Auditory free classification—a task in which listeners classify auditory samples into unconstrained groups—has provided insight into perceptual representation and categorization for several sources of speech variability including U.S. regional dialects, nonnative accents, and foreign languages. Within these studies, phonological markedness and geography have emerged as central organizing principles. However, previous studies were limited by including only one source of variability. To address this gap, the perception of U.S. regional dialects, international English dialects, and nonnative accents was investigated within one classification task. Listeners categorized talkers based on perceived location of origin. Cluster analysis demonstrated a perceptual divide between native and nonnative talkers. Native talkers were further delimited by geographic proximity into Southern Hemisphere, U.S., and United Kingdom groups. One exception was the consistent grouping of Southern U.S. talkers with talkers from England. Nonnative talkers were grouped into three major branches: French and German, Asian, and other. The "other" branch primarily consisted of less familiar accents. The results suggest that native and nonnative accents are perceived as separate categories regardless of accent markedness. Additionally, when listeners are presented with a wide range of dialects and accents, geography remains an important organizing principle. [Research support by IU Hutton Honors College.]
INTRODUCTION

The speech signal is extremely variable; the same word can sound quite different depending on the person producing it. One source of variability is a speaker’s dialect or accent. Variations of a language specific to a certain region within a nation constitute regional dialects (e.g., Southern American English or New England English), and international dialects are variations of a language across nations with the same native language (e.g., Australian English or South African English). Thus, both regional dialects and international dialects are native dialects. In contrast, foreign-accented speech is produced by talkers who are not native speakers of the target language (e.g., native speakers of Spanish or Korean speaking in English). The current study investigated how native listeners perceive and categorize speech stimuli that include variability stemming from these three sources.

The auditory free classification task is a method that has been used to investigate the perception of speech variability (Bradlow, Clopper, Smiljanić, and Walter, 2010; Clopper and Bradlow, 2008, 2009; Clopper and Pisoni, 2007). This method allows listeners to freely categorize speech samples based on their perception without experimenter-determined structures. Analyses of the classification behavior can indicate the perceptual similarity structure of the speech samples and the salient perceptual dimensions of the speech stimuli. Previous work with regional dialects of American English using this technique demonstrated that listeners make more fine-grained categories than can be predicted by forced-choice tasks (Clopper and Pisoni, 2007). Further, three dimensions – markedness (i.e., the number of phonological variants that differ from other dialects), geography, and gender – emerged as perceptually salient for the categorization of dialects (Clopper and Pisoni, 2007). For nonnative speech varieties, markedness and talker gender were also two central perceptual dimensions (Atagi and Bent, 2011). It is not known how classification behavior and salient perceptual dimensions may change when multiple sources of variability – such as both regional dialects and foreign accents – are examined within the same free classification task. Furthermore, the perception of international dialects has not been examined with this methodology. The current study addressed these gaps.

Numerous studies have found that the presence of an unfamiliar dialect or foreign accent can cause decrements in intelligibility and slower processing (Adank, Evans, Stuart-Smith, and Scott, 2009; Clopper and Bradlow, 2008; Munro and Derwing, 1995). However, very few studies have directly compared the perception and processing across different speech varieties. The two studies that have investigated this question have found that nonnative accents, ethnic dialects, and international dialects are generally harder to process and comprehend than an unfamiliar regional dialect or the standard dialect, particularly in difficult listening conditions (Adank, Evans, Stuart-Smith, and Scott, 2009; Major, Fitzmaurice, Bunta, and Balsubramanian, 2005). For example, Standard British English listeners were slower to complete truth verification judgments when the sentences were produced by a Spanish-accented speaker than when the sentences were produced by a speaker of an unfamiliar native dialect (i.e., Glaswegian English). The listeners were fastest when the speaker shared the native dialect of the listener (Adank et al., 2009). Previous studies have been limited by including only a small range of linguistic varieties (i.e., 3 to 5). Therefore, our understanding about how listeners perceive and represent the full range of variation present in a language is severely limited.

While previous studies have typically investigated only one type of speech variability, the current study employed a free classification task containing U.S. regional dialects, international English dialects, and nonnative accents simultaneously. The current study used six U.S. dialect regions similar to Clopper and Pisoni (2007), expanded on the foreign accents in Atagi and Bent (2011), and added international dialects of English. Simultaneously investigating three types of speech variation in a single task allowed for the investigation of how perception and categorization are influenced when the range of speech variability is broad and includes three distinct types of variation. Further, this study contributes to the understanding of the psycholinguistic distance among different linguistic varieties (Cristia, Seidl, Vaughn, Schmale, Bradlow, and Floccia, 2012). Specifically, this study addressed the question of whether native varieties, which are phonologically quite distinct from the listener’s dialect, are perceived similarly to nonnative varieties (Perceptual Distance Hypothesis) or whether there is a fundamental divide in the cognitive processing of native and nonnative varieties (Different Processes Hypothesis) (Goslin, Duffy, and Floccia, 2012).
METHODS

LISTENERS

Fifty monolingual, native speakers of American English (15 males, 35 females) between the ages of 18 and 30 (mean age: 21) participated in the study. All listeners passed a hearing screening and reported no history of speech or hearing impairments. Forty-four of the participants grew up in the Midwest (Indiana, Illinois, Ohio, Iowa) including 31 participants from Indiana. Four additional participants were tested, but their data was excluded from the analysis because they were bilingual.

STIMULI

The speech samples were selected from the Speech Accent Archive from George Mason University (Weinberger, 2013). Since gender is perceptually salient even when listeners are told to ignore it, only male talkers were included to avoid listeners focusing on talker gender rather than dialect/accent (Clopper and Pisoni, 2007). Each speech sample was a different talker reading the same two sentences: “Please call Stella. Ask her to bring these things with her from the store.” Six regional dialects of English, six International English varieties, and twelve foreign accents were each represented by three talkers each for a total of 72 speech samples. The United States regions were divided into North, South, West, Mid-Atlantic, New England, and Midland (Clopper and Pisoni, 2006; Labov, Ash and Boberg, 2006). The 18 International English talkers were from Australia, New Zealand, England (London), Ireland, Scotland, and South Africa. The foreign-accented talkers had the following native language backgrounds: French, German, Spanish, Russian, Thai, Korean, Japanese, Mandarin, Gujarati, Arabic, Somali, and Swahili. The three talkers within each dialect, accent, or language background were chosen from similar geographic areas.

PROCEDURES

Participants were seated at a computer in a sound-attenuated booth, and the speech stimuli were presented through headphones. The experiment was composed of four free classification tasks. For each task, there was a 22x22-cell grid presented on the screen with squares representing each talker to the left of the grid. Each square contained two arbitrary letters and participants were told that the letters were not significant. Listeners could hear a sample by clicking on a square, and could listen as many times as they liked. The listeners were instructed to put talkers into groups on the grid according to where they thought the talkers were from, with talkers from the same place in the same group. Listeners were told that they could make as many groups as they wished, and there could be as many squares per group as they liked. They could also move the squares around as many times as needed.

The first classification task contained the 18 talkers representing the six American English dialects and the second task contained all 36 native English speakers, American and International. The third task contained the 36 nonnative speakers, and the final classification included all 72 talkers. At the end of the last free classification task, participants reported where they thought the talkers were from in free response format. Only the results for the final free classification task with all 72 talkers are described below.

RESULTS

To analyze listeners’ classification behavior, the summed similarity matrix of all participants’ classification of the 72 talkers was submitted to an additive clustering analysis (Gorter, 1982). The results showed that there was a divide between native talkers and nonnative talkers, and there were further divisions within these two groups.

The main cluster containing the native talkers was divided into three sub-branches. One of these was made up of accents from the Southern Hemisphere: Australia, New Zealand, and South Africa. A second branch was composed of talkers from the United Kingdom (i.e., England, Scotland, and Ireland) as well as talkers from the Southern US regional dialect region. The talkers from Ireland were split up between two different branches, with two of the talkers being grouped with the United Kingdom branch and one in the Southern Hemisphere branch. The third native branch included all the talkers from US regional dialect regions, with the exception of the talkers from the South.

The nonnative talkers were divided into four clusters. The first two clusters consisted of French and German talkers. One cluster had two French and one German talker and the other cluster had two German and one French talker. The second cluster was composed of talkers from Asia and included the Korean-, Japanese-, Mandarin-, and
Thai-accented talkers. The third, “other” cluster was composed of an assortment of accents that are less familiar to the participants, such as Arabic, Gujarati, Swahili, Somali, and Russian, as well as Spanish. Arabic and Spanish were on the same sub-branch within this third cluster.

**DISCUSSION**

The current study investigated the categorization of multiple linguistic varieties – regional dialects, international dialects, and foreign accents – within a single task. Previous studies have compared limited varieties of native and nonnative accents, and many have focused on processing costs related to comprehension. This study used a large number of speech samples and used listeners’ categorization patterns to examine the psycholinguistic distance among talkers with a wide range of nonnative and native accents. Specifically, the current study aimed to determine whether there is a perceptual divide between native and nonnative varieties (Different Processes Hypothesis) or whether the perception of these varieties would be better described along a continuum (Perceptual Distance Hypothesis) (Goslin, Duffy, and Floccia, 2012). The Perceptual Distance Hypothesis refers to the idea that regional accents and foreign accents are both represented in terms of their acoustic distance from the listener’s native variety of speech. Thus, regional and foreign accents could be placed on a continuum with foreign accents at the far end of the continuum. Under this hypothesis, the same mechanism is used to process both native and nonnative varieties. The Different Processes Hypothesis suggests that regional and foreign accents are processed using different mechanisms or strategies. In the current study, clustering analysis of the free classification data revealed that there was a perceptual divide between native and nonnative talkers, supporting the Different Processes Hypothesis.

Supporting evidence for this hypothesis has also been observed with an event-related potentials study which found that regional accents were normalized immediately at the pre-lexical level, while foreign accents required additional processing beyond the pre-lexical level (Goslin, Duffy, and Floccia, 2012). In the current study, all native talkers – including the International talkers whose accents were quite phonologically distinct from the US talkers – were in the same main cluster. This separation between the native and nonnative talkers suggests that listeners are sensitive to native status, regardless of how marked or distinct the accent is. However, there were still subgroupings within the native cluster indicating the possibility of a continuum of variability within this category.

As has been observed in previous free classification studies that included more limited speech varieties, geography remained a central organizing principle when listeners grouped a large range of dialects and accents. The divisions within the native and nonnative branches were at least partially based on geography. For example, with the nonnative talkers, there were western European (i.e., French- and German-accented talkers) branches and an Asian branch, which was composed of Korean-, Japanese-, Mandarin-, and Thai-accented talkers. Similarly, the native branch was divided into American, Southern Hemisphere, and United Kingdom sub-branches.

The listeners in the current study also likely attended to acoustic-phonetic similarities of the talkers when making their groups. Although these acoustic-phonetic features tend to correlate with the geographic origin of the talker, there may be similarities between languages that do not have any relation to each other, geographically or historically, because there is a limit to the number of possible sounds that can be produced in any human language (Bradlow et al., 2010). Similarities in acoustic-phonetic features could explain the “other” nonnative branch containing Gujarati, Arabic, Spanish, Swahili, Somali, and Russian. For example, Arabic-accented talkers and Spanish-accented talkers were in the same subgroup within this “other” nonnative branch. Though they are not in the same language family (Afro-Asiatic vs. Indo-European), their production patterns have similar phonetic characteristics. For example, both the Arabic- and Spanish-accented talkers substituted an alveolar trill for the target English /l/.

Familiarity emerged as another possible organizing principle in this study. The consistent grouping of the Southern US talkers with the British English talkers may be explained by accent familiarity together with markedness. The accents of the American Southern talkers and the British English talkers were marked for the listeners, but were still highly familiar. The confluence of markedness and familiarity may have led listeners to group these talkers together. Familiarity leads to increases in accent identification (Bayard, Weatherall, Gallois, and Pittam, 2001) and may also influence categorization performance. Most of the language backgrounds in the “other” nonnative cluster (Gujarati, Arabic, Spanish, Swahili, Somali, and Russian) are likely to be less familiar to most listeners, and thus, listeners may have grouped the talkers with those language backgrounds based on that factor. It also remains possible that shared acoustic-phonetic features of these accents were responsible for the categorization behavior.

In summary, these results demonstrated that listeners are sensitive to the differences between native and nonnative accents, even when the range of variation within the native talkers is substantial. The results provided support for the hypothesis that there are fundamental differences in the processing of native and nonnative accents.
(i.e., the Different Processes Hypothesis), but the possibility of listeners’ using a continuum-like strategy when categorizing talkers within those two categories remains. Furthermore, listeners appear to use many of the same organizing principles when categorizing a small number of linguistic varieties or a large range of varieties, as was employed in the current study. Accent familiarity also emerged as a possible additional perceptual dimension that may influence categorization of linguistic varieties.

ACKNOWLEDGMENTS

This research was funded by a grant from the Indiana University Hutton Honors College. The authors would like to thank Charles Brandt for his technical assistance with programming.

REFERENCES