Perceptual Dialectology

Oxford Handbooks Online

Perceptual Dialectology
Jennifer Cramer

Abstract and Keywords

This chapter introduces the topic of Perceptual Dialectology (PD), an area of sociolinguistics concerned with how nonlinguists understand dialectal variation. The chapter provides a brief history of the field and explores the ways in which the perceptions and language attitudes of nonlinguists have typically been elicited in research conducted within the modern tradition of PD with a particular focus on mental maps. Additionally, this chapter identifies ways in which these methods have been improved upon, specifically through the use of geographic information systems (GIS) tools. As an illustration of both the typical tools used in PD research and these recent advances in data analysis, a research project on the perceptions of dialectal variation within and across the state of Kentucky is presented.

Keywords: Perceptual Dialectology, mental maps, nonlinguists, language attitudes, Kentucky, geographic information systems, dialectal variation

Introduction

Perceptual Dialectology (PD) is the study of how nonlinguists understand dialectal variation. This field of inquiry seeks to include what nonlinguists think about linguistic practices, including where they think variation comes from, where they think it exists, and why they think it happens, in holistic examinations of variation that incorporate aspects of both linguistic production and perception. It is a branch of folk linguistics, one that does not, as in the American structuralist tradition of Leonard Bloomfield, eschew the attitudes of nonlinguists in favor of “real” data, which have been the lexical and phonetic variables of traditional dialectology (Preston, 1989; Niedzielski and Preston, 2000). As such, PD is located at the intersection of fields as diverse as linguistics, sociolinguistics, anthropology, sociology, psychology, and cognitive science.

Nonlinguists’ perceptions of dialectal variation have been, in many linguistic studies, not only ignored but also considered nonscientific. This type of data has been called impoverished, as the folk are thought to have an inadequate vocabulary with which to
discuss their thoughts on language variation, and inaccessible, as the connection between
speech and attitudes toward speech has been seen as unclear (Niedzielski and Preston,
2000). Yet countless studies in PD have revealed that speakers of a given language are
very aware of the language within which they live, and they are quite willing to report on
this awareness. Such data presents an opportunity for linguists to bridge the gap
between language production and language perception, and PD seeks, as a field, to
address such a gap.

Thus, the inclusion of folk beliefs in linguistic studies can offer important insights into the
realities of language variation and change. The work of PD offers the ability to make
larger connections between how language is actually produced and how nonlinguists
perceive it. It also attends to the level of impact linguistic beliefs have on an individual’s
performance of language (i.e., is the speaker likely to avoid certain linguistic behaviors
because of the social stigma they perceive to be connected with such behaviors?).
Additionally, the knowledge gained in studies of linguistic perceptions can inform policies
developed by educators and politicians, so as to make such policies account for the
attitudes of the speakers to whom the policies apply.

In this chapter, a brief history of the field of PD is provided, and the varying ways in
which the perceptions of nonlinguists have been unearthed are explored. As an
illustration of both the typical tools used in PD research and certain recent advances in
data analysis, a recent research project on the perceptions of dialect variation within and
across the state of Kentucky is presented. This analysis includes the use of geographic
information systems (GIS) tools1 that allow the analyst to connect mental maps drawn by
participants in PD studies to real world maps and coordinates through a process of
goreferencing, thus intertwining the perceptions held by nonlinguists and the world in
which those perceptions are enacted.

The Roots of Perceptual Dialectology

To say that sociolinguists have completely overlooked the importance of language
attitudes in the context of their research programs would be an overstatement. William
Labov’s (1972) early work with African Americans in New York City, for example, drew
insights from research participants to explore connections between language use and
language attitudes. However, such research has been conducted within the framework of
the Observer’s Paradox, which claims that “underlying attitudes toward language are
evoked more accurately when the subject doesn’t realize that language is in
question” (Labov, 2006: 324). Thus, while interest in the overt opinions of nonlinguists
has been low, a wealth of information on the covert attitudes of nonlinguists has been
uncovered, especially in domains like the sociology of language (e.g., Fishman et al.,
1971) and social psychology (e.g., Lambert et al., 1960; Ryan and Giles, 1982).

However, in many traditional dialectology studies, past and present, overt attitudes and
beliefs about linguistic varieties have often been disregarded in favor of the phonetic and
lexical variables of linguistic production that have long been a staple of the field. Yet, in
the 1960s, despite the prevalence of the ideology that attitudes were secondary to production data, a perspective highlighted by the American structuralist movement often associated with Bloomfield (e.g., 1944), Henry Hoenigswald (1966) encouraged linguists to reconsider this position. He claimed that

we should be interested not only in (a) what goes on (language), but also in (b) how people react to what goes on (they are persuaded, they are put off, etc.) and in (c) what people say goes on (talk concerning language). It will not do to dismiss these secondary and tertiary modes of conduct merely as sources of error.

(Hoenigswald, 1966: 20)

With this statement, Hoenigswald established reinvigorated interest in an area referred to as folk linguistics. Dennis Preston, the father of the modern American tradition in this field, has further stressed the importance of perception in the study of linguistic variation, presenting the image in figure 1 as a way of envisioning how these components listed by Hoenigswald are interrelated.

![Figure 1. Connections between language use and reactions to language use](image)

With this statement, Hoenigswald established reinvigorated interest in an area referred to as folk linguistics. Dennis Preston, the father of the modern American tradition in this field, has further stressed the importance of perception in the study of linguistic variation, presenting the image in figure 1 as a way of envisioning how these components listed by Hoenigswald are interrelated.

Hoenigswald’s call, however, was not the establishment of the field of PD. In Preston (1999), the history of the field is explored, revealing that the beginnings of this type of analysis of language attitudes come from the Netherlands and Japan. Early Dutch surveys asked respondents to identify areas that were similar and areas that were different from their own way of speaking (Rensink, [1955] 1999). This earliest foray into nonlinguists’ perceptions of language variation used a method called Pfeilchenmethode, or “little-arrow method,” which had been developed by Antonius Weijnen (1946) to connect “a respondent’s home area to another that the respondent asserts to be linguistically similar to the locations they described as linguistically similar” (Preston, 1999: xxvi). These early studies showed that while some participants’ perceptions aligned with the traditional production maps of the Netherlands, others perceived the dialect landscape in ways that differed from what production data had revealed.
A different tradition was established in Japan for exploring the beliefs of nonlinguists. Sibata ([1959] 1999) asked respondents to list villages that had a variety different from their own. Instead of the "little-arrow method" used in the Netherlands, of which Sibata was uninformed, the analysis involved the use of lines of varying thickness to delineate where respondents perceived dialect boundaries. Like the work in the Dutch tradition, Sibata found a mismatch between perceptual and production boundaries.

This mismatch between production and perception has, in many cases, served as the justification for exploring the issues addressed in PD. But these early researchers were disheartened with the discrepancy. Sibata claimed his own data to reveal uninteresting findings. Like Bloomfield, Willem Grootaers (1959, [1964] 1999) suggested that the perceptions of nonlinguists were subjective and worthless. Furthermore, despite the success of his "little-arrow method," even Weijnen ([1968] 1999) heightened the controversy surrounding the mismatched research findings.

While this controversy caused some disruption in the history of the field of PD, it ultimately spawned a new style of research, referred to by Preston as the "modern" trend in PD, that, following from the early research conducted by Preston himself, continues to explore the intersection of language production and perception. The methods of the modern tradition (described below) have been more expansive than the earlier Dutch and Japanese traditions, in terms of both geographical reach, having been employed in numerous locations around the globe, for example, Brazil (Preston, 1989), France (Kuiper, 1999), Germany (Dailey-O'Cain, 1999; Diercks, 2002; Palliwoda and Schöder, 2016), Korea (Long and Yim, 2002; Jeon and Cukor-Avila, 2016), United Kingdom (Montgomery, 2007, 2016; Braber, 2016), and in terms of the various types data collected.

Preston’s (1989) book *Perceptual Dialectology: Nonlinguists’ Views of Areal Linguistics* serves as a compilation of several of his preliminary folk linguistic studies. Later, a two-volume handbook on the field (Preston, 1999; Long and Preston, 2002) was released, thus providing an introduction to the methods and findings of early and more recent research in the area. The vast majority of studies conducted in the modern tradition have focused on linguistic perceptions in the United States, much like Preston’s own research. While most studies have asked respondents about their perceptions of the United States as a whole (cf. Preston, 1989; Fought, 2002; Fridland, Bartlett, and Kreuz, 2004; Hartley, 2005; Fridland and Bartlett, 2006), more recent studies have turned the attention to how a single state perceives its own location both state-internally and within the larger linguistic landscape of the United States, as with research conducted in Ohio (Benson, 2003), California (Bucholtz et al., 2007; Bucholtz et al., 2008), Washington (Evans, 2011a, 2011b, 2016), and Texas (Cukor-Avila et al., 2012).

In this modern tradition, perceptual dialectologists have highlighted the importance of including the perceptions of nonlinguists in the research conducted by linguists concerned with production and attitudes. Because, as Preston proposed,
Perceptual Dialectology

[w]ithout knowledge of the value-ridden classifications of language and language status and function by the folk, without knowledge of where the folk believe differences exist, without knowledge of where they are capable of hearing major and minor differences, and, most importantly, without knowledge of how the folk bring their beliefs about language to bear on their solutions to linguistic problems, the study of language attitudes risks being: 1) a venture into the investigation of academic distinctions which distort the folk reality or tell only a partial truth or, worse, 2) a misadventure into the study of theatrically exaggerated speech caricatures.

(Preston, 1993b: 252)

Therefore, research in the modern tradition of PD seeks to fill the gap left by linguists and social psychologists in sidestepping the natural and necessary connection between linguistic ideologies, linguistic practices, and the appearance of such ideologies and practices in everyday social settings (Milroy and Preston, 1999). The conduct of such research requires the use of specific tools and methods, and the following section outlines the tools that have been most useful in these endeavors.

Tools of the Trade

As the modern paradigm has been the leading methodology used within PD, this section provides an introduction to the tools typically used in data collection in the projects following Preston’s methods. Preston proposed five types of data to be collected; these sources of data were developed over the course of Preston’s own long research program beginning in the 1980s. He offered the following five-point approach to the study of language perceptions in PD:

1. **Draw a map.** Respondents draw boundaries on a blank (or minimally detailed) map around areas where they believe regional speech zones exist ...
2. **Degree of difference.** Respondents rank regions on a scale of one to four (1 = same, 2 = a little different, 3 = different, 4 = unintelligibly different) for the perceived degree of dialect difference from the home area ...
3. “**Correct**” and “**pleasant.**” Respondents rank regions for correct and pleasant speech ...
4. **Dialect identification.** Respondents listen to voices on a “dialect continuum” [and] ... are instructed to assign each voice to the site where they think it belongs.
5. **Qualitative data.** Respondents are questioned about the tasks they have carried out and are engaged in open-ended conversations about language varieties, speakers of them, and related topics. (Preston, 1999: xxxiv)

This methodology provides a well-rounded view of nonlinguists’ visions of their dialect landscape, including not only a representation of where they believe the physical boundaries of certain varieties exist but also ratings of, accuracy of identification of, and their qualitative beliefs and attitudes toward the varieties they perceive. Such a
methodology allows researchers to go beyond the question of whether differences are perceived between varieties in order to systematically examine how those perceptions are operationalized.

**Drawing Mental Maps**

The “draw a map” task asks participants to draw (and label) the dialect boundaries that they perceive on a map of a given location. Such a map might best be described as a *mental map*, which, as a theoretical construct, is conceived of as the image one has in his or her mind about a certain place. Work in cultural geography (e.g., Gould and White, 1986) has indicated that getting people to draw these maps can give us some insight into how they see their world. As Gould and White discuss, mental maps can aid, for example, in town planning, as in Goodey’s (1971) work in Birmingham, England. In such projects, people share their varying perceptions of area landmarks (Lynch, 1960) or their neighborhoods (Ladd, 1967; Orleans, 1967) to reveal certain underlying sentiments, which can inform city planners in constructing paths, sidewalks, and so on.

To that end, Gould and White (1986) explored the mental maps of people in Britain, the United States, and elsewhere, to discover which areas of a country were the most desirable places to reside. Respondents were asked, if given free choice, where they would choose to live. In Britain, the national sentiment appeared to be in favor of living in the south of the country; local preferences for the home area, however, were prevalent the further north respondents lived. In the United States, opinions of northerners and westerners were in union, in that the South, specifically the Southern Trough (Mississippi, Alabama, and parts of South Carolina and Georgia), was seen as the least desirable place to live. For Alabamans, the picture was quite the opposite. They made more distinctions within the South, preferring their own state but giving harsh ratings to their neighbor, Mississippi. These maps reveal the precise distinctions one can make about one’s home area and shows that certain social, economic, and other factors might impinge upon people’s perceptions.

A great deal of research conducted within the modern tradition of PD has focused on nonlinguists’ production of these kinds of mental maps, with a focus shifted from residential preference to the existence of regional dialectal variation, particularly in the United States. In these studies, respondents are asked to “draw maps of the areas of the United States where people speak differently”² (Preston, 1982, as cited in Niedzielski and Preston, 2000: 46). For linguists, individual mental maps and their corresponding composite maps can give linguists clues about subjects’ perceptions of space, which provides added ethnographic detail of the group under examination. Additionally, studies of folk beliefs can enhance our understanding of linguistic variation, in that it is unlikely that nonlinguists experience linguistic change in a way completely unrelated to the ways traditional dialectologists have described it (Niedzielski and Preston, 2000). Although perceptual and production maps often yield similar results, this need not be the case (Benson, 2003).
These maps, which typically present an image of the entire United States, usually include state borders, as early studies (e.g., Preston, 1989, 1993a) indicated that people have great difficulty with completely blank maps due to a general lack of knowledge about American geography. Some studies have included additional information, such as city locations, topographical information, other major landmarks, and so on, though such practices have yet to reveal any systematic differences in the type or quality of data collected (but see Bounds and Hettel, 2014). If anything, the inclusion of additional information, even state lines, provides participants with predefined foci for their drawings, and participants have often found it difficult to “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” (Preston, 1989: 25).

Such individually created maps can provide valuable, anecdotal visions of a particular linguistic landscape, but it is not possible to generalize from single maps. Thus, one of the earliest goals in Preston’s work with mental maps was to create composite maps that characterized how many people from a place perceived the linguistic landscape under consideration. Specifically, “[t]he power of an aggregate is that it gives a generalized ‘picture’ of perception which has more explicative power than single images of mental maps produced by individual respondents” (Montgomery and Stoeckle, 2013: 52).

Early attempts at creating composite maps consisted of considering each respondent’s delimitations to be isoglosses, and, if a bundle of isoglosses appeared when all maps were overlaid, such bundles became the boundaries of a region for the entire study population. Recent advances in the use of GIS tools in PD have made this process of isogloss identification much easier (described later).

One of the earliest studies to explore the use of mental maps in this kind of research focuses on the perceptions of regional variation from the perspective of native Hawaiians. The goal of this early study was “to determine a method for producing a generalized map from a number of individual, hand-drawn ones” (Preston, 1989: 25). In this early study, Preston determined that a composite representation could be created by drawing perceptual isoglosses based on the lines of greatest agreement between respondents. Regions are included in composite maps if a large number of respondents used a similar label for a region, and the boundaries of that region are determined based on where the most respondents drew their boundaries. As evidenced by these composite and individual maps, especially when compared to recent production maps, like Labov, Ash, and Boberg’s (2006) Atlas of North American English (ANAE), the facts of perception do not necessarily coincide with facts of production. In the Hawaiian example mentioned above, a Midwestern region appears, which differs, both in terms of label and geographic scope, from the dialect area most closely associated with this part of the country, namely the Midland dialect. Additionally, the composite map reveals that Hawaiians perceive California and Texas to exist as stand-alone dialects, the Southern dialect region is not thought to extend as far north as in production maps, and certain parts of the country (parts of Kentucky, West Virginia, and Pennsylvania, among others) appear to receive no
regional affiliation, despite certainly having specific linguistic features associated with these regions.

Preston (1999) has made the generalization that, in map-drawing activities, respondents tend to draw first the most stigmatized areas of a country, while then further providing added detail to their local areas. It also appears that locals make more distinctions in their home area than outsiders do, as was the case with Gould and White’s (1986) work on residential preference. Similar results have been found in more recent studies where maps of smaller portions of the country were given to respondents with the same goal of discovering perceived dialect boundaries. For example, the map in figure 2 (Cramer 2010) reveals an urban/(mountain) rural distinction perceived by Kentuckians within their state (explored further in Cramer, 2016a, 2016b) while the composite map in figure 3 suggests that some of the most commonly delimited areas of this region for Louisville respondents are ones that have typically been stereotyped in negative ways (i.e., Cajun/Creole, Appalachia, Southern).

In addition to creating composite maps, researchers have also found it useful to explore the individual labels given to regions by the respondents. In these analyses, the goal is to see how respondents label varieties in ways that reveal larger ideological responses than the typically more neutral labels given to varieties by linguists. For example, Evans’s (2011a) exploration of perceived variation in Washington state reveals that the myth of a homogeneous
Washington dialect is accurately perceived as a myth by Washingtonians, and her analysis of the labels used by respondents suggests that the categories of “farmers,” “Latinos/Latinas,” and “gangsters” are important classes of types of Washingtonians, all of which are perceived to also have their own varieties of English. These analyses also typically reveal where nonlinguists perceive standard and nonstandard varieties to exist. In Bucholtz et al. (2007), the labels “standard” and “normal” are typically used to describe Northern California, while “broken” is used most often to describe San Diego. Finally, respondents also often use rather derogatory labels for a given variety; Cramer (2014, 2016a) found that, when referring to the Appalachian region of Kentucky, Appalachians and non-Appalachians alike use words like “hick,” “hillbilly,” and “redneck,” all terms that typically carry negative connotations.3

The analysis of mental maps provides evidence that goes beyond the typical language attitudes studies that employed matched-guise techniques (e.g., Lambert et al., 1960; Ryan and Giles, 1982), as such studies did not take into account whether respondents could place regional voices. But maps alone, however, tell us little about speaker attitudes toward different varieties. To understand this part of the question, perceptual dialectologists have employed ratings tasks, wherein participants rate how different, correct, and pleasant they believe certain varieties to be.

Rating Degree of Difference, Pleasantness, and Correctness

In addition to these mental maps, participants in PD studies are also often asked to rate varieties in terms of how different a specific variety is with respect to the participant’s own variety. The focus in the map-drawing activity is on locating difference, not on exploring the intensity of that difference. Thus, many researchers employing PD methods also ask respondents to describe how different varieties are from one another. The most common way in which this has been accomplished in studies of the United States has been through surveys asking respondents to rate each state4 on a scale of 1 (no difference) to 4 (unintelligibly different). Mean scores are calculated for each state, and maps like the one in figure 4 are created to exhibit where participants indicate the least and most different varieties. This map, from Cramer (2016a), shows the mean degree of difference perceived by Louisville respondents. It indicates that Indiana and Ohio are seen as very similar to a Louisville way of speaking. A comparison of these maps to the composite maps created using individual mental maps reveals that the two types of difference being evaluated may not be exactly the same.
Another level of rating takes place when respondents are asked more specifically to address certain social characteristics of language varieties, like how correct or pleasant such varieties are perceived to be. Such research begins to address what Preston (2013) has more recently referred to as language regard, or the metalinguistic perceptions and beliefs that nonlinguists attach to language varieties. This task, like the degree of difference task, typically asks respondents in studies about the United States to rate the states with respect to level of correctness and level of pleasantness, though the scale on which to rate the states tends to be larger, typically from 1 (least correct/pleasant) to 9 (most correct/pleasant). The map in figure 5 from Cramer (2016a), indicates that Louisvillians perceive Maine, Maryland, Ohio, Colorado, Nevada, and those in the Pacific Northwest, as well as Washington, DC, to be beacons of correctness, while most of the Southern states, West Virginia, and New Jersey are rated rather poorly in terms of correctness, a result that has been mirrored in many other PD studies of the United States.

Similar ideas about correctness and pleasantness emerge in the labels individual respondents give to varieties on their mental maps, where labels like “correct” or “standard” are associated with some areas while “friendly” and “pleasant” are associated with others. Preston’s generalizations of the results of these correctness and pleasantness ratings are that respondents who have high levels of what Labov (1966) called linguistic insecurity, “rate the local area as most pleasant, but they rate a number of areas as most correct” (Preston, 1999: xxxiv), while those who are more linguistically secure, like Preston’s Michigan participants, tend to rate their own local area as the sole bastion of correctness while encompassing a somewhat larger area around their local area for pleasantness.

Recent work by Cramer (e.g., 2010, 2016a) has explored making the connections between the map drawing task and the ratings task more explicit. In her studies, after completion of the “draw a map” task, respondents were asked to complete a language attitudes survey, where they listed the labels used on their maps and, using a four-point scale, rated these varieties in terms of the following social characteristics: difference (with
respect to their own variety), correctness, pleasantness, standardness, formality, beauty, and education. Thus, each participant received one survey sheet per variety delimited. That is, if a participant circled four regions, he or she received and completed four survey sheets, one about each of the varieties he or she designated. This method has proven successful in exploring which dialect areas respondents hold in high esteem and which areas are seen as least desirable, as it makes a more complete connection between the dialects they perceive and how they perceive them. It also more closely aligns with the bottom-up approach of dialect identification that is of interest to perceptual dialectologists.

Such ratings can be analyzed using statistical methods as well. While research like that of Fridland and Bartlett can utilize standard statistical tests (like repeated measures ANOVAs, t-tests), the analysis brought about from research like Cramer’s requires post hoc analyses because of the lack of pre-existing categories. In this research, Cramer employed a Tukey Honestly Significant Difference (HSD) post hoc analysis to analyze and compare the perceptions and attitudes respondents provided in the surveys.

**Placing Voices**

Another tool used in Preston’s PD studies involves placing voices. Given voices on a “dialect continuum” (in scrambled order) and a forced-choice set of locations, respondents are asked to determine the home location of a voice that they hear. This tool is seemingly the most problematic in terms of value, as varying studies have shown different levels of accuracy for respondents completing this type of task. For instance, while Preston (1993b) maintains that subjects do relatively well at placing voices, a notion that has been supported by work on the perception of vowels associated with the Southern Vowel Shift as Southern in Memphis (Fridland, Bartlett, and Kreuz, 2004; 2005) and by work on ethnic dialect identification (Purnell, Idsardi, and Baugh, 1999), others have found that respondents perform rather poorly on this task (e.g., Williams, Garrett, and Coupland, 1999; Clopper and Pisoni, 2004; Cramer, 2010).

Despite this possible controversy, this type of research is deemed important by Preston, as one cannot make connections between regional dialects and language attitudes without knowing the respondent’s ability to identify where a voice belongs. That is,

> unless we ask (and surprisingly few studies of language attitude have), we do not know where the respondents believe a voice is from. A report might accurately state that respondents had certain attitudes towards a South Midland voice sample, but the respondents might have gone home believing that they had heard an Inland North one.

(Preston, 1993b: 193)

Preston’s most famous example of this task (discussed in Niedzielski and Preston, 2000: 82–95, and elsewhere) involves the placement of voice samples from nine middle-class, middle-aged male speakers from nine different locations on a north–south continuum.
Perceptual Dialectology

from Saginaw, Michigan, to Dothan, Alabama. In addition to showing that nonlinguists are relatively good at placing voices on this north–south dimension, Preston showed that respondents seem to experience certain “minor” and “major” dialect boundaries that often resemble dialect boundaries described in traditional dialectology. For instance, while his Michigan respondents hear a major north–south boundary between sites #4 and #5 (Muncie, Indiana, and New Albany, Indiana, respectively), which approximately coincides with the traditional boundary of the South that runs alongside the Ohio River (e.g., Labov, Ash, and Boberg 2006), they also hear a minor boundary between sites #6 and #7 (Bowling Green, Kentucky, and Nashville, Tennessee, respectively), which might replicate a distinction between what Carver (1987) called “Upper South” and “Lower South” dialects.

On the other hand, Preston’s Indiana respondents heard two major dialect boundaries: one between sites #2 and #3 (Coldwater, Michigan, and South Bend, Indiana, respectively), and another between sites #3 and #4 (South Bend, Indiana, and Muncie, Indiana, respectively). The first of these boundaries resembles Carver’s line between “Upper North” and “Lower North,” while the second seems to align with the ANAE division between “The North” and “The Midland,” where a section of northern Indiana exists inside “The North” but outside “Inland North.” The results perhaps indicate a certain level of super-awareness of northernness associated with their sense of linguistic insecurity. They also hear the boundary between sites #4 and #5 (Muncie, Indiana, and New Albany, Indiana, respectively), though for Indiana respondents, this is a minor boundary.

Cramer (2010) also conducted a study of the accuracy with which participants could identify the home location of a specific voice. In this research, the goal was to find a speaker that participants in Louisville, Kentucky, would identify as most representative of the local Louisville dialect. Respondents were given ten segments of audio taken from a reality television show called Southern Belles: Louisville, five of which came from actual Louisvillians (target) and five of which belonged to people from elsewhere (filler). Respondents were given a map of a small region of the United States with major cities marked (see figure 6) and were asked to select any point on the map to indicate where they believed a speaker to be from. Results revealed that participants had quite a difficult time placing voices very consistently. For example, as can be seen in table 1, of the five Louisville women, Emily was selected as the speaker who was from closest to Louisville, at an average rate of more than 226 miles away from the city.

<table>
<thead>
<tr>
<th>Table 1 Mean distances from Louisville</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target</strong></td>
</tr>
<tr>
<td>Emily</td>
</tr>
<tr>
<td>Hadley</td>
</tr>
</tbody>
</table>
Interestingly, one of the fillers, Girl in Las Vegas, with an average distance of about 160 miles, is placed much closer to Louisville than any of the target speakers. All of the other fillers were approximately equally distant from Louisville as the target speakers. However, the numbers seem to indicate that the subjects actually place no speaker very close to Louisville. For example, the average of all of these mean distances is about 233 miles. Cities that are approximately this far from Louisville include: St. Louis, Missouri (265 miles), Gary, Indiana (265 miles), Columbus, Ohio (209 miles), Charleston, West Virginia (246 miles), Knoxville, Tennessee (246 miles), and Huntsville, Alabama (285 miles). The map in figure 6 reveals the points of origin selected by participants in this study for Emily. As is evident, respondents were fairly certain that Emily did not come from south of Nashville, Tennessee. Many respondents placed Emily in the state of Kentucky, with little clusters around both Louisville and Lexington. Some respondents, however, picked locations for Emily as far away from Louisville as Kansas City, Missouri, Virginia Beach, Virginia, and Toledo, Ohio. While the clustering around Louisville likely led to the selection of Emily as the closest to Louisville, the varied selections indicate that the certainty with which Louisvillians categorized this speaker was perhaps low. The maps for the other Louisville natives were similarly varied.

Research in this vein aligns closely with some of the goals of sociophonetics, the intersection between sociolinguistics and phonetics research. While this term has typically been used to refer to acoustic (vowel) analysis, it has recently taken a more broad definition so as to include not only production studies but perception ones as well. Research in placing voices, with particular attention to the phonetic variables available in the voice samples, can be used to examine not only the regional location of such voices, as in these studies, but can also indicate further social meaning attributed to certain speech sounds.

<table>
<thead>
<tr>
<th>Name</th>
<th>Latitude</th>
<th>Relation</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Julie</td>
<td>236.9597826</td>
<td>Shea’s fiancé</td>
<td>234.162016</td>
</tr>
<tr>
<td>Kellie</td>
<td>231.9585662</td>
<td>Shea’s dad</td>
<td>228.9426451</td>
</tr>
<tr>
<td>Shea</td>
<td>231.8052375</td>
<td>Hadley’s boyfriend</td>
<td>233.9634885</td>
</tr>
</tbody>
</table>

Source: Cramer, 2010: 212.

Figure 6. Map of point of origin selections for Emily (Cramer, 2010: 216)
including race (e.g., Purnell, Idsardi, and Baugh, 1999) and sexuality (e.g., Levon, 2006), among others.

The main generalization Preston (1999) has drawn from this kind of research has to do with the already mentioned level of accuracy of respondents. But he also suggests that respondents from different places hear the voices presented to them in varying ways. More specifically, the “minor” and “major” boundaries perceived by individuals are different for different groups of respondents from different locales, and it is suggested that respondents hear these boundaries with different levels of intensity, “more often making more distinctions closer to the local area and fewer in areas farther from the local area” (Preston, 1999: xxxv).

**Qualitative Data Analysis**

Finally, though briefly, it is important to acknowledge how the qualitative data that often accompany many PD studies can enhance our understanding of how nonlinguists evaluate the linguistic varieties they recognize. In many ways, this kind of information can be more insightful than the quantifications of the other tools presented here. In many of Preston’s studies, respondents have been asked to answer questions about the tasks they have completed. These questions often result in open-ended conversations about language and variation, including discussions of the people who speak certain varieties. Examples of this conversational data are discussed at length in Niedzielski and Preston, (2000).

While it is often difficult to make generalizations about this type of data, Preston (1999: xxxv) has noted some general trends in this type of conversational data. He claims that: (1) people mention face-to-face encounters with people who speak other varieties more often than they mention popular culture (i.e., television, movies, etc.) depictions of said varieties; (2) people often have trouble explicitly detailing phonological (and other) features of certain dialects, and imitations of said dialects can be accurate or inaccurate in many ways (Preston, 1992); and (3) people tend to be very concerned with correctness.

**Recent Advances in the Field**

One major way in which the techniques used in PD (specifically mental maps) have changed over the years includes a shift from more manual processes to those incorporating digital tools. As mentioned above, early work with composite maps required the physical overlay of individually, hand-drawn maps for the discovery of similarly marked regions. The tedious nature of this early process meant that only a very small number of maps could be combined and that it was impossible “to investigate gender, generation, class, or ethnic differences” (Niedzielski and Preston, 2000: 49) of the respondents who drew the individual maps without repeating the same manual process multiple times.
Later, Preston and Howe (1987) experimented with the first computer-created generalized maps. In this process, maps were individually traced onto a digitizing pad that fed the individual shapes into a program containing a base map. This program recorded one hit for each pixel enclosed in or touching the respondent’s boundaries. Maps were created, though they were not created automatically and were sometimes uninformative and unwieldy, wherein it was difficult to determine patterns of agreement. It was at this stage that composite maps were often created that represented patterns of agreement where regions were represented if outlined by half of the respondents. Unlike the early manual work with dialect boundaries on composite maps, this technique allowed for the creation of maps not only of larger numbers of respondents but also of specific portions of the study population.

In the 1990s, Onishi and Long (1997) created the Perceptual Dialect Quantifier (PDQ) program using a Windows-based system. In this program, after digitizing individual maps, dialect areas were analyzed in ways similar to Preston and Howe, though dialect areas were calculated not only by counting pixels but also by coloring areas that were most often encircled in a darker color, thus revealing better patterns of agreement.

While PDQ is still available for use on a small number of aging computers in Japan, ultimately these techniques became outdated, as aggregate maps could no longer be reliably created using such obsolete hardware and software, and the maps produced by such technology were typically of rather low quality and resolution. Additionally, these approaches utilized graphical tactics in presenting the data (that is, they considered the data to be separate from the map underlying it), thus neglecting the spatial dimension of such perceptions. Recent advances in the use of GIS tools may provide the answer to both of these issues.

A GIS “lets us visualize, question, analyze, interpret, and understand data to reveal relationships, patterns, and trends” (ESRI, 2014). Tools like ArcGIS (produced by ESRI) as well as freely available (open source) GIS tools like QGIS allow perceptual dialectologists to integrate hardware, software, and data in ways that connect what map-drawers draw to real world maps and coordinates through a process of georeferencing, which “uses coordinate systems in order to tie data to a set position on the earth’s surface” (Montgomery and Stoeckle, 2013: 58). Such a move away from graphical representations allows researchers to connect nonlinguists’ perceptions to the world in which those perceptions are enacted, providing further detail as to the types of information that might be influencing such perceptions. For example, in a GIS, it is possible to overlay many different types of information (i.e., cities, topography, census data, linguistic production isoglosses), such that the researcher can connect the perceptions elicited from respondents to the other information on which they may be relying when providing their perceptions (even when the maps used in the task do not contain such information).

GIS offers many ways to geoprocess and visualize linguistic data. Despite its offerings, few linguists to date have utilized the tools. Kirk and Kretzschmar (1992) and Pederson...
(1993) were some of the earliest linguists with interests in dialect geography to acknowledge the possibilities of using GIS to explore linguistic variation. In PD, however, more researchers are beginning to utilize these tools. For example, Evans (2011b) used GIS tools to create the composite map in figure 7, which shows that the areas outlined by respondents in Washington state tend to coincide with larger urban areas like Seattle and Spokane. Others (e.g., Cramer, 2010, 2016a, 2016b; Cukor-Avila et al., 2012; Evans, 2011a, 2016; Jeon and Cukor-Avila, 2016) have also used these tools, and an edited volume by Cramer and Montgomery (2016) highlights some of this more recent work in PD utilizing GIS tools for the analysis of perceptual data.

The steps for using GIS in PD research have been clearly laid out in Montgomery and Stoeckle (2013). The flow chart provided by those authors, recreated here in figure 8, shows the steps one might take to pursue such data analysis using ArcGIS. It is clear, from Montgomery and Stoeckle’s description, that GIS offers many advantages to perceptual dialectologists: allows for better processing and visualization of data; provides the ability to map PD data alongside other kinds of data, like geopolitical boundaries, topography, and so on; can combine several defined areas on one map, which was difficult in previous attempts to use computational methods with PD data; includes many different possibilities in terms of calculations and statistics previously unavailable.
A recent PD project on intrastate linguistic variation in Kentucky can serve as a useful example for understanding not only how GIS tools have been incorporated in such an analysis, but to also connect such recent advances to the tools that have traditionally been used in PD research. As mentioned earlier, recent research on the perceptions of nonlinguists in the United States have turned to explorations within and about a single state (i.e., Benson, 2003; Bucholtz et al., 2007; Bucholtz et al., 2008; Evans, 2011a, 2011b, 2016; Cukor-Avila et al., 2012). This current project considers Kentuckians’ perceptions about the varieties spoken in the state to discover how and where Kentuckians see linguistic variation. Kentucky is presented as an important region for the examination of nonlinguists’ perceptions because of its position at a major linguistic, perceptual, and cultural border. Because of its indeterminate position in the linguistic landscape of the United States, it is hypothesized that Kentuckians’ experience confusion of perceptions, such that their perceptions reveal their split perception of themselves as both Northern/Midwestern and Southern. In this project, respondents were asked to complete the “draw-a-map” task, using a map that only included the outline of the state of Kentucky, as well as a language attitudes survey about each area delimited (Cramer, 2010). This survey focused on the typical correctness and pleasantness ratings of earlier PD studies while also adding other social characteristics such as beauty, education, formality, and standardness. Participants were asked to rank the ways of speaking on a scale of 1–4, where a lower score indicates a lower value, on each of the social characteristics given.

In total, 250 maps were collected, scanned, and digitized using ArcGIS 10. A sample individual map can be seen in figure 9. An initial analysis of the labels used by respondents indicated that the most frequently used labels for defining the linguistic areas of the state were directional in nature (i.e., northern, eastern, southern, western, and central Kentucky). These labels correspond to common labels used for other descriptions of the state in the popular discourse, such as in the media and in tourist bureau websites.
These directional labels were selected as the overarching category labels for dialect areas within the state in the process of creating composite maps. An example of one such composite map, made in ArcGIS and showing 30 percent agreement, can be seen in figure 10. At this level of agreement, there is considerable overlap between what is perceived to be a Southern Kentucky\textsuperscript{10} variety and the Eastern and Western Kentucky varieties. A map showing 50 percent agreement levels, as in figure 11, reveals smaller regions (as expected), but there is much less overlap, indicating the most prominent areas associated with these varieties.

ArcGIS was also used to create heat maps, or maps that show the extent of all areas encircled within a given overarching category. An example of one of these maps can be seen in figure 12. Of the total 250 maps, there were 188 regions that were classified as Eastern Kentucky. This region was the most commonly delimited region within the state, a result which is not surprising given that this is likely the most stigmatized variety in the state. Through the use of varying color densities, this map shows the levels of overlap of these 188 regions, indicating that the easternmost (darkest) portion of the state is the most agreed upon portion of the state that belongs to this dialect area.
(2010), the ratings for each overarching category label were averaged in each of the social categorizations evaluated (correctness, pleasantness, standardness, formality, beauty, and education), and those averages were ranked in relation to one another, as in table 2. This table reveals that Kentuckians perceive a Northern Kentucky dialect to be the most correct dialect in the state while Eastern Kentucky is rated lowest.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Mean (Correct)</th>
<th>Rank (Correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Kentucky/Bluegrass</td>
<td>119</td>
<td>1.89916</td>
<td>2</td>
</tr>
<tr>
<td>Eastern Kentucky</td>
<td>188</td>
<td>2.93617</td>
<td>5</td>
</tr>
<tr>
<td>Northern Kentucky</td>
<td>154</td>
<td>1.876623</td>
<td>1</td>
</tr>
<tr>
<td>Southern Kentucky</td>
<td>91</td>
<td>2.450549</td>
<td>4</td>
</tr>
<tr>
<td>Western Kentucky</td>
<td>117</td>
<td>2.222222</td>
<td>3</td>
</tr>
</tbody>
</table>

After ranking the dialects, a pairwise comparison was made using a Tukey Honestly Significant Difference (HSD) statistical test, yielding the statistical significance of the rankings for the different language attitudes categories. Pairs are created by taking the lowest rated variety and comparing it first to the highest rated variety until the lowest rated variety has been compared to all varieties. The comparison continues between the variety rated next-to-last and the highest rated variety, and so on, until a non-significant result is obtained. What is shown in table 3 is the Q score that was calculated for each pair in the analysis of level of correctness. If the Q score is higher than the Q-critical score of 3.86, the difference between the pair of varieties is significant (indicated in the table as TRUE in the last column). As table 3 shows, every dialect is rated as statistically significantly more correct than the eastern Kentucky dialect.
Table 3 Tukey HSD results for level of correctness

<table>
<thead>
<tr>
<th>Pairs</th>
<th>n</th>
<th>mean</th>
<th>Q</th>
<th>Q-critical</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Kentucky</td>
<td>188</td>
<td>2.93617</td>
<td>19.07227</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Northern Kentucky</td>
<td>154</td>
<td>1.876623</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern Kentucky</td>
<td>188</td>
<td>2.93617</td>
<td>17.31898</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Central Kentucky/Bluegrass</td>
<td>119</td>
<td>1.89916</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern Kentucky</td>
<td>188</td>
<td>2.93617</td>
<td>11.86163</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Western Kentucky</td>
<td>117</td>
<td>2.222222</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eastern Kentucky</td>
<td>188</td>
<td>2.93617</td>
<td>7.43962</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Southern Kentucky</td>
<td>91</td>
<td>2.450549</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern Kentucky</td>
<td>91</td>
<td>2.450549</td>
<td>8.492</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Northern Kentucky</td>
<td>154</td>
<td>1.876623</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern Kentucky</td>
<td>91</td>
<td>2.450549</td>
<td>7.746388</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Central Kentucky/Bluegrass</td>
<td>119</td>
<td>1.89916</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Classification</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td>Southern Kentucky</td>
<td>91</td>
<td>2.450549</td>
<td>3.195918</td>
<td>3.86</td>
<td>FALSE</td>
</tr>
<tr>
<td>Western Kentucky</td>
<td>117</td>
<td>2.222222</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western Kentucky</td>
<td>117</td>
<td>2.222222</td>
<td>5.513113</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Northern Kentucky</td>
<td>154</td>
<td>1.876623</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western Kentucky</td>
<td>117</td>
<td>2.222222</td>
<td>4.854595</td>
<td>3.86</td>
<td>TRUE</td>
</tr>
<tr>
<td>Central Kentucky/</td>
<td>119</td>
<td>1.89916</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bluegrass</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central Kentucky/</td>
<td>119</td>
<td>1.89916</td>
<td>0.361236</td>
<td>3.86</td>
<td>FALSE</td>
</tr>
<tr>
<td>Bluegrass</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern Kentucky</td>
<td>154</td>
<td>1.876623</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ultimately, the eastern Kentucky variety, which is also the portion of the state where one would find the Appalachian Mountains, an area of the country subject to some of the harshest negative stereotypes, was ranked as lowest in each of the social categorizations provided on the attitudes survey (correctness, pleasantness, standardness, formality, beauty, and education), while the more urban areas of Central and Northern Kentucky were rated highest. The results of both the mental maps and the language attitudes study suggest a (mountain) rural/urban divide within the state, one that coincides well with broader cultural impressions and stereotypes in the American linguistic landscape.

As these methods and technologies continue to develop, researchers interested in PD, and many other areas of linguistic variation, will have numerous ways to process and visualize data. It is hoped that the illustration of these tools provided here will encourage more PD studies that not only incorporate these tools but also help advance them. There is much more to be done to fully understand how GIS can be most effectively used in (socio)linguistic research, and such an understanding can only be gained through the utilization and adaptation of these tools to the specific research questions we have. PD research has been revitalized because of these advances, and it seems reasonable that similar results would be obtained in other areas if these methods become part of the general toolkit of linguists.

References


Perceptual Dialectology


Palliwoda, Nicole, and Schöder, Saskia (2016). “Perceptual Dialectology, Speech Samples, and the Concept of Salience: Initial Findings from the DFG-Project ‘Lay Linguists’ Perspective on German Regional Varieties: Reconstructing Lay Linguistic Conceptualizations of German in a Perceptual Dialectology Approach.” In *Cityscapes and Perceptual Dialectology: Global Perspectives on Non-linguists’ Knowledge of the Dialect*
**Perceptual Dialectology**


Perceptual Dialectology


Notes:

(1) For more on these and other PD methods, see Montgomery and Cramer, 2016.

(2) Typically, the word “dialect” is explicitly avoided in research on nonlinguists’ perceptions, as the word (as used by nonlinguists) tends to necessarily carry negative connotations.

(3) It is possible, however, that Appalachian participants have reappropriated (Chen, 1998) such terms for more positive or acceptable usage.

(4) New York City and Washington, DC, have also often been examined separately so that respondents also were asked to rate them.

(5) Earlier attempts at gauging these attitudes required participants to rank each state plus New York City and Washington, DC, from 1 to 52.

(6) Most of the studies in the United States have revealed that labels of the first kind are typically attributed to Northern varieties while labels of the second kind are typically attributed to Southern varieties. As we will see, similar results are found in the quantitative study of these attributes.

(7) For more details on sociophonetics, see Baranowski, 2013.
(8) Parts of this research project have been presented at the American Dialect Society Annual Meeting (Cramer and Hardymon, 2014) and the Southeastern Conference on Linguistics (Cramer, 2014). The data also has been considered in Cramer (2016a).

(9) For further discussion of Louisvillians’ split perceptions, see Cramer, 2013.

(10) It is possible that the Southern Kentucky region represents a conflation of two types of “Southern”: one that indicates direction in the state and another that indicates the larger cultural region in the United States with which they associate this portion of the state. Such a conflation could create the results found here.

(11) For a more thorough description of the Tukey HSD method used here, see Cramer, 2010.

(12) For more on Appalachian stereotypes, see Billings, Norman, and Ledford, 1999.

Jennifer Cramer

Jennifer Cramer is Assistant Professor in the Department of Linguistics at the University of Kentucky. She received her Ph.D. from the University of Illinois at Urbana-Champaign in 2010. Her main research focus is on variation in the perception and production of regional linguistic identities, with a specific focus on the varieties of English spoken in Kentucky.