Perceptual Similarity of Unfamiliar Regional Dialects: 
Some Preliminary Findings¹

Terrin N. Tamati²

Speech Research Laboratory
Department of Psychological and Brain Sciences
Indiana University
Bloomington, Indiana 47405

¹ This work was supported by NIH-NIDCD Training Grant T32-DC00012 and NIH-NIDCD Research Grant R01-DC000111. I would like to thank Robert Nosofsky for his help with data analysis. I would also like to thank Luis Hernandez for his helpful advice and assistance with this project.

² Also, Department of Linguistics, Indiana University, Bloomington, IN
Perceptual Similarity of Unfamiliar Regional Dialects: Some Preliminary Findings

Abstract. Linguistic experience has been shown to affect the perception of regional dialect variation. This pilot study examined the effect of familiarity on the perceived similarity of regional dialects. Native speakers of American English completed a paired comparison perceptual similarity rating task with a group of unfamiliar regional dialects. Speakers of American English were asked to make explicit judgments about the similarity of a set of unfamiliar talkers from the United Kingdom and Ireland on the basis of regional dialect. Results showed that listeners judged pairs of talkers from the same dialect region as more similar than pairs of talkers from different dialect regions. A multidimensional scaling analysis revealed two dimensions of perceptual dialect similarity, which reflected the geographic locations of the cities of origin of the talkers. These findings suggest that, despite being unfamiliar with the regional dialects in the study, listeners were able to use dialect-specific differences in the acoustic signal to make judgments of the perceptual similarity of talkers based on their regional dialect.

Introduction

Studies on the perception of regional dialect variation have shown that naïve listeners can perceive dialect-specific differences and can use stored knowledge of such differences to make reliable judgments about the regional background of talkers. Van Bezooijen and Gooskens (1999) and Van Bezooijen and Ytsma (1999) examined the categorization of regional varieties of Dutch in a multi-level forced-choice categorization task. They found that naïve listeners, when presented with audio recordings of talkers from different regions of the Netherlands and Belgium, were able to categorize talkers by regional dialect. Van Bezooijen and Gooskens (1999) also observed similar results for regional varieties of English spoken in the United Kingdom.

More recently, Clopper and Pisoni (2004b) used a forced-choice categorization task to investigate the perceptual categorization of regional varieties of American English. In their study, naïve listeners were able to place talkers into dialect categories. An analysis of the acoustic properties of the speech samples used in the study, the regional dialect of the talkers, and the responses of listeners in the categorization task showed that listeners used dialect-specific acoustic-phonetic properties in making their judgments. In addition, the listeners’ response errors in the categorization task were consistent, revealing patterns of perceptual similarity among dialects.

Other studies have observed similar findings. Clopper and Pisoni (2007) asked naïve American English listeners to group unfamiliar talkers by region of origin in an auditory free classification task. Listeners were able to use dialect-specific phonetic differences to group talkers into regional dialect categories. Using a paired comparison perceptual similarity rating task, Clopper, Levi, and Pisoni (2006) also found that naïve listeners can make explicit judgments about the similarity of talkers’ voices based on regional origin. In both studies, listeners’ judgments reflected the perceived similarity of regional dialects of American English.

Linguistic experience and a listener’s past developmental history has been shown to influence the perception of regional dialects. Williams, Garrett, and Coupland (1999) looked at the categorization of six regional varieties of English in Wales by adolescent males. The listeners were from the same six regions under study. In a forced-choice categorization task, overall accuracy was fairly low (30%).
However, a more detailed analysis of the various listener groups revealed that listeners were more accurate at identifying the region of origin of talkers from the same region (45%) than talkers that were from other regions (24%). Similarly, Baker, Eddington, and Nay (2009) investigated the effects of region of origin and amount of time spent in Utah on listeners’ ability to identify talkers from Utah. In their study, listeners who were from Utah were better at identifying Utah talkers than listeners from other regions of the United States and listeners from the Western United States were more accurate at identifying Utah talkers than listeners from other regions. Experience also influenced performance, as listeners who had spent more time in Utah were better able to identify the Utah talkers. In addition, region of origin, but not experience, determined the features the listeners used to identify the talkers in the task.

Other studies have investigated how residential history influences the perceived similarity among dialects in further detail. Analyzing the error patterns in the perceptual categorization task, Clopper and Pisoni (2004b) found significant differences in the perceived similarity of dialects among three listeners groups (Northern Indiana, Southern Indiana, and Out-of-State). In another study, Clopper & Pisoni (2004a) examined the effect of geographic mobility on dialect categorization in American English. In that study, a group of mobile listeners, who had lived in at least three different states, and a group of non-mobile listeners, who had only lived in Indiana, performed a six-alternative forced-choice categorization task. Mobile listeners were more accurate than the non-mobile listeners in the categorization task and mobile listeners who had lived in a particular dialect region were better at categorizing talkers from that region than mobile listeners who had not lived in that specific region. Clustering analyses on the categorization responses also revealed that the perceived similarity of the regional dialects differed for the mobile and non-mobile listeners.

In another study, Clopper and Pisoni (2006) explicitly looked at the effect of region of origin on dialect categorization. Four groups of listeners who differed with respect to region of origin (North or Midland dialect regions) and geographic mobility (mobile or non-mobile) performed a forced-choice categorization task. While listeners did not differ in accuracy, geographic mobility and region of origin affected the perceived similarity of the regional dialects. Clopper and Pisoni (2007) found similar listener background effects using an auditory free classification task.

Clopper and Bradlow (2007) examined how native language status (native or non-native) affects the perceived similarity of regional dialects. In their study, despite limited experience with American English regional dialects, non-native listeners were able to use variation in the acoustic signal to classify talkers by regional dialect of American English in an auditory free classification task. The non-native and native listeners formed a similar number of talker groups and showed a similar perceptual similarity space. However, the type and amount of experience the listeners had with English influenced some aspects of perceptual dialect classification. Non-native listeners were less accurate and less consistent than native listeners in dialect categorization and, while they had a similar dialect similarity structure, the talker groups were more clearly defined in the native similarity space. Furthermore, the non-native listeners relied on dialect-specific phonetic features in grouping the talkers to a greater extent than the native listeners, who were able to use their knowledge of regional dialect variation in American English in completing the task.

Taken together, these earlier studies establish an important role for linguistic experience in the perception of regional dialect variation. The findings suggest that the more experience one has with a particular regional variety, the better one will be at identifying the regional variety of unknown talkers. In addition, these studies show that experience with regional varieties affects the perceived similarity among regional varieties. In other words, the more familiar the listener is with a given variety, the more
distinct that variety will seem, perhaps resulting in easier identification and greater perceived distinctiveness among regional dialects.

Based on the previous studies, then, it would be expected that listeners would have difficulty in differentiating unfamiliar regional dialects. This prediction seems to be supported by recent findings from Ikeno and Hansen (2007). In their study, they looked at the effect of listener background on the comprehension and identification of regional dialects of English in the United Kingdom. Native speakers of British English, native speakers of American English, and non-native speakers of English living in the United States completed a forced-choice dialect categorization task. These groups represented three levels of familiarity with the regional dialects: familiar with both English and the regional dialects (British English group), familiar with the language but not the regional dialects (American English group), and limited familiarity with both English and the regional dialects (non-native group). Listeners classified talkers as having one of three different types of accents, which were introduced in a pre-test familiarization period. Talkers were from Belfast (Northern Ireland), Cambridge (England), and Cardiff (Wales).

The British English listeners were much more accurate (83%) than the American English listeners (56%) and the non-native listeners (45%). The listener groups also showed different confusion patterns in their responses, suggesting that the perceived similarity of the dialects was different for each listener group. Thus, the groups who were not familiar with the three dialects in the study (the American English listeners and the non-native listeners in the United States) had more difficulty in categorizing unfamiliar talkers by regional dialect. However, these groups did perform quite well with 56% correct for the American English listeners and 45% for the non-native listeners, suggesting that they were able to reliably detect differences between the three unfamiliar dialects.

One of the limitations of the Ikeno and Hansen (2007) study is that they used a three-alternative forced-choice categorization task. Providing the dialect regions and labels imposes response constraints on the listeners in completing the task (Clopper & Pisoni, 2007). Given that the listeners were unfamiliar with the regional dialects in the study, the forced-choice categorization task would likely have been very difficult. The auditory free classification task (Clopper & Pisoni, 2007) and the paired comparison similarity rating task (Clopper et al., 2006) have recently been used to examine the perceived similarity of regional dialects of American English. Since both the auditory free classification task and the paired comparison similarity rating task allow listeners to directly compare regional varieties without imposing experimenter-based categories, they might be better suited for studies on how familiarity affects the perception of regional dialect variation.

The purpose of the current pilot study was to evaluate, as a basis for future studies on the effect of familiarity of the perception of regional dialect variation, the perceived similarity of a group of unfamiliar regional dialects using the direct two-interval dialect similarity rating task adopted from Clopper et al. (2006). Native speakers of American English were presented with a set of talkers from five different regions of the United Kingdom and Ireland and asked to rate the similarity of the talkers based on regional dialect. Based on previous studies, it was expected that listeners would have a great deal of difficulty in distinguishing the regional dialects. It was unclear if listeners would be able to perceive talkers from the same dialect region as more similar than talkers from different dialect regions or if similarity judgments would reflect the regional background of the talkers.
Method

Listeners

Ten native speakers of American English participated in the study. Listeners had diverse residential histories but none reported having lived in the United Kingdom, Ireland, or any other English-speaking country outside of the United States. Three listeners reported being fluent in at least one other language. All of the nine participants were between the ages of 22 and 27 at the time of testing. One participant reported a history of a speech or hearing disorder.

Talkers

Twenty talkers were selected from the IVie Corpus (Grabe, Post, & Nolan, 2001). The IVie Corpus consists of audio recordings of twelve to fourteen talkers for each of nine cities in the British Isles. Half of the talkers for each city were female and half were male. Cities represented in the corpus include Belfast, Bradford, Cambridge, Cardiff, Dublin (Malahide), Leeds, Liverpool (Scouse), London, and Newcastle upon Tyne. All talkers were around the age of sixteen and enrolled in urban high schools at the time of recording. Talkers from Bradford, Cardiff, and London were ethnic minorities. Talkers from Bradford were bilingual Punjabi-English speakers, talkers from Cardiff were bilingual Welsh-English speakers, and talkers from London were of Jamaican descent. Materials for each talker include a conversation, a map task, a story told from memory, read sentences, and a read passage. This corpus was designed for studies on variation in the prosodic systems of different regional dialects.

The twenty talkers used in the current study were females around the age of sixteen at the time of recording. Four talkers were chosen from each of the following five cities: Belfast, Dublin (Malahide), Cambridge, Leeds, and Newcastle upon Tyne. These five cities were chosen for their geographic locations, as they were, geographically, distant from one another. Talkers from Bradford, Cardiff, and London were not used in the study in order to avoid introducing additional within-dialect variation from varying degrees of bilingualism or contact with other varieties of English, such as English spoken in India or Jamaica. The talkers from each city were selected based on the quality of the recording available for the read passage and the fluency of their speech in the target phrase.

Stimulus Materials

Stimulus materials consisted of a single short phrase for each talker. The phrase was selected from the read passage in the IVie Corpus (Grabe et al., 2001). An attempt was made to select a phrase that would include a variety of speech sounds that could be used to distinguish the regional dialects in the study. The phrase used throughout this study is reported in (1):

(1) the beauty who had stolen his heart

For each talker, the target phrase was extracted from the read passage and saved in an individual sound file. When necessary, the sound file was edited so that there would be no noticeable speech errors.

Procedure

Listeners were seated in individual booths in a quiet testing room. Each sat in front of a computer equipped with a keyboard and headphones. On each trial, listeners heard the target phrase produced by two different talkers separated by 500 ms of silence. The phrase also appeared visually on the computer.
screen. Listeners were asked to judge the similarity of the accents of the two talkers on a scale from 1 (“very different”) to 7 (“very similar”). This way, they were asked to make an explicit judgment on the similarity of the unfamiliar dialects, with higher ratings reflecting greater dialect similarity than lower ratings. For each trial, listeners had seven seconds from the onset of the second production of the target phrase to respond. Listeners responded by pressing the button on the keyboard that corresponded to the similarity rating for a particular talker pair. If the listener failed to give a response in that time, the next trial would begin. They received no feedback during the experiment.

Each listener heard all twenty talkers paired with all other talkers two times throughout the study, for a total of 380 trials. Of the total number of trials, 60 trials included two talkers from the same city (“same-dialect” trials) and 320 included two talkers from different cities (“different-dialect” trials). For each talker pair, both possible orders of presentation of the talkers (AB and BA) were included. The order of presentation of the trials was determined randomly. The experiment was divided into two halves, with a short break provided after the first 190 trials. It took approximately 45 minutes to complete. The similarity ratings and response times for the ratings were collected. All trials for which no response was recorded were removed (N = 20). Only the similarity ratings will be discussed here.

Results

To see whether listeners were able to make explicit judgments on dialect similarity for unfamiliar dialects, responses for same-dialect and different-dialect pairs were examined. Table 1 shows the mean similarity ratings for the same-dialect and different-dialect pairs.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same-dialect</td>
<td>5.40 (1.50)</td>
</tr>
<tr>
<td>Different-dialect</td>
<td>3.85 (1.76)</td>
</tr>
</tbody>
</table>

Table 1: Mean and standard deviation of responses for same-dialect and different-dialect trials.

Overall, listeners showed higher similarity ratings for same-dialect pairs than different-dialect pairs. To assess the difference between ratings for same-dialect and different-dialect pairs, a paired t-test was carried out on similarity ratings with dialect match (whether it was a “same-dialect” or “different-dialect” trial) as the factor. The difference in responses for same-dialect and different-dialect trials was highly significant (t(9) = 9.064, p < .001). Thus, same-dialect pairs were rated significantly higher than different-dialect pairs.

A multidimensional scaling analysis was then carried out to obtain measures of the underlying similarity space. A 20 x 20 talker similarity matrix was constructed from the similarity ratings given to a particular talker pair averaged across all listeners. An ALSCAL analysis was performed on the similarity matrix. One-, two-, three-, and four-dimensional solutions were obtained. Mean stress values for each solution were .296, .162, .103, and .085, respectively. These values suggest an “elbow” at the two-dimensional solution, as stress was greatly reduced from the one- to two-dimensional solution and the addition of the third dimension did not result in a large decrease in stress. Given these results and the high interpretability of the two-dimensional solution, the two-dimensional solution was selected for discussion.
The resulting two-dimensional multidimensional scaling solution from the ALSCAL analysis was plotted. The first dimension appeared to separate the talkers from England on the left and the talkers from Ireland on the right and the second dimension appeared to separate the talkers from the North of each country from talkers from the South of each country. Since the dimensions seemed to roughly correspond to the geographic locations of the cities, in order to better visualize the results, the plot was rotated, with the geographic coordinates of the cities as the normative solution, using the ROTAT program. The resulting plot after rotation is shown in Figure 1. Each talker is represented by the first letter of her city. The four Bs represent the four talkers from Belfast, Northern Ireland, the four Cs represent the four talkers from Cambridge, England, the four Ds represent the four talkers from Dublin, Ireland, the four Ls represent the four talkers from Leeds, England, and the four Ns represent the four talkers from Newcastle upon Tyne, England.

![Figure 1](image)

**Figure 1.** Rotated two-dimensional multidimensional scaling solution. Talkers are represented by the first letter of their home city: B for Belfast, C for Cambridge, D for Dublin, L for Leeds, and N for Newcastle upon Tyne.

This plot shows that the talkers from Dublin and Belfast are separated from the talkers from Cambridge, Leeds, and Newcastle along the first dimension and that the talkers from Dublin, Cambridge, and Leeds are separated from the talkers from Belfast and Newcastle along the second dimension. In other words, the first dimension distinguishes the talkers from the West from the talkers from the East (or Ireland from England) and the second dimension distinguishes the talkers from the North of each country from the talkers from the South of each country. These two underlying perceptual dimensions can be interpreted, then, in terms of geography: West versus East and North versus South.
Discussion

The results of the two-interval similarity rating task show that, given the phrase *the beauty who had stolen his heart*, listeners judged the similarity of the accents of two talkers based on their regional background. The analysis of the same dialect and different dialect ratings showed that listeners rated pairs of talkers from the same city as more similar than pairs of talkers from different cities.

In addition, the multidimensional scaling analysis of the similarity ratings revealed two dimensions of perceptual dialect similarity. These dimensions reflected the geographic location of the cities; the first dimension was interpreted as West versus East and the second dimension was interpreted as North versus South. These perceptual dimensions produced a space that roughly corresponded to the geography of England and Ireland. As shown in Figure 1, the Belfast talkers are in the Northwest quadrant (top left in Figure 1), the Dublin talkers are in the Southwest quadrant (bottom left in Figure 1), the Newcastle talkers are in the Northeast quadrant (top right in Figure 1), and the Leeds and Cambridge talkers are in the Southeast quadrant (bottom right in Figure 1), with the Leeds talkers located slightly above the Cambridge talkers.

The listeners’ ability to judge the similarity of the talkers’ accents based on regional background is consistent with the earlier results of Clopper *et al.* (2006). In that study, native speakers American English made explicit judgments about the dialect of talkers, who were also native speakers of American English, based on their regional background. However, unlike Clopper *et al.* (2006), where listeners would have had varying degrees of experience with the different American English regional dialects, in the current study, listeners were unfamiliar with the regional dialects that they rated. That is, listeners had very little experience or familiarity with the English and Irish dialects.

As mentioned above, none of the listeners lived in England, Ireland, or any other English-speaking country outside the United States. Some of them, however, had acquaintances from England and Ireland, and, given the linguistics background of some of the participants, it is also likely that a few of the listeners were familiar with studies on regional dialects in England and Ireland. In a post-test questionnaire, most listeners reported that they recognized the talkers as being from England or Ireland, but none of the listeners were able to identify explicitly the exact city of origin of any of the talkers. In a few instances, listeners erroneously thought that some of the talkers were speakers of American English or non-native speakers. Only two of the listeners were able to identify the broad regions of origin of some of the talkers. One listener indicated that some of the talkers were from Northern Ireland and another indicated that some of the talkers were from Southern England.

Thus, despite having little direct experience with the regional dialects in the study, listeners were able to make explicit judgments about the perceptual similarity of unfamiliar talkers’ voices that reflected regional dialects of Ireland and the United Kingdom. Since the results of this pilot study cannot be compared to perception data from listeners familiar with all, or some, of the regional dialects, they cannot yet be interpreted in terms of the effects of familiarity on the perceptual similarity of regional dialects. Still, the performance of the speakers of American English raises the issue of how familiarity with regional dialects influences the perceived similarity of those dialects.

Previous studies suggest that, because they were unfamiliar with the regional dialects, listeners should have had difficulty judging the similarity between the regional accents. Since they have had such little experience with the dialects, they would not have knowledge of the phonetic differences that are important for distinguishing the dialects nor would they have knowledge of variation irrelevant to distinguishing the dialects. Nevertheless, listeners were able to identify dialect-specific phonetic features
in the speech of the unfamiliar talkers and use these features, while disregarding other irrelevant variation, to make judgments about the similarity of talker pairs based on regional dialect. The multidimensional scaling solution obtained from the similarity data shows that talkers were clearly grouped based on their city of origin along geographic dimensions.

So, given how well the American English listeners performed, could a listener familiar with the regional dialects in the study do any better? A listener who is familiar with the regional dialects would be influenced by their linguistic experience, including their own dialect, mobility, and social background. This linguistic experience would shape the way that they hear the differences between the dialects. However, as the results of the current study suggest, the “perceptual warping” of the dialect similarity space resulting from linguistic experience would not necessarily result in dialects being heard as more distinct from one another. Experience may actually lead to regional dialects being perceived as more similar, rather than different. This interpretation is consistent with previous studies that have shown that non-mobile American English listeners from the Northern dialect region actually perceive their own dialect as more similar to the Midland dialect or a supraregional standard variety (Clopper & Pisoni, 2006; Clopper & Pisoni, 2007; Niedzielski, 1999). To examine this in further detail, another study should also be carried out with listeners with different degrees and types of familiarity with the regional dialects.

Conclusions

In this study, a paired comparison perceptual similarity rating task was used to examine the role of familiarity on the perceptual similarity space of regional dialects. Speakers of American English made judgments of the similarity of the accents of unfamiliar talkers from various regions of the United Kingdom and Ireland. Results showed that listeners were generally successful at judging the similarity of the accents of two talkers based on their regional background. Listeners rated pairs of talkers from the same city as more similar than pairs of talkers from different cities. Furthermore, the multidimensional scaling analysis of the similarity ratings revealed two underlying dimensions of perceptual dialect similarity, which reflected the geographic location of the cities. These findings differ from previous studies on the effect of familiarity on regional dialect identification or classification, which predicted that listeners with little or no experience with a set of regional dialects should be very poor at differentiating such dialects. The results of this study, however, showed that listeners were successful at distinguishing the British and Irish regional dialects, suggesting that experience with regional dialect variation may not necessarily result in greater perceived differences between regional dialects.

References


