The current study used a hand-drawn map task, a dialect difference rating task, and a dialect classification task to explore the relationship between participants’ ideologies about dialect differences and their classification of authentic talkers from six regional varieties in Enshi Prefecture, China. The talkers frequently mistaken for each other in the dialect classification task were those who came from counties that were perceived to have similar dialects in the hand-drawn map task and the dialect difference rating task. Participants showed a positive response bias for the Enshi dialect in classifying talkers, corresponding to the dialect difference ratings that Enshi was rated as least different. Thus participants’ classification of real talkers was largely consistent with their ideologies about differences among “imagined” dialects. Participants’ ideologies about dialect differences were shaped by their home county, and their classification performance was affected by their home county and the talker’s social background.

1. Introduction

In perceptual dialectology, a set of methods have been developed to elicit nonlinguists’ perceptions of and beliefs about regional varieties (Benson, 2003; Evans, 2011; Hartley, 1999, 2005; Preston, 1986, 1989). These methods, including hand-drawn maps, are typically based on nonlinguists’ mental representations of regional varieties that are stored in long-term memory. In dialect identification and categorization studies, listeners are presented with real speech stimuli and are asked to categorize the talkers in terms of region of origin (Clopper & Pisoni, 2004a, 2004b, 2006; Niedzielski & Preston, 2000; Van Bezooijen & Gooskens, 1999). By examining behavioral responses to speech stimuli, researchers are able to determine how listeners identify talkers based only on the talkers’ speech characteristics.

Despite substantial previous research using perceptual dialectology and dialect identification and categorization methods, few studies have combined these methods to investigate the relationship between nonlinguists’ prior conceptions of dialect differences and their actual classification of talkers when presented with real linguistic stimuli (cf. Montgomery, 2007; Preston, 1993a). Moreover, although previous studies have examined regional dialect perception and categorization in countries such as the United States, Japan, and France (Inoue, 1999; Kuiper, 1999; Long, 1999a, 1999b; Niedzielski & Preston, 2000; Preston, 1989), little attention has been paid to the dialects in China. The current study investigated the perceptual dialect categorization of regional varieties in Enshi Prefecture, Hubei Province, China. This study aimed to examine how participants’ home county affected their perceptions of dialect differences across counties and whether their perceptions differed for different regional dialects. Participants’ perceptions of dialect differences were elicited by a hand-drawn map task and a dialect difference rating task. The effects of participants’ home county, talker’s county origin and talker’s urban/rural origin on participants’ classification of real talkers were also explored in a dialect classification task, in which participants were presented with authentic speech samples and asked to identify the county origin of talkers. Participants’ ideologies about dialect differences were compared with their actual classification of real talkers who spoke these dialects to reveal whether their classification performance matched their ideologies.

1.1 Perceptual Dialectology

Inspired by work in cultural geography (Gould & White, 1986), perceptual dialectology has employed the hand-drawn map task (Preston, 1986) to reveal nonlinguists’ perceptions of regional variation. In it, respondents are given a relatively blank map of a country or region and are asked to draw lines to indicate where they think people speak differently. Perceptual dialectology research has also used a dialect difference rating task to explore how nonlinguists perceive other regional varieties to be different from their own variety (Preston, 1993b). In the dialect difference rating task, respondents are presented with a list of regions and are asked to rate individual regions on a scale (e.g., a five-point scale) to indicate how the speech of each region is different from their own speech. The hand-drawn map task and the dialect difference rating task have been used to examine nonlinguists’
perceptions in a wide range of locations around the world, including Great Britain (Inoue, 1996), Germany (Dailey-O’Cain, 1999), France (Kuiper, 1999), Turkey (Demirci & Kleiner, 1999), Wales (Coupland, Williams, & Garrett, 1999; Williams, Garrett, & Coupland, 1999) and Japan (Inoue, 1999; Long, 1999a, 1999b), among many others. Similar approaches have also been employed to explore nonlinguists’ perceptions of regional varieties within the United States. Work in this vein includes exploration of dialect perceptions by respondents from Hawaii, Michigan, New York (Preston, 1989), Oregon (Hartley, 1999), California (Fought, 2002), Boston (Hartley, 2005), and Nevada (Fridland & Bartlett, 2006). Some recent research has focused on nonlinguists’ perceptions of smaller regions, for example, dialect variation within a single state. Research has been conducted in Ohio (Benson, 2003; Campbell-Kibler, 2012; Campbell-Kibler & Bauer, under review), California (Bucholtz, Bermudez, Fung, Edwards, & Vargas, 2007; Bucholtz, Bermudez, Fung, Vargas, & Edwards, 2008), and Washington (Evans, 2011), providing a more local understanding of nonlinguists’ perceptions of dialect variation than studies involving a whole country.

Previous research has also demonstrated that both linguistic and extralinguistic factors, such as the geographical and cultural environment of the region as well as personal experience and exposure to dialect variation, play a role in shaping perceptions of regional dialects. On the one hand, nonlinguists’ conceptions of dialect variation are influenced by the linguistic landscape of the area. Mase (1992) and Lance (1999) have shown that perceived dialect boundaries parallel linguistic divisions to some degree in Japan and the United States, respectively. Pearce (2009) demonstrated that in northeast England, participants responded to real linguistic variation in their judgments of similarity and difference of dialects rather than simply relying on broader nonlinguistic perceptions, such as geographical or cultural proximity. Some other studies, on the other hand, have found that subjective dialect division is greatly influenced by extralinguistic factors such as real-life experience and geographical demarcation. For example, Benson (2003) reported that people from different parts of Ohio categorized varieties spoken within the state differently, partly depending on their exposure to speakers from other parts of the state. Inoue (1996) found that in Great Britain, students’ subjective dialect division was influenced by their geographical conceptions. The labels given by students, such as “northern,” “southern,” and “midland,” corresponded well with the commonly used areal divisions in school educational maps. More recently, Bucholtz et al. (2007) found that in California, the perception of dialect variation is affected by highly salient social groups, an important part of the cultural landscape of California. Finally, Evans (2011) found that participants’ perceptions of dialect variation in Washington were influenced by standard language ideology (e.g., beliefs about where standard and nonstandard English is spoken) and an urban-rural dichotomy (e.g., “hicks” and “farmers” were most frequently associated with the category country) (2011: 403). Thus, the patterns of dialect perception reflect linguistic, historical, and social knowledge of the relevant region.

As noted above, compared with perceptual dialectology work that focuses on an entire country, work that elicits perceptions in smaller regions provides a highly detailed picture of local scenes. The current study was conducted to examine nonlinguists’ perceptions of dialect variation in a small area in China that we know little about, and to explore the effect of home county on the perception of dialect differences.

1.2 Dialect Identification And Categorization

Perceptual dialectology research reveals how nonlinguists conceptualize and represent different regional varieties in long-term memory. This line of research, however, does not address their ability to identify the dialect of a talker in response to real speech stimuli. One of the earliest attempts to get at this level of dialect perception is work by Bush (1967), who presented listeners with utterances and asked them to identify the talkers’ dialect in a three-alternative forced-choice task: American English, British English, and Indian English. Later, Preston (1993a) explored perceptual categorization of regional varieties of American English. In a forced-choice dialect categorization experiment, listeners were presented with short utterances and asked to identify the talkers’ region of origin. A north-south continuum from Dothan, Alabama to Saginaw, Michigan was provided. The results suggested that while listeners could identify Northern and Southern talkers, they were not able to tell exactly where the talkers came from.

In a dialect classification study on Chinese dialects, Blum (2004) used talkers from various language backgrounds from China: three were speakers of Kunminghua (Kunming dialect), four of Putonghua (Standard Mandarin; these speakers came from various places in China with different accents), and several speakers of other dialects. The listeners were either from Kunming or from other places in Yunnan Province. Most of them (84%) were college students studying in Kunming. The listeners were asked to identify the place of origin of the talker. The results showed that listeners’ responses to local Kunming talkers were accurate, with a success rate of 63% for an older talker, and 86% and 94% for the two younger talkers. They were also
accurate in identifying Putonghua talkers, with a success rate between 55%-86% across the four talkers. These results demonstrate that Chinese listeners can accurately identify the local variety and the national language.

Researchers working on dialect categorization and classification have shown that listeners’ region of origin has a substantial impact on their dialect categorization performance. For example, Williams et al. (1999) and Baker, Eddington, and Eddy (2009) found that listeners more accurately identified talkers from their own region than those from other regions. Niedzielski and Preston (2000) reported that in a dialect classification task, the perceived dialect boundaries were different for different groups of respondents depending on their region of origin. Using a forced-choice categorization task, Clopper and Pisoni (2004b) found greater discrimination between local dialects among local residents than nonlocal residents, supporting the view that the ability to accurately perceive dialect differences was affected by where the listeners came from. In addition to listeners’ region of origin, the effect of listeners’ geographic mobility has also been documented. Clopper and Pisoni (2004a) found that “army brats,” who had lived in at least three different states, demonstrated better overall performance on a dialect categorization task than “homebodies,” who had lived only in one state.

Taken together, listeners’ linguistic experience as measured by their region of origin and geographic mobility affects performance on a perceptual dialect categorization task. Specific exposure to a dialect leads to better categorization performance for talkers from that dialect as a result of experience with the phonetic and phonological variation in that particular dialect; general exposure to dialect variation leads to overall more accurate categorization, presumably due to the raised perceptual distinctiveness of different varieties. The present study examines whether participants are able to accurately classify talkers in terms of county origin when they are presented with real speech samples produced by urban and rural talkers from each county, and how their classification performance is affected by the participants’ region of origin, the talker’s county origin, and the talker’s urban/rural origin.

1.3 The Current Study

The main goal of the current study was to examine the relationship between nonlinguists’ ideologies about dialect differences and their classification of real talkers who speak those dialects. Their perceptions of difference and similarity between regional varieties were elicited by a hand-drawn map task and a dialect difference rating task. Their reactions to real talkers from different varieties were revealed in a dialect classification task, in which they were asked to identify the county origin of each of the twelve talkers. To obtain the most authentic speech from speakers of different regional varieties, two speakers from each variety (one urban, one rural) were used as the talkers in the dialect classification task. These speakers were long-term residents of the area and their speech displayed multiple regional dialect features documented in previous production work in Enshi Prefecture (Chao, Ding, Yang, Wu, & Dong, 1946; Yang, 2011; Yuan, 2001), confirming that their speech can be considered typical of the local varieties (see the discussion below for the dialect features present in their speech samples). The current design allowed us to directly address whether nonlinguists are able to accurately identify regional dialects, and whether their classification of talkers in the dialect classification task is consistent with their prior conception of dialect differences revealed in the hand-drawn map task and the dialect difference rating task.

2. About Enshi

Enshi Tujia and Miao Autonomous Prefecture (Enshi Prefecture) is located in the southwest region of Hubei Province, China. Enshi Prefecture comprises eight counties: Enshi, Jianshi, Lichuan, Xianfeng, Lai Feng, Badong, Hefeng, and Xuanen. Enshi County is the economic, cultural, and political center of Enshi Prefecture. The maps in Map 1 (from right to left) show Hubei’s location in China, Enshi’s location in Hubei Province, and the eight counties in Enshi Prefecture. The current study covers six counties (Enshi, Jinshi, Badong, Hefeng, Xuanen, and Lai Feng). The two western counties (Lichuan and Xianfeng) were not included due to fieldwork time limitations.

As in most parts of southwest China, residents of Enshi Prefecture speak local dialects that belong to Southwestern Mandarin. Although the dialects in Enshi Prefecture are conventionally classified into eight main dialect groups based on geographic/county divisions (e.g., the dialect spoken in Jianshi county is named Jianshi dialect), there are a larger number of dialect varieties spoken in the territory and many more fine-grained divisions could be made. The dialects spoken in the urban areas are different from the ones in rural areas. Within the rural areas, the dialect differences between two villages 20 miles away from each other can be considerable, although these differences typically do not lead to difficulties in communication. Within the urban areas, distinct varieties coexist. For example, two varieties, laocheng hua ‘old city speech’ (老城话) and xin cheng hua ‘new city speech’ (新城话) coexist in urban Enshi. Old city speech is almost only used in the Liujaoting District (六角亭), historically called
“old city”, and new city speech is used in other districts. In urban Jianshi, two local varieties are distinguished by the glide-dropping phenomenon. The syllable-medial glide /j/ is maintained in one variety, while in the other variety the syllable is simplified and the glide is absent. Since the focus of this study is not to describe dialects in Enshi Prefecture based on production data, I adopt the county-based dialect names that are conventionally used but include both urban and rural speakers in the dialect classification task.

Based on the traditional dialect classification of Mandarin (Guo, 2009), Enshi, Jianshi, and Badong are grouped into Danyi pian ‘Danjiangkou and Yichang subarea’ (丹江口), Xuanen and Laifeng are grouped into Yunzhu pian ‘Yunxi and Zhushan subarea’ (云溪), and Hefeng belongs to Changhe pian ‘Changde and Hefeng subarea’ (常德). This division is based on tonal development, as well as other phonological characteristics, such as the phonetic realization of the consonants in /ʂu/ ‘book’ and /ɕu/ ‘weak’, and the vowels in /tu/ (a surname) and /tʂu/ ‘help’.

As documented in previous production work in the prefecture (Chao et al., 1946; Yang, 2011), differences in phonology, lexicon, and grammar abound between Enshi dialects and Putonghua, and within various dialects spoken within Enshi Prefecture. A number of noticeable features distinguish Enshi dialects from Putonghua. For example, the initial consonants /n/ and /l/ merge into [n] in the majority of Enshi dialects, thus /lan/ ‘basket’ becomes [nan], and /lian/ ‘face’ becomes [nian]. The merger of /x/ and /t/ is another salient feature of Enshi dialects, although the exact merging pattern differs across dialects. In the Badong and Enshi dialects, /x/ and /t/ merge into [x] before /ɑŋ/, and into [f] before other rhymes (Yuan, 2001). The retroflexes /ts, tsʰh, s/ are often realized as alveolars [ts, tsʰ, s] in urban Enshi and Badong. In Laifeng, the retroflexes /ts, tsʰ, s/ become palatals [tc, tcʰ, c] before /u/, and become alveolars [ts, tsʰ, s] elsewhere.

Vowel variation is also observed across the dialects in Enshi Prefecture. The Standard Mandarin /ɤ/ is realized as [o] in the Enshi dialect. /uo/ becomes [ue] in some subvarieties of the Jianshi dialect and [o] in the Enshi dialect. In the Laifeng dialect, the Standard Mandarin /u/ is realized differently depending on the preceding consonant: it becomes [ou] when following alveolars and [y] when following retroflexes. In addition, /uei/ is realized as [ei] after alveolars in several varieties (e.g., Laifeng and Xuanen). Lastly, /uan/ is realized as [an] after alveolars in the Laifeng dialect, and /ue/ becomes [io] after palatals in some Enshi and Jianshi subvarieties.

There are four tonal categories in most varieties in Enshi Prefecture: yīnpíng 55, yángpíng 13, shǎng 53, and qù 214. These four tonal categories are identical to those in Standard Mandarin, but with different tonal contours except yīnpíng. The three other tonal contours in Standard Mandarin can be described as yīnpíng 35, shǎng 214, and qù 51. The tone system of the Hefeng variety is quite different from other varieties; even within Hefeng county, tonal differences are found between urban and rural varieties. The variety spoken in urban Hefeng has four tones: yīnpíng 45, yīnpíng 11, shǎng 51, and qù 214, whereas the variety spoken in a rural town (Zouma town 卓玛乡) has five tones: yīnpíng 55, yīnpíng 24, shǎng 51, qù 44, and rù 35 (Yang, 2011).

The current study aims to reveal how local residents perceive these differences between their own dialect

Map 1. (from right to left) Hubei’s location in China, Enshi’s location in Hubei Province, and the eight counties in Enshi Prefecture.
and the dialects spoken in other counties, and whether they are able to accurately identify the county of origin of an unfamiliar talker based only on short speech samples. The methods this study employs include hand-drawn maps, a dialect difference rating task, and a dialect classification task.

3. Methodology

3.1 Participants

Participants were recruited through a friend-of-a-friend method and advertisements posted in shopping malls and main streets. The data for this study came from 120 participants, balanced for gender (male, female), education level (high school or lower education, college and higher education), and county affiliation (Enshi, Jianshi, Badong, Hefeng, Laifeng, and Xuanen). There were five participants in each gender, education level, and county affiliation cell. Information about the urban/rural origin of most (82%) of the participants was not obtained because they only provided county names for their birthplace and the places they have lived, without specifying the town or village. The fieldwork was conducted in 4-9 sites in each county, including both urban and rural locations. Based on these fieldwork sites, there were roughly equal numbers of participants of urban and rural origin in each county.

All participants were monolingual native speakers of Mandarin Chinese (although not necessarily the standard variety) and most of them spoke only one local Mandarin variety. The participants who were bidialectal typically had parents who came from two different counties. Speaking Standard Mandarin (Putonghua) was not required to participate in this study. All tasks and interviews were conducted in Enshi Mandarin (the author’s native dialect), not in Standard Mandarin (Putonghua). The participants who completed junior high and higher education were generally able to speak Standard Mandarin, but none of them were native speakers since Standard Mandarin was only acquired at school through formal education. All participants spent all or nearly all of their lives in Enshi Prefecture. Of those who had not lived their entire lives in Enshi, they left either for college or for work for a short period of time. Participants were paid RMB36 ($5) for participating.

3.2 Procedures

3.2.1 Hand-drawn Map Task

Employing the technique of mental mapping (Preston, 1982, 1986, 1988), the hand-drawn map task aimed to reveal nonlinguists’ knowledge about regional variation. In this task, participants were given a blank map of Enshi, which showed county boundaries, county names, and the names of contiguous regions (Yichang and Hunan). The use of a blank map without any boundary lines may cause considerable confusion for participants who do not have adequate geographical knowledge of the region (Preston, 1993a: 335). Thus a map with county boundaries was used so that participants had a concrete map with county boundaries to consult. Participants were asked to draw circles or lines on the map to indicate the areas “where people talk alike.” They were encouraged to make comments either orally or by writing them down. The map presented to participants is shown in Map 2.

The hand-drawn map task provides insights about participants’ mental representations of place-based regional varieties in Enshi and elicits participant’s perceptions of language variation without exposing them to any real speech, therefore establishing their beliefs about the existence of regional dialects and their boundaries.

3.2.2 Dialect Difference Rating Task

Mental representations of regional varieties involve not only where participants believe dialect boundaries
exist, but also how they believe other varieties to be different from their own variety. The rating task, administered following the hand-drawn map task, aimed to measure the perceived degree of difference of nearby dialect varieties from participants’ own variety. In this task, participants were given the names of the six counties in Enshi Prefecture included in the map task. They were asked to consider each local dialect and assign a score from 1 to 5 based on the perceived degree of difference of that dialect from their own (1 = same or very similar, 2 = a little different, 3 = quite different, but I can still understand, 4 = I can barely understand that dialect, and 5 = I cannot understand that dialect at all). It is possible that the hand-drawn map task may prime the participants’ dialect difference ratings. To minimize this potential priming effect, different instructions were given in the two tasks. In the hand-drawn map task, participants were asked to find areas “where people talk alike” without having to consider their own dialect, while in the dialect difference rating task, they were explicitly asked to evaluate the degree of difference of each dialect compared with their own dialect. Thus, the two tasks differed both in the explicit focus (dialect similarity vs. difference) and in the relevance of the native dialect.

The rating task provides clues about nonlinguists’ perceptions of different regional varieties in Enshi and how the perception of dialect difference varies from individual to individual. The rating task, combined with the hand-drawn map task, uncovers Enshi natives’ ideological representations of place-based regional varieties.

3.2.3 Dialect Classification Task

The dialect classification task followed the rating task. The goal was to examine the ability of Enshi natives to accurately identify the home county of an unfamiliar talker based on a short voice sample. A comparison between the participants’ difference ratings for each regional variety and the classification of each talker in this task can also shed light on the relationship between their beliefs about dialect differences and their responses to real speech.

The perceptual stimuli in the dialect classification task consisted of excerpts from stories recorded by twelve non-mobile, older (above 35 years old) male talkers who were selected to represent the authentic local variants (Chambers & Trudgill, 1998:29; Orton & Dieth, 1962:15; Preston, 1989:128). Two talkers from each county were recorded, including one of urban origin and one of rural origin. Map 3 shows the town/village where each talker was from. In this map, black stars indicate urban talkers, and gray dots represent rural talkers.

Each talker was recorded telling the story “The Emperor’s New Clothes.” The length of the original recordings varied from talker to talker, with an average length of 150 seconds (range: 88-226 seconds). The talkers were asked to tell the story as if they were telling stories to their children or family members and to speak as naturally as possible. Short clips of each recording were selected for the speech samples played to participants in the dialect classification task. Each sample was approximately 30 seconds in length to provide listeners adequate information to make their judgments. The speech samples were relatively coherent in content. Participants listened to the samples one at a time in a random order and were asked to determine the county origin of each talker.

The story “The Emperor’s New Clothes” was selected because it is familiar to Chinese children and adults so that talkers did not have to read a script, and it contained multiple segmental features that differed across varieties and thus could be potential cues for dialect classification. The linguistic cues contained in the speech samples that might influence participants’ judgments of the talkers included segmental features, suprasegmental features and lexical items. Segmental
features included initial consonants and syllable rhymes (glides, vowels, and codas), suprasegmental features mainly involved lexical tones, and lexical features included dialect-specific lexical items. For example, syllable final /ŋ/ was realized as [n] in the speech of the Enshi, Jianshi, Xuanen urban, and Enshi rural talkers. Syllable-medial glide /j/ was dropped in the Jianshi urban talker’s speech but maintained in the other talkers’ speech. Syllable initial /t/ was realized as [tʰ] in the Laifeng, Xuanen, Hefeng urban, and Laifeng, Badong rural talkers’ speech, but was deleted in the Enshi rural talker’s speech. The realization of /x/ as [f] as in [fan] ‘change’ and [fa] ‘or’, and the realization of /tɕʰ/ and /tʂ/ as [ʦʰ] and [ʦ] as in [ʦʰa] ‘match’ and [ʦu] ‘knit’, respectively, are salient characteristics of the Laifeng urban talker’s speech. The Laifeng rural talker also exhibited several local features, including the realization of /ʦ/ as [ʦ], /s/ as [s], and /ʨ/ as [k]. The vowel /ɤ/ was realized as [o] in the Laifeng urban and rural, Jianshi rural, and Badong urban talkers’ speech. The realization of /uo/ as [ue] is a feature of Jianshi urban talker’s vowels; for Enshi talkers, /uo/ becomes [o].

3.2.4 Interview

A short semi-structured interview followed the dialect classification task to elicit more comments from participants.

3.3 Statistical Analysis

For the hand-drawn maps, the percentage of participants who drew boundaries to indicate the areas “where people talk alike” was calculated to find out the main perceptual dialect regions. The percentages of participants from each of the six counties who identified each of the main perceptual dialect regions were also calculated.

For the dialect difference rating task, a one-way ANOVA on the difference ratings with rated county (Enshi, Jianshi, Badong, Hefeng, Laifeng, Xuanen) as the factor was conducted to find out whether the dialect difference ratings were significantly different for the six counties. Post hoc $t$ tests were used to reveal any significant differences in the difference ratings between six rated counties. A series of one-way ANOVAs on the difference ratings for each county with participants’ home county (Enshi, Jianshi, Badong, Hefeng, Laifeng, Xuanen) as the factor was then conducted to explore whether participants from different counties gave significantly different difference ratings for each county. Post hoc Tukey tests were conducted to reveal whether there were significant differences in the difference ratings for each county by participants from different counties.

For the dialect classification task, to reveal the effect of the talker’s country of origin and urban/rural origin, and the participant’s home county on classification accuracy, a repeated measures ANOVA was calculated using talker’s county origin (Enshi, Jianshi, Badong, Hefeng, Laifeng, Xuanen) and talker’s urban/rural origin (urban, rural) as within-subject factors and participant’s home county (Enshi, Jianshi, Badong, Hefeng, Laifeng, Xuanen) as a between-subject factor for participant’s classification accuracy. Post hoc Tukey and $t$ tests were conducted to reveal whether participant’s classification accuracy was significantly different for urban and rural talkers, talkers of different county origins, and participants from different counties.

4. Results and Discussion

The analysis of the hand-drawn maps and the dialect difference rating task will reveal how participants’ home county affected their perceptions of dialect similarities/differences across counties and whether different regional dialects were perceived differently in terms of the degree of difference. The analysis of the dialect classification task will show the effects of participants’ country of origin, talker’s county origin, and talker’s urban/rural origin on participants’ classification of real talkers. Participants’ hand-drawn maps and their dialect difference ratings will be compared with their classification of real talkers to reveal whether participants’ classification performance was consistent with their ideologies about dialect differences.

4.1 Hand-drawn Maps

To determine the main perceptual dialect regions in Enshi, the percentage of participants who drew boundaries to indicate any particular regions where people talked alike was calculated. The majority of participants drew lines or circles along county boundaries; a few drew lines within the territory of one county. In the latter case, the grouping of only part of one county with another county was not included in the estimation of the main perceptual regions. Very few participants grouped three or more counties together. Since only a small number of participants made comments in the hand-drawn maps, the comments they provided are not discussed here.

Images 1 and 2 in Map 4 show the two main perceptual dialect regions designated by participants. The first region is comprised of Xuanen and Laifeng, which were identified as sharing a dialect by 68.3% of all participants. The second most frequently designated region was Enshi and Jianshi, recognized by 36.7% of all participants. Images 3 and 4 in Map 4 show two other less frequently identified regions: Jianshi and Badong,
and Enshi and Xuanen, designated by 30.0% and 23.3% of participants, respectively. Hefeng was rarely grouped with any other counties, and thus forms its own region as shown in image 5.

Table 1 shows the recognition percentages of each of the main perceptual dialect regions by participants from the different counties. Participants who grouped Xuanen and Laifeng as a single region were almost evenly distributed across counties. Thus, it appears that regardless of home county, participants shared the consensus that the Xuanen and Laifeng dialects are similar. The participants who grouped Enshi and Jianshi together were mainly from these counties, whereas participants from other counties identified Enshi and Jianshi as a region less frequently. Therefore, local participants contributed to the Enshi-Jianshi grouping more than nonlocal participants, where “nonlocal” is defined as participants whose home county is not the county in question. A similar pattern was found for the Jianshi and Badong grouping: local Jianshi and Badong participants grouped Jianshi and Badong together more frequently than nonlocal participants. As for the Enshi and Xuanen region, an unbalanced contribution of local participants was found: whereas 42.9% of the participants who grouped Enshi and Xuanen were from Xuanen, only 10.7% were from...
Categorization of Enshi Mandarin Dialects

Table 1. Percentages of participants who identified each of the four main dialect regions by participant’s home county. For each dialect region, percentages for the local participants are in bold.

<table>
<thead>
<tr>
<th>Dialect regions</th>
<th>Participant’s home county</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enshi</td>
</tr>
<tr>
<td>Xuanen-Laifeng</td>
<td>14.6%</td>
</tr>
<tr>
<td>Enshi-Jianshi</td>
<td>22.7%</td>
</tr>
<tr>
<td>Jianshi-Badong</td>
<td>5.5%</td>
</tr>
<tr>
<td>Enshi-Xuanen</td>
<td>10.7%</td>
</tr>
</tbody>
</table>

Enshi. This result suggests that more Xuanen participants and fewer Enshi participants perceived the Xuanen and Enshi dialects to be similar. This unbalanced perception of dialect difference can be partially attributed to the fact that the Enshi variety was perceived to be the most correct one in Enshi prefecture while Xuanen was perceived to be much less correct (Yan, under review). To align themselves more closely with the more perceptually correct Enshi variety, Xuanen participants grouped Enshi County with their own county to form a single dialect region.

The results of the hand-drawn maps correspond with the dialect subarea division based on production data (Guo, 2009). The first main region identified by the majority of participants, Xuanen and Laifeng, belongs to Yunzhu pian ‘Yunxi and Zhushan subarea.’ Hefeng, which was rarely grouped with other counties, belongs to Changhe pian ‘Changde and Hefeng subarea.’ The other counties are included in Danyi pian ‘Danjiangkou and Yichang subarea.’ This similarity between the hand-drawn maps and the dialect divisions based on production data indicates that participants are able to perceive some noticeable differences within the prefecture, and their mental representations of regional dialects and dialect boundaries reflect the actual patterns of dialect variation to some extent. This result is consistent with previous findings that nonlinguists’ perceptions of dialect boundaries parallel real linguistic variation (Inoue, 1972; Lance, 1999; Mase, 1992).

4.2 Dialect Difference Rating Task

In the rating task, participants considered each local dialect and assigned a value of 1 to 5 based on the perceived degree of difference of that dialect from their own. A high score for a particular county suggests that, overall, participants view the variety in that county as being different from their own; in contrast, a lower score suggests similarity. The mean scores of the degree-of-difference ratings for the six counties are represented in Table 2. In this table, the columns represent the home county of the participants, and the rows represent the county being rated. The highest mean score for each county is in bold, and the lowest mean score for each county by nonlocal participants is in italics.

As shown in Table 2, the difference ratings by local participants (the diagonal) are all very close to 1, confirming that participants were doing the task correctly. As for the ratings by nonlocal participants for each county, Enshi was perceived to be most different by Laifeng participants. Jianshi was rated highest on difference by Hefeng and Laifeng participants. The highest difference ratings for Badong, Hefeng, and Laifeng came from Enshi participants. Xuanen was perceived to be most different by Badong participants. Most counties were rated highest on difference by participants from a county which did not border the counties in question (except Hefeng participants rating Jianshi), suggesting that geographical remoteness may have led to higher perceived dialect difference. With regard to how nonlocal participants perceived a dialect to be least different, Enshi was rated lowest on difference by Xuanen participants, Jianshi by Badong participants, Badong by Hefeng participants, Hefeng by Jianshi participants, Laifeng by Xuanen participants, and Xuanen by Laifeng participants. Thus, a dialect was likely to be perceived as similar by participants from nearby counties.

An ANOVA was conducted to examine whether the six counties were rated significantly differently on the degree of difference. The result showed a significant effect of rated county ($F(5,714) = 9.49, p < 0.001$). Post hoc paired-sample t tests were conducted to examine the effect of rated county and revealed that the degree of difference ratings for Enshi were significantly lower than those for Xuanen, Laifeng, Badong, and Hefeng (all $p < 0.05$). The difference between Enshi and Jianshi was just beyond the level of significance ($p = 0.06$). Jianshi was rated significantly lower than Laifeng, Badong, and Hefeng (all $p < 0.01$), but not significantly lower than Xuanen. Xuanen was rated significantly lower than Laifeng, Badong, and Hefeng (all $p < 0.05$). Thus, overall, the Enshi variety is perceived as the least different variety by participants from other counties.
and Jianshi is the second least different variety, followed by Xuanen. The difference ratings for Enshi, Jianshi and Xuanen are in line with the results in the hand-drawn maps, where Enshi and Jianshi were identified as one dialect region by 36.7% of participants, and Xuanen was grouped with Enshi by 23.3% of participants. Since Enshi is rated as least different, it is not surprising that Jianshi and Xuanen, which are frequently grouped with Enshi to form dialect regions, are perceived as less different than Laifeng, Badong, and Hefeng.

The low difference ratings for Enshi may reflect true similarities between the Enshi variety and the other varieties, prestige associated with the Enshi variety, or both. Based on participants’ evaluations of each variety (see Yan, under review) and their comments in the interview, it is likely that the prestige associated with Enshi accounts for the low difference ratings it received from nonlocal participants. Most participants want to align themselves more closely with the Enshi variety than with any other variety. As a result, the Enshi variety is perceived to be the “neutral” dialect which is relatively similar to everyone’s own dialect.

To examine the effect of participant’s home county on the dialect difference ratings, a series of ANOVAs was conducted. The detailed results of the ANOVAs are summarized in Table 3. Participants’ home county was a significant contributor to the difference ratings for all six counties. Post-hoc Tukey tests confirmed that the difference ratings given to a particular county by non-local participants were significantly different from those given by local participants.

The ANOVA on the effect of rated county reveals that in contrast to the low difference ratings for Enshi, Jianshi, and Xuanen, the mean scores for Laifeng, Badong, and Hefeng by nonlocal participants were significantly higher (Tables 2 and 3). The ANOVA on the effect of participant home county shows that Laifeng received high difference ratings from Enshi, Badong, Hefeng, and Jianshi participants, but not from Xuanen participants (Table 3). Again, these difference ratings correspond to the results of the hand-drawn maps. In the hand-drawn map task, Hefeng was seldom grouped with any other counties; Badong, though grouped with Jianshi by 30.0% of participants, was not...
recognized as being similar to any other varieties; Laifeng was frequently grouped with Xuanen to form a single region in the map task, but was not frequently grouped with any other counties. Thus, it is not surprising that Hefeng was regarded to be highly different by nonlocal participants, that Badong was viewed as being quite different, and that Laifeng was also perceived to be different by nonlocal participants except for Xuanen participants.

Post hoc Tukey tests further confirmed the high degree of perceptual similarity between Laifeng and Xuanen. As shown in Table 3, for Laifeng, there was not only a significant difference between the ratings by local Laifeng participants and those by Enshi, Jianshi, Badong, and Hefeng participants, but also a significant difference between the ratings by Xuanen participants and those by Enshi, Jianshi, Badong, and Hefeng participants. A similar pattern was observed for Xuanen, for which a significant difference in rating was found not only between local Xuanen participants and Enshi, Jianshi, Badong, and Hefeng participants, but also between Laifeng participants and participants from the four other counties. These results suggest that Laifeng and Xuanen participants perceived each other’s variety to be highly similar to their own.

Although Laifeng and Xuanen participants perceived their varieties to be alike, they differed in their ratings for Enshi. Laifeng participants rated Enshi much higher than Xuanen participants did, indicating that they perceived the Enshi variety to be more different from their own variety than Xuanen participants did. One possible explanation for this difference is that Xuanen borders Enshi while Laifeng is farther away. This geographical proximity may have led Xuanen participants to regard the Enshi variety as more similar to their own variety than the Laifeng participants did. In the hand-drawn maps, 23.3% of all participants grouped Xuanen and Enshi together, whereas no participant grouped Laifeng and Enshi as a dialect region, indicating that the Xuanen and Enshi varieties are perceptually more similar than the Laifeng and Enshi varieties.

Unlike participants from other counties who perceived at least one other variety to be similar to their own speech, Hefeng participants rated all other varieties higher than 1.90, indicating that they perceived other varieties to be different from their own. Interestingly, while Hefeng participants viewed Enshi as being only “a little different” from their own variety, Enshi participants did not return the favor; they viewed the Hefeng variety as the most distinct one, with a mean score of 2.80. This difference in perception highlights the attempts of Hefeng participants to align themselves with the more prestigious Enshi variety and confirms the status of the Enshi variety as the regional “norm” that participants from other varieties want their speech to be like. A similar discrepancy in perception between speech varieties was reported by Blair (1990) and Grimes (1995) for mutual intelligibility. Some varieties are inherently intelligible to some extent; however due to social and historical reasons (e.g., political conflict), or different cultural and historical attitudes held by speakers of different varieties, the intelligibility is not mutually perceived.

In summary, the results of the dialect difference rating task show that while the difference ratings for any particular variety by local participants were significantly different from those given by outsiders, participants from different parts of Enshi Prefecture believed the Enshi variety to be similar to their own way of speaking, even when Enshi is not their neighbor geographically, reflecting its position as the local norm. Jianshi was rated as the second most similar variety. Hefeng and Badong were perceived to be quite different by nonlocal participants; Laifeng was viewed as being rather distinct except by Xuanen participants.

The hand-drawn map task and the dialect difference rating task revealed participants’ ideologies about difference and similarity between regional varieties. The finding that Enshi, Jianshi, and Xuanen were rated low on overall difference in the dialect difference rating task is consistent with the result of the hand-drawn maps in which Enshi and Jianshi, and Enshi and Xuanen were identified as overlapping regions. The fact that Hefeng was perceived as most different by nonlocal participants corroborated the hand-drawn map task result that Hefeng was rarely grouped with any other counties. In addition, the fact that Xuanen and Laifeng participants perceived each other’s variety to be highly similar verified what was revealed in the hand-drawn maps.

Participants’ home county shapes how they perceive regional dialect difference. In the hand-drawn map task, dialect regions were generally more frequently designated by local participants than by nonlocal participants (e.g., Enshi-Jianshi and Jianshi-Badong), consistent with the results reported by Benson (2003) and Preston (1986). In the dialect difference rating task, local participants perceived the local dialect to be most similar. A dialect was perceived to be more similar by participants from neighboring counties than those from more remote counties.

4.3 Dialect Classification Task
4.3.1 Classification Performance

Participants were able to correctly classify the talkers by regional dialect with 56% accuracy. This overall success rate was significantly higher than statistical chance (chance performance being 17%; t(119) = 2.07, p < 0.05), and also higher than the 30% overall success rate reported by Williams et al. (1999) and Clopper and
Pisoni (2004a, b). Moreover, given that this study was conducted in a small geographic region (the area of Enshi Prefecture is equal to one fifth of the state of Ohio), participants may find it difficult to accurately identify the county origin of talkers. Nonetheless, dialect varieties in this small region are quite distinct partly due to long-term isolation between different counties, and even villages. Enshi Prefecture is a mountainous area and the high mountains have served as geographical barriers for thousands of years. Thus, distinct regional dialects have developed and their differences are well maintained today.

Although overall classification accuracy was 56%, participants varied considerably in their ability to identify the talkers’ county origin. Five participants correctly classified eleven talkers out of twelve (success rate 92%), whereas two participants were only able to identify one talker (success rate 8%). 110 participants (92% of all participants) correctly identified at least four talkers (success rate above 33%).

A repeated measures ANOVA with talker’s county origin and talker’s urban/rural origin as within-subject variables and participant’s home county as a between-subject variable revealed a significant main effect of talker’s urban/rural origin (F(1,1438) = 8.09, p < 0.01), a significant talker’s urban/rural origin x talker’s county origin interaction (F(5,1434) = 28.05, p < 0.001), a significant participant’s home county x talker’s county origin interaction (F(25,1414) = 4.11, p < 0.001), and a significant participant’s home county x talker’s urban/rural origin x talker’s county origin interaction (F(25,1414) = 2.22, p < 0.001). The effect of participant’s home county was just beyond the level of significance (p = 0.052). 3

Participants’ classification performance in response to real talkers revealed their different perceptions of urban and rural varieties, which was not obtained in the hand-drawn maps and dialect difference ratings. Figure 1 shows the proportion of correct responses for urban and rural talkers from each variety, collapsed across participant groups. Post hoc Tukey tests revealed that performance on rural talkers was better than performance on urban talkers overall (p < 0.05), but this pattern was not observed for all dialects. Performance on the Badong and Hefeng rural talkers was significantly better than performance on the Badong and Hefeng urban talkers, respectively (both p < 0.001), but the Enshi and Jianshi urban talkers were classified more accurately than the Enshi and Jianshi rural talkers, respectively (p < 0.01 and p < 0.001, respectively). No significant difference was found between the classification accuracy of the Laifeng urban and rural talkers, or the Xuanen urban and rural talkers.

The focus of the participant’s home county x talker’s county origin interaction confirms that the performance on talkers from any given county differed by participant’s home county. The significant participant’s home county x talker’s urban/rural origin x talker’s county origin interaction further indicates that the performance on talkers from any given county differed not only by participant’s home county, but also by talker’s urban/rural origin. The results of the paired sample t tests for the participant’s home county x talker’s county origin interaction are shown in Table 4, and the results of the paired sample t tests for the participant’s home county x talker’s urban/rural origin x talker’s county origin interaction are presented in the third column of Table 5.

Consistent with previous research (Baker et al., 2009; Clopper & Pisoni, 2004b; Williams et al., 1999), local participants performed better overall on talkers from the same county than nonlocal participants in terms of classification accuracy, as shown in Table 4. Some nonlocal participants also performed well. For example, Badong participants classified the Enshi talkers significantly more accurately than Laifeng and Xuanen participants, and Enshi and Hefeng participants classified the Badong talkers significantly more accurately than Laifeng participants. Laifeng participants’ performance on the Enshi, Jianshi, Badong, and Hefeng talkers was significantly worse than the local Enshi, Jianshi, Badong, and Hefeng participants’ respectively. Given the geographical location of Laifeng as the southernmost county in Enshi Prefecture and only bordering Xuanen, Laifeng participants’ difficulty in classifying the Enshi, Jianshi, Badong, and Hefeng talkers was likely due to their unfamiliarity with the regional varieties spoken in more remote areas.

The fact that local participants performed better overall on local talkers than nonlocal participants does not mean that they performed equally well on local

Figure 1. Proportion of correct responses for urban and rural talkers from each variety, collapsed across participant groups. Error bars indicate standard error.
urban talkers and rural talkers. As shown in Table 5, Enshi participants performed better than Laifeng and Xuanen participants on the Enshi urban talker, but not on the Enshi rural talker. Similarly, Laifeng participants classified the Laifeng urban talker, not the Laifeng rural talker, more accurately than nonlocal participants. Badong participants only classified the local rural talker more accurately than Laifeng participants. Only Jianshi and Hefeng participants performed better on both local urban and rural talkers than nonlocal participants.

The Jianshi and Enshi urban talkers were classified most accurately among urban talkers. The status of the Enshi urban variety as the regional norm which was least accented may be responsible for the high success rate for the Enshi urban talker. In fact, there was a response bias for Enshi. Table 6 shows the proportion of correct responses to each of the six talker groups, collapsed across urban/rural talkers from each of six counties and 120 participants. The proportion was calculated by dividing the number of actual responses/classifications by the total number of responses to the urban and rural talkers from each county (i.e., 240). The Jianshi, Badong, Hefeng, and Laifeng talkers were classified as Enshi more frequently than the Enshi talker was classified as Jianshi, Badong, Hefeng, and Laifeng, suggesting a positive response bias for Enshi. In particular, participants chose Enshi as a response more often than other response alternatives for the Badong and Hefeng talkers, indicating a strong bias toward Enshi responses. For the Jianshi, Laifeng, and Xuanen talkers, Enshi was the second most frequent incorrect response.

### 4.3.2 Perceptual Similarity Between Dialect Varieties

In the dialect classification task, each talker was classified to different county origins since not every participant gave the correct answer (see Table 6). To explore how perceptually similar the talkers sounded to participants, and to determine the effect of participants’ home county on perceptual dialect similarity, an Individual Differences Scaling (INDSCAL) analysis (Carroll & Chang, 1970) was used. A $6 \times 6$ county dissimilarity matrix was constructed.
for each participant group. When the talkers from one county were correctly classified, the value of the corresponding cell was set to 0. When the talkers from one county were incorrectly classified as talkers from other counties, the value of the corresponding cell was set to 1. One participant group corresponds to participants from one county, thus a total of six dissimilarity matrices were entered into the INDSCAL analysis. The perceptual similarity space calculated over all of the participants regardless of their home county is shown in Figure 2. In this figure, each label represents one talker group, collapsed across urban/rural talkers from the same county. Talkers that are perceived to be similar are plotted near each other, and those perceived to be different are plotted far away from each other (Hartley, 2005).

In the perceptual similarity space shown in Figure 2, the first perceptual dimension divides the Xuanen and Laifeng talkers from talkers from other counties. The Xuanen and Laifeng talkers are located in the right-hand portion of the space, clearly distinguished from the other talkers who are in the left-hand region. The Xuanen and Laifeng talkers cluster together because they were most often mistaken for each other in the dialect classification task. The proportion of Xuanen responses for the Laifeng talkers and the proportion of Laifeng responses for the Xuanen talkers are much higher than the incorrect response proportion for other talkers (Table 6), suggesting that the Xuanen and Laifeng dialects are perceptually most similar among the dialects in Enshi prefecture. Importantly, the proportion of the Xuanen or Laifeng talkers being classified as Enshi, Jianshi, Badong, or Hefeng is smaller than 0.08, indicating that Xuanen and Laifeng dialects as a whole are highly linguistically marked and can be easily distinguished from other dialects. Thus, the first perceptual dimension is related to perceived dialect markedness. In the hand-drawn map task, Laifeng and Xuanen stood out as the most frequently identified dialect region. In the dialect difference rating task, Laifeng and Xuanen participants perceived their varieties to be alike. The results of the dialect classification task further confirm that these two varieties are highly similar such that it is hard to distinguish one from the other. These results also suggest that the ideologies about the similarity between the Laifeng and Xuanen dialects uncovered by the hand-drawn maps and the dialect difference rating task are consistent with participants’ responses to actual linguistic input.

The second perceptual dimension divides the Badong and Enshi talkers from other talkers. This dimension is related to the perceived dialect correctness/standardness, with the more correct and standard dialects on the positive side of Dimension 2 and the less correct and standard dialects on the negative side. The correctness/standardness of a dialect is evaluated relative to Standard Mandarin (Putonghua). In a dialect rating task, Enshi and Badong were rated as the most correct dialects in Enshi prefecture (Yan, under review). Xuanen, Laifeng, and Hefeng were rated to be less correct than Enshi, Badong, and Jianshi. Although Jianshi was ideologically perceived to be more correct than Xuanen, Laifeng, and Hefeng, the Jianshi talkers are located on the negative side and next to the Hefeng talkers. The location of the Jianshi talkers in the second dimension can be explained by the classification of the Jianshi rural talker, who was correctly identified only 32.5% of the time and was misidentified as a Hefeng talker 45.8% of the time. The Jianshi rural talker was born and grew up in Guandian village which was near the border of Jianshi and Hefeng county (Map 3). The Guandian dialect is highly similar to the Hefeng dialect such that many Hefeng

<table>
<thead>
<tr>
<th>Talker</th>
<th>Enshi</th>
<th>Jianshi</th>
<th>Badong</th>
<th>Hefeng</th>
<th>Laifeng</th>
<th>Xuanen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enshi</td>
<td>.58</td>
<td>.14</td>
<td>.09</td>
<td>.05</td>
<td>.04</td>
<td>.10</td>
</tr>
<tr>
<td>Jianshi</td>
<td>.16</td>
<td>.53</td>
<td>.04</td>
<td>.23</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td>Badong</td>
<td>.16</td>
<td>.10</td>
<td>.61</td>
<td>.03</td>
<td>.05</td>
<td>.05</td>
</tr>
<tr>
<td>Hefeng</td>
<td>.15</td>
<td>.10</td>
<td>.06</td>
<td>.60</td>
<td>.03</td>
<td>.06</td>
</tr>
<tr>
<td>Laifeng</td>
<td>.08</td>
<td>.01</td>
<td>.01</td>
<td>.02</td>
<td>.58</td>
<td>.30</td>
</tr>
<tr>
<td>Xuanen</td>
<td>.07</td>
<td>.04</td>
<td>.06</td>
<td>.03</td>
<td>.25</td>
<td>.55</td>
</tr>
</tbody>
</table>

Table 6. Proportion of responses from each of the six response alternatives to each of the six talker groups in the dialect classification task. The correct responses for each talker group (the diagonal) are in bold.
participants classified this talker as being a Hefeng talker. Although the Jianshi urban talker was correctly classified 73.3% of the time and was classified as an Enshi talker 13.3% of the time, the high proportion of Hefeng responses for the Jianshi rural talker decreases the distance between the Jianshi and Hefeng talkers, and increases the distance between the Jianshi and Enshi talkers. The discrepancy between the perceived similarity between the Enshi and Jianshi variety as shown in the hand-drawn maps and the perceived similarity between the Jianshi and Hefeng talkers in the dialect classification task does not imply that participants’ classification performance contradicts their ideologies about dialect differences. Since the hand-drawn maps and dialect difference ratings only elicited participants’ ideologies about the differences between various imagined place-based dialects without specifying urban/rural varieties, it is likely that participants’ ideologies about the similarity between Jianshi and Enshi was based on the urban variety, rather than the rural variety. In fact, participants’ responses to the Jianshi urban talker support their prior perceptions of the similarity between Jianshi and Enshi varieties which are evident in the Enshi-Jianshi dialect region designated in the hand-drawn maps and the similar low difference ratings for Enshi and Jianshi in the dialect difference rating task. Thus, the urban varieties may play a larger role in constructing linguistic ideologies than rural varieties, at least for the participants from Enshi prefecture. The finding that participants’ ideologies about dialect differences are generally consistent with their responses to real talkers demonstrates their knowledge about dialect variation; the fact that their classification of some rural talkers does not exactly match their prior conceptions of dialect differences suggests that their ideologies associated with imagined place-based dialects are likely based on urban varieties rather than rural varieties.

In addition to the interpretable two dimensions, the distance of a particular talker from the center of the plot can be viewed as an indication of the perceived difference between that talker and the Enshi variety, the regional “norm” which is regarded as most correct and standard in Enshi Prefecture (Yan, under review). The Enshi talkers are located around the zero point of both dimensions and thus near the center of the plot. The further away from the center, the more different a talker is perceived to be from the Enshi norm. The Hefeng, Laifeng, and Badong talkers are quite far from the center, suggesting that these talkers are perceptually distinct from the Enshi norm. The Jianshi and Xuanen talkers are closer to the center than the Hefeng, Laifeng, and Badong talkers, indicating that the former is perceptually more similar to the Enshi norm than the latter. In the hand-drawn map task, Enshi and Jianshi, and Enshi and Xuanen are delimited as the second and third main dialect regions, respectively, indicating that the varieties spoken in these three counties are alike. In addition, Enshi, Jianshi, and Xuanen are rated as less different than Badong, Hefeng, and Laifeng in the dialect difference rating task. Thus, participants’ classification of real talkers is consistent with their ideological perception of dialect difference as revealed in the hand-drawn maps and the dialect difference ratings.

The INDSCAL analysis also returned weights for the two dimensions for each input matrix, and the weights are given in Table 7. Dimension 1 was weighted more heavily than Dimension 2 for all six participant groups, suggesting that the perceived dialect markedness is more relevant in assessing talker similarity than the perceived dialect correctness/standardness. Jianshi, Enshi, and Hefeng participants showed more attention to the dialect markedness dimension than other participants, and Badong participants were least attentive to this dimension. Badong is located in the northeastern corner of Enshi prefecture, far away from Xuanen and Laifeng. Due to Badong’s geographic location, Badong participants may be less sensitive to the difference between Xuanen/Laifeng dialect and other dialects than other participants. As for Dimension 2, Hefeng and Laifeng participants, who came from the counties that were perceived to be less correct and standard than the Enshi norm, were most attentive to the dialect correctness/standardness dimension. Jianshi participants showed the least attention to this dimension. Participant group differences are slightly larger in Dimension 2 than in Dimension 1, indicating that participants from different counties differed more in the attention to the dialect correctness/standardness dimension than the attention to the dialect markedness dimension.

The results of the INDSCAL analysis revealed perceptual similarity between dialect varieties, indicating that although participants varied in their ability to correctly identify the talkers’ county origin, their classification performance was not random. Participants more often
mistook the talkers from varieties that were highly similar and much less frequently confused talkers from counties with distinct varieties. The INDSCAL analysis also revealed two perceptual dimensions: the perceived dialect markedness and correctness/standardness. Participants from different counties differed in perceiving dialect similarity. Jianshi, Enshi, and Hefeng participants showed more attention to the dialect markedness than Laifeng, Xuanen, and Badong participants; Hefeng and Laifeng participants paid more attention to dialect correctness/standardness than Xuanen, Enshi, and Jianshi participants. Although a discrepancy was observed between the ideological similarity between the Enshi and Jianshi variety and the perceived similarity between the Jianshi and Hefeng talkers in the dialect classification task, the overall classification pattern demonstrates that participants’ responses to actual linguistic stimuli in the dialect classification task are largely consistent with their ideologies about dialect differences uncovered by the hand-drawn map task and the dialect difference rating task. The observed discrepancy was driven by the Jianshi rural talker, whose speech was highly similar to the Hefeng dialect. The classification of the Jianshi urban talker matched participants’ ideologies about the similarity between the Enshi and Jianshi variety. These results suggest that participants’ ideological representations of regional dialects are based more on urban varieties than rural varieties.

4.4 Interview

Participants were also asked about how they judged the county origin of the talkers in the dialect classification task. Participants’ responses to this question can be divided into two categories: those based on linguistic features and those based on paralinguistic features. Linguistic features mainly included phonetic features, intonation, and lexical items. Participants identified a range of specific phonetic features associated with various dialect varieties. For example, several features frequently noted were glide-dropping in the Jianshi urban variety, the realization of /f/ as [x] in the Enshi urban variety and the Hefeng urban and rural varieties, the pronunciation of /ŋ/ as [n] in the Enshi rural variety and the Jianshi urban variety, and the realization of /tʰŋ/ as [tʰn] or [tsʰn] in the Laifeng varieties. Glide-dropping for the Jianshi urban talker was the phonetic feature most frequently commented on by participants from throughout the prefecture. Many participants mimicked the glide-dropping pronunciations such as [pʰn], [mɛn], and [tʰɛn] (the underlying forms are /pʰɛn/, /mɛn/, and /tʰɛn/, respectively) and remarked that such pronunciations were not standard. Participants’ responses revealed their sensitivity to the acoustic-phonetic characteristics of the different dialects.

In addition to individual phonetic features, participants also relied on intonation to classify talkers. One participant, in describing how he classified the Jianshi rural talker, stated that “I can tell by his intonation; nobody from other places speaks with that intonation” [Jianshi male]. Another participant pointed out that “there is not much intonation fluctuation in Enshi dialect. It is simply flat” [Xuanen female]. One participant suggested that “Hefeng residents prolong their final intonation” [Badong male].

Dialect-specific lexical items were also commonly used to identify the county origin of the talkers. For example, one participant claimed that “some particular words, such as me de gẹ ‘what’ and guai ɛ ‘nice, good-looking,’ are sufficient to tell the talker came from Hefeng” [Enshi male], and another participant pointed out that “zen ger ɛŋ jì, ‘why’ is commonly used in Badong dialect” [Hefeng male]. Another participant noted that “nowadays the younger generation in Hefeng prefers to add war ʊ jì, ‘child’ to people’s names to show intimacy” [Badong male].

The paralinguistic features participants listed ranged from the attributes of a dialect variety to its comprehensibility. In discussing the Enshi variety, one participant remarked that “the Enshi dialect is affectedly sweet, particularly when spoken by young females” [Badong male]. Others suggested that “the Badong dialect is soft and gentle” [Laifeng female], and “the Hefeng dialect is coarse and the most accented dialect in Enshi Prefecture” [Xuanen female], while another pointed out that “Hefeng natives change the tone on the last word, which makes their speech sound more friendly” [Jianshi female]. One participant claimed that “Laifeng residents have a very peculiar way of pronouncing words such that it is not easy for me to understand” [Jianshi female]. The fact that participants pointed to both linguistic and paralinguistic features and that they provided specific and broad details about dialect variation revealed their knowledge about dialect differences in the prefecture and their sensitivity to different linguistic levels of variation.

5. Conclusions

Combining methods developed in the fields of perceptual dialectology and dialect identification and categorization, the current study aimed to address an area that was understudied: the relationship between non-linguists’ ideologies associated with dialect differences and their classification of real talkers from these varieties based on authentic voice cues, and the effect of
participant’s home county on the perception of dialect differences and classification of real talkers.

Participants’ classification of real talkers in the dialect classification task was largely consistent with their ideologies about differences among dialects as uncovered in the hand-drawn maps and dialect difference ratings. The talkers who were most frequently mistaken for each other in the dialect classification task were those who came from counties that were perceived to have similar dialects in the hand-drawn map task and the dialect difference rating task. The positive response bias for Enshi in the dialect classification task corresponds to the finding that the Enshi dialect was perceived to be least different in the dialect difference rating task.

The participant’s home county is an important factor in shaping participants’ perceptions of regional dialects. In the hand-drawn map task, how participants grouped counties as dialect regions was affected by where they came from. In the dialect difference rating task, local participants consistently rated the local dialect as most similar, as expected. Finally, in the dialect classification task, local participants performed better on local talkers than nonlocal participants overall. The INDSCAL analysis showed that participants from different counties differed in their attention to two perceptual dimensions: the perceived dialect markedness and correctness/standardness.

The results of the hand-drawn maps, the dialect difference ratings, and the dialect classification performance are not only revealed how nonlinguists divide geographic regions on a linguistic basis, but also what they think about various regional varieties. For example, in the hand-drawn map task, the Enshi-Xuanen dialect region was identified much more frequently by Xuanen participants than Enshi participants, suggesting an unbalanced perception of how similar Enshi and Xuanen dialects were by Enshi and Xuanen participants. In the dialect difference rating task, the low difference ratings Enshi received from nonlocal participants can be largely attributed to the prestige associated with Enshi throughout the prefecture. Thus most nonlocal participants (particularly Hefeng participants) wanted to align themselves with the Enshi variety more than with any other variety. In addition, in the dialect classification task, the positive response bias for Enshi further confirms the status of the Enshi variety as the regional “norm.” The current results uncovered participants’ implicit desire to affiliate with the more prestigious variety, thus providing a deeper understanding of the linguistic landscape of the Enshi community from the perspective of the community members.

Two questions remain regarding the perceptual salience of linguistic features and the perception of urban/rural varieties. First, it is unclear which linguistic features are perceptually salient for participants in classifying unfamiliar talkers. Participants’ responses in the interview indicated that some linguistic features are highly salient, and the recognition of these features alone can lead to high classification accuracy. For example, glide-dropping in the Jianshi urban talker’s speech, intonation in the Jianshi rural talker’s speech, and several particular lexical items used by the Hefeng and Badong talkers are salient properties for some participants. Additional research is needed to determine the complete set of linguistic features that are perceptually salient for listeners, and the relative salience of various linguistic features in dialect classification. In the context of Mandarin regional dialects, it is also necessary to explore the role of lexical tone and intonation.

Second, although the current study used one urban talker and one rural talker from each county to capture the difference between urban and rural dialects, this design does not fully represent the dialect variation in Enshi Prefecture. Dialect differences within the rural areas in one county are considerable; even in the urban areas, different varieties are used. Although the hand-drawn maps and dialect difference ratings were not designed to elicit participant’s perceptions of urban and rural varieties, the results of the dialect classification task showed that talker’s urban/rural origin interacted with participant’s home county and talker’s county origin to affect participant’s classification accuracy. Enshi, Laieng, and Badong participants performed better on either local urban or rural talkers than nonlocal participants, indicating that the “local advantage” may not be equally applied to talkers even from the same county. Additionally, participant’s urban/rural origin was not collected in this study, thus the effect of participant’s urban/rural origin on their perceptions and classifications of dialects is not clear. More research is needed to explore participants’ perceptions of various varieties in urban areas and those in rural areas within one county, and how participants’ urban/rural origin affects these perceptions.

In summary, this study provides a local understanding of folk perceptions of dialect varieties in an under-studied region, Enshi Prefecture in China. Through an examination of hand-drawn maps, dialect difference ratings, and dialect classification performance, I have shown that nonlinguists are able to identify perceptual dialect regions and explicitly judge the degree of dialect difference in an area much smaller than a single province. They are also able to accurately classify unfamiliar talkers by county origin based on short speech samples, and importantly, their classification performance was largely consistent with their prior conception of differences among the dialects. These results indicate that nonlinguists’ beliefs about dialect variation have a linguistic basis; the dialect differences
are not only represented in the minds of nonlinguists, but also commented on by them in the interview.

This study adds to the discussion of perceptual dialectology and dialect categorization in China, where more similar studies are needed. Although substantial production work has been conducted in China to investigate dialect variation, very few perceptual dialectology or dialect categorization studies have been done to examine how nonlinguists perceive their own dialects and other dialects spoken in China. The present study shows that the perception of dialect variation can be examined within small geographical regions and can yield results that complement production studies.

Acknowledgements
I would like to thank Cynthia Clopper for her support at various stages in this project and insightful comments on the earlier drafts of this paper. I also wish to thank Kathryn Campbell-Kibler, Don Winford, and Marjorie Chan for their help and feedback. My gratitude also goes to the participants who contributed to this work. This project was funded by The Ohio State Linguistics Department.

Notes
1 The retroflexes become palatals before /u/ in the Laifeng dialect. This /u/ does not correspond to the /u/ in Standard Mandarin. The Standard Mandarin /u/ is realized differently in the Laifeng dialect. The Standard Mandarin /u/ corresponds to [y] when it follows retroflexes in the Laifeng dialect.

2 There were several Enshi, Badong, and Xuanen participants who assigned score 2 to their own variety, which is reflected in the mean scores in the corresponding cells for Enshi, Badong, and Xuanen in Table 2. In the dialect difference rating task, a rating of 1 indicated the dialect in question was the same or very similar to participants' own dialect, and 2 indicated "a little different." When participants evaluated the degree of difference of their own variety, in principle the rating should be 1. The ratings of 2 given by some participants may reflect their perception that their own dialect was a little different from the "imagined" dialect spoken in the county where they lived.

3 It is possible that the classification accuracy was also affected by participant's urban/rural origin. However, this information was not collected from every participant; thus, the effect of listeners' urban/rural origin on their classification performance cannot be determined.

References


Chao, Yuanren, Shengshu Ding, Shifeng Yang, Zongji Wu & Tongsu Dong. 1946. Hubei fangyan diaocha baogao [A report on Hubei dialects]. Shanghai: Shangwu yingshuguan.


Yan, Qingyang (under review). The perceptual evaluation of Enshi Mandarin regional varieties. Submitted to Language in Society.

Dialect change and its consequences for the Dutch dialect landscape. How much is due to the standard variety and how much is not?

Wilbert Heeringa,¹* and Frans Hinskens²

¹ Institute of German studies, Carl von Ossietzky University of Oldenburg, Germany
² Meertens Institute, The Netherlands

We recorded older male speakers and younger female speakers of 86 local dialects of Dutch. Using these data, we analyze and visualize the influence of standard Dutch on apparent time changes in these dialects. Focusing for the most part on variation in the sound components, we test whether (I) dialect change is mainly the result of convergence to standard Dutch, (II) sound changes in two dialects which make them converge to standard Dutch also make them more similar, and (III) sound changes in two dialects which make them diverge from standard Dutch also make them less similar. We used three-dimensional (first hypothesis) and five-dimensional Levenshtein distance implementations (second and third hypothesis). These implementations are a novel step in dialectometry and in the study of ongoing processes of language change and their consequences for the dialect landscape. The findings corroborate all three hypotheses.

1. Introduction: Terminology, Research Questions and Hypotheses

In the Oxford English Dictionary the term ‘dialect’ is defined as “a variety of a language that is a characteristic of a particular group of the language’s speakers.” Usually the speakers are found in a geographically limited part of a language area which is ‘roofed’ by a structurally related standard variety. Several geographically adjacent dialects typically form a ‘dialect continuum’, which is described by Chambers & Trudgill (1998:5) from the perspective of a traveler:

If we travel from village to village, in a particular direction, we notice linguistic differences which distinguish one village from another. Sometimes the differences will be larger, and sometimes smaller, but they will be cumulative. The further we get from our starting point, the larger the differences will become.

Like all living language systems, dialects are constantly changing (Auer, Hinskens & Kerswill, 2005), and consequently, dialect landscapes do so as well. Both the mutual relationship between dialects and their relationship to the overarching standard variety may change. Dialect convergence can be defined as the increase in similarity between dialects, with “partial similarities increasing at the expense of differences” (Weinreich, 1954:395). Processes leading to convergence affect the structure of (dialect or standard) varieties and thus that of the diasystem and linguistic repertoires they are part of; these processes result in unification, focusing and homogenization of the linguistic repertoire. Dialect divergence is a decrease in similarity between dialects, which amounts to linguistic diversification, growing diffusion and heterogenization—although it may lead to focusing in a repertoire, making the surviving varieties more distinct from each other. Convergence and divergence are relational concepts, affecting the degree of structural distance between dialects.

A koine is a lingua franca which incorporates features of various dialects; koines mostly develop through dialect mixing, simplification and reduction (cf. Hinskens, Auer & Kerswill, 2005:11; Kerswill, 2002; Siegel, 1985, 2001; Trudgill, 1986). A koine is thus a compromise dialect. In the processes which give rise to koines, reduction consists of the elimination of the most peculiar features of the constituent dialects; typically these are local features, but of course features can also be involved which are specific to different community types, such as neighborhoods. Siegel (2001) distinguishes between regional koines and immigrant koines. The notion of regiolect was coined in 1983 by the Dutch dialectologist Hoppenbrouwers. In Hoppenbrouwers’ conception, a regiolect is a continuum of subtly different intermediate varieties in the structural space between the traditional dialects and the standard variety (cf. Bellmann’s (1996) diaglossia); the various intermediate varieties form a continuum between the traditional dialects and standard Dutch. Meanwhile, the notion of regiolect has become important in European sociodialectological studies, though many use it to refer...
to the result of specifically cross-dialectal convergence, a regional koine. In German dialectology, the notion ‘Ausgleichdialekt’ is sometimes used to refer to this type of variety.

There exist several quantitative studies which focus on dialect change, which is frequently expressed in terms of convergence to the standard variety. Examples for the Dutch language area are Giesbers (2008), Heeringa & Hinskens (2014), Heeringa & Nerbonne (2000) and Heeringa, Nerbonne, Niebaum, Nieuweboer & Kleiweg (2000).

Heeringa & Nerbonne (2000) used material collected by Winkler in 1874 and Scholtmeijer in 1996 on a range of Dutch dialects, and found that 23 of the 41 varieties studied have converged towards standard Dutch at the level of the sound components, i.e. phonetics, phonology and morpho-phonology.

Heeringa et al. (2000) studied eight varieties spoken in or close to the German county of Bentheim, along with nine neighbouring Dutch varieties, using data from the Reeks Nederlandse Dialectatlassen ‘Series of Dutch dialect atlases’ (Blancquaert & Pée, 1925-1982) that was gathered in 1974-1975, as well as new data collected in 1999. These 17 varieties were compared to standard Dutch and standard German at the level of the sound components. All of the Dutch dialects were found to be converging towards standard Dutch, while all German dialects studied appeared to be converging towards standard German.

Giesbers (2008) studied the Kleverland dialect continuum, which extends from Duisburg in Germany to Nijmegen in the Netherlands. In the early 19th century, the Dutch-German national border was drawn through this dialect continuum, and Giesbers studied five dialect pairs, each of which consisted of a Dutch and a German local dialect. Giesbers found that Dutch varieties have converged more strongly towards standard Dutch than the German ones have towards standard German, both at the lexical and sound components levels. Particularly on the lexical level, German speakers seem to preserve more old dialect forms than their Dutch counterparts.

Giesbers (2008) studied the Kleverland dialect continuum, which extends from Duisburg in Germany to Nijmegen in the Netherlands. In the early 19th century, the Dutch-German national border was drawn through this dialect continuum, and Giesbers studied five dialect pairs, each of which consisted of a Dutch and a German local dialect. Giesbers found that Dutch varieties have converged more strongly towards standard Dutch than the German ones have towards standard German, both at the lexical and sound components levels. Particularly on the lexical level, German speakers seem to preserve more old dialect forms than their Dutch counterparts.

Heeringa & Hinskens (2014) measured dialect change in apparent time. Between 2008 and 2011, they compiled a large corpus database of dialect recordings for 86 local dialects spoken in the Netherlands and the northern part of Belgium (‘Flanders’) and of Standard Netherlandic Dutch and Standard Belgian Dutch. In each of the 86 locations, two older male speakers and two younger female speakers were recorded. Comparing the dialect use of the older male speakers with that of the younger females, they measured dialect change in apparent time, focusing on lexis, morphology and sound components. They found that (a) dialect change is a geographically capricious process, and (b) the lexical level has been affected the most, while the morphological level is the most stable. While dialects in the Netherlands appear to be converging significantly towards standard Dutch, this is not the case for the Belgian dialects. Dialects have in general converged towards each other. The number of different dialect groups has slightly increased at the lexical level, and decreased at the morphological sound component levels.

In the studies mentioned here, dialects are shown to converge to their overarching standard varieties. Although the studies make clear that dialect change is at least partly the result of convergence to the standard variety, they do not establish to what extent dialect change is a matter of convergence to the standard variety. Furthermore, it remains unclear to what extent dialect change involved divergence from the standard variety.

In Heeringa & Hinskens (2014), dialects have already been shown to converge to each other. The question may arise whether there exists a relationship between convergence to the standard variety and convergence between dialects. To what extent does convergence to the standard variety make dialects more similar to each other?

Whereas dialect continua are the residue of language history (some of which is still visible in the present-day dialect landscape), the dynamics in the relationships between the standard variety and the dialects on the one hand and among the dialects on the other hand are occurring today, and they constitute the object of the present study.

The current paper builds further on the work of Heeringa & Hinskens (2014), using the same database which contains recordings of 86 Dutch local dialect varieties. In this paper we analyze and visualize the influence of standard Dutch on these dialects. We will focus on variation in the sound components and test the following hypotheses:

1. Dialect change is mainly the result of convergence to standard Dutch.
2. Sound changes in two dialects which make them converge to standard Dutch make them also closer to each other.
3. Sound changes in two dialects which make them diverge from standard Dutch make them also more distant from each other.

We will test these hypotheses using quantitative methods. In Section 2, we describe the data that form the basis of this study. In Section 3, we will test the hypotheses, and in Section 4 we draw our conclusions and present some proposals for future studies.

2. Collecting and Analyzing the Data

2.1 Collecting

Between 2008 and 2011, we compiled a large corpus database of dialect recordings for 86 local dialects and
three standard varieties, namely Standard Netherlandic Dutch, Standard Belgian Dutch, and Afrikaans; the latter will not concern us in the present paper. The dialects are evenly spread over the Dutch and Frisian language areas and represent the major dialect regions (see Map 1). The Dutch language area comprises both the Netherlands and the northern part of Belgium (or ‘Flanders’). Five of the recordings have been made by Stichting Ons Bildt, and represent the varieties of Frisian spoken in Westhoek, Sint Jacobiparochie, Nij Altoenea, Vrouwenparochie and Oudebildtzijl. Along with Sint Annaparochie, they are shown as a dense cluster of dots in the north-west of Frisia, close to the coast.

Dialect change in this paper is measured in apparent time. For this purpose, at least two male speakers aged 60 or older, and two or more female speakers aged between 20 and 40 were recorded in each of the 86 locations. The males represent the older phase of the dialect variety at issue and the females the newer phase. The reasoning for this is that we assume that there is a scale of conservativeness, where older males are the most conservative speakers, followed by the older females, who are in turn followed by the younger males. The younger females, meanwhile, are the least conservative speakers. In general, the speech of young speakers tends to be more innovative than that of older members of their speech community. In addition, according to Romaine (1984:113), the speech of females is usually more innovative than that of males: “women consistently produce forms which are nearer to the prestige norm more frequently than men” (see also Labov, 1990:206; Chambers, 1995:102-103; for the Dutch language area e.g. Boves & Gerritsen, 1995:195-226).² At first sight, it may look as if the dimensions of age and gender are intertwined, but they can in fact be reduced to

Map 1. Distribution of 86 Dutch dialect varieties. The Dutch provinces are shaded in light gray marble and the Belgian provinces are shaded in dark gray pine.
just one: innovativeness, with older males at one extreme and younger females at the other.

Using the approach set out above, we abstract from potentially socially bound dialect-internal variations in order to broaden the study in dialect geographical respects. This enables us to gain greater insight into the present-day dynamics of the Dutch dialect landscape at large.

An episode of the Charlie Chaplin movie ‘The Kid’ served as the basis of the recordings we made. This part of the film focuses on a neighborhood where many windows suddenly get broken. By accident (or so it seems), a glazier is walking around in the same area and is very keen to carry out the necessary repairs. Meanwhile, a policeman tries to find out why so many windows were broken in such a short period of time. At some point, he sees a little boy who is just about to throw a stone at a window. The policeman then realizes that the child is doing this on the orders of the glazier; the policeman tries to catch both protagonists but is unable to do so.

The story was presented to our dialect speakers by way of stills from the movie as well as in narrative form, presented in written form in the standard variety. The episode can be regarded as a cross-section of plain, simple daily spoken language, and consists of 23 sentences, each containing an average of 7.6 words. We used a selection of 13 sentences for this study, which include a maximum of 125 words in the written standard Dutch version of the text. 90 different word types (lexical forms) are represented. Both the older male and the younger female speakers operated in small groups. The number of informants per group varied between two and four, but in the overwhelming majority of cases a small group consisted of two people. When a small group was being recorded, the individuals were asked to write down a translation of the text in their own dialect, independently of each other. Then, they compared their translations and discussed the differences between them. For each difference, they had to decide which alternative was the better version. They were then asked to write a new dialect translation together, which might be seen as a consensus version upon which both of them agreed; this approach helped us to counter the observer’s paradox as well as to reduce potential idiolectal noise as much as possible. Finally, they both (or all three or four) read this third version of the text aloud.

Additionally, we created recordings in Standard Netherlandic Dutch (read out by Maartje van Weegen, host of the Dutch national classical radio station), Standard Belgian Dutch (by Martine Tanghe, Belgian presenter of the Flemish public broadcast station, VRT) and Standard Afrikaans (by Mariëtta Kruger, presenter at the South African television channel, SABC 2). The Dutch and Flemish standard speakers read the standard Dutch text aloud, and the South-African standard speaker read the Afrikaans text aloud. They were all aged between 50 and 60 when they were recorded, thus being in an intermediate position between the older male and the younger female dialect speakers. In their respective speech communities, the three presenters from the national public radio/TV are generally considered as models for ‘correct’ Dutch/Afrikaans. In this paper, the Standard Netherlandic Dutch recording is used only; we refer to it as standard Dutch throughout the paper. The present study is thus based on one older male consensus dialect version and one younger female dialect version of the story for each single local dialect, plus the standard variety version.

Subsequently, the first author made phonetic transcriptions of the recordings in order to measure dialect change, changes in the relationships among dialects, and changes in the relationship to standard Netherlandic and Belgian Dutch. Usually, two recordings of the consensus dialect version of the story were produced for both the older males and the younger females. Since phonetic transcription is time-consuming, only one recording per group was transcribed. When selecting which recording to use, we had a preference for the speaker who was the most autochthonous (established through their geographical background and that of their parents), had the clearest voice, and read the text most fluently. The transcriptions were made in IPA and digitized in X-SAMPA.

2.2 Analyzing

In Section 1, we hypothesized that dialect change is mainly the result of convergence to standard Dutch. In order to test this hypothesis, a further analysis of the dialect change measurements is required. We need to split the measurements into 1) a part which is the result of sound changes which make a dialect converge to standard Dutch, and 2) a part which is the result of sound changes which make a dialect diverge from standard Dutch. When the first part is significantly larger than the second part, we have proved that our hypothesis is true.

In Section 2.2.1, we explain how we measure dialect change. In Section 2.2.2, we discuss the methodology with which we split the dialect change measurement into two parts. In Section 2.2.3, we present some further considerations concerning the complexity of the methodology. In Section 2.2.4, we measure changes of the relationships between dialects and again show how to split those measurements into two parts.

2.2.1 Measuring Dialect Change

The extent to which a given dialect has changed in regards to the sound components is measured by...
comparing the transcriptions of the old male speakers with the transcriptions of the young female speakers. Comparisons are made with the aid of the Levenshtein distance metric (Levenshtein, 1966). This algorithm was introduced into dialectology by Kessler in 1995. The Levenshtein distance between two phonetic strings is calculated as the ‘cost’ of the total set of insertions, deletions and substitutions needed to transform one string into another (Kruskal, 1999). In the original, simplest form of the algorithm, all operations have the same cost, e.g. 1. We illustrate this with an example. In Diepenbeek the older male speaker pronounced st ꞏ r ± a t ‘street’ as [strɔːdə] and the younger female speaker pronounced the same word as [strɔət]. When ignoring potential suprasegmentals and diacritics, the Levenshtein algorithm will take the alignment as shown in Table 1 as a basis for the distance calculation. We obtain a total cost of 4 operations (two substitutions, an insertion and a deletion), and the alignment length is equal to 7.

In this paper, with respect to the cost of the operations needed to establish segment distances, we used graded weights that are effectively segment distances, meaning that the pair [i, ɒ] is seen as being more different than the pair [i, ɪ]. The segment distances are obtained on the basis of acoustic representations of a canonical set of IPA samples (see Heeringa, 2004:79-119).

A restricted set of diacritics was used in the transcriptions for this study and processed by the distance measure. We did this for the palatalization of final /n/ ([ɲ]), the verlarization of initial l ([l̥]) and nasalization (for example [ɻ]). When a given segment was compared to a palatalized segment, the segment distance was averaged by the distance between the segments in question and [j]. So, for example, the distance between [t] and [ɲ] is equal to the average of the distance between [t] and [n] and between [t] and [j]. In the case of a comparison with verlarized and nasalized segments, the distances were averaged by the distances to [y] and [n], respectively.

The Levenshtein algorithm is adapted when it comes to dealing with syllabification in words, meaning that only a vowel can be matched with a vowel and a consonant with a consonant. The glides [j] or [w] may also match with a vowel (or vice versa), and [i] or [u] with a consonant (or vice versa). A central vowel (in our research only schwa) may be matched with a sonorant (or vice versa) or a full vowel (since schwa is the reduced vowel). In this way, unlikely matches (e.g. [p] with [a]) are precluded.

Distances are calculated between the members of a pair of variants that are lexically and morphologically identical. We used normalized distance measures, calculating the aggregated distance between two dialects as the sum of a maximum of 125 word pair distances divided by the sum of the alignment lengths that correspond to the word pairs. We illustrate this with an example in which we measure dialect change in the dialect of Diepenbeek by comparing a subset of four word realizations by the older male speaker with a set containing the corresponding word realizations by the younger female speaker (see Table 2).

In this example, for didactic reasons we use binary (rather than scalar or gradient) operation weights: insertions, deletions and substitutions count as 1 and matches count as 0. The Levenshtein distances are found in the fifth column and the corresponding alignment lengths in the sixth. The sum of the Levenshtein distances of the four word pairs is 6, and the sum of the alignment lengths is 19. The amount of dialect change is calculated as a percentage: 6 divided by 19, multiplied

| Table 1. Alignment of the realization of the dialect variant of ‘street’ of an older male speaker with the realization of the younger female speaker. The lowest line give the operations needed to change the first realization into the second one. |
|---|---|---|---|---|---|---|
| 1 2 3 4 5 6 7 |
| older male | s t r o d ə |
| younger female | s t r ə ə t |
| sub. ins. sub. del. |

| Table 2. Calculation of dialect change in Diepenbeek on the basis of four word pairs. |
|---|---|---|---|---|
| standard Dutch orthography | English orthography | transcription older male speaker | transcription younger female speaker | distance | alignment length |
| straat | street | stʃɔːdə | stʃət | 4 | 7 |
| politie | police | polis | polis | 1 | 5 |
| goed | good | ʧut | ʧut | 0 | 3 |
| juist | just | ʧyst | ʧyst | 1 | 4 |
| 6 | 19 |
by 100 = 31.6%. However, rather than using binary operation weights, we use graded operation weights throughout this paper, as mentioned above. These weights vary between 0 and 1.

Dialect change measurements are based on 125 word pairs (fewer if words were missing). We found that all 125 words vary in terms of the sound components across the dialects. Several words appeared in the text more than once; for example, the word ‘straat’ ‘street’ appeared three times. So, when calculating dialect change, for ‘straat’ each of the corresponding word pair distances counted for one third. In this way, each word pair was weighed. The sum of the weights was 90, which is the number of the different word types.

Dialectometry and Levenshtein distances have meanwhile become common in some lines of socio-linguistic literature; evidence supporting the overall usefulness of these methods in comparing dialects includes the match between Levenshtein measurements with the outcomes of perception tests, as in e.g. Gooskens & Heeringa (2004).

Using the Levenshtein distance, the apparent time change of a particular dialect in the sound components was measured by comparing the phonetic realization of the words by the older male informants with the realization of the same words by the younger female participants. Word pairs are only considered when the realizations are lexically and morphologically the same. Besides, the realizations should be lexically and morphologically identical to the corresponding word used in the standard Dutch text. This is done so in order to obtain results which are comparable in every relevant dimension.

The results are presented in Map 2. Individual local dialects are represented by dots in each of the maps. The darker a dot, the more the dialect has changed. It is difficult to recognize particular patterns in and similarities across the dialects. Indeed, it seems that dialect change is a capricious process, and that no particular regional area has changed more than others. The average dialect change is equal to 13.3% with a standard deviation of 3.8%.

2.2.2 Relating Dialect Change to Change in the Relationship to Standard Dutch

The Levenshtein distance in its original form is two-dimensional. The algorithm, explained above in 2.2, compares two strings with each other and finds the least costly set of operations which map the one string onto the other. We explained the algorithm briefly in Section 2.2.1. When measuring dialect change, word pairs are compared to each other by the algorithm. Each word pair consists of the realization of an older male and a younger female.

In this study we want to distinguish between sound changes which cause a dialect to converge to standard Dutch, sound changes which cause a dialect to diverge from standard Dutch, and sound changes which do not affect the relationship to standard Dutch. In order to make this three-way distinction, we need to take into account a third realization which is aligned with the other two, namely the realization in standard Dutch. In other words: we need a three-dimensional Levenshtein distance, since three realizations need to be aligned to each other: older male vs. younger female vs. standard Dutch. We illustrate this by a hypothetical example. Assume in some dialect the standard Dutch word ‘hart’ is pronounced as [hɛrt] by the older male speaker and as [ʔɑrt] by the younger female speaker. The standard Dutch pronunciation is [hɑrt]; although the variants of /r/, which mainly differ in their place of articulation, are not mutually ranked in terms of prestige or ‘standardness.’ The three realizations need to be aligned as is shown in Table 3.

Table 3. Alignment of three realizations: older male vs. younger female vs. standard Dutch. The operation weights are given on the bottom line.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>older male</td>
<td>h</td>
<td>e</td>
<td>r</td>
<td>t</td>
<td>0</td>
</tr>
<tr>
<td>younger female</td>
<td>ʔ</td>
<td>a</td>
<td>r</td>
<td>t</td>
<td>0.67</td>
</tr>
<tr>
<td>standard Dutch</td>
<td>h</td>
<td>a</td>
<td>r</td>
<td>t</td>
<td>0.67</td>
</tr>
</tbody>
</table>

0.67 0.67 1.00 0 0.67
The alignment consists of five slots. The operation weights of each of the five slots are calculated by considering all pairs of segments within the slot. Per slot there are three pairs: older male versus younger female, older male versus standard Dutch, and younger female versus standard Dutch. When using binary weights, we get the comparisons in the first slot as shown in Table 4.

The sum of the weights is equal to 2. Since there are three segment pairs, we calculate the average operation weight of the first slot as \( 2/3 = 0.67 \). For the sake of simplicity, we used binary weights here, but as explained in Section 2.2.1 the results in this paper are obtained on the basis of measures which use graded weights, varying between 0 and 1.

When inspecting the alignment in Table 3, we find two sound changes which cause the dialect to converge to standard Dutch (slots 2 and 5), one sound change which causes the dialect to diverge from standard Dutch (slot 1), and one sound change which does not affect the relationship to standard Dutch (slot 3, see Table 5). The amount of change as the result of convergence to standard Dutch is \( 0.67 + 0.67 = 1.34 \) (cf. Table 3), the amount of change as the result of divergence from standard Dutch is 0.67, and the amount of change not affecting the distance to standard Dutch is 1.00.

Just as for the dialect change measurements in Section 2.2.1, we use normalized distance measures. We calculate the aggregated sound change due to convergence to standard Dutch as found for the maximally 125 word pairs and divide them by the sum of the alignment lengths that correspond to the word pairs. The same procedure applies mutatis mutandis (the same) for dialect measurements on the basis of sound changes which cause a dialect to diverge from standard Dutch or which do not affect the relationship to standard Dutch.

### 2.2.3 Complexity

As we mentioned above, the original Levenshtein distance, as considered by Vladimir Levenshtein in 1965, is two-dimensional, since it compares two strings with each other. Given string 1 and string 2, there are three possible edit operations: deletion of an element in string 1, insertion of an element in string 2, and substitution of an element in string 1 by an element in string 2. This is schematically shown in Table 6; ‘s’ stands for (sound) segment. The number of operations is equal to \( 2^n - 1 \), where \( n \) is the number of strings to be compared. When two strings are compared, the number of operations is \( 2^2 - 1 = 3 \).

In this paper we use a three-dimensional variant of Levenshtein distance. The number of operations is \( 2^3 - 1 = 7 \). The operations are schematically shown in Table 7. Each operation is actually a combination of three operations, where each of the three operations is an insertion, deletion or substitution. For example, in the seventh slot element \( s1 \) in string 1 is substituted by element \( s2 \) in string 2, and by element \( s3 \) in string 3 and element \( s2 \) in string 2 is substituted by element \( s3 \) in string 3. Therefore the operation for this slot consists of three substitutions. In the fourth slot element \( s1 \) in string 1 is substituted by element \( s2 \) in string 2, and

<table>
<thead>
<tr>
<th>Table 4.</th>
<th>Comparisons in the first slot. The sum of the weights is 2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>versus</td>
</tr>
<tr>
<td>h</td>
<td>versus</td>
</tr>
<tr>
<td>?</td>
<td>versus</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5. Two sound changes cause the dialect to converge to standard Dutch; one sound change causes the dialect to diverge from standard Dutch, and one sound change does not affect the relationship to standard Dutch.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>older male</td>
<td>h</td>
<td>r</td>
<td>r</td>
<td>t</td>
<td>o</td>
</tr>
<tr>
<td>younger female</td>
<td>?</td>
<td>a</td>
<td>a</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>standard Dutch</td>
<td>h</td>
<td>a</td>
<td>r</td>
<td>t</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6. Two-dimensional Levenshtein distance has three operations: an element in string 1 (s1) can be deleted, an element in string 2 (s2) can be inserted, and an element in string 1 (s1) can be substituted by an element in string 2 (s2).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>string 1</td>
<td>s1</td>
<td>0</td>
<td>s1</td>
</tr>
<tr>
<td>string 2</td>
<td>0</td>
<td>s2</td>
<td>s2</td>
</tr>
</tbody>
</table>

### Table 7. Three-dimensional Levenshtein has seven operations.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>string 1</td>
<td>s1</td>
<td>0</td>
<td>0</td>
<td>s1</td>
<td>s1</td>
<td>0</td>
<td>s1</td>
</tr>
<tr>
<td>string 2</td>
<td>0</td>
<td>s2</td>
<td>0</td>
<td>s2</td>
<td>0</td>
<td>s2</td>
<td>s2</td>
</tr>
<tr>
<td>string 3</td>
<td>0</td>
<td>0</td>
<td>s3</td>
<td>0</td>
<td>s3</td>
<td>s3</td>
<td>s3</td>
</tr>
</tbody>
</table>

s1-s2 del. ins. sub. del. ins. sub. s1-s3 del. ins. del. sub. ins. sub. s2-s3 del. ins. del. ins. sub. sub.
both element \( s_1 \) and element \( s_2 \) are deleted in string 3. The operation includes one substitution and two deletions. As we saw in Section 2.2.1 we calculate the average weight as the average of the three operation weights.

When considering multidimensional Levenshtein distance, we find that the number of operations increases exponentially, as shown in Figure 1. Four-dimensional Levenshtein distance has 15 operations and five-dimensional Levenshtein distance has 31 operations.

2.2.4 Measuring Changes of the Relationships between Dialects

In Section 1, we hypothesized that sound changes in two dialects which make them converge to standard Dutch, make them also grow closer to each other (second hypothesis), and that sound changes in two dialects which make them diverge from standard Dutch, make them also more distant from each other (third hypothesis). In order to test these hypotheses we measure convergence/divergence due to sound changes in the dialects—the sound changes in A corresponding with the ones in B, i.e. each of the sound changes in A are found in the same word at the same position as the sound changes in B—which cause both A and B either to converge to standard Dutch (measurements of second hypothesis) or to diverge from standard Dutch (measurements of third hypothesis).

When considering dialect pair \( A/B \), per word under consideration we need to align the realizations of the older male speaker of dialect \( A \), the younger female speaker of dialect \( A \), the older male speaker of dialect \( B \), the younger female speaker of dialect \( B \), and standard Dutch. For this purpose we use a five-dimensional Levenshtein distance, since five realizations need to be aligned to each other.

We illustrate the use of the five-dimensional Levenshtein distance by means of a hypothetical example. We measure the amount of change in the relationship between dialects \( A \) and \( B \) which amounts to the convergence to standard Dutch of both \( A \) and \( B \), and we measure the amount of change in the relationship between dialects \( A \) and \( B \) which diverge from standard Dutch of both \( A \) and \( B \). Assume the following realizations:

<table>
<thead>
<tr>
<th></th>
<th>older male speaker of dialect ( A )</th>
<th>younger male speaker of dialect ( A )</th>
<th>older male speaker of dialect ( B )</th>
<th>younger female speaker of dialect ( B )</th>
<th>standard Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z ) ( \chi ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
</tr>
</tbody>
</table>

The five-dimensional Levenshtein will align the realizations as follows:

<table>
<thead>
<tr>
<th></th>
<th>older male speaker of dialect ( A )</th>
<th>younger female speaker of dialect ( A )</th>
<th>older male speaker of dialect ( B )</th>
<th>younger female speaker of dialect ( B )</th>
<th>standard Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
</tr>
</tbody>
</table>

When focusing on dialect \( A \) we find one sound change which causes divergence from standard Dutch, and three sound changes which cause convergence to standard Dutch:

<table>
<thead>
<tr>
<th></th>
<th>older male speaker of dialect ( A )</th>
<th>younger female speaker of dialect ( A )</th>
<th>standard Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
</tr>
</tbody>
</table>

When focusing on dialect \( B \) again we find one sound change which causes divergence from standard Dutch, and three sound changes which cause convergence to standard Dutch:

<table>
<thead>
<tr>
<th></th>
<th>older male speaker of dialect ( B )</th>
<th>younger female speaker of dialect ( B )</th>
<th>standard Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
<td>( \chi ) ( \varepsilon ) ( r ) ( d ) ( z )</td>
</tr>
</tbody>
</table>
When comparing dialect A with dialect B we find three slots where dialects A and B converge to each other and one slot where dialects A and B diverge from each other:

<table>
<thead>
<tr>
<th>older male speaker of dialect A</th>
<th>χ</th>
<th>e</th>
<th>r</th>
<th>d</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>older male speaker of dialect B</td>
<td>h</td>
<td>e</td>
<td>r</td>
<td>s</td>
<td></td>
</tr>
<tr>
<td>younger female speaker of dialect A</td>
<td>a</td>
<td>r</td>
<td>t</td>
<td></td>
<td></td>
</tr>
<tr>
<td>younger female speaker of dialect B</td>
<td>a</td>
<td>r</td>
<td>t</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We can summarize the results as follows:

<table>
<thead>
<tr>
<th>dialect A versus dialect B</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The amount of change in the relationship between dialects A and B is calculated as the sum of the weights divided by the length of the alignment: 1/5 = 0.2 or 20%. This is a positive percentage which means that dialects A and B converge to each other on average. So in this imaginary example, the overall convergence of dialects A and B to the standard variety amounts to cross-dialectal convergence or koineization, but their overall divergence from the standard variety does so as well.

In Section 2.2.1 we explained that distances between dialects are calculated as the aggregate of maximally 125 word pair distances. Similarly, the amount of change in the relationship between two dialects (amounting to either convergence to standard Dutch or divergence from standard Dutch) is obtained on the basis of maximally 125 sets of five realizations. Usually the number of words will be smaller since, for the sake of comparison, all of the words in a set need to be cognates of each other—which is not always the case. We implicitly normalize over the number of realized sets, since the amount of change in the relationship between two dialects (resulting in either convergence to standard Dutch or divergence from standard Dutch) is calculated as the sum of the weights (which are calculated as explained in this section) divided by the sum of the alignment lengths of the words in the word sets.

3. Testing the Hypotheses

In Section 3.1, we focus on the change of dialects where we test the first hypothesis which we mentioned in Section 1. In Section 3.2 we study the change of relationships between dialects and test the second and third hypothesis.

3.1 Testing the First Hypothesis

In Section 2.2.1, we explained that dialect change is measured as a percentage. A percentage gives the average degree to which segments in the realizations of words of a dialect spoken by an older male have changed, resulting in the segments in the realizations of words of the same dialect spoken by a younger female (cf. Table 2).
Dutch, change not affecting the relationship to standard Dutch, and change resulting in divergence from standard Dutch.

The results are shown in Map 3. The figure shows dialect change measured on the basis of all sound changes (top left), dialect change on the basis of sound changes which do not affect the distance to standard Dutch (top right), dialect change on the basis of sound changes which cause dialects to converge to standard Dutch (bottom left) and on the basis of sound changes which cause dialects to diverge from standard Dutch (bottom right). The intensity of blue in a dot represents the extent to which a variety has changed.

<table>
<thead>
<tr>
<th></th>
<th>all changes</th>
<th>convergence</th>
<th>neutral</th>
<th>divergence</th>
</tr>
</thead>
<tbody>
<tr>
<td>all changes</td>
<td>0.76****</td>
<td>0.36**</td>
<td>0.71****</td>
<td></td>
</tr>
<tr>
<td>convergence</td>
<td>0.08</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>neutral</td>
<td></td>
<td></td>
<td>0.31**</td>
<td></td>
</tr>
<tr>
<td>divergence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The intensity of blue in a dot represents the extent to which a variety has changed. Table 8 shows the correlations between the different types of dialect change measurements. ** means: p < 0.01, *** means: p < 0.001, **** means: p < 0.0001.
dialects to diverge from standard Dutch correlate significantly with the pattern on the basis of all sound changes. Furthermore, the ‘neutral pattern’ and the ‘divergence pattern’ significantly correlate to each other.

In Section 2.2.1, we found that the average dialect change is equal to 13.3%. This means that on average the dialects studied remain 86.8% the same. The average change resulting in convergence to standard Dutch is equal to 6.8%, the average change not affecting the relationship to standard Dutch is 0.6%, and the average change resulting in divergence from standard Dutch is equal to 5.9%. Therefore, the greater part of change results in convergence to standard Dutch: 6.8/13.3 = 51.1%. Of course, this does not mean that divergence would be any less important; in fact, the third hypothesis (which will be tested below) concerns the divergence from standard Dutch.

In Section 1, we hypothesized that dialect change is mainly the result of convergence to standard Dutch. In order to test this hypothesis, we need to find out whether the average change resulting in convergence to standard Dutch is significantly higher than the two other types of dialect change (‘neutral’ and ‘divergence’). We used a paired-samples $t$-test and found that change due to convergence to standard Dutch is significantly larger than both ‘neutral change’ ($t = 21.932, p < 0.0001$) and change due to divergence from standard Dutch ($t = 2.644, p < 0.01$). Therefore, we conclude that dialects change mainly as the result of convergence to standard Dutch.

Dialect change not affecting the distance to standard Dutch is extremely small (0.6%) and not only significantly smaller than change due to convergence to standard Dutch, but also smaller than change due to divergence from standard Dutch ($t = 22.162, p < 0.0001$).

3.2 Testing the Second and Third Hypothesis

The previous section focused on aggregated change of individual dialects. The findings discussed in that section indicate that the relevant hypothesis is correct. In this section we focus on change in relationships between dialects, especially on measurements of the change in the distance between dialect pairs as the result of convergence to standard Dutch. These measurements may be either positive or negative. When the distance change of a dialect pair is positive, this means that dialects have converged to each other on average. When the distance change of a dialect pair is negative, the dialects have diverged from each other on average. The results are shown in Figure 2. The interpretation of the maps requires some caution. The maps do not tell us whether dialects mainly converge to or diverge from standard Dutch. For example, the map in the center tells us whether sound changes which cause two dialects to converge to standard Dutch also cause them to converge to each other. If so, this does not necessarily imply that each of the two dialects have mainly converged to standard Dutch. The change of a dialect may involve both sound changes which make realizations of some words closer to the standard Dutch realizations, and sound changes which make realizations of some words more distant to standard Dutch realizations. When most sound changes make a dialect linguistically closer to standard Dutch, we would say that the dialect has converged to standard Dutch. This, however, is not shown in the maps; we simply distinguish between two types of sound change—change causing convergence to standard Dutch and change causing divergence from standard Dutch—and the consequences of both types of sounds changes for the mutual relationships between dialects (see Section 2.2.3).

![Figure 2](image-url)

Figure 2. Convergence/divergence between dialects measured on the basis of all sound changes (left), on the basis of sound changes which cause dialects to converge to standard Dutch (center) and on the basis of sound changes which cause dialects to diverge from standard Dutch (right). Red lines indicate convergence and blue lines indicate divergence; the intensity represents the degree of convergence (red shade) or divergence (blue shade).
The results indicate that both the second and third hypotheses are correct. The map in the middle shows that when both members of a dialect pair converge to standard Dutch, they usually also converge to each other (second hypothesis). The map on the right shows that when both members of a dialect pair diverge from standard Dutch, they usually also diverge from each other.

We had a closer look at the change measurements obtained on the basis of all sound changes as shown in the left picture in Figure 2. We correlated them with geographic distances and found a weak correlation of \( r = -0.14 \) \( (p < 0.001) \). This means that geographically close dialects converge to each other and geographically distant dialects diverge from each other. The average geographic as-the-crow-flies distance of converging dialect pairs is 143 km, and the average geographic distance of diverging dialect pairs is 162 km. Boxplots of geographic distances of both converging and diverging dialect pairs are shown in Figure 3. The geographic distances of converging dialect pairs vary from 2.2 km (between Sint Annaparochie and Nij Altenae, in the northwest of Frisia) to 392.0 km (between Uithuizen, in northern Groningen, and Poperinge, in southwestern West Flanders). The geographic distances of diverging dialect pairs vary between 1.9 km (between Kampen and Ijsselmiiden) and 396.3 km (between Finsterwolde and Poperinge).

We compared the two groups of dialect pairs by means of the Welch's \( t \)-test and found that the geographic distances of converging pairs is significantly smaller than the geographic distances of diverging pairs \( (t = -6.84, df = 3569, p < 0.0001) \).

Figure 4 shows only the changes in relationships between neighboring dialects. The picture obtained on the basis of all sound changes (left) gives the impression that dialects in the west converge to each other and dialects in the north and the east mainly diverge from each other. The pictures obtained on the basis of sound changes resulting in convergence to standard Dutch (middle) and divergence from standard Dutch (right) show a remarkably large number of white lines. A white line between two dialects means that the relationship has hardly changed. When comparing the pictures in Figure 4 with the corresponding ones in Figure 2, we notice that changes in relationships resulting in convergence to or divergence from standard Dutch mainly occur between geographically more distant dialects, rather than between neighboring dialects.

We test the hypothesis that sound changes in two dialects which make them converge to standard Dutch, make them also become more similar to each other. This hypothesis is tested by testing the null hypothesis that measurements of the change in the distance between dialect pairs are zero on average, i.e. convergence to standard Dutch does not cause dialects either to mainly converge to each other or to mainly diverge from each other. We applied a right-sided one-sample \( t \)-test to the measurements of the change in the distance between

---

**Figure 3.** Boxplots showing the distributions of geographic as-the-crow-flies distances of converging dialects pairs (left) and diverging dialect pairs (right).

**Figure 4.** Convergence/divergence between neighboring dialects measured on the basis of all sound changes (left), on the basis of sound changes which cause dialects to converge to standard Dutch (center) and on the basis of sound changes which cause dialects to diverge from standard Dutch (right). Red lines indicate convergence and blue lines indicate divergence; the intensity represents the degree of convergence (red shade) or divergence (blue shade).
dialect pairs and found $t = 23.259$ and $p < 0.001$. We reject the null hypothesis. The large positive $t$ statistic shows that the measurements are larger than zero, i.e. mainly represent convergence. We conclude that sound changes in two dialects which make them converge to standard Dutch, make them also converge to each other, i.e. koineize, possibly resulting in regiolectal varieties.

Our third hypothesis states that sound changes in two dialects which make them diverge from standard Dutch, make them also more distant to each other. We applied a left-sided one-sample $t$-test on the measurements of the change in the distance between dialect pairs as the result of divergence to standard Dutch, testing again the null hypothesis that distance change measurements are zero on average, i.e. divergence from standard Dutch does not cause dialects either mainly to converge to each other or mainly to diverge from each other. We found $t = -18.144$ and $p < 0.001$. We reject the null hypothesis. The large negative $t$ statistic indicates that the measurements are smaller than zero, i.e. mainly represent divergence. We conclude that sound changes in two dialects which make them diverge from standard Dutch, make them also diverge from each other.

4. Conclusions

In this study, we used three- and five-dimensional Levenshtein distances in order to study dialect change in 86 different local dialects of Dutch and relating this change to processes of convergence to and divergence from standard Dutch.

We found that dialects do not change for the larger part, and the average dialect change is 13.3%. When focusing on this change, we found that 51.1% results from divergence to standard Dutch. This is the larger part and also significantly larger than change which does not affect the relationship to standard Dutch, and change due to divergence to standard Dutch.

We also studied the change in the relationships between dialects. We related this change to processes of convergence to or divergence from standard Dutch. We found that convergence to standard Dutch usually goes hand in hand with the convergence between dialects. Divergence between dialects as the result of convergence of dialects to standard Dutch is exceptional. Divergence from standard Dutch usually goes hand in hand with the divergence between dialects. Convergence between dialects as the result of divergence of dialects from standard Dutch is exceptional.

Several questions remain unanswered. Can dialects converge to or diverge from each other while the relationships to standard Dutch remain unchanged? How is the convergence/divergence between dialects resulting in convergence to/divergence from standard Dutch related to the convergence/divergence between dialects which does not result in convergence/divergence with respect to standard Dutch? In so far as our data allow, we will try to address these questions in a future study.

Acknowledgements

The maps in this paper are created with RuG/L04 – software for dialectometrics and cartography. We thank Peter Kleiweg for developing and making available this software. Especially on our request, he modified one of his programs so that we could create the maps in Figure 3. Additionally, we would like to thank the anonymous reviewers as well as the editors for their valuable remarks and suggestions. We alone are responsible for any remaining shortcomings.

Notes

1 See: http://www.stichtingonsbildt.nl/. The foundation aims to preserve the regional language and culture of Het Bildt, a small area in the northwest of the province of Frisia in the Netherlands.

2 Boves & Gerritsen summarize older dialectological and younger sociolinguistic research of language variation relevant to the gender variable and they also present and discuss a range of explanations for gender differences and specifically for the persistent finding that women tend to be oriented towards prestige/(perceived) standard norms much more obviously than men.

3 The 13 sentences still represent the story in an understandable way.

4 The two latter recordings have not been used in this study, but are mentioned here for the sake of completeness.

5 Since in each of the 86 dialect locations, at least four informants were involved in the recordings, and another three provided us with recordings of three standard languages, more than 347 informants have made the research presented in this paper possible. We would like to thank all of them. Furthermore, we are grateful to Peter Kleiweg, whose RuG/L04 package was used to create the maps presented in this work.

6 Although this is an imaginary form, Van Oostendorp (2007) has documented and analyzed cases of incomplete final devoicing in a small group of Dutch dialects.

References


Perceptual regions in Poland: An investigation of Poznań speech perceptions

Paulina Bounds*

Tennessee Tech University

In this paper, ideas from perceptual dialectology, linguistics of speech, and cognitive science are drawn upon to explain the perceptions of others’ speech. The perceptual map task, based on Preston’s “Draw-a-map methodology,” was collected from 215 respondents in Poland and transformed into result maps. The second step in the analysis of the perceptual maps was to collect all of the labels that were assigned to the perceived speech varieties around the country. Those two facets of the data show how the idea of connecting speech with a specific locality can be observed quantitatively. The results propose an explanation of the distribution of speech perceptions in Poland. The shape of the perceptions is created through activating gestalt processes to arrive with an observational artifact. Those are based mostly on the concept of geography facilitated by shared cultural schemas. All of those steps lead to a creation of multidimensional perceptual regions.

Keywords: perceptual dialectology, cognitive structure of linguistic perception, gestalt, schema, multidimensional region, perceptions, Poznań, Poland

1. Introduction

1.1 Perceptual Dialectology

In our daily lives, we encounter speech in almost every minute of our day. By the time we are adults, we are exposed to an enormous amount of speech data that is somehow stored and processed in our brains. Our perceptions about speech originating both close and far away from us offer one approach to these cognitive activities. Following the methods of modern perceptual dialectology (Preston, 1989, 2004; Niedzielski and Preston, 2000), we can gather information about speech perceptions and map their geographical location. The participants of perceptual dialectology studies were asked to draw areas of speech varieties on maps, assign a label to those speakers, and give examples of their speech. The results of such studies were tabulated to show regions that received the most recognition by the respondents to have a speech variety attached to them (Preston, 2010). Interestingly, the studies found that respondents in the United States could not agree where the Southern dialect’s boundaries are, although most of them indicated that such a speech variety is present in their perception. Moreover, in a close analysis of the results of the perceptual maps, Preston notices two opposite notions appearing together: areas speakers assigned a “dialect” according to caricaturistic linguistic features, and large areas unaccounted for by any of the speakers. For the latter, Preston notes, “This space suggests that respondents have no experience with an area, that an area has no caricaturistic linguistic features or stereotypes, or that an area has no popular culture notoriety” (1989:121). The notoriety that certain regions—such as the American South—receive are labeled oftentimes based on the ideas people have about the area’s social, economic, political, and environmental standing (Preston, 2010; Gould and White, 1986). Those labels frequently have little to do with the linguistic repertoire of the speakers, and for the South, they range from “gentlemanly and ‘everyday’ to such negative labels as ‘scratch and claw, hillbilly, damn Yankees, annoyingly nasal’, and ‘spoken mainly by ignorants’” (Preston, 2010). The prominence of labels not connected directly with speech gives us an insight into what the speakers reach for when creating perceptual maps: stereotypes.

As shown in the perceptual dialectology research, “local identity is not strong unless the area has some linguistic or other caricature which helps promote such identification” (Preston, 1989:118). The stereotype gives us an easy shortcut to access when we are asked about our perceptions of others. This idea is also present in the research done by Diercks (2002:67) that showed that even if we do not ask respondents for stereotypes, but instead play dialect recordings for them, they still access their preconceived ideas about others to locate and identify speech varieties. In the process, they activate mental maps to serve as the basis for their judgments. The mental map concept explored in depth through the work of Gould and White (1986), provided groundwork for areas of perceptual research. Gould and White investigated where people most desire to live in the United States. The mental maps produced showed that
for most of the informants, their local area was the most desirable place to live. In addition, the state of California showed up consistently throughout different groups of subjects as a place they would like to live. Desire for both local areas and those quite far from respondents was explained by the process of activation of stereotypes and caricatures (just as described by Preston above) emerging from collapsed ideas about the environmental, social, economic, and political factors pertaining to various places. Interestingly, in the research done by Fought (2002), when Californians were asked about their perceptions of their own state, responses varied greatly.

It seems that California is a state caught between a general aura of desirability and a specific association with negative linguistic stereotypes, and this dual identity can be seen on the perceptual maps of its residents (133). This phenomenon is also described in Preston (1989) as the “prescriptive orientation” (123) of labels on maps. The prescription oriented labels show us that the caricatures are present not only for the speakers themselves, but their language as well. This way, we can see that the positive stereotype of the speakers does not have to necessarily exclude a caricature of their speech.

Preston (1989) observes that even when the geographical distance between the respondents and the speech areas they indicated was the same, some regions were not indicated similarly. For example, “Canada was mentioned by the Michigan respondents, but not by the western New Yorkers (who live as close to the border)” (Preston, 1989:118). This phenomenon was also seen in the work of Gould and White (1986:143-146), where children living an equal distance from each other in Sweden and Norway had different perceptions of each other. Swedish children did not have much of a mental map of the bordering Norway, while Norwegian children seemed to have a distinct mental representation of the neighboring area of Sweden. As the research indicates, this perceptual variation occurred mostly as a result of the type of information children receive through schooling and popular media. Further, such variation contributes to the fact that schooling is where shared cultural ideas on the regional or national level are created.

The idea of mental proximity that we create as members of various communities translates into how we perceive speech communities. Although Scandinavian children are geographically close, their perceptions of community are distant and distinct. In the research performed by Kerswill and Williams (2002), informants judged speaking samples to help researchers describe language variation. Their findings reflect the idea that has been seen in the research discussed above: people who are deeply rooted in their community will more readily recognize local speech, indicating a short mental distance between speakers and their location. As the researchers indicate, “It is almost a tautology to say that accents which are familiar to the judge will be better recognised than those which are not... However, the range of factors contributing to familiarity is wide” (Kerswill et al., 2002:200). This research confirms that people care and know more about the places close to them and those they have direct experiences with. The further we travel mentally and geographically, the less perceptions are based on experience and the more they are grounded in information fragments or no information at all.

The ideas presented above bring to the forefront various issues connected to our perceptions. Each documented study sheds new light on the creation and activation of perceptions, contributing insights about dialect boundaries, mental maps, proximity, and stereotypes. At the same time, the findings offer possible explanations for one or some of the issues, not all of them. An approach called linguistics of speech and ideas developing from this theory provide additional opportunities to explore the mechanisms behind our perceptions.

1.2 Linguistics of Speech

The linguistics of speech has its foundation in parole, described by Saussure in the 1900s. As Kretzschmar points out,

The choice of linguistic structure is not inevitable, not ‘natural’ in the sense that it corresponds to an inborn faculty or property of species; it is the nucleus of an argument to create a science of linguistics, one based on a model with particular premises and with a definite arrangement of its variables (2009:38).

Now when we are equipped with technology solutions that Saussure could only dream about, we are capable of investigating speech as a behavior with all its overwhelming frequencies and inherent variation.

Under the linguistics of speech, what people say is the only subject of study. Terms such as “dialect” or “language” are not considered to have any boundaries but instead live on a continuum. Therefore, as Saussure points out, “between dialects and languages there is a difference of quantity, not of nature” (1916:43).

Since then, scientific research in various fields, such as cognitive anthropology, neuroscience, or psychology, have arrived at new solutions and compelling evidence for the explanation of world phenomena. One of those notions is complexity theory, a notion used in physical science to describe the workings of emergent order in non-equilibrium systems. The notion of complex adaptive systems is presented here in opposition to an equilibrium system, which is a closed, low-energy
system. As an example of an equilibrium system, we can put a ball into a big bowl, and it will roll around for some time and then rest down on the bottom when the energy is exhausted. The order of the ball’s position has been established, and “it has become a static, low-energy system.” However, “nonequilibrium systems by definition are open, and exchange energy and matter in a dynamic fashion. They very often show order” (Kretzschmar, 2009:178). This can be best illustrated by the example given by Kretzschmar (2009) of a creation of a whirlpool when a bathtub drains. The whirlpool will be there while the drain remains open and we add water to the tub, which asserts its openness as a system. Also, the self-organizing order that emerges will be there regardless of the circumstances; it does not matter whether it is Monday or Tuesday in Poland or the United States. The whirlpool will be there if there is water in the tub and the drain is open. Moreover, we do not need to stir the water to create the whirlpool. If we provide the conditions required for a complex system to operate, it will behave accordingly.

Moreover, “complex systems, also known as complex adaptive systems, share a number of characteristics besides being open, dynamic, and not at equilibrium” (2009:147). Among those characteristics is the idea that complex systems contain a large number of components, and they show self-organizing emergent order. Thus speech is a perfect example of use for complex system theory as it possesses an immense number of elements—not only the segments of speech categorized by linguists, as for example phonemes, morphemes, words, or sentences, but also the number of speakers inherently variable in their linguistic behavior. However, speakers are agents, so they facilitate the complex system without being a part of it in the same way that pronunciation, words, etc. are in the system. In other words, speech is a kind of tool humans can operate. Speakers are agents who use speech for their own purposes and exercise control bias in what they choose to say, when, and how. Therefore, they are part of the complex system, but not in the same way speech is. Speakers put the speech into motion depending upon what type of control bias they exercise.

The self-organizing emergent order comes from the operation of chance among the elements and interactions between them. We need to take note that the chance here is considered to be a formal idea of randomness. Therefore, random processes happening in the complex adaptive systems result in emergence of patterns and clusters (Kretzschmar, 2009:179).

Now, one more factor needs to be added to the analytical choices that we make, and it is perception. Based on how we perceive speech, we arrive with conclusions about linguistic features, and “thus perception is a necessary element of speech, because without it there could be no linguistic features” (Kretzschmar 2009:46). Taking perception into account opens another explanation of how we can deal with perceptual dialectology under the linguistics of speech:

Since there are no natural dialects, then, the inventories of linguistic features that we collect constitute dialects because we so name them, and they are useful because they help us to conceive of “the primary and natural phenomenon of differentiation into independent areas” (2009:48)

Such a statement has its foundation in the notion previously discussed in which the categories that we create do not occur naturally in the world around us, but it is us who create them to comprehend the world. Such a statement leads us to the foundational statement for linguistics of speech, in which the relation between what we know about speech and what we perceive of it is established. Speech is seen as a continuum without boundaries; the boundaries between languages, or in other words linguistic systems, are only a perception of our minds. What is really happening is that from locality to locality people share some linguistic features and differ in others. The further away we are geographically or socially, the less we might have in common when it comes to the linguistic features. However, it will never be the case that speakers on one side of a line will speak in exactly one way and those on the other side differently. The boundaries, or isoglosses, are generalizations created by our perceptions and facilitated by the methodologies used to study speech.

Perceptions are still an important component of our linguistic behavior because they not only constrain our choices in speech, but also reveal how our lack of information is filled with perceptions to create gestalts. Our mind prefers patterns and complete entities. Thus, one of the mechanisms with which it complies with the information received from the world is creating cognitive wholes—gestalts. We are then able to create a perceptual, finite concept out of interrupted and incomplete information. Such a method explains, for example, why we are able to have an opinion about the speech of speakers whom we have never heard or seen before (Kretzschmar, 2009:188). When cognitive anthropology comes to play, Kretzschmar adopts the schema theory into the linguistics of speech approach. Therefore, “schema theory is not about objects with particular, established characteristics (of which an individual is a concrete example, and a prototype is an abstract example), but about abstract specifications for what might be relevant in what comes to be recognized as a category of experience” (Kretzschmar, 2009:188).

This concept, known also as a “frame,” “scene,” “scenario,” or “script,” has been used for a long time; the first mentions of such a thought can be attributed to Kant (1781 [1998]). More recently, Mandler (1984)
described schema as “abstract representations of environmental regularities” (1984:55). Each experience that we gain in our lives leaves a mark and helps to formulate such a schema. Therefore, we recognize the world in the realms of schemas. Gould and White (1986:150) showed us that the information that we gain with age behaves in a predictable manner:

Until children are about seven or eight their preference signals are so small that we cannot really detect them against the background noise; but then as sudden flood of information impinges upon them, and the preference surfaces begin to crystallize out of the mental flux. By adulthood the surfaces have set pretty firmly with a high degree of agreement about them.

This relationship may be seen as a base for how our schemas are formulated, especially those cultural schemas that we gain through schooling. Moreover, schema is not only a frame for various types of experiences of our life, but at the same time, it is a mechanism with which we parse and designate to a specific schema all the input of our existence. “Most, if not all, of the activation processes occur automatically and without awareness on the part of the perceiver-comprehended” (Mandler, 1984:56). On top of that, schemas have “slots” for features which can be filled in with concrete detail. The number of slots depends on the speaker and their type of previous experiences, and to some extent the information for one slot can determine the rest of the slots.

Each speaker develops their own individual schemas, but through shared experiences of common culture, as a community, they also develop shared cultural schemas (Kretzschmar, 2009:188). Because both types of schemas are present in the minds of the respondents, when asked about their perceptions about speech, part of the results will be different from the rest, as every individual experience is different. But there will be some overlap, as we do share cultural schemas to some degree. Individual schemas are made out of slots for characteristics out of which a pattern is created. Cultural schemas “‘average’ the ratings by the individuals... except that now... slots for relevant characteristics within schema, are the target for analysis, and not fixed characteristics themselves” (Kretzschmar, 2009:189). Such an approach to the analysis is more flexible than the prototype approach. In any schema, we develop “slots” through our experiences.

For example, through participating in schooling we may develop a schema called “a classroom.” In such a case, we may have slots for “the teacher,” “students,” “seating organization,” “time period,” “recess,” “textbooks,” and so on. Those slots allow us to fill them up with specific information pertaining to a Biology class, as opposed to an English class, as opposed to a Chemistry class. As a result, each experience of that schema will be slightly different, but it will use common slots to organize it, recognize it, and maybe even experience it. When using schemas, we do not always have all the information needed to fill in the slot with a certain feature, so we do need a mechanism to help fill in the gaps. Such a mechanism is described as gestalt, originally used by the Berlin School in psychology in the late 19th and early 20th century. Since then, it has been a part of various disciplines, such as psychology, anthropology, or cognitive science. It also found its way into linguistics, especially cognitive linguistics, in the work of Lakoff (1987), Evans and Green (2006), or Grice (1989). The use of gestalt mechanism that is adapted by linguistics of speech is described by Evans and Green (cited in Kretzschmar, 2009) as important “because they allow unconscious perceptual mechanisms to construct wholes or ‘gestalts’ out of incomplete perceptual input” (2009:186).

Gestalt plays an important role in the research conducted on speech perception, since it is a way to account for lack of experience and information to create a definite perception. As explained by Günther (et al., 1996), an observational artifact can be created using the gestalt mechanism, and at the same time it invokes the notion that an object can be created out of ideas or perceptions. In this case, subjects created an observational artifact named “speech variety,” out of incomplete information, based on the scarce linguistic items, and other experiences in the community. In this way “speech variety” becomes an object that can be put on a map, as was seen in the results of the perceptual maps. In this frame of mind we can think of places that have become observational artifacts as multidimensional regions: “Region is then necessarily a complex multidimensional construct, and the physical and behavioral characteristics of regional culture always exist in a dynamic, self-aware relationship with the perceptions of participants” (Kretzschmar, 2011:190). In this way, we can start to explain how those multidimensional regions are being constructed and how much and what type of input is needed for such a region to appear on a map.

Using the ideas presented above, I will investigate the distribution and emergence of perceptions of speech in Poland created by Polish participants. I will argue that the perceptual, multidimensional regions that appear in “draw-a-map” result from the process of activating gestalt processes to create an observational artifact based mostly on the concept of geography that is facilitated by a shared cultural schema.

2. Methods

In my research, I used Dennis Preston’s “draw-a-map” methodology. Respondents from Preston’s study (1989) were given a map of the U.S. with state boundaries on it
and asked to circle areas where people speak in various ways and to add labels to them (Preston, 1989). The perceptual map task presented here was a part of a bigger project investigating the speech production and perception of Poznań speech. I wanted to see if the respondents would perceive the speech of Poznań as distinct from other places; therefore, it was necessary to have the city on the map. As Preston points out, the amount of information provided on the map is crucial to the type of responses that we want to receive. He reports:

[M]any could not escape the notion that state lines were dialect boundaries, a fact which supports the conclusion that nonlinguists’ impressions of the position of dialect boundaries are historical-political, not linguistic... Perhaps a map with major rivers, cities, and mountains, would have prevented this sort of response. (Preston, 1989:25)

Therefore, I have decided that it might be more useful for the proposed research to indicate cities on the map instead of other administrative divisions. Map 1 displays the map used for the task with 13 major cities used.

Map 1. The map of Poland used for the ‘Perceptual map of Poland’ task.

All of those cities are distributed in a regular pattern across the country.

The maps were black and white for technical reasons; a lot of details would potentially disrupt the comprehension of the task. I wanted to keep a balance by providing enough detail for easy geographical orientation and not obscure it with too much information.

The informant sample was collected by convenience through snowball sampling in 2008. All of the respondents were students at Adam Mickiewicz University in Poznań and had no previous training in linguistics. The respondents filled in a paper version of the task, asking them (following Preston, 1989) to circle areas where people speak in various ways and to add labels to them. This way of conducting and distributing the task proved to be productive as it resulted in 215 completed perceptual maps.

All of the respondents were in the 18-30 age group, and mostly indicated their occupations as students, with the occasional exception those indicating themselves as white collar workers. Moreover, the respondents had the same level of education (high school), as they were all in the process of obtaining a higher education. I did not ask for demographic information of their parents. Thus, out of all the collected categories, gender and place of birth/upbringing were the only ones that showed variation in the answers. Table 1 shows the distribution of the informants, according to the gender and the place of upbringing.

Once I collected the maps, I processed them so I could observe patterns emerging from my data. First, I scanned the maps. In Adobe Photoshop CS2, I colored the circled areas and erased any other information from the scan. Then, I used the PICtoASCII program, which converts bitmaps into ASCII symbols. This program allows saving such a converted bitmap in a .txt format in which various colors correspond to different symbols, and white space is white space. Then, using the Find and Replace function in Microsoft Word, I inserted tab delimitations in order to be able to open such maps in a Microsoft Excel spreadsheet. I used formulas to count all the symbols in the cells. This way I was able to add all the maps together. This type of formula resulted

All of the respondents were in the 18-30 age group, and mostly indicated their occupations as students, with the occasional exception those indicating themselves as white collar workers. Moreover, the respondents had the same level of education (high school), as they were all in the process of obtaining a higher education. I did not ask for demographic information of their parents. Thus, out of all the collected categories, gender and place of birth/upbringing were the only ones that showed variation in the answers. Table 1 shows the distribution of the informants, according to the gender and the place of upbringing.

Once I collected the maps, I processed them so I could observe patterns emerging from my data. First, I scanned the maps. In Adobe Photoshop CS2, I colored the circled areas and erased any other information from the scan. Then, I used the PICtoASCII program, which converts bitmaps into ASCII symbols. This program allows saving such a converted bitmap in a .txt format in which various colors correspond to different symbols, and white space is white space. Then, using the Find and Replace function in Microsoft Word, I inserted tab delimitations in order to be able to open such maps in a Microsoft Excel spreadsheet. I used formulas to count all the symbols in the cells. This way I was able to add all the maps together. This type of formula resulted

Table 1. The perceptual maps respondents.

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Poznań residents</td>
<td>32</td>
<td>69%</td>
<td>14</td>
</tr>
<tr>
<td>Wielkopolska region residents</td>
<td>74</td>
<td>69%</td>
<td>33</td>
</tr>
<tr>
<td>Other residents</td>
<td>38</td>
<td>61%</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
<td>71%</td>
<td>71</td>
</tr>
</tbody>
</table>

P. Bounds
in a spreadsheet containing numbers distributed in the shape of Poland, in which each number corresponds to the number of respondents who indicated some sort of speech in that particular cell, corresponding to a respective area on the map. This map has been transformed into a 2D graph to show the results.

3. Results

3.1 Speech Perceptions

The result map of all of the responses indicating perceptions on the map of Poland is displayed in Map 1.

The darkest areas of the map are where the highest numbers were located, and white areas are where 10% or fewer of the informants indicated anything on the map. We can see that there are four main areas where at least 40% of the subjects circled something on the map; those areas are in the North, South, East, and West of Poland. There is a vast surface around the epicenters with less than 30% of the respondents perceiving any speech variety there. There is a dramatic difference between the little amount of area that many subjects indicated, and a large surface that not many of them agreed upon. It seems that Poznań, in the eyes of the respondents, has a particular type of speech. Another city clearly pointed out is Katowice in the South. Warszawa, although visible in the East, does not receive the same level of agreement as the previous two. In the North, the situation seems to be different. The area specified the most is not around any major city. It may be an indication of a speech pattern present in the area but not strongly associated with the surrounding towns. A similar situation can be observed in the South, where the mountain region of Poland received the most recognition out of all. The most important observation that those perceptual maps show, is that people perceive locations to have speech varieties; however, respondents do not agree where exactly those places are located, indicating that this is more of a mental concept.

3.2 Labels on Perceptual Maps

While the respondents circled the areas on the maps, they also labeled the areas they indicated. When I collected all of the labels from the maps, four semantic categories emerged from the data:

Geographical names, in which the subjects labeled the area according to the name of the region, for example Wielkopolska; or used town/city names, such as Poznań, Kielce; or some other description of the region, as for example góry “mountains,” wschód Polski “East of Poland.”

Names of people, in which informants gave names of groups of people living in an area, very often based on the geographical names, for example poznaniancy “Poznanians,” śląscy “Silesians”; or gave characteristics describing the people, for example szaleni kierowcy “crazy drivers” (for people in the West), and głupi warszawiacy “stupid Warszawians.” Another strategy was to use nicknames for people living in certain areas, for example pyry, pyrusy, pyrole “potato people” (for people from Poznań), scyzorki “pocket knives” (for people from Kielce), legioniści “legion people” (people from Warszawa, fans of local soccer team named Legia Warszawa “Warszawa legion”).

Dialect labels. All of the instances in which the words dialect, patois, language, or accent were used are included in this group, as were the names for speech varieties, like góralsczyzna “the speech of mountain highlanders” and ślaski “Silesian language.”

Features of speech, in which respondents gave a characteristic of the local speech: for example należalości z rosyjskiego “influences from Russian” (about east region), zaciaganie “drawl” (about east region).

Those four categories provide an insight into the type of perceptions that the respondents have about where, how, and by whom speech is used in Poland. Each of the 215 maps that had some speech perception data on it was separated into as many as four individual spreadsheet maps, each holding only areas labeled by one of the semantic categories. This way I was able to add all maps with the same type of label together. Maps 3 through 6 show those results below.

First of all, not a single category appeared on all of the 215 maps. The label that appeared on the highest number of maps was the “Names of People” category (on 161 maps), then “Geographical Names” (on 123 maps), “Dialect Labels” (on 107 maps), and “Features of Speech” (on 34 maps). The areas covered by those distributions are similar, but they are not exactly the same. The more frequent categories cover more areas, but all of the peaks for all of the categories correspond to the areas around Poznań and in the mountains in the South. This reinforces the idea that respondents not only perceive those regions as possessing some sort of speech variety, but also reflect it in the highest number of labels.

The number of maps that included labels gives us one view, but when we look at the number of tokens for each category, we can see that the first three categories are also overwhelmingly numerous, and the “Features of Speech” category is very small compared to them. We can see the number of tokens for each type of semantic category from all of the 215 maps in Table 2.

The average number of labels per map is 6.7. This tells us that on average the subjects put 6 or 7 labels on a map, and they must have been from different categories of labels, because none of the category averages came close to 7. The “Names of People” category has an average of 3 labels per map, “Geographical Names” 1.9,
“Dialect Labels” 1.5, and “Features of Speech” 0.3. The various ways to describe people and where they live seems to be the most frequent category, pointing to the fact that it might be the most important way to describe a locality. However, as shown by the aerial distribution on the maps, this locality does not have to be physically and correctly connected with a geographical location. It is more of an approximation; the places that we label are more of a mental concept than concrete realities. We will come back to this issue in the later discussion.

Another revealing facet of the way people see others in their speech communities is shown when we look at the combinations of categories of labels occurring together. As we have seen earlier, most of the time respondents put several labels on a single map. Map 7 gives us an example showing a hand-drawn map with several types of labels. Only aggregating and analyzing a number of maps like that shows how those choices were structured, and possibly gives us insight into the way perceptions are organized.

First of all, only 5% of the respondents did not use labels, and only 7% put all four categories of labels on their map. We do not know the reasons why some respondents did not label their maps. It might be that they did not have labels to put on, but it might be as well that they did not follow the instructions. The respondents who put all types of labels on the map were not numerous (7%), but they show us that some people can pull together all of the facets of the perceptions that they have. Their behavior indicates how complex our perceptions can be.

The next group includes all the maps for which only one category of label was chosen. This was the second most popular choice. As the results show us (in Table 3), if we were to put just one type of label on the map, our first choice is most likely to name the people; next, the variety of speech they use; and finally where they are located. Such an order shows us that the non-linguistic features are first to be chosen to show our perception of others.

That leads us to the group of two label categories on a map. This was the most popular choice, as it covers 45% of the maps (see Table 3). Out of those, only 4% take the “Features of Speech” category as a partner. The overwhelming majority is made out of a combination of the first three categories, “Names of People,”
“Geographical Names,” and “Dialect Labels.” So, if we were to put more than one type of label on a map, chances are it would be a description of the people living there, and we would add information about where they live, and what to call their speech. However, we would be least likely to point out any linguistic features of their speech, and the linguistic features would not be combined with a geographical name.

The relationship between non-linguistic and linguistic labels is similar in the group consisting of three different label categories appearing together on the map. When people put three different types of labels on the map, the set of the non-linguistic labels covered 16% of data out of the 21% total for this group (see Table 3). The remaining 5% of maps on which linguistic features do appear is divided into the three possible category combinations, each of which occurs at a low level of these; the combinations that include “Dialect Labels” and “Features of Speech” occur at the lowest level.

4. Discussion

When we look at the relationships between the labels that are described in the context of the people’s most frequent choices, we see a disjunction between the “Features of Speech” category and the rest of the categories. Moreover, the relationship between “Geographical Names,” “Dialect Labels,” and “Names of People” is reinforced by the fact that, almost exclusively for the “Names of People” category and exclusively for “Dialect Labels” category, the concept of geography is present within the labels; for example: Kashubians.
people from Poznań, or speech of Wielkopolska, dialect of Silesia, language in Warszawa. Therefore, the concept of geography seems conceptually to be the common thread for all the labels, except the linguistically driven “Features of Speech” category. That shows us that it is the concept of geographical names of regions that is the common denominator. However, it seems that geography alone is not enough because we saw that it is not the most frequent label when in isolation; the “Names of People” category was. If we assume, following Niedzielski and Preston (2000:18), that “the communicative function of language is so strong that it overcomes the ability to give an accurate report of performance whether of self or of others and whether of general or restricted phenomena,” then we can start considering the basis of those stereotypes based on the labels provided by the respondents. This way, we can offer a possible process in

Table 3. Label categories distribution.

<table>
<thead>
<tr>
<th>Number of label categories on a map</th>
<th>Percentage</th>
<th>Break Down Within the Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Label</td>
<td>26 %</td>
<td>Names of People 14%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dialect Labels 9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geographical Names 3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Features of Speech 0 %</td>
</tr>
<tr>
<td>2 Labels</td>
<td>45 %</td>
<td>Geographical Names + Names of People 24%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Names of People + Dialect Labels 11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geographical Names + Dialect Labels 6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Names of People + Features of Speech 3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dialect Labels + Features of Speech 1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geographical Names + Features of Speech 0%</td>
</tr>
<tr>
<td>3 Labels</td>
<td>21 %</td>
<td>Geographical Names + Names of People + Dialect Labels 16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geographical Names + Names of People + Features of Speech 3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Names of People + Dialect Labels + Features of Speech 1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Geographical Names + Dialect Labels + Features of Speech 1%</td>
</tr>
</tbody>
</table>

Map 6. Distribution of perceptions labeled in the category of “Features of Speech”

Map 7. A hand-drawn perceptual map showing multiple labels from “Dialect Labels,” “Geographical Names,” and “Names of People” categories.
which, when a mental concept of a place is created, the ideas of people, and the name of the speech that they use is very important, but the underlying concept of geography is what connects them. It is not to say that the speakers indicate an exact location as the geographical map would tell us. Instead, they are using an approximation of where such a place is located (as discussed above in the research of Gould and White, 1986). The main purpose is to describe the mental meaning of this place as reflected in the labels. The linguistic features that linguists ask about do not rise to be the main concern of the respondents, and even when they do put them on, they reflect a folk understanding of speech with descriptions such as: they say “Tej” and “nie,” Poznań pronunciations, they sing when they speak, they have a lot of German words (all examples about speech in the Poznań area). Therefore, not only is the linguistic description superficial, but it is also not frequent. On the other hand, what we see is a great emphasis on the geographical cognitive concepts present in the labels in addition to people living there and the names for their speech.

The fact that nobody used linguistic features to stand on their own emphasizes even more that we do not rely primarily on linguistic knowledge in our perceptions. Following Preston’s (2010) assertion mentioned earlier that “nonlinguists seem to be using factors other than the perception of purely linguistic differences,” we can see this in the fact that the linguistic features are not the first thing subjects consider when creating the mental image of a place. They are last. Moreover, the lack of the relationship between the linguistic and geographical features suggests that there is something inherently different about our perceptions of speech and the concept of geography.

This distinction between the linguistic and nonlinguistic foundations leads to one more important point in the quest to describe the way those perceptions were created. As Preston (2010) observes in the U.S. research:

[The] labels suggest that a regard for language correctness is a dominating one... Correctness is not, however, the only theme to emerge. The respondents were clearly distinguishing between correct and pleasant varieties. Language attitude studies confirm, then, that regional varieties are not all equal, and such findings help further establish the basis for another perspective on varieties: an account of what speakers of various regions (and classes, sexes, ethnic groups, age groups, and so on) believe about dialect variety.

This quote suggests that the two dimensions of the labels attached to perceptual maps in the U.S. research have to do with correctness and pleasantness. However, when we look at the labels present on the Polish maps, those dimensions do not seem to be prevalent. Instead, the underlying concept of geography seems to be connecting and influencing the shape of the labels attached to the perceptions. Both the U.S. and Polish data seem to be mostly based on a nonlinguistic basis, but that base seems to be different. The question remains, how then does the process of creating perceptions seem to be similar while at the same time, the basis for those perceptions is so different?

One possible explanation can be found in the approach presented by Kretschmar (2009), which suggests that the results indicated that people did not keep information about language separate from nonlinguistic information: linguistic and non-linguistic schemas that overlap arise because a majority of individual speakers are participating in a shared culture. Each experience that we gain in our lives leaves a mark and helps to formulate such a schema. Therefore, we recognize the world in the realm of schemas. Schemas are “organized frameworks of objects and relations which has yet to be filled in with concrete detail” (D’Andrade, 1995:124). With that said, throughout our lives, we experience the world in many dimensions, and through those experiences patterns emerge, in which we organize and categorize the world. With time, we create schemas that provide us with structure of slots that we can fill in with individual pieces of information. Moreover, schema creation and processing is not only a framework for various types of experiences of our lives, but at the same time it is a mechanism with which we parse and designate to a specific schema all the input of our existence (D’Andrade, 1995:122). The highest frequency labels associated with the cultural schemas present on the map of Polish speech varieties suggest that what people know about speech is not based on linguistic knowledge. Instead, the concept of geography, likely gained in school, or more generally through living in Poland, is the most common concept evoked when “speech variety” schema is triggered.

This observation opens the door for an interpretation in which subjects created schemas for the labeling of speech varieties using geographical knowledge, as their awareness of linguistic details was scarce or non-existent. However, the issue of very limited input needed to create a schema of “Poznań speech” can be explained by the use of gestalt mechanisms based on the low frequency linguistic items present on the perceptual maps. When a schema is invoked, it is then perceived as an observational artifact of “Poznań speech.” In such an observational artifact of “Poznań speech” the assumption is that there must also exist speakers, location, vocabulary, and other items. Therefore, it only takes a label or two to indicate it on a map, and possibly it only takes a word or two of those considered to be Poznań-specific uttered by a speaker for that person to
be perceived by others as belonging to the group of “Poznań speakers.” The types of labels found on the maps suggest that the concept of geography is enough to create an observational artifact that can also be seen as a multidimensional region (Kretzschmar, 2011).

In such a region, geography seems to be reflected in the labels as the common denominator. However, it is not enough. People do not pinpoint geographical locations exactly, because people care about their local surroundings much more than anywhere else, and they do not have a comprehensive knowledge about places far away (Kretzschmar, 2009, 2011). Moreover, “A region is not merely some arbitrary tract of geography, but instead a location in time and space in which people behave in some particular way, and in which we may find physical evidence related to that behavior” (Kretzschmar, 2011:187). In this light, we can see Poznań or any other location that became an observational artifact to be such a region, in which we have people, location, and their speech intermingled. Labels that people assign to the observational artifact created for those regions show that what we care about is where people are from, since the majority of the labels employ the concept of geography. Therefore, the location or the region is what we care about the most and what shows up the most on the perceptual maps. This idea shows in contrast to Preston’s (2010) U.S. findings, where the idea of correctness and pleasantness of speech seemed to be more important as the basis for the creation of schema. Within the concept of region we can see “the necessity for merged geographical/social categories like region, because putting location and other information together is the only way to achieve coherence, by imposing order on apparently miscellaneous perceptions” (Kretzschmar, 2011:193-194). Thus, the idea of folding the geographical and social information into one and putting it on the map seems to explain why the majority of labels are associated with those concepts and not linguistic features. We have seen that two out of the four highest levels of agreement about speech perceptions were located around a specific point on the map—Poznań or Warszawa. Heaving those cities on the map definitely helped to anchor the observational artifact on the map. However, in the two other observational artifacts present on the perceptual maps, in the South and North of Poland, there is no city to attach the artifact to. Instead, for the South, the area indicated seems to show that people are not really sure where the highlanders do live, but they live “somewhere in the South.” The same situation is evident in the North in connection to the Kashubian speech. The region itself is designated in various ways by the people living there, speaking the language, and performing research on the language. That lack of sharply-defined boundaries of the region is also evident in the distribution of the perceptions. Many of the respondents know about it and have a general location in the North for this speech variety, but they do not know “exactly” where it is. People just do not have enough information to give us a comprehensive description of features of speech varieties in a country. Instead, the common denominator that might be the easiest to reach for is a geographical concept in the form of a region. Such a region does not have to be based on an abundance of information, because speakers can activate the gestalt mechanism to fill in the gaps. The product of such a process is a multidimensional region in the form of an observational artifact to be readily drawn on a map.

Therefore, perceptions about speech end up not being solely about what and how we speak, but more about who and where we are. The research presented above points to the intricacies of how we create our mental maps of speech varieties. Our experiences with the world beyond speech largely determine how we recreate the world around us. This observation accounts for our ability to generate ideas about people and places spatially far away from us out of “thin air,” without requiring any informational or experiential input. Using tools from perceptual dialectology, linguistics of speech, cognitive linguistics, and linguistic geography, we can investigate the web of connections that speakers develop to create mental maps of their worlds. By utilizing the expertise from the aforementioned research, we can start building a methodology that can accommodate the complicated workings of human perception. Once we can better describe perception and how it works, we might be able to use perception to investigate the connection between speech perception and speech production. Links between perception and production might in turn reveal how much our perception influences the production of our speech and how much the research on speech production has been reflecting perceptions of respondents.

Notes

1 For more detailed information on the perceptual task part of this research, refer to: Bounds, Paulina. 2010. Perception of Polish speech varieties. Poznań Studies in Contemporary Linguistics 46(2). 155-176.
2 The exact wording of the instructions was following Preston’s (1989): “People in various areas of Poland speak in different ways. How do you call those areas? How do you call the people living there and their way of speaking? If you have multiple ways to describe the area, the people, and their speech, list them all. If this map is not detailed enough, use the last page to draw a more detailed map of Poland or its regions. You can leave any additional comments on the last page as well.”
3 The categories asked: Age, Gender, Education, Occupation, Where were you born? Where did you spend most of your childhood?
References


