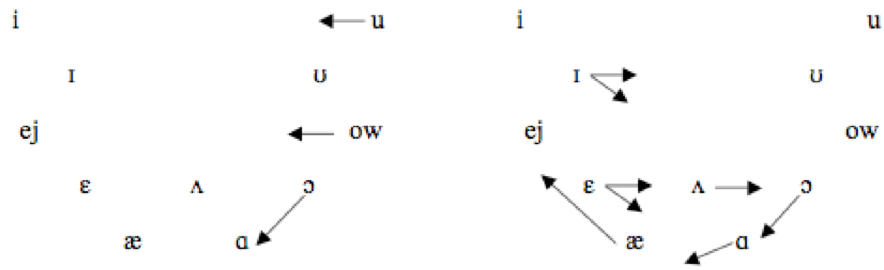

Figure 2. Schematics of the Midland (left) and Northern (right) vowel systems.

Based on this characterization of these two dialects, a set of predicted phonological confusions was established. The baseline for comparison for these predicted confusions was assumed to be the "standard" variety of American English, depicted in Figure 2 by the location of the phonetic vowel symbols. This strong assumption is based the results of previous cross-dialect speech perception studies, which suggest that listeners use the standard variety as their perceptual filter in speech processing tasks. Clopper and Bradlow (2008) and Floccia et al. (2006) observed a processing benefit for the standard variety in their sentence transcription and lexical decision tasks, respectively. In addition, Labov and Ash (1997) found that even native speakers of Southern American English could identify extremely Southern-shifted vowels with only moderate accuracy. Thus, the predicted phonological confusions in the current study were based on a comparison to the standard description of the American English vowel system.

A summary of the predicted phonological vowel confusions is shown in Table 1. For the Midland dialect, /ɔ/ was predicted to be confusable with /ɑ/, and /ow/ was predicted to be confusable with /ʊ/. While /u/ is fronted in the Midland dialect, it is typically not fronted enough to be confusable with /i/ or /ɪ/. Thus, no confusions were predicted for the Midland fronted /u/. For the Northern dialect, /ɪ/ was predicted to be confusable with /ɛ/, /ɛ/ with /æ/, /æ/ with /ɛ/, /ɑ/ with /æ/, /ʌ/ with /ɔ/, /ɔ/ with /ɑ/, and /ʊ/ with /ʌ/. Note that the predicted phonological confusions were based on comparisons of only first and second formant frequencies and that vowel duration and formant trajectory were not taken into consideration. Thus, these predictions are based on a very rough assessment of phonetic similarity across vowel categories. This approach was based on the previous literature on regional vowel variation in American English, which tends to focus almost exclusively on first and second formant frequencies for describing variation, with much less attention paid to duration and formant trajectory variation across dialects or vowels (e.g., Clopper et al. 2005; Labov et al. 2006; Thomas 2001). In addition, by setting aside considerations of duration and formant trajectory, we were able to obtain a broad overview of the role of dialect variation in producing vowel confusions in word recognition in noise.

Table 1. Predicted phonological vowel confusions for the Midland and Northern dialects relative to the “standard” variety of American English.

Midland	North
ɔ → ɑ	ɪ → ɛ
ow → ʊ	ɛ → æ
	æ → ɛ
	ɑ → æ
	ʌ → ɔ
	ɔ → ɑ
	ʊ → ʌ

2.3. Talkers

Six female talkers were selected for the word recognition experiment from the Indiana Speech Project Corpus (Clopper et al. 2002). The Indiana Speech Project corpus contains recordings of five college-aged women from each of six regions of Indiana. Each of the women was recorded producing isolated words and nonwords, sentences, passages of connected speech, and in an interview. The talkers selected for the word recognition study were all monolingual native speakers of American English and ranged in age from 19 to 21 years old at the time of recording. Three of the talkers were classified as representatives of the Northern dialect. They had lived exclusively in the Northern dialect region until age 18 and both parents of each of the Northern talkers were also raised in the Northern dialect region. The other three talkers were classified as representatives of the Midland dialect. They had lived exclusively in the Midland dialect region until age 18 and both parents of each of the Midland talkers were also raised in the Midland dialect region.

2.4. Stimulus materials

Fifty-five real English monosyllabic words were selected as stimulus materials for the word-competitor condition. Five different words were selected for each of 11 vowels (i, ɪ, ej, ɛ, æ, ɑ, ɔ, ʌ, ow, ʊ, u). For the eight vowels for which phonological confusions were predicted based on the descriptions of the Midland and Northern dialects (ɪ, ɛ, æ, ɑ, ɔ, ʌ, ow, ʊ), the target words were selected so that the predicted phonological confusions resulted in real English words. For example, all of the selected words containing /ɛ/ had a real minimal pair competitor in English containing /æ/ (e.g., *bet* and *bat*). The competitor words did not appear in the stimulus list. Target words were selected to control for mean lexical frequency and familiarity across vowels and between target words and their phonological competitors. Target and competitor words were not significantly different in mean lexical frequency or familiarity, except for target /ɔ/ words, which were less frequent and less familiar on average than their /ɑ/ competitors, given the relatively small number of CɔC words in English and the relatively low familiarity and frequency of those words. Overall, the target words were highly familiar (mean on a 7-point scale = 6.94; Nusbaum et al. 1984) and moderately frequent (mean log frequency per million words = 2.38). The initial consonant of the target words varied, and included stops, affricates, fricatives, nasals, liquids, and glides. The final consonant of the target words was restricted to obstruents, to avoid known vowel variation in pre-nasal and pre-lateral position. Due to lexical gaps in the English lexicon, two of the target /ow/ words (*bowl*, *pole*) included a final /l/ because no other /ow/~/ʊ/ minimal pairs met our lexical frequency and familiarity requirements. A complete list of the word-competitor condition target words, and their lexical frequency, familiarity, and density, as well as the competitors, and their lexical frequency and familiarity is provided in Tables A1and A2 in the Appendix.

Forty additional real English monosyllabic words were selected for the nonword-competitor condition. For the three vowels for which phonological confusions were not predicted across the two dialects (i, ej, u), the same words were used in the nonword-competitor condition and the word-competitor condition. For the eight vowels for which phonological confusions were predicted across the two dialects, five new words containing each vowel were selected for the nonword-competitor condition so that the predicted phonological confusions did not result in real English words. For example, all of the selected words containing /ɛ/ did not have a real minimal pair competitor in English containing /æ/ (e.g., *chess* and **chass*). As in the word-competitor condition, the target words in the nonword-competitor condition were highly familiar (mean on a 7-point scale = 6.92) and moderately frequent (mean log frequency per million words = 2.15). Independent sample t-tests confirmed that the stimulus materials in the two conditions did not differ significantly in terms of lexical frequency or familiarity. However, the target words in the nonword-competitor condition had significantly fewer neighbors on average than the target words in the word-competitor condition ($t(78) = -4.5$, $p < .001$). In the nonword-competitor condition, all of the final consonants in the target words were obstruents. A complete list of the nonword-competitor condition target words, and their lexical frequency, familiarity, and density is provided in Tables A1 and A3 in the Appendix.

All of the selected target words in each condition were produced by all three talkers from both dialects, for a total of 330 stimulus items per condition. The stimulus materials were segmented into individual digital sound files at a sampling rate of 22050Hz with 16-bit resolution. The sound files were mixed with speech-shaped white noise at a signal-to-noise ratio of +2dB for presentation to the listeners.

2.5. Acoustic analysis of the stimulus materials

The first and second formant frequencies of the vowel in each target word produced by each of the six talkers were extracted from the temporal midpoint of the vowel. The mean formant frequencies for each vowel category are shown in Figure 3 for each talker dialect (circles for Midland and triangles for North) and each lexical condition (open for the nonword-competitor condition, filled for the word-competitor condition, and gray for the target words that occurred in both conditions). The Northern talkers produced somewhat lowered and backed /ɪ, ɛ/, raised /æ/, mildly backed /ʌ/, and lowered /ʊ/ relative to the hypothetical "standard" variety of American English. The Midland talkers produced raised /ʌ/, mildly fronted /ow/, and lowered /i, ej/ relative to the hypothetical "standard" variety. Talkers from both dialects produced fronted /uw/.

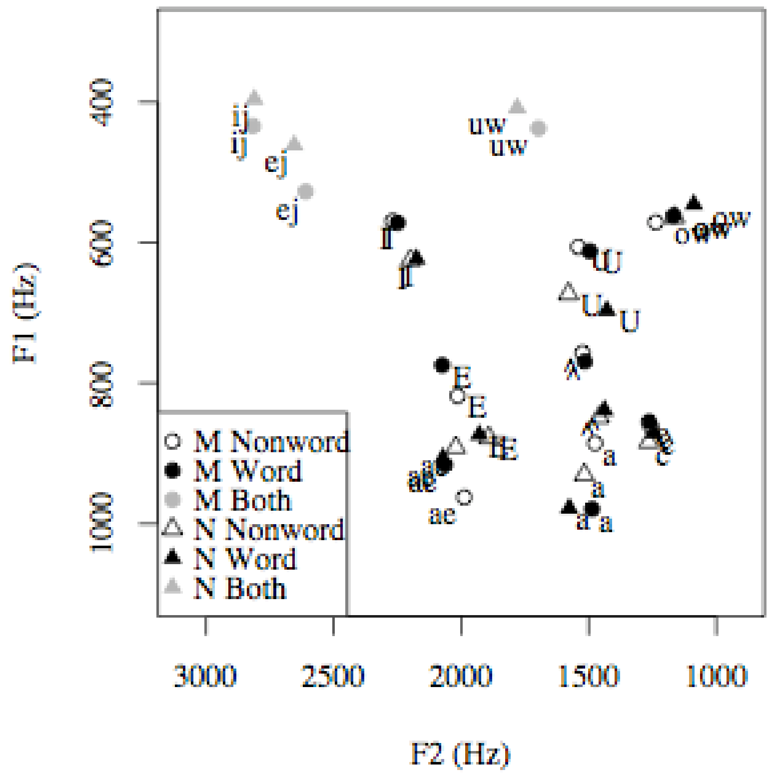

Figure 3. Mean first and second formant frequencies of the target vowels in the stimulus words in each condition for each talker dialect. The non-IPA vowel symbols correspond to IPA vowel symbols as follows: I = ɪ, E = ɛ, ae = æ, a = ɑ, c = ɔ, ^ = ʌ, U = ʊ.

2.6. Word recognition procedure

The listeners were seated at personal computers equipped with headphones and a keyboard. In each of the two conditions, the listeners were presented with each of the 55 target words produced by each of the six talkers, for a total of 330 trials[2]. They were instructed to listen to each word and to type the word that they heard onto the computer screen. They were permitted to listen to each stimulus item exactly one time before responding and were asked to make their best guess on every trial. The stimulus materials were presented in a different fully randomized order for each listener.

2.7 Statistical analysis of the word recognition results

The responses were scored for both correct word and correct vowel. For example, the response *code* to the target word *code* was scored as correct for both word and vowel, whereas the response *coat* for the target word *code* was scored as incorrect for word but correct for vowel. Homophones (e.g., *need* for *knead*) and obvious typographical and spelling errors were corrected (e.g., *yaght* for *yacht*). Multisyllabic words, nonsense words, heteronyms (e.g., *read*), and words with multiple possible pronunciations (e.g., *mauve*) were scored as incorrect because the vowel intended by the listener could not be uniquely determined.

Word and vowel accuracy were analyzed by subjects and by items using repeated measures ANOVAs, with four independent factors in each analysis: lexical condition (word-competitor, nonword-competitor), talker dialect (Midland, North), listener dialect (General American, North, Mobile), and target vowel (i, ɪ, ej, ɛ, æ, ɑ, ɔ, ʌ, ow, ʊ, u). In the subject analyses, lexical condition and listener dialect were treated as between-subject variables and talker dialect and target vowel were treated as within-subject variables. Given that some items occurred in both conditions, and other items occurred in only one condition, the item analyses were split into within-item analyses and between-item analyses. In the within-item analyses, target vowel (i, ej, uw) was treated as a between-item factor and lexical condition, talker dialect, and listener dialect were treated as within-item factors. In the between-item analyses, target vowel (ɪ, ɛ, æ, ɑ, ɔ, ʌ, ow, ʊ) and lexical condition were treated as between-item factors and talker dialect and listener dialect were treated as within-item factors. The alpha level for the four-factor analyses was set at $\alpha = .05$, and the alpha level for the post-hoc analyses of the significant four-way interactions was set at $\alpha = .005$ to correct for multiple comparisons. When either the subject or item analysis was significant at the $\alpha = .005$ level, results are also reported for the other analysis at the less conservative $\alpha = .05$ level.

The statistical power of the item analyses is low, with only five target words per vowel. However, substantially increasing the item sample size was not possible for many of the phonological vowel confusions. As shown in Table 2, the total number of CVC minimal pairs for three of the predicted phonological vowel confusions is less than 20 when all final consonants are included. When the final consonant is limited to obstruents, the total population from which to draw is even smaller.

Table 2. CVC minimal pair counts in English for each of the predicted phonological confusions for all final consonants and for final obstruents only.

Phonological Confusion	All Final Consonants	Final Obstruents Only
ɪ-ɛ	52	24
æ-ɛ (= ɛ-æ)	40	30
ʌ-ɔ	28	16
ɔ-ɑ	15	12
ɑ-æ	41	39
ow-ʊ	11	5
ʊ-ʌ	7	7

The correlation between the lexical frequency of the target word and word recognition accuracy was not significant in either competitor condition ($r^2 = .14$ in the word-competitor condition, $r^2 = .04$ in the nonword competitor condition). A repeated measures ANOVA on lexical frequency with lexical condition as a between-subject variable and word type (target vs. response) as a within-subject variable revealed a significant main effect of word type ($F(1, 105) = 26.7, p < .001$). The responses had significantly higher lexical frequencies on average than the target words. However, the main effect of lexical condition and the interaction were not significant. Therefore, lexical frequency was not included as a covariate in the subject and item analyses.

In addition to the subject and item analyses of word and vowel accuracy, the error patterns were also analyzed to explore the specific phonological confusions that the listeners exhibited. First, the actual phonological neighbors of the target words were examined to determine the proportion of one-consonant-substitution neighbors and one-vowel-substitution neighbors. Responses that differed by one segment from the target words were then analyzed using goodness-of-fit chi-square tests to determine how the proportions of vowel and consonant confusions in the responses compared to the expected values based on the English lexicon. Second, for each lexical condition, stimulus-response vowel confusion matrices were constructed for each talker dialect and each listener dialect based on the target and response vowels for each listener. Vowel error patterns that were significantly different from chance out of all scored responses were identified using a binomial distribution analysis.

3. Results

Average word and vowel accuracy scores for each talker dialect in each lexical condition are shown in Table 3. At the word level, performance was more accurate for the Northern talkers than the Midland talkers across both lexical conditions. In addition, word recognition performance was more accurate in the nonword-competitor condition than the word-competitor condition for both talker dialects. The main effects of talker dialect and lexical condition were significant for word accuracy in the subject analysis (talker dialect: $F1(1, 33) = 16.2$, $p < .001$, lexical condition: $F1(1, 33) = 18.7$, $p < .001$) and the between-item analysis (talker dialect: $F2$(1, 64 = 4.5, $p = .04$, lexical condition: $F2(1, 64) = 5.5$, $p = .02$). The effects of talker dialect and lexical condition were not significant in the within-items analysis, and the lexical condition x talker dialect interaction was not significant in any of the three analyses.

Table 3. Percent correct words and vowels in each condition for each talker dialect.

	Words		**Vowels**	
Condition	**Midland**	**North**	**Midland**	**North**
Word-competitor	64	67	82	81
Nonword-competitor	71	74	87	88

At the vowel level, performance did not differ across the two talker dialects in either condition. However, vowel recognition performance was more accurate in the nonword-competitor condition than the word-competitor condition for both talker dialects. The main effect of talker dialect was not significant for vowel accuracy in the subject analysis or either of the item analyses, but the main effect of lexical condition was significant for vowel accuracy in the subject and between-item analyses ($F1(1, 33) = 32.5$, $p < .001$, $F2(1, 64) = 8.6$, $p = .005$). The effect of lexical condition was not significant in the within-item analysis. The lexical condition x talker dialect interaction was not significant in any of the three analyses.

The significant main effect of talker dialect on word recognition performance, but not vowel recognition performance, suggests that while the Northern talkers were more intelligible overall, this intelligibility benefit may have been limited to consonants. The significant effect of lexical condition on both word and vowel recognition performance is consistent with the lexical neighborhood effect. When the predicted phonological confusions were real words, listeners could respond with either the target word or the confusable minimal pair neighbor. When the predicted phonological confusions were not real words, however, the number of potentially competing responses was more limited and the listeners were more likely to respond with the target word. Crucially, the lexical condition effect was not significant in the within-item analysis that compared performance on the same lexical items across the two lexical conditions.

In the nonword-competitor condition, 77 of the 604 responses that differed by one phoneme from the target word involved a vowel substitution, which is significantly fewer vowel substitutions than would be expected based on the actual neighbors of the target words ($\chi^2 = 48.4$, $df = 1$, $p < .001$). Given that the noise in the word recognition experiment was speech-shaped white noise, which may disproportionately mask stop burst and fricative information relative to vowel information, it is not surprising that more consonant errors were observed than expected. However, in the word competitor condition, 323 of the 914 responses that differed by one phoneme from the target word involved a vowel substitution, which is significantly more vowel substitutions than would be expected based on the actual neighbors of the target words ($\chi^2 = 12.4$, $df = 1$, $p < .001$). This finding provides further evidence for the specificity of the lexical competition effect. When the predicted vowel confusions were real words, listeners were more likely than expected to produce a minimal pair differing by one vowel. When the predicted vowel confusions were not real words, however, listeners were more likely than expected to produce a minimal pair differing by one consonant.

The results of the subject and item analyses on vowel accuracy also suggest that this effect of lexical condition is mediated by effects of talker dialect, listener dialect, and target vowel. In the subject analysis, the main effect of target vowel ($F1(10, 330) = 86.4$, $p < .001$), the lexical condition x target vowel interaction ($F1(10, 330) = 13.1$, $p < .001$), the talker dialect x target vowel interaction ($F(10, 330) = 14.9$, $p < .001$), the lexical condition x talker dialect x target vowel interaction ($F1(10, 330) = 6.1$, $p < .001$), the talker dialect x listener dialect x target vowel interaction ($F1(20, 330) = 2.0$, $p = .007$), and the lexical condition x talker dialect x listener dialect x target vowel interaction ($F1(20, 330) = 1.6$, $p = .046$) were all significant. No other main effects or interactions were significant in the subject analysis. In the between-item analysis, the main effect of target vowel ($F2(7, 64) = 7.1$, $p < .001$), the main effect of listener dialect ($F2(2, 128) = 6.4$, $p = .002$), the lexical condition x listener dialect interaction ($F2(2, 128) = 10.0$, $p < .001$), the lexical condition x listener dialect x target vowel interaction ($F2(14, 128) = 2.0$, $p = .02$), and the lexical condition x talker dialect x listener dialect x target vowel interaction ($F2(14, 128) = 2.1$, $p = .02$) were all significant. No other interactions were significant in the item analysis. In the within-item analysis, no significant main effects or interactions were observed.

The significant main effect of target vowel in the subject and item analyses suggests that some vowel contrasts are more confusable overall than others. This finding is consistent with previous research on vowel perception in American English (e.g., Hillenbrand et al. 1995; Peterson and Barney 1952), and will not be analyzed further. Post-hoc paired-sample t-tests on the listener dialect effect in the item analysis revealed that performance by the Northern listeners was significantly better than performance by either the General American ($t(159) = 2.7$, $p = .007$) or the Mobile ($t(159) = 2.9$, $p = .004$) listeners. The difference between the General American and Mobile listeners was not significant. However, the significant four-way interaction in the subject and between-item analyses suggests that these main effects of target vowel and listener dialect were also mediated by other factors. To unpack the four-way interaction, post-hoc subject and item repeated measures ANOVAs were conducted for each target vowel with lexical condition, talker dialect, and listener dialect as factors.

3.1. Main effects of lexical condition

The main effect of lexical condition was significant in the subject analysis for /ɪ/ ($F1(1, 33) = 129.6$, $p < .001$), /ɛ/ ($F1(1, 33) = 36.3$, $p < .001$), /æ/ ($F1(1, 33) = 12.6$, $p = .001$), /ɔ/ ($F1(1, 33) = 19.8$, $p < .001$), and /ow/ ($F1(1, 33) = 28.8$, $p < .001$). The main effect of lexical condition was significant in the item analysis at the less conservative $\alpha = .05$ level for /ɪ/ ($F2(1, 8) = 7.3$, $p = .03$), /ɛ/ ($F2(1, 8) = 9.0$, $p = .02$), and /ow/ ($F2(1, 8) = 6.7$, $p = .03$). For all five of these vowels, accuracy was higher in the nonword condition than the word condition. Thus, the main effect of lexical condition was driven primarily by these five vowels. Given that the target words for /i, ej, u/ did not differ across the two lexical conditions, the lack of a significant effect of lexical condition for these target vowels is not surprising. However, the remaining three vowels /ɑ, ʌ, ʊ/ were predicted to exhibit the lexical condition effect. As discussed below, the lexical condition x talker dialect interaction was significant for /ɑ/, which may explain the lack of a main effect of condition for that vowel. The results of the acoustic analysis of the stimulus materials shown in Figure 3 reveal that lexical condition had very little effect on the production of /ʌ/ for both talker dialects or the production of /ʊ/ for the Midland talkers. These minimal acoustic differences may have led to reduced lexical condition effects for these target vowels.

3.2. Main effects of talker dialect

The main effect of talker dialect was significant in the subject analysis for /ej/ ($F1(1, 33) = 28.2$, $p < .001$), /ɛ/ ($F1(1, 33) = 37.6$, $p < .001$), /ɑ/ ($F1(1, 33) = 16.2$, $p < .001$), and /ow/ ($F1(1, 33) = 41.1$, $p < .001$). The main effect of talker dialect was also significant in the item analysis for /ɛ/ ($F(1, 8) = 17.0$, $p = .003$). Accuracy was higher for the Midland talkers than the Northern talkers for /ɛ/ and /ɑ/, but accuracy was higher for the Northern talkers than the Midland talkers for /ej/ and /ow/. As shown in Figure 3, /ɛ/ is lower for the Northern talkers than the Midland talkers in both lexical conditions, consistent with the Northern Cities Chain Shift and the predicted phonological confusion of /ɛ/ with /æ/. The analysis of the error patterns revealed significantly more /ɛ/ to /æ/ confusions for the Northern talkers (N = 75) than would be expected by chance, whereas the errors for /ɛ/ for the Midland talkers were more randomly distributed, and included only three /æ/ responses. The main effect of talker dialect was mediated by lexical condition for /ɑ/, and will be discussed in more detail below. The main effect of talker dialect for /ej/ was not predicted based on previous descriptions of these dialects, and the error patterns do not reveal any significant confusions for either dialect. However, an inspection of Figure 3 suggests that the Midland /ej/ may be more confusable with /ɪ/ than the Northern /ej/, and more /ej/ to /ɪ/ confusions were observed for the Midland talkers (N = 25) than the Northern talkers (N = 2). Finally, the main effect of talker dialect for /ow/ is consistent with the predicted phonological confusion due to Midland back vowel fronting. However, as shown in Figure 3, fronting of /ow/ among the Midland talkers was fairly modest, and across the two lexical conditions, no significant error patterns were observed for either the Midland or the Northern talkers.

3.3. Lexical condition x talker dialect interactions

The lexical condition x talker dialect interaction was significant in the subject analysis for /ɛ/ ($F1(1, 33) = 10.4$, $p = .003$) and /ɑ/ ($F1(1, 33) = 13.0$, $p = .001$). This interaction was not significant in any of the post-hoc item analyses, due to the low overall power of the lexical condition factor by item. As shown in Figure 4, the effect of lexical condition is stronger for the Northern talkers than the Midland talkers for /ɛ/. In addition, vowel accuracy for /ɛ/ is essentially at ceiling for the Midland talkers in both lexical conditions and for the Northern talkers in the nonword-competitor condition. This finding is consistent with our hypothesis that the acoustically more confusable Northern /ɛ/ will lead to more phonological confusions when the competitor is a real word neighbor than when it is a nonword, despite very similar acoustic profiles for the Northern /ɛ/ in the two lexical conditions. An inspection of the error patterns provides further evidence for this interpretation: for the Northern talkers, a significant number of /ɛ/ to /æ/ confusions was observed only in the word-competitor condition (N = 63). The errors in the nonword-competitor condition were more randomly distributed, and included only 12 /æ/ responses.

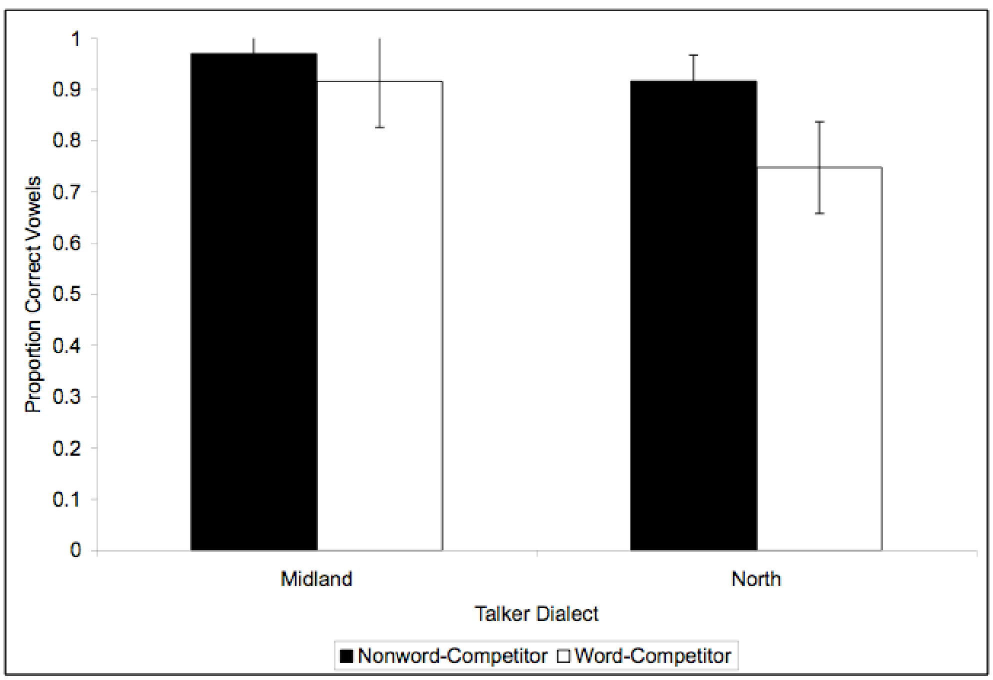

Figure 4. Mean vowel accuracy for target /ɛ/ for each talker dialect in each lexical condition.

The lexical condition x talker dialect interaction for /ɑ/ is shown in Figure 5. For this vowel, the effect of lexical condition is stronger for the Midland talkers than the Northern talkers and is in the opposite direction than what was predicted: performance was better in the word-competitor condition than the nonword-competitor condition. The source of this unexpected result can be seen clearly in Figure 3: Midland /ɑ/ in the word-competitor condition was lower (and more acoustically different from /ʌ/) than Midland /ɑ/ in the nonword-competitor condition. An inspection of the error patterns reveals significant /ɑ/ to /ʌ/ confusions for the Midland talkers in both lexical conditions (N = 28 in the word-competitor condition, N = 31 in the nonword-competitor condition) and for the Northern talkers in the word-competitor condition (N = 48), as well as significant /ɑ/ to /æ/ (N = 27) and /ɑ/ to /ɑj/ (N = 34) confusions for the Northern talkers in the nonword-competitor condition. These confusions between the low back vowels /ɑ, ʌ/ are consistent with the results of previous research on the perception of American English vowels (e.g., Hillenbrand et al. 1995; Peterson and Barney 1952), which also revealed more perceptual confusions among these vowels than for vowels in other parts of the vowel space. However, the different pattern of errors for the Northern talkers in the nonword-competitor condition suggests that overall error rates are not always sufficient to capture the effects of dialect variation and lexical condition on vowel recognition performance.

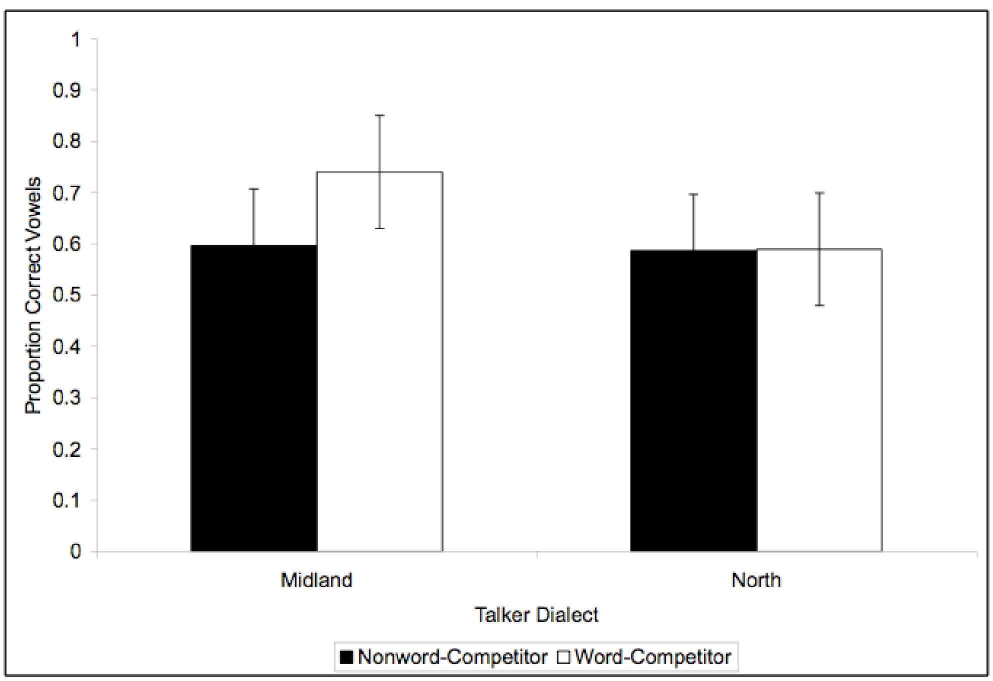

Figure 5. Mean vowel accuracy for target /ɑ/ for each talker dialect in each lexical condition.

3.4. Interactions with listener dialect

The lexical condition x listener dialect interaction was significant for /ɔ/ in the item analysis (*F2*(2, 16) = 10.8, *p* = .001) and at the less conservative α = .05 level in the subject analysis (*F1*(2, 33) = 3.6, *p* = .04). As shown in Figure 6, performance was better for the Northern listeners than the General American and Mobile listeners in the word-competitor condition, but performance was similar across the three listener dialects in the nonword-competitor condition. Given that the Northern listeners are likely to maintain a distinction between /ɔ/ and /ɑ/ in production (Labov et al. 2006), it is not surprising that they would be more accurate than the General American and Mobile listeners, who may have merged productions of /ɔ/ and /ɑ/, in perceptually differentiating target /ɔ/ from competitor /ɑ/ in the word-competitor condition. In the nonword-competitor condition, performance by all three listener groups was good, because the potential lexical competitors with /ɑ/ were not real words.

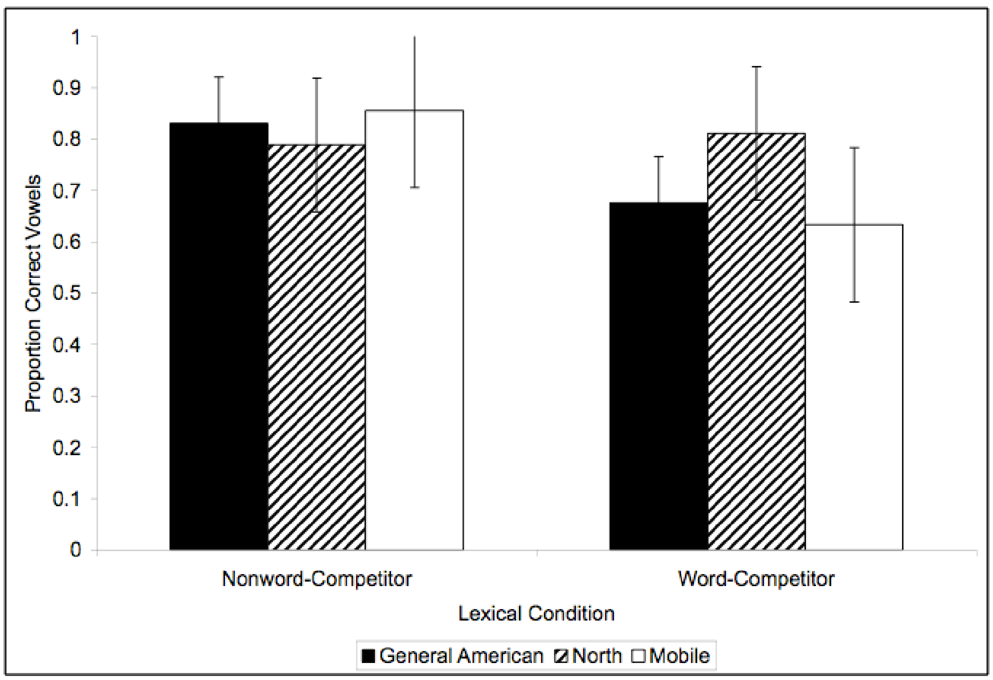

Figure 6. Mean vowel accuracy for target /ɔ/ for each listener dialect in each lexical condition.

Finally, the talker dialect x listener dialect interaction was significant in the subject analysis for /ʌ/ (*F1*(1, 37) = 12.2, *p* = .001) and in the item analysis for /ej/ (*F2*(2, 8) = 11.1, *p* = .005). As shown in Figure 7, the Mobile listeners performed better on /ʌ/ for the Northern talkers than the Midland talkers, whereas performance was similar across the two talker dialects for the General American and Northern listeners. As shown in Figure 3, the Northern /ʌ/ is mildly backed relative to the Midland /ʌ/, consistent with the Northern Cities Chain Shift. The better performance by the Mobile listeners for the Northern talkers suggests that they were less susceptible to the predicted /ʌ/ to /ɔ/

phonological confusion. This interpretation is supported by the error patterns, which reveal fewer /ʌ/ to /ɔ/ errors for the Northern talkers by the Mobile listeners (N = 8) than by the General American (N = 41) or Northern (N = 16) listeners.

As shown in Figure 8, the General American listeners performed worse on /ej/ for the Midland talkers than the Northern talkers, whereas performance was similar across the two talker dialects for the Mobile and Northern listeners. In addition, vowel accuracy for /ej/ was essentially at ceiling for the Mobile and Northern listeners for both talker dialects and for the General American listeners for the Northern talkers. As noted above, Midland /ej/ was acoustically more similar to /ɪ/ than the Northern /ej/, and an inspection of the error patterns reveals that most (N = 23) of the /ej/ to /ɪ/ confusions were produced by General American listeners. These results are inconsistent with our prediction that more familiar dialects would lead to fewer phonological confusions than less familiar dialects, but are consistent with previous research showing that listeners can be less sensitive to variation in their own variety than in other varieties (e.g., Niedzielski, 1999), and that Mobile listeners with exposure to multiple different varieties can outperform local listeners on some dialect perception tasks (e.g., Clopper and Pisoni 2004).

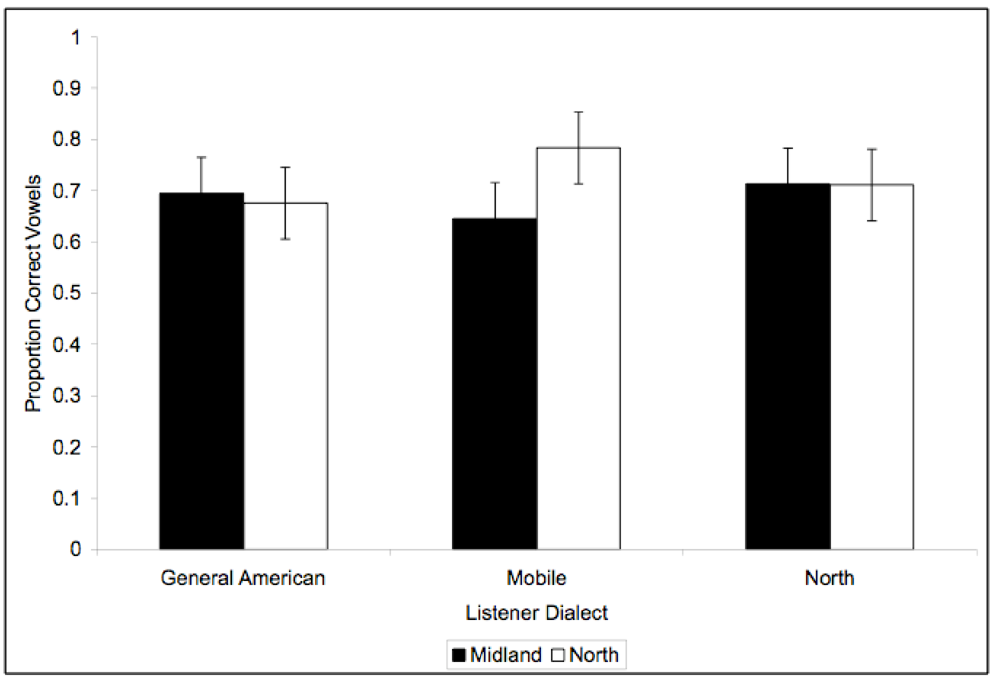

Figure 7. Mean vowel accuracy for target /ʌ/ for each talker dialect and each listener dialect.

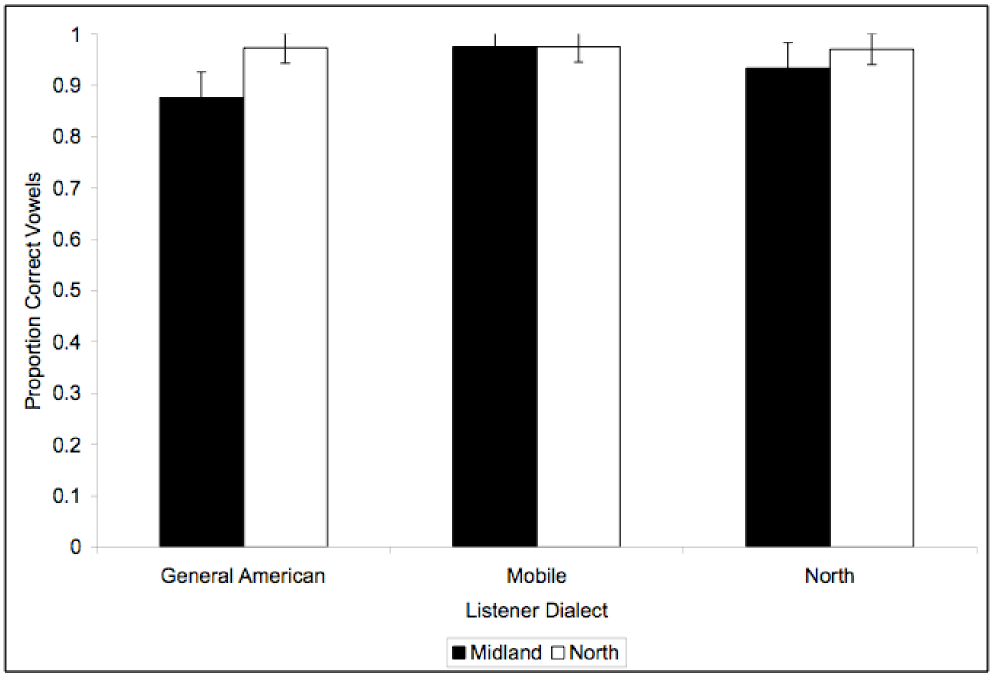

Figure 8. Mean vowel accuracy for target /ej/ for each talker dialect and each listener dialect.

4. General discussion

The goal of the current study was to examine the interaction between the lexical neighborhood density effect and dialect variation in an open-set word recognition task in noise. The predicted interaction between lexical condition and talker dialect was observed for /ɛ/ and /ɑ/. In the case of /ɛ/, the interaction revealed significantly more phonological confusions between /ɛ/ and /æ/ for the Northern talkers only in the word-competitor condition only, as expected. When /ɛ/ was phonetically more similar to /æ/ (in the Northern dialect) and when the target /ɛ/ words had real minimal pair neighbors with /æ/ (in the word-competitor condition), listeners were more likely to respond with /æ/ than when /ɛ/ was phonetically more different from /æ/ and/or when the target /ɛ/ words did not have real minimal pair neighbors with /æ/. In the case of /ɑ/, the same overall pattern was observed: /ɑ/ was most phonetically similar to /ʌ/ in the nonword-competitor condition for the Midland talkers, and it was in that condition that performance was poorest. These findings are consistent with the results of previous studies suggesting that lexical competition is driven by phonological confusions between phonetically similar segments (e.g., Benkí 2003; Wang and Bilger 1973), and that vowels may be particularly confusable in English due to their role in marking sociolinguistic categories such as regional background (e.g., van Ooijen 1996).

The main effect of lexical condition was observed for /ɪ, æ, ɔ, ow/, but the lexical condition x talker dialect interaction was not observed. For /ɪ/, performance was near ceiling for both talker dialects in the nonword-competitor condition, and significantly poorer in the word-competitor condition. While the overall accuracy for the two talker dialects in the word-competitor condition was similar, the error patterns reveal an effect of dialect on phonological confusions: for the Northern talkers, /ɪ/ was misidentified as /ɛ/ significantly more often than chance (N = 28), whereas for the Midland talkers, the errors for /ɪ/ were more randomly distributed, and only 11 /ɛ/ responses were obtained. Thus, as expected, /ɪ/ was more confusable with /ɛ/ for the Northern talkers only in the word-competitor condition only. For /ɔ/, both talker dialects were expected to exhibit confusions with /ɑ/. Thus, the lack of a significant interaction between talker dialect and lexical condition for this vowel is not surprising, and the main effect of lexical condition merely confirms the lexical neighborhood density effect.

The lack of a significant lexical condition x talker dialect interaction for /æ, ow/ cannot be explained by the error patterns produced by the listeners. For /ow/, the error patterns revealed significant /ʊ/ responses to /ow/ targets for both the Midland and the Northern talkers in the word-competitor condition. However, these errors were overwhelmingly the result of *pole* → *pull* and *bowl* → *bull* confusions. For many Ohioans, /ow/ and /ʊ/ are merged in pre-lateral position, which may account for the high error rate for these two target words, and the lack of a lexical condition x talker dialect interaction for /ow/. For /æ/, no significant error patterns were produced for either talker dialect in either of the two lexical conditions. However, the acoustic analysis of the stimulus materials suggests that duration or formant trajectory differences between /æ/ and /ɛ/ may explain the lack of perceptual confusability of /æ/ and /ɛ/. In the stimulus materials in the word-competitor condition, the Northern /æ/ was too long or too diphthongal to be confusable with /ɛ/ (see also Hillenbrand et al. 1995). Thus, how vowel duration and formant trajectory affect perceptual vowel similarity is an important area for future research, and an empirical measure of phonetic similarity or category overlap that incorporates both spectral and temporal information is needed to make more accurate and more comprehensive predictions about phonological confusions across dialects.

Finally, the main effect of talker dialect was observed for /ej, ow/, but the lexical condition x talker dialect interaction was not observed. As noted above, the lack of an interaction for /ow/ was most likely the result of the inclusion of *pole* and *bowl* in the stimulus list in the word-competitor condition. The same target words for /ej/ were used in both conditions, so no main effects or interactions with lexical condition were expected for this vowel. The main effect of dialect was surprising, however, given that previous research has not indicated that /ej/ varies significantly across these two varieties (e.g., Clopper et al. 2005; Labov et al. 2006). Additional research is therefore needed to explore potential variation in /ej/ in the Midland and Northern dialects.

Listener dialect was also found to affect performance in the word recognition task. The lexical condition x listener dialect interaction observed for /ɔ/ is consistent with the ongoing merger of /ɑ/ and /ɔ/ in many varieties of American English. In the word-competitor condition, the General American and Mobile listeners, who were more likely to exhibit this merger in their own speech, were more likely to respond with /ɑ/ than the Northern listeners, who were more likely to maintain a distinction between these two vowels in production. Given that /ɔ/ words tend to be less frequent and familiar than their

/ɑ/ minimal pairs, it is also not surprising that this effect was unidirectional: a similar interaction was not observed for target /ɑ/.

The talker dialect x listener dialect interactions for /ʌ/ and /ej/ are consistent with previous work showing that local listeners are not always sensitive to phonetic variation in their own variety (e.g., Labov and Ash 1997; Niedzielski 1999). The Mobile listeners were most able to correctly identify the Northern backed /ʌ/, and the General American listeners were least able to correctly identify the Midland lowered /ej/. Thus, while a match between the talker's and the listener's dialects may facilitate vowel recognition performance for some categories, talker-listener dialect matches do not necessarily lead to more accurate vowel recognition performance. Additional research with balanced listener groups and more locally-oriented talker and listener populations is needed to further examine the potentially important roles of listener background and talker-listener dialect match or mismatch in cross-dialect speech perception and processing.

The predicted phonological confusions in the current study were based on a comparison to an idealized "standard" variety of American English. This comparison to the standard produced a reasonably good set of predicted phonological confusions for the current study, but comparisons to an idealized standard may not be appropriate in all cases. Experimental work is needed to determine what listeners use as their baseline phonological system in laboratory perception tasks. For example, they may rely on a standard variety (Clopper and Bradlow, 2008; Floccia et al. 2006), their own native variety (Floccia et al. 2006; Labov and Ash 1997), or some other variety triggered by the experimental environment (Hay et al. 2006a; Niedzielski 1999). If it is determined that listeners tend to rely on standard varieties in experimental tasks, additional research will be needed to establish criteria for determining what the standard variety of a given language is in a given region.

In the current study, the stimulus materials were presented in a fully randomized order and the listeners were not told in advance that the stimulus materials were produced by talkers from different dialects. The error patterns produced by the listeners suggest that the experimental design prevented them from adapting to the dialect differences across the talkers. However, we might predict that the lexical condition x talker dialect interaction would be reduced if listeners were able to adapt to talker dialect before, or during, the word recognition task. Online perceptual adaptation to dialect variation has been reported for longer utterances, including sentences and passages (Floccia et al. 2006; Maye et al. 2008, but see Clopper 2009 where adaptation did not occur in a blocked sentence intelligibility task). Thus, additional research is needed to explore the time course of dialect adaptation and the extent to which dialect adaptation reduces the effect of phonetic similarity on phonological confusions in cross-dialect word recognition.

The results of the current study contribute to a growing literature showing that social indexical information can affect other well-established linguistic processes, including semantic priming (Hay et al. 2006b) and speeded lexical classification (Clopper 2007). In speech production, social indexical information interacts with other well-known sources of linguistic variability, such as semantic context (Clopper and Pierrehumbert 2008) and lexical neighborhood density (Munson 2007). In the current study, word recognition performance was affected by the interaction between acoustic-phonetic dialect variation and the structure of the English lexicon (i.e., the presence vs. absence of minimal pairs). Recent models of speech processing based on exemplar theories of perception and memory can account for the interactions between phonetic, lexical, and social information that have been observed across these speech perception and production tasks (e.g., Goldinger 1998; Johnson 1997; Pierrehumbert 2002). In these models, individual acoustically-detailed utterances are stored in long-term memory. Thus, lexical representations include acoustic information and it is not necessary to posit autonomous acoustic-phonetic and lexical levels of processing (c.f., Norris et al. 2000). In an exemplar-based model, the observed interaction between lexical and acoustic information in phonological processing is an inherent component of the model. Thus, the results of the current study also point to the need for models of speech processing that can account for the interactions between social and linguistic sources of variability (Pierrehumbert 2006).

Acknowledgements
This work was supported by the James S. McDonnell Foundation Award 21002061 to Northwestern University and the Department of Linguistics at the Ohio State University. The authors would like to thank John Pate for assistance with data collection and Luis Hernandez for assistance in compiling the CVC minimal pair counts in Table 2.

Appendix
Table A1. Target words that occurred in both conditions of the word recognition task. Log frequency, familiarity, and density values for each word were obtained from the Hoosier Mental Lexicon (Nusbaum et al. 1984).

Target Vowel	Target Word	Log Frequency	Familiarity	Density
i	meek	2.00	6.33	19
i	geese	1.48	7.00	7
i	sheep	2.36	7.00	19
i	knead	3.56	6.92	16
i	beef	2.51	7.00	13
ej	shake	2.23	7.00	22
ej	date	3.01	7.00	20
ej	tape	2.54	7.00	14
ej	cave	1.95	7.00	19
ej	pave	1.30	7.00	17
u	tube	2.49	7.00	8
u	juice	2.04	7.00	11
u	boot	2.11	7.00	26
u	moose	1.00	7.00	18
u	suit	2.68	7.00	24

Table A2. Target and competitor words in the word-competitor condition. Log frequency, familiarity, and density values for each word were obtained from the Hoosier Mental Lexicon (Nusbaum et al. 1984).

Target Vowel	Target Word	Log Frequency	Familiarity	Density	Competitor	Log Frequency	Familiarity
ɪ	pig	1.90	7.00	18	peg	1.60	6.17
ɪ	rid	2.28	7.00	25	red	3.29	6.50
ɪ	lit	2.23	6.75	26	let	3.58	6.92
ɪ	knit	2.00	7.00	27	net	2.53	6.92
ɪ	pick	2.74	7.00	27	peck	1.70	6.75
ɛ	dead	3.24	7.00	22	dad	2.18	7.00
ɛ	head	3.63	7.00	22	had	4.71	7.00
ɛ	guess	2.75	7.00	9	gas	2.99	7.00
ɛ	pet	1.90	7.00	25	pat	2.54	7.00
ɛ	set	3.62	7.00	24	sat	3.18	7.00
æ	bad	3.16	7.00	27	bed	3.10	7.00
æ	sad	2.54	7.00	21	said	4.29	7.00
æ	bag	2.62	7.00	26	beg	2.04	7.00
æ	mass	3.04	7.00	20	mess	2.34	7.00
æ	bat	2.26	7.00	33	bet	2.30	7.00
ɑ	knob	1.30	7.00	18	nab	1.00	6.75
ɑ	cop	2.18	7.00	24	cap	2.43	7.00
ɑ	mop	1.48	7.00	15	map	2.11	7.00
ɑ	sock	1.60	7.00	22	sack	1.90	6.67
ɑ	shock	2.49	7.00	20	shack	1.00	7.00
ɔ	chalk	1.48	7.00	12	chock	1.00	4.55
ɔ	hawk	2.15	7.00	14	hock	1.00	5.17
ɔ	caught	2.99	7.00	20	cot	1.00	7.00
ɔ	sought	2.74	6.58	19	sot	1.00	4.08
ɔ	taught	2.82	7.00	19	tot	1.00	5.92
ʌ	dug	2.20	7.00	21	dog	2.88	7.00
ʌ	hug	1.48	7.00	21	hog	1.48	7.00
ʌ	bus	2.54	7.00	17	boss	2.30	7.00
ʌ	buck	2.30	7.00	26	balk	1.00	6.18
ʌ	cuff	1.00	6.25	16	cough	1.85	7.00
ow	code	2.60	7.00	23	could	4.20	6.42
ow	hoed	1.00	6.58	23	hood	1.85	6.75
ow	showed	3.46	7.00	23	should	3.95	7.00
ow	bowl	2.36	7.00	28	bull	2.15	7.00

ow	pole	2.43	6.92	29	pull	2.71	7.00
ʊ	hoof	1.30	6.50	5	huff	2.00	6.25
ʊ	look	3.60	7.00	17	luck	2.67	7.00
ʊ	took	3.63	7.00	18	tuck	1.30	6.83
ʊ	put	3.64	7.00	14	putt	1.85	6.83
ʊ	shook	2.76	6.75	15	shuck	1.00	7.00

Table A3. Target words in the nonword-competitor condition. Log frequency, familiarity, and density values for each word were obtained from the Hoosier Mental Lexicon (Nusbaum et al. 1984).

Target Vowel	Target Word	Log Frequency	Familiarity	Density
ɪ	tic	1.48	7.00	26
ɪ	hitch	1.70	6.75	17
ɪ	dig	2.00	6.92	17
ɪ	kiss	2.23	7.00	13
ɪ	fish	2.54	7.00	9
ɛ	chess	1.48	7.00	9
ɛ	wedge	1.60	7.00	11
ɛ	fetch	1.78	7.00	8
ɛ	deaf	2.08	7.00	11
ɛ	jet	2.46	7.00	17
æ	sash	1.48	6.50	16
æ	pad	1.90	6.83	22
æ	calf	2.04	6.58	17
æ	gap	2.23	7.00	19
æ	match	2.61	7.00	14
ɑ	wad	1.00	6.92	17
ɑ	yacht	1.60	6.75	16
ɑ	mob	2.00	7.00	14
ɑ	dock	2.45	7.00	21
ɑ	shop	2.80	7.00	15
ɔ	gawk	1.00	6.25	12
ɔ	cough	1.85	7.00	7
ɔ	log	2.04	6.73	12
ɔ	wash	2.57	7.00	4
ɔ	fought	2.66	7.00	17
ʌ	tug	1.48	7.00	18
ʌ	lush	1.70	6.75	12

ʌ	duck	1.95	6.75	22
ʌ	hut	2.11	7.00	26
ʌ	rough	2.61	7.00	17
ow	poke	1.00	7.00	22
ow	toad	1.60	7.00	16
ow	loaf	1.60	7.00	14
ow	boat	2.86	7.00	27
ow	hope	3.25	6.92	19
ʊ	hood	1.85	6.75	13
ʊ	cook	2.67	7.00	13
ʊ	foot	2.85	7.00	9
ʊ	good	3.91	7.00	10
ʊ	should	3.95	7.00	10

Notes

[1] The General American varieties of English are characterized by the merger of the low back vowels /ɑ/ and /ɔ/, and include the regional varieties of New England, the Midland, and the Western United States (Labov 1998). The vowel systems of these varieties are highly similar (e.g., Clopper et al. 2005; Labov 1998), and dialect classification judgments of talkers from these dialects reveal that they are also highly perceptually similar (e.g., Clopper and Pisoni 2007).

[2] Due to an oversight in setting up the experiment, the target word *hug* was presented to listeners in the word-competitor condition four times (by one Midland talker and by all three Northern talkers), for a total of 328 trials in the word-competitor condition.

References

Adank, Patti & James M. McQueen. 2007. The effect of an unfamiliar regional accent on spoken word comprehension. *Proceedings of the 16th International Congress of Phonetic Sciences*, 1925-1928.

Benkí, José R. 2003. Analysis of English nonsense syllable recognition in noise. *Phonetica* 60. 129-157.

Best, Catherine T. 1995. A direct realist view on cross-language speech perception. In Winifred Strange (ed.), *Speech perception and linguistic experience*, 171-204. Timonium, MD: York Press.

Bonatti, Luca L., Marcela Peña, Marina Nespor & Jacques Mehler. 2005. Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science* 16. 451-459.

Clopper, Cynthia G. 2007. Effects of dialect variation on speeded word classification. *Journal of the Acoustical Society of America* 121. 3189.

Clopper, Cynthia G. 2009. Effects of semantic context and regional dialect variation on speech intelligibility in noise. *Journal of the Acoustical Society of America* 125. 2575.

Clopper, Cynthia G. & Bradlow, Ann R. 2008. Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech* 51. 175-198.

Clopper, Cynthia G. Allyson K. Carter, Caitlin M. Dillon, Luis R. Hernandez, David B. Pisoni, Connie M. Clarke, James D. Harnsberger, & Rebecca Herman. 2002. The Indiana Speech Project: An overview of the development of a multi-talker multi-dialect speech corpus. *Research on Spoken Language Processing Progress Report No. 25* (Speech Research Laboratory, Indiana University, Bloomington). 367-380.

Clopper, Cynthia G. & Janet B. Pierrehumbert. 2008. Effects of semantic predictability and regional dialect on vowel space reduction. *Journal of the Acoustical Society of America* 124. 1682-1688.

Clopper, Cynthia G. & David B. Pisoni. 2004. Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change* 16. 31-48.

Clopper, Cynthia G. & David B. Pisoni. 2007. Free classification of regional dialects of American English. *Journal of Phonetics* 35. 421-438.

Clopper, Cynthia G., David B. Pisoni & Kenneth de Jong. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America* 118. 1661-1676.

Clopper, Cynthia G., David B. Pisoni & Adam T. Tierney. 2006. Effects of open-set and closed-set task demands on spoken word recognition. *Journal of the American Academy of Audiology* 17. 331-349.

Cutler, Anne, Andrea Weber, Roel Smits & Nicole Cooper. 2004. Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America* 116. 3668-3678.

Dirks, Donald D., Sumiko Takayanagi & Anahita Moshfegh. 2001. Effects of lexical factors on word recognition about normal-hearing and hearing-impaired listeners. *Journal of the American Academy of Audiology* 12. 233-244.

Floccia, Caroline, Jeremy Goslin, Frédérique Girard & Gabrielle Konopczynski. 2006. Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance* 32. 1276-1293.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105. 251-279.

Hay, Jennifer, Aaron Nolan & Katie Drager. 2006a. From fush to feesh: Exemplar priming in speech perception. *Linguistic Review* 23. 351-379

Hay, Jennifer, Paul Warren & Katie Drager. 2006b. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34. 458-484.

Hillenbrand, James, Laura A. Getty, Michael J. Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97. 3099-3111.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (eds.), *Talker variability in speech processing*, 145-166. San Diego: Academic Press.

Kirk, Karen Iler, David B. Pisoni & Miyamoto, R. Christopher. 1997. Effects of stimulus variability on speech perception in listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research* 40. 1395-1405.

Kraljic, Tanya, Susan E. Brennan & Arthur G. Samuel. 2008. Accommodating variation: Dialects, idiolects, and speech processing. *Cognition* 107. 54-81.

Labov, William. 1998. The three dialects of English. In Michael D. Linn (ed.), *Handbook of dialects and language variation*, 39-81. San Diego: Academic Press.

Labov, William & Sharon Ash. 1997. Understanding Birmingham. In Cynthia Bernstein, Thomas Nunnally, & Robin Sabino (eds.), *Language variety in the south revisited*, 508-573. Tuscaloosa, AL: University of Alabama Press.

Labov, William, Sharon Ash & Charles Boberg. 2006. *Atlas of North American English*. New York: Mouton de Gruyter.

Luce, Paul A. & David B. Pisoni. 1998. Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing* 19. 1-36.

Mason, Harry M. 1946. Understandability of speech in noise as affected by region of origin of speaker and listener. *Speech Monographs* 13(2). 54-68.

Maye, Jessica, Richard N. Aslin & Michael K. Tanenhaus. 2008. The Weckud Wetch of the Wast: Lexical adaptation to a novel accent. *Cognitive Science* 32. 543-562.

Miller, George A. & Patricia E. Nicely. 1955. An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* 27. 338-352.

Munson, Benjamin. 2007. Lexical characteristics mediate the influence of sex and sex typicality on vowel-space size. *Proceedings of the 16th International Congress of Phonetic Sciences,* 885-888.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18. 62-85.

Norris, Dennis, James M. McQueen & Anne Cutler. 2000. Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23. 299-370.

Nusbaum, Howard C., David B. Pisoni & Christopher K. Davis. 1984. Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No.10* (Speech Research Laboratory, Indiana University. Bloomington). 357–376.

Peterson, Gordon E. & Harold L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24. 175–184.

Pierrehumbert, Janet B. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 101-140. Berlin: Mouton de Gruyter.

Pierrehumbert, Janet B. 2006. The next toolkit. *Journal of Phonetics* 34. 516-530.

Rakerd, Brad & Bartek Plichta. (2003). More on perceptions of /ɑ/ fronting. Paper presented at New Ways of Analyzing Variation 32, Philadelphia, PA, October 9-12.

Sommers, Mitchell S., Karen Iler Kirk & David B. Pisoni. 1997. Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners I: The effects of response format. *Ear and Hearing* 18. 89-99.

Stevens, Michael A., James M. McQueen & Robert J. Hartsuiker. 2007. No lexically-driven perceptual adjustments of the [x]-[h] boundary. *Proceedings of the 16th International Congress of Phonetic Sciences*, 1897-1900.

Sumner, Meghan & Arthur G. Samuel. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60. 487-501.

Takayanagi, Sumiko, Donald D. Dirks & Anahita Moshfegh. 2002. Lexical and talker effects on word recognition among native and non-native listeners with normal and impaired hearing. *Journal of Speech, Language, and Hearing Research* 45. 585-597.

Thomas, Erik R. (2001). *An acoustic analysis of vowel variation in New World English*. Durham, NC: Duke University Press.

van Ooijen, Brit. 1996. Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory and Cognition* 24. 573-583.

Wang, Marilyn D. & Robert C. Bilger. 1973. Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustical Society of America* 54. 1248-1266.