

INDIVIDUAL AND GROUP DIFFERENCES IN THE PERCEPTION OF REGIONAL
DIALECT VARIATION IN A SECOND LANGUAGE

Terrin N. Tamati

CHAPTER 1

PERCEPTION OF REGIONAL DIALECT VARIATION

Introduction

In everyday situations, listeners will encounter many talkers with diverse developmental, linguistic, and social backgrounds. Highly detailed indexical variability that conveys information about talker-specific and group characteristics is encoded in the speech signal, along with the abstract symbolic linguistic content of the utterance (Abercrombie, 1967). Although much attention has been previously given to the processes underlying the perception of the linguistic content of an utterance, such as sound segments, syllables, and words, indexical variability also plays an important role in speech perception processes. The listener must simultaneously process the linguistic and indexical information in the speech signal to be able to arrive at the intended meaning of the utterance, while also extracting information about the talker who produced the utterance and how it was produced (e.g., Johnson & Mullennix, 1997; Pisoni, 1997). Previous studies have shown that the listener may use this information to make judgments about a talker, including his or her identity (e.g., Van Lancker, Kreiman, & Emmorey, 1985; Van Lancker, Kreiman, & Wickens, 1985), gender (e.g., Lass, Hughes, Bowyer, Waters, & Bourne, 1976), age (Ptacek & Sander, 1966), and region of origin and background (e.g., Labov, 1972). Furthermore, the learning of indexical patterns in the speech signal also is crucial in understanding speech. Listeners are able to learn talker- and group-specific indexical patterns to facilitate speech recognition (e.g., Nygaard, Sommers, & Pisoni, 1994; Nygaard & Pisoni, 1998). Thus, dealing with indexical variability from a variety of sources is part of normal, everyday speech communication.

Indexical Variability in Second Language Acquisition

Natural variability in the speech signal from talker and group differences may contribute to communication difficulties in a second language (L2). Language-specific indexical patterns in speech may present a challenge to L2 listeners because non-native listeners lack detailed, native-like knowledge of the L2 phonological system and often are not exposed to substantial variability in their learning environment. Several studies have shown that non-native listeners may have difficulty processing indexical information in a second language. Goggin, Thompson, Strube, and Simental (1991) examined monolingual English and monolingual Spanish listeners' ability to identify talkers in a familiar and an unfamiliar language. The Spanish listeners were significantly better at identifying voices in their own language than in the unfamiliar language. Thompson (1987) also found that monolingual English listeners were better at identifying voices in English than Spanish-accented English and Spanish, but that they were also better for Spanish-accented English than Spanish. Given that listeners in those studies were better able to identify voices in their native language than an L2 and better able to identify voices in an L2 of talkers who share the same native language, processing of language-specific indexical information relies heavily on the listeners' detailed knowledge of the phonological system of the language.

For listeners who lack native-like knowledge of an L2, successfully identifying and/or discriminating indexical variability in the L2 may be a challenging task. However, previous studies on cross-language perception of indexical variability suggest that listeners may also draw on more general language-independent knowledge. Winters, Levi, and Pisoni (2008) carried out talker identification and discrimination tasks with listeners using a familiar (English) and unfamiliar (German) language. Listeners were either trained to identify bilingual German-English speakers in English or in German and were then tested on their ability to identify the

talkers in both the trained and non-trained languages. Like Goggin et al. (1991) and Thompson (1987), Winters et al. (2008) found that listeners were, overall, better at identifying the talkers when they used the familiar English language than the unfamiliar German language. However, when the listeners had to identify the talkers in the language they were not explicitly trained on, the German-trained listeners were better able to generalize the talkers' voices in the non-training language than the English-trained listeners, suggesting that the German-trained listeners may have attended more to language-independent talker features, which allowed them to form more robust and generalizable representations of the listeners while the English-trained listeners relied more on language-dependent talker features. Interestingly, Winters et al. (2008) also found that English-trained listeners used language-specific segmental information even in the generalization test in German because listeners performed poorly on words that contained unfamiliar German vowels.

Winters et al. (2008) also carried out a talker discrimination task in which listeners heard either English or German words (German-German, English-English, German-English, and English-German) and indicated whether the two words were produced by the same talker or a different talker. In that task, listeners performed significantly above chance for all language pairs (German-German, English-English, German-English, and English-German), but they were more accurate in matched language pairs, especially the English-English pair. This pattern suggests that listeners were not only using language-dependent information to make their decisions, but they were also better at processing language-dependent information in the familiar language. Across listener groups, English-trained and untrained listeners were more likely to give the same-talker response on matched language pairs than on mismatched language pairs. The German-trained listeners did not show the same tendency, suggesting that they were not

conflating linguistic and indexical information by relying on the linguistic content of the utterance as well as the talker-specific indexical cues to the same extent as the English-trained listeners. Winters et al. (2008) concluded that processing of indexical information relies on language-dependent information when the listeners understand the language. This allows listeners to use integrated indexical and linguistic information to identify talkers, resulting in more accurate identification in the familiar language but less generalizable representations for an unfamiliar language. Processing of indexical information relies on language-independent information when listeners do not understand the language, resulting in poorer recognition performance but more generalizable representations.

Although the Winters et al. (2008) study does not give much insight into second language speech perception, the findings suggest that the extent to which language-dependent information is used in processing is strongly related to the listener's knowledge of and familiarity with the sound system of the language. Other cross-language talker identification studies also provide support for this interpretation. Köster, Schiller, and Kunzel (1995) found that English listeners with knowledge of German performed better on a talker identification task in German than listeners with no knowledge of German. Similarly, Köster and Schiller (1997) found that Spanish and Chinese listeners with knowledge of German performed better than Spanish and Chinese listeners with no knowledge of German. Later, Sullivan and Schlichting (2000) found improvement in voice recognition after initial study of English. However, they did not find any later improvement after the second semester of study. Nevertheless, taken together, these studies indicate that improvement in talker identification and processing of indexical information occur with increased familiarity with the second language, and, perhaps, with increased mastery of the sound system of the second language.

One source of variation that likely plays an important role in the acquisition of a second language is regional origin. Both native and non-native listeners frequently encounter talkers with diverse regional backgrounds in a second language environment. Regional dialect variation in speech perception may be particularly challenging for non-native listeners, because the ability to adapt to and use this source of information in speech requires both sensitivity to dialect-specific features in the speech signal and signal-independent sociolinguistic knowledge of regional dialect characteristics and stereotypes (e.g., Tamati & Pisoni, in press). Non-native listeners not only have imperfect knowledge of the second language but they also have less experience with sociolinguistic variation in the second language. Thus, in addition to acquiring knowledge of the second language sound system, they must also acquire and form representations of regional dialect categories, along with the social and linguistic knowledge associated with those categories. This knowledge serves as a foundation for developing a native-like capability of quickly adapting to different talkers of different language backgrounds to understand the utterance and use of indexical information to make judgments about the talker.

However, acquiring knowledge of regional dialect differences in an L2 may be difficult since non-native listeners are not exposed to the variation in the second language to the extent of a native speaker and they receive little explicit feedback as to the region of origin of talkers in day to day communication. Furthermore, at early stages of L2 learning, more cognitive effort may be placed on understanding the linguistic content of an utterance rather than making judgements about the talker's voice. Because of these factors, non-native listeners would not be expected to have detailed knowledge of regional dialect variation in the second language and, as such, would likely to be more susceptible to regional dialect effects in comprehension tasks (Nathan, Wells, & Donlan, 1998). Furthermore, they would likely be less to have knowledge of

which sounds pattern together in different dialects, making dialect recognition or discrimination a difficult task (Clopper & Bradlow, 2009). Consequently, native-like communication ability may be impeded in the acquisition of an L2.

Speech Recognition and Regional Dialects

Previous studies have demonstrated that regional dialect variation interacts with the listener's unique developmental linguistic history to influence his or her ability to successfully recognize cross-dialect sounds, words, and sentences. Regional dialects that are familiar to a listener or more standard dialects are easier to understand (e.g., Mason, 1946). In a study on cross-dialect intelligibility, Labov and Ash (1997) found that listeners who shared a dialect with the talkers were more accurate than listeners who did not share a dialect with talkers in a transcription task under good listening conditions. Listeners from Birmingham, Alabama were more accurate in their responses than listeners from Philadelphia and Chicago when they were presented with phrase-length utterances from a Birmingham talker. Floccia, Goslin, Girard, and Konopczynski (2006) found that listeners responded more quickly in a lexical decision task to talkers from the same region as well as talkers who spoke the standard Parisian dialect compared to unfamiliar regional French dialects. Similarly, Clopper and Bradlow (2008) found that the General American talkers were the most intelligible for listeners of all dialect backgrounds at three signal-to-noise ratios (SNRs). Taken together, these studies suggest that native listeners benefit substantially from previous experience and are able to more easily understand familiar local and standard dialects in the native language, but they may still have difficulty in understanding unfamiliar or nonstandard dialects.

Non-native listeners often have more difficulty understanding speech than native listeners, especially in more challenging listening conditions (for a review, see García Lecumberri, Cooke, & Cutler, 2010). Non-native listeners may also have more difficulty recognizing speech produced by talkers from different regional dialects in their L2. Cooke, García Lecumberri, and Barker (2008) found that the intelligibility of multiple talkers in a variety of listening conditions depended on the native language of the listener (native (English) or non-native (Spanish)). Individual and group talker characteristics likely played a large role because non-native listeners had particular difficulty with the nonstandard Scottish variety, likely due to lack of exposure or familiarity with that particular variety.

Ikeno and Hansen (2007) also found that non-native listeners who lived in the United States were much poorer at understanding dialects of British English in a transcription task compared to British English and American English listeners. Additionally, the intelligibility of three dialects differed for the native and non-native listeners. While both British and American listeners found Cambridge and Cardiff talkers to be more intelligible than Belfast talkers, non-native listeners found only Cardiff talkers to be the most intelligible and performed similarly for Cambridge and Belfast talkers. As such, the comprehension of the speech was severely impacted by the dialect variation for non-native speakers as compared to American English listeners who also had very little experience with the varieties. These findings suggest that, besides lack of experience with regional dialects, not sharing the phonological system with the target language substantially influences the intelligibility of the dialects in an L2. Similarly, Major, Fitzmaurice, Bunta, and Balasubramanian (2005) found that L2 listeners with different native language backgrounds living in western United States were better able to understand native speakers of a supraregional (standard) American English variety and the Southern American English variety

compared to African American Vernacular English, Indian English, and Australian English. Although the effects of the different talker dialects in that study may be related to a variety of factors, including dialect familiarity and exposure, the findings suggest that L2 listeners have more difficulty understanding some regional dialects than others and that the primary regional dialect of instruction or exposure may be the variety that is most intelligible.

While longer length of residency has been found to improve general speech recognition abilities in the L2 (e.g., Oyama, 1976; Mackay, Flege, Piske, & Schirru, 2001; Mackay, Meador, & Flege, 2001; Ingvalson, McClelland, & Holt, 2011), exposure to local or standard varieties has also been found to influence L2 speech recognition abilities. Through exposure, L2 learners may acquire dialect-specific phonological knowledge that would facilitate speech understanding for familiar regional dialects. Baker and Smith (2010) found that learners of European and Quebec French showed dialect-specific perception and production of French vowels, reflecting their language learning histories. Similarly, Escudero and Boersma (2004) found that the perception of the English vowels /i/ and /I/ by Spanish learners of English reflected both influence from their native language and the realization of these vowels in the target dialect (Scottish or Southern British). Acquiring dialect-specific phonological knowledge in the L2 may directly improve listening comprehension for that particular dialect.

Schmidt (2009) examined Spanish learners' listening comprehension performance before and after a three-week study abroad in the Dominican Republic. Native speakers of English significantly improved in their ability to understand Dominican Spanish after the short exposure to that variety during the study abroad. Although general speech understanding other varieties of Spanish also improved slightly, the gains observed for comprehension of the Dominican variety were disproportionately greater and therefore not related to a more general improvement in

Spanish (Schmidt, 2009). Taken together, these studies suggest that a non-native listener may acquire dialect-specific knowledge of the L2 phonology after brief to extended exposure to those dialects in the L2, improving the ability to discriminate sounds and understand speech of the L2 regional dialect.

However, other studies indicate that cross-dialect speech recognition in an L2 may be influenced by a variety of other factors. Fox and McGory (2007) found that native English listeners and native Japanese learners of English, regardless of location of residence (Ohio or Alabama), performed more accurately in a vowel identification task for the General American English vowels than Southern American English vowels, suggesting a processing benefit for the standard variety. However, unlike the native listeners from Ohio who showed better performance than listeners from Alabama with talkers from Ohio, the native Japanese learners living in either Alabama or Ohio did not exhibit a processing benefit for the local variety. Eisenstein and Verdi (1985) also found that for English learners in New York who were very familiar with General American English, African American English, and New York English, African American English was significantly less intelligible than both General American and New York English. The findings of these studies suggest that non-native listeners exhibit a processing benefit for increased exposure to L2 speech, and more specifically to standard varieties, but they may or may not receive a similar processing benefit from exposure to other familiar, local nonstandard varieties. Regional dialects may pose a significant challenge to successful non-native speech recognition, but performance on any particular regional variety may reflect the interaction of both the intelligibility of the variety for the non-native listeners as well as the amount and type of previous exposure to the variety.

Regional Dialect Identification and Discrimination

Previous studies have also found that naïve listeners can perceive dialect-specific differences in the speech signal and can use stored knowledge of such differences to make judgments about the regional background of talkers in their native language. These studies have shown that listeners are able to categorize talkers by region of origin (e.g., Van Bezooijen & Gooskens, 1999; Clopper & Pisoni, 2004b), group talkers by region of origin (e.g., Clopper & Pisoni, 2007), or make explicit judgments about the similarity of talkers' voices based on regional dialect (e.g., Clopper, Levi, & Pisoni, 2006). In Clopper and Pisoni (2004b), listeners used dialect-specific properties in the acoustic signal to identify unfamiliar talkers' regions of origin in a forced-choice dialect classification task using six American English regional dialects. Similarly, using an auditory free classification task, Clopper and Pisoni (2007) found that naïve listeners were sensitive to dialect-specific phonetic differences and were able to use this information to group talkers into broad regional dialect categories. Clopper et al. (2006) also found that naïve listeners can make explicit judgments about the similarity of talkers' voices based on their region of origin. Findings from these studies suggest that naïve listeners have knowledge of dialect-specific phonological variation in their native language and are able to use this knowledge to make reliable judgements about talkers based on broad regional dialect categories in a variety of experimental tasks.

Many studies have also shown that dialect identification and discrimination abilities are influenced by the prior linguistic experience and history of the listener. Greater familiarity with regional dialects, for example, from living in or near specific geographic regions, results in easier identification and greater perceived distinctiveness of dialects (e.g., Clopper & Pisoni, 2004a; Clopper & Pisoni, 2004b; Clopper & Pisoni, 2006). Clopper and Pisoni (2004a) examined the

effects of residential history on the perception of regional dialect variation in American English. In their study, a group of mobile listeners (“Army Brats”), who had lived in at least three different states, and a group of nonmobile listeners (“Homebodies”), who had only lived in Indiana, performed a six-alternative force-choice dialect categorization task. Mobile listeners were more accurate than the nonmobile listeners in the categorization task and mobile listeners who had lived in a particular dialect region were better at categorizing talkers from that region than mobile listeners who had not lived in that specific region. Clustering analyses from the categorization task for the different listener groups also revealed that the perceived similarity of the regional dialects differed for the mobile and nonmobile listeners. Clopper and Pisoni (2006) explicitly looked at the effect of region of origin on dialect categorization. Four groups of listeners who differed in terms of region of origin (North or Midland dialect regions) and geographic mobility (Mobile or Non-mobile) performed a forced-choice categorization task. While listeners did not differ in overall categorization accuracy, geographic mobility and region of origin affected the perceived similarity of the regional dialects. Clopper and Pisoni (2007) found similar listener background effects using an auditory free classification task. Taken together, these studies suggest an important role for linguistic experience in the perception of regional dialect variation.

Recent studies have shown that while non-native listeners have knowledge of (Clopper & Bradlow, 2009) and are capable of accommodating to and learning (Ikeno & Hansen, 2007) new patterns of regional dialect variation in a second language, their performance is not as accurate as native listeners of that language. Non-native listeners not only lack experience with different regional dialects in their L2 compared to native listeners, but they also do not share the same phonological system with the target language. The second language phonological system

influences the non-native listener's ability to perceive and use dialect-specific information in the speech signal to identify a talker's region of origin. Ikeno and Hansen (2007) found that after a brief familiarization with the British English dialects, non-native listeners who lived in the United States were able to perform slightly above chance in a three forced-choice dialect categorization task with single words and phrases, but they did not perform nearly as well as British English listeners or American English listeners who also were not familiar with British English dialects.

Using an auditory free classification task, Clopper and Bradlow (2009) found that non-native listeners were able to use variation in the acoustic signal to classify talkers by regional dialect of American English. While both native and non-native listeners showed similar classification strategies, native listeners were more accurate overall than non-native listeners. The difference in performance was attributed to non-native listeners being less able to use distributional information about the features of particular dialects when faced with within-dialect variation (Clopper & Bradlow, 2009). Tamati and Pisoni (in press) also found that non-native listeners with little experience in the U.S. were near chance at identifying talkers' regions of origins in a forced-choice dialect categorization using American English dialects. Although there was a great deal of listener variability in accuracy across dialects and overall, listeners in their study did not have robust knowledge of dialectal differences in American English that they could draw upon to successfully complete the task. These earlier studies suggest that while the non-native listeners are able to perceive the variation in the speech signal, they lack signal-independent knowledge of the distribution of reliable dialect-specific features to be able to accurately identify or discriminate regional dialects in the L2.

However, exposure to regional dialect variation in a second language may improve non-native ability to use dialect-specific information to identify or discriminate an unfamiliar talker's region of origin. Eisenstein (1982) found that beginner and intermediate learners of English in New York City were less accurate than native listeners on a dialect discrimination task with several types of English dialects, but advanced learners of English performed as well as native English listeners. Stephan (1997) also found that native German learners of English were not as accurate as native English listeners in identifying varieties of world English in an open-set dialect identification task. While the non-native listeners were much better at identifying more familiar or salient varieties, like Southern American English, they were very poor at identifying other less familiar varieties of English, like South African English. The identification performance on the dialects in the study may reflect the listeners' familiarity with those varieties. The more familiar the listeners were with a particular variety, the better they were at identifying it (Stephan, 1997). Studies exploring the role of linguistic experience in dialect identification or discrimination by native listeners suggest that, while non-native listeners are less accurate overall at identifying regional dialects, experience with a variety or varieties in a second language leads to more robust representations for regional dialects in the L2 and better processing skills in perceiving and using dialect-specific information encoded in the signal in speech communication tasks.

Individual Listener Variability in Speech Perception and L2 Acquisition

While there have been many studies investigating the perception of regional dialect variation using a variety of linguistic and indexical tasks, there have been few, if any, studies that have looked at the performance of individual listeners. Previous studies have found that even

listeners with highly similar language backgrounds vary substantially in their ability to adapt and adjust to multiple sources of variability in the speech signal. In a recent study, Gilbert, Tamati, and Pisoni (2013) examined speech recognition under high-variability listening conditions. In their study, young, normal-hearing adults who were monolingual native speakers of American English completed a sentence recognition task using sentences produced by multiple female and male talkers from various U.S. dialect regions from the PRESTO test (Perceptually Robust English Sentence Test Open-set) in multi-talker babble. They found that the listeners varied greatly in their ability to understand the highly variable PRESTO sentences. Performance of 121 normal-hearing native speakers ranged from approximately 40% to 76% key word recognition on the high-variability sentences across four different SNRs (+3 dB, 0 dB, -3 dB, and -5 dB SNR). This wide range of sentence recognition performance demonstrates that some listeners were better able to rapidly adapt and adjust to the variability in the test materials to understand the sentences, despite sharing the same native language and having a normal developmental and hearing history. Other studies on speech recognition in a variety of listening conditions have also reported substantial individual variability in listener performance (e.g., Neff & Dethlefs, 1995; Richards & Zang, 2001; Wightman, Kistler, & O'Bryan, 2010).

Variability in non-native speech perception performance, even among listeners of the same native language background, has also been widely documented in the L2 literature. Sources of group differences in L2 speech perception abilities among non-native listeners with the same native language include age and length of exposure to the L2 (e.g., Oyama, 1976; Mackay et al., 2001a; Mackay et al., 2001b; Ingvalson et al., 2011), and frequency and amount of L2 use (e.g., Oyama, 1976; Flege, Bohn, & Jang, 1997; Flege & MacKay, 2004; Ingvalson et al., 2011). However, even when the language backgrounds of the non-native listeners are controlled,

individual listeners vary greatly in how quickly and how reliably the phonological system of an L2 is acquired. Indeed, listener variability in L2 perception and production has been a major issue in the field of second language acquisition and learning (e.g., Skehan, 1991; Dörnyei, 2005; Golestani, Molko, Dehaene, LeBihan, & Pallier, 2007).

While individual differences in a variety of L2 speech perception and production tasks have been addressed in the literature, little attention has been given to non-native speech perception in highly variable conditions or on indexical perception tasks. Tamati and Pisoni (in press) found that non-native listeners display a great deal of variability in adapting to high-variability listening conditions. Using the same PRESTO sentence materials as Gilbert et al. (2013), Tamati and Pisoni (in press) found that native speakers of Mandarin, who had lived in the U.S. for 27 months or less, had great difficulty understanding the high-variability PRESTO sentences. Additionally, the non-native listeners showed large differences on the task with performance ranging from approximately 12% to 35% key word accuracy across the four SNRs (+3 dB, 0 dB, -3 dB, and -5 dB SNR). Furthermore, these differences were not entirely accounted for by indices of language background or experience. Gaining a better understanding of individual differences in speech perception in highly variable or challenging listening conditions, including listening environments involving multiple talkers and several regional dialects, may contribute to our understanding of the factors that underlie speech perception in an L2 speech environment. Furthermore, by identifying factors that contribute to individual differences in L2 speech recognition ability, sources of difficulty for poor L2 users may be more easily recognized and addressed both within and outside the L2 classroom.

Improving Knowledge of L2 Regional Dialects

Another unexplored issue surrounding research on regional dialect perception in an L2 is how to improve learners' knowledge of regional dialect variation. Although experience and exposure to an L2 has been shown to improve the ability to adapt and adjust to regional dialect variation in both a native or second language, many non-native listeners have little, if any, experience in the L2 environment and therefore must deal with this source of variability in speech without sociolinguistic or grammatical knowledge of their L2. Without basic sociolinguistic knowledge of the L2, regional dialect variation may pose a particularly significant challenge to speech communication. Given these problems, non-native learners without substantial experience in the L2 environment would likely benefit from targeted training on L2 regional dialects. To deal with L2 regional dialect variation, learners must develop robust perceptual skills that are necessary for successfully adapting to the variability in the speech signal. These include sensitivity to sublexical and subphonemic variability and knowledge of reliable dialect-specific acoustic-phonetic characteristics (Clopper & Bradlow, 2009). Therefore, any effective training regimen that is designed to improve L2 sociolinguistic knowledge of regional dialect variation must be able to improve the learner's sensitivity to these dialect-specific features and help these listeners learn how features group together for each dialect.

Previous research suggests that high-variability regional dialect training with explicit feedback may provide one method to improve knowledge of L2 regional dialects. High-variability perceptual training has been successful in promoting learning of specific speech sounds or sound contrasts in a second language (see, for example, the r/l studies of Bradlow, Pisoni, Yamada, & Tohkura, 1997; Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994). This type of training presents listeners with multiple exemplars of a

particular category, with immediate explicit feedback as to the correct category, in order to build robust perceptual categories. Adapting and using this approach for high-variability L2 regional dialect training, listeners may gain knowledge of L2 regional dialects and form robust, talker-independent perceptual dialect categories. Exploring the potential use of high-variability training to improve L2 sociolinguistic knowledge of regional dialects to facilitate dialect identification and comprehension in the L2 may provide insight on the processes underlying the perceptual analysis of variability in speech, as well as the acquisition of sociolinguistic categories in the L2. Additionally, research on regional dialect categorization training may help to develop adequate tools for improving perceptual abilities in daily communication in an L2.

New Directions

Despite advances in our understanding of how regional dialect variation is perceived and processed, it remains unclear what specific factors contribute to non-native listeners' ability to perceive, encode, and store robust dialect-specific information in day-to-day life, which they would be able to draw upon in various communicative environments. While earlier research suggests that the listener's native language and previous experience to native or L2 regional dialects affects the perception of dialectal variations in dialect identification task and the intelligibility of regional dialects in linguistic tasks, it is not clear how these factors contribute to and interact with the developing second language phonological system and underlying cognitive functions. To address these issues, the research reported in this dissertation expands on the previous work dealing with perception and acquisition of regional dialect variation by investigating: 1) non-native ability to understand as well as identify or discriminate regional

dialects of American English, 2) sources of individual differences in these perceptual abilities, and 3) the benefits of perceptual regional dialect categorization training in a second language.

Chapter 2 presents the results of three perceptual tests that were designed to investigate non-native ability to recognize speech produced by talkers from different regions of origin within the United States, as well as identify and discriminate regional dialects of American English. Native speakers of American English, Korean non-native speakers of American English who resided in the United States between 6 and 18 months (KST), and Korean non-native speakers of American English who have resided in the United States for more than three years (KLT) completed a multi-dialect sentence recognition task, a forced-choice regional dialect categorization task, and an auditory regional dialect free classification task. Each task used four regional dialects of American English: General American, Northern, Mid-Atlantic, and Southern. Chapter 3 discusses the individual differences found in in multi-dialect sentence recognition task and forced-choice regional dialect categorization task discussed in Chapter 2, and their relations to listener performance on some traditional measures of second language phonological and lexical knowledge. Chapter 4 also examines individual differences but focuses on underlying neurocognitive factors potentially contributing to differences in performance. Chapter 5 reports a regional dialect training study in which native and non-native listeners completed a high-variability regional dialect categorization training task with explicit feedback about the unfamiliar talker's region of origin. Finally, Chapter 6 provides a summary of findings from these studies and discusses several potential implications for speech science and second language phonological acquisition.

References

- Abercrombie, D. (1967). *Elements of general phonetics*. Chicago: Aldine.
- Baker, W., & Smith, L. C. (2010). The impact of L2 dialect on learning French vowels: Native English speakers learning Québécois and European French. *The Canadian Modern Language Review*, 66, 711-738.
- Bradlow, A. R., Pisoni, D. B., Yamada, R. A., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310.
- Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech*, 51, 175-98.
- Clopper, C. G. & Bradlow, A. R. (2009). Free Classification of American English dialects by native and non-native listeners. *Journal of Phonetics*, 37, 436-451.
- Clopper, C. G. & Pisoni, D. B. (2004a). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change*, 16, 31-48.
- Clopper, C. G. & Pisoni, D. P. (2004b). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111-140.
- Clopper, C. G., & Pisoni, D. B. (2006). Effects of region of origin and geographic mobility on perceptual dialect categorization. *Language Variation and Change*, 18, 193-221.
- Clopper, C. G. & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics*, 35, 421-438.
- Clopper, C. G. Levi, S. V., & Pisoni, D. B. (2006). Perceptual similarity of regional dialects of American English. *Journal of the Acoustical Society of America*, 119, 566-574.

- Cooke, M., García Lecumberri, M. L., Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *Journal of the Acoustical Society of America*, 123, 414-427.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Eisenstein, M. (1982). A study of social variation in adult second language acquisition. *Language Learning*, 32, 367-391.
- Eisenstein, M., & Verdi, G. (1985). The intelligibility of social dialects for working-class adult learners of English. *Language Learning*, 35, 287-298.
- Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26, 551-585.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25, 437-470.
- Fox, R. A., & McGory, J. T. (2007). Second language acquisition of a regional dialect of American English by native Japanese speakers. In O.-S. Bohn &
- Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? Human Perception and Performance. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1276-1293.
- García Lecumberri, M. L., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, 52, 864– 886.
- Gilbert, J. L., Tamati, T. N., & Pisoni, D.B. (2013). Development, Reliability and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *Journal of the American Academy of Audiology*, 24, 1-11.

- Goggin, J. P., Thompson, C. P., Strube, G., & Simental, L. R. (1991). The role of language familiarity in voice identification. *Memory & Cognition*, 22, 448-458.
- Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, 17, 575-582.
- Ikeno, A., & Hansen, J. H. L. (2007). The Effect of Listener Accent Background on Accent Perception and Comprehension. *EURASIP Journal on Audio, Speech, and Music Processing*, 2007, 1-8.
- Ingvalson, E. M., McClelland, J. L., & Holt, L. L. (2011). Predicting native English-like performance by native Japanese speakers. *Journal of Phonetics*, 39, 571-584.
- Johnson, K., & Mullennix, J. W. (1997). Complex representations used in speech processing: Overview of the book. In K. Johnson & J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 1-8). San Diego, CA: Academic Press.
- Köster, O., & Schiller, N. O. (1997). Different influences of the native language of a listener on speaker recognition. *Forensic Linguistics*, 4, 18-28.
- Köster, O., Shiller, N. O., & Kunzel, H. J. (2005). The influence of native-language background on speaker recognition. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the Thirteenth International Congress of Phonetic Sciences*, 4, pp. 306-309. Stockholm.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W., & Ash, S. (1997). Understanding Birmingham. In T. Bernstein, T. Nunnally, & R. Sabino (Eds.), *Language Variety in the South Revisited* (pp. 508-573). Tuscaloosa, AL: Alabama University Press.

- Lass, N. J., Hughes, K. R., Bowyer, M. D., Waters, L. T., & Bourne, V. T. (1976). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *Journal of the Acoustical Society of America*, *59*, 675-678.
- Lively, S.E., Logan, J.S., & Pisoni, D.B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, *94*, 1242-1255.
- Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y. I., Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, *96*, 2076–2087.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, *36*, 173–190.
- Mackay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001a). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America*, *110*, 516–528.
- Mackay, I. R. A., Meador, D., & Flege, J. E. (2001b). The identification of English consonants by native speakers of Italian. *Phonetica*, *58*, 103–125.
- Mason, H. M. (1946). Understandability of speech in noise as affected by region of origin of speaker and listener. *Speech Monographs*, *13*, 54-68.
- Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language*, *25*, 343-365.

- Neff, D.L., & Dethlefs, T.M. (1995). Individual differences in simultaneous masking with random-frequency, multicomponent maskers. *Journal of the Acoustical Society of America*, 98, 125–134.
- Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific perceptual learning in speech perception. *Perception & Psychophysics*, 60, 355-376.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42-46.
- Oyama, S. (1976). A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research*, 5, 261-283.
- Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 9-32). San Diego, CA: Academic Press.
- Ptacek, P. H. & Sander, E. K. (1966). Age recognition from voice. *Journal of Speech and Hearing Research*, 9, 273-277.
- Richards, V.M., & Zeng, T. (2001). Informational masking in profile analysis: Comparing ideal and human observers. *Journal of the Association for Research in Otolaryngology*, 02, 189-198.
- Schmidt, Lauren B. (2009). The effect of dialect familiarity via a study abroad experience on L2 comprehension of Spanish. In J. Collentine et al. (Eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla.
- Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition*, 13, 275-298.

- Stephan, C. (1997). The unknown Englishes? Testing German students' ability to identify varieties of English. In E. W. Schneider (Ed.), *Englishes around the world* (pp. 93-108). Amsterdam: John Benjamins.
- Sullivan, K.P.H. and Schlichting, F. 2000. Speaker discrimination in a foreign language: first language environment, second language learners. *Forensic Linguistics*, 7(1), 95-111.
- Tamati, T. N., Gilbert, J. L., Pisoni, D. B. (2013). Some Factors Underlying Individual Differences in Speech Recognition on PRESTO: A First Report. *Journal of the American Academy of Audiology*, 24, 616-634.
- Tamati, T. N., & Pisoni, D. B. (in press). Non-native speech recognition in high-variability listening conditions using PRESTO. Submitted to the *Journal of the American Academy of Audiology*.
- Thompson, C. P. (1987). A language effect in voice identification. *Applied Cognitive Psychology*, 1, 121-131.
- Van Bezooijen, R., & Gooskens, C. (1999). Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology*, 18, 31-48.
- Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: Patterns and parameters: Part I. Recognition of backward voices. *Journal of Phonetics*, 13, 19-38.
- Van Lancker, D., Kreiman, J., & Wickens, T. (1985). Familiar voice recognitions: Patterns and parameters: Part II. Recognition of rate-altered voices. *Journal of Phonetics*, 13, 39-52.
- Wightman, F. L., Kistler, D. J., O'Bryan, A. (2010). Individual differences and age effects in a dichotic informational masking paradigm. *Journal of the Acoustical Society of America*, 128, 270-279.

Winters, S. J., Levi, S. V., & Pisoni, D. B. (2008). Identification and discrimination of bilingual talkers across languages. *Journal of the Acoustical Society of America*, 123, 4524-4538.

CHAPTER 2

NON-NATIVE PERCEPTION OF AMERICAN ENGLISH REGIONAL DIALECTS

Introduction

Although previous studies have examined the ability to recognize speech produced by talkers from diverse regions of origin and the ability to identify regional dialects in a second language (L2), little is still known about the factors that contribute to an individual L2 listener's ability to adapt to and use regional dialect information in speech perception tasks. The ability to process both linguistic and indexical information in the speech signal to recover the intended meaning of the utterance and information about how the utterance was said and who said it is a fundamental part of speech perception processes (e.g., Nygaard, 2005). However, it is widely attested that individual listeners vary greatly in performance on both speech recognition tasks and indexical categorization tasks involving different sources of talker and regional dialect variability (e.g., Tamati, Gilbert, & Pisoni, 2013). Previous studies on the perception of indexical information in speech have suggested that an individual listener's ability to recognize highly variable speech signals and identify and discriminate sources of variability in the speech signal may be closely related. Recently, Tamati, Gilbert, and Pisoni (2013) examined individual differences in recognition performance exhibited by native listeners on a high-variability sentence recognition test. They found that differences in recognition performance were related to individual differences in the processing of indexical information. Listeners who had good sentence recognition performance were also better at discriminating talkers by gender and categorizing talkers by regional of origin. In another study, Tamati and Pisoni (in press) found

that key word accuracy on the high-variability sentences was related to the non-native listeners' processing skills in identifying the region of origin of unfamiliar talkers.

In two pioneering studies that examined the effects of talker familiarity training on intelligibility, Nygaard, Sommers, and Pisoni (1994) and Nygaard and Pisoni (1998) found that previous experience learning a novel talker's voice and vocal source characteristics increased word recognition in noise. Listeners in these studies were trained to identify a set of voices. After training, a transcription task was carried out using the same talkers and new talkers. While the trained listeners were more accurate in recognizing speech produced by familiar talkers, they also showed substantial individual differences in learning to recognize the talkers' voices. Furthermore, those listeners who were poor at learning to explicitly categorize the novel voices showed no benefit in recognizing speech produced by familiar talkers in the transcription task, while those listeners who were better at learning the voices showed a significant benefit in spoken word recognition for familiar talkers.

In a follow-up study, Levi, Winters, and Pisoni (2011) investigated whether the familiar talker advantage (i.e., better intelligibility for familiar talkers than unfamiliar talkers) could be obtained cross-linguistically using the same talkers speaking two languages. Listeners were trained to identify bilingual talkers using German or English utterances. After the voice training, listeners performed a word recognition task in English. Again, there was much variability in the ability of the listeners to learn to identify the voices, and not all listeners were able to reach the 70% accuracy criteria on the last day of training in either the German-trained group or the English-trained group. Unlike the previous studies, however, while the familiar talker advantage was found for good English-trained listeners, the effect was not observed for the German-trained listeners. The results suggested that listeners had acquired knowledge of language-dependent

indexical properties to get the familiar talker advantage. Furthermore, finding good and poor voice learners regardless of the training language suggests that whatever makes a listener good or bad at learning voices is language-independent and may be linked to performance on other speech related tasks.

Research on individual differences in deaf children with cochlear implants further suggests that the perception of indexical information in the acoustic signal is related to speech recognition. Cleary and Pisoni (2002) examined the ability of deaf children with cochlear implants (mean age of 8.76 years, range of 7.92 to 9.91 years) to discriminate female voices using a talker discrimination task. While Cleary and Pisoni (2002) found that overall the children had difficulty with the task, there was much listener variability in performance. The children who were better able to discriminate the talkers' voices were also better able to recognize spoken words. In a similar study, Cleary, Pisoni, and Kirk (2005) examined talker discrimination abilities of deaf children with cochlear implants (mean age of 8.05 years, range of 5.50 to 12.58 years) using a task measuring how acoustically different voice cues must be to result in the perception of two different talkers. Cleary et al. (2005) found that there was much variability in performance among the children with cochlear implants, and that many of the children had great difficulty making talker discrimination judgements. However, some of the deaf children with cochlear implants showed patterns of performance that were similar to normal-hearing children, and those children were also better able to identify key words from spoken sentences. The results of these two studies show that the deaf children with cochlear implants who were better at the linguistic tasks were also more sensitive to talkers' voices in the indexical tasks.

Studies examining the perception of linguistic and nonlinguistic information in the speech signal suggest that the processing of these two channels of information in speech may be

linked or coupled for an individual listener. Individual listeners may benefit from more robust, highly detailed representations and memory codes in long-term memory for both signal and context information in processing linguistic and indexical information, as in the sentence recognition or indexical identification or discrimination tasks. However, the perception of sociolinguistic variation, in particular regional dialect variation, in both linguistic and indexical tasks has yet to be investigated for either native or non-native listeners. Given the findings from previous studies on individual differences in speech perception, the individual listener's ability to recognize words produced by talkers from different regions of origins and his/her ability to discriminate and identify regional dialects are likely to be closely related.

The Current Study

The current study investigated the perception of regional dialect variation in American English by non-native listeners in order to: 1) assess non-native speech recognition across four U.S. dialects; 2) evaluate the perceptual categorization and classification of regional dialects by non-native listeners; 3) investigate the relation between speech recognition across U.S. regional dialects and the identification of regional dialects for non-native listeners; and 4) examine group (length of residency) and individual differences among non-native listeners in the processing of linguistic and indexical information in speech. Non-native speakers of American English residing in Bloomington, IN completed three speech perception tasks: 1) a sentence recognition task using multiple American English regional dialects, 2) a forced-choice regional dialect categorization task, and 3) an auditory regional dialect free classification task. To limit the effect of native language and language background on performance, all non-native listeners were native speakers of Korean; the amount of exposure to American English in the second language setting

was also limited. Non-native participants had either lived in the U.S. for a relatively short (6-18 months) or long period of time (more than 3 years).

The sentence recognition task was designed to measure speech recognition across several different U.S. dialect regions. This task assessed the non-native listeners' ability to recognize words in sentences produced by unfamiliar native speakers of English from four different U.S. dialect regions (General American, Northern, Mid-Atlantic, and Southern) in multi-talker babble. Given demographic characteristics of Bloomington, IN, the primary U.S. residence for the non-native listeners, all the listeners were expected to have been the most familiar with the General American variety, somewhat familiar with the Northern and Southern varieties spoken by many Indiana University students and in nearby regions, and least familiar with the Mid-Atlantic variety, which is not well-represented in Bloomington and the surrounding areas.

Two different tasks were used to investigate non-native regional dialect classification. The first task used a four-alternative forced-choice regional dialect categorization test. This categorization task required listeners to perceive dialect-specific information in the speech signal and use stored knowledge of regional dialect variation to select an unfamiliar talker's region of origin from a closed set of four response alternatives. The four response alternatives were General American, Northern, Mid-Atlantic, and Southern, which were the same regional dialects used in the sentence recognition task. Unlike the forced-choice categorization task, the free response task did not require the listener to explicitly identify any of the regional dialects. In this task, listeners sorted talkers from the same four U.S. dialect regions into groups based on region of origin. Listeners were not required to provide any category names, and they were able to listen to the talkers as many times as they desired in any order. Thus, the free classification task

allowed us to measure the perceived similarity of the regional dialects without imposing explicit category labels (Clopper, 2007), which may be unfamiliar to the non-native listeners.

Given previous findings, non-native listeners were expected to recognize fewer words in the sentence recognition task than native listeners. Additionally, if direct experience with dialectal variation in the L2 improves speech recognition performance, non-native listeners would also be expected to have more difficulty with the nonstandard regional dialects in the task (i.e., Northern, Mid-Atlantic, and Southern) given their lack of exposure to these varieties. Similarly, for the dialect classification tasks, non-native listeners would be expected to be less accurate at using dialect-specific features in the speech signal to identify the region of origin of a talker or classify the talkers by region of origin, especially for relatively unfamiliar dialects. However, the free classification task should have also been easier for the listeners compared to the forced-choice task because its design gave listeners more flexibility in listening to the stimuli and in responding.

Among the non-native listeners, it was expected that the listeners who had lived in the U.S. for a longer period of time would have had more exposure to regional dialect variation in American English and would therefore have acquired more sociolinguistic knowledge about dialect variation. However, the effects of length of residency may be manifested in several different ways. Non-native listeners with relatively longer lengths of residency may be better at recognizing words produced by talkers, or identifying the region of origin of talkers, from 1) only the familiar or local dialect region because they had a greater amount of exposure to that specific variety, 2) only unfamiliar (and nonstandard) dialect regions because they had much more time to encounter speakers from other regions of the country, or 3) both familiar and unfamiliar dialect regions. Furthermore, on the indexical processing tasks, the benefit of living in

the U.S. for a longer period of time may be more pronounced in the forced-choice categorization task since the non-native listeners would have had more opportunities to acquire specific knowledge of the category labels and regional dialect boundaries used in the task.

Beyond length of residency, individual listeners were expected to vary greatly in performance on all three tasks. If greater knowledge of regional dialect variation in American English benefits listeners in both recognizing speech produced by talkers from different dialect regions and identifying different regional dialects, listeners should show a general tendency to perform well or poorly on all tasks (i.e., the sentence recognition task, the dialect categorization task, and the free classification task). The relation between speech recognition and dialect identification may also be dialect-specific and not generalizable to all dialects, i.e., performance on the different tasks will be related only for specific dialects and not across dialects. However, performance may also reflect access and use of a more general perceptual similarity space for regional dialects in long-term memory, or the perceived similarity of the regional dialects, and the tasks will show broader relations with each other across regional dialects. In this case, performance on a single dialect would not only be related across tasks, but would also be related to other dialects across tasks.

Methods

Listeners

Fifty-nine young adults participated in the current study. Of these 59, 21 participants were native speakers of American English (native American English group = NAE), 23 participants were native speakers of Korean who had lived in the United States for approximately three years or more (long-term Korean group = KLT), and 15 participants were native speakers

of Korean who had lived in the United States for approximately one year or less (short-term Korean group = KST). The 21 NAE participants included 15 females and 6 males, aged 18-27 years ($M = 21.0$; $SD = 1.8$ years). All NAE participants grew up in a region of the United States where General American English is spoken, i.e., North Midland, West, and parts of New England. The 23 KLT participants included 8 females and 15 males, aged 19-32 years ($M = 24.0$; $SD = 3.9$ years). All KLT participants were native speakers of Korean and had spent 3 years or more in the United States (LOR, $M = 5.1$; $SD = 1.5$ years). The 15 KST participants included 9 females and 6 males, aged 19-32 years ($M = 25.3$; $SD = 4.3$ years). All KST participants were native speakers of Korean and had spent a total of approximately 0.5 to 1.5 years in the United States (LOR, $M = 0.8$; $SD = 0.4$ years).

All participants were self-reported normal-hearing with no significant history of hearing or speech disorders at the time of testing. They received \$20 for 120 minutes of participation (rate of \$5 per 30 minutes), which included completing several other speech perception and production tasks, neurocognitive tasks, and self-report questionnaires.

Materials and Procedures

Procedures are identical for each of the three participant groups except where explicitly mentioned. Participants were seated in front of a personal computer equipped with headphones (BeyerDynamic DT100), a keyboard, and a mouse. Participants were tested individually in a quiet testing room, where they completed all of the experimental tasks. The output level of the target sentences for all the perception tasks were calibrated to be approximately 65 dB SPL. Procedures and materials for all three tasks are described below.

Sentence Recognition Task

Multi-Dialect Sentence Recognition: Participants completed a multi-dialect sentence recognition task. Six talkers (3F/3M) from each of four dialect regions (General American, North, South, and Mid-Atlantic) were selected from the NSP corpus (Clopper & Pisoni, 2006), for a total of 24 talkers. Materials consisted of 4 unique high-probability (HP) sentences per talker, for a total of 96 sentences. All HP sentences were from the Speech in Noise (SPIN) test (Kalikow, Stevens, & Elliott, 1977). Two randomly selected sentences per talker were mixed with 6-talker multi-talker babble at +5 signal-to-noise ratio (SNR), and two randomly selected sentences per talker were mixed with babble at +2 dB SNR.

On each trial, participants were presented with one sentence produced by a single talker over the headphones. They were asked to type the words they recognized into a dialog box on the computer screen. All subjects were encouraged to give partial answers or to guess. The experiment was self-paced and participants proceeded to the new trial by either pressing the return button or clicking on an on-screen button. Subjects did not receive any feedback about the accuracy of their response and they only listened to each sentence once. Sentences were pseudo-randomly presented, where a talker could appear once in each of four blocks of the task, to attempt to reduce the likelihood that the same talker would be appear twice within a few trials. The participants were presented with the two blocks of sentences at the easier (+5 dB) SNR first and then with the two blocks of sentences at the more difficult (+2 dB) SNR second. Within each block of trials, sentences were randomly presented. Responses were scored by the number of content words correct, which consisted of all nouns, verbs, adjectives, and adverbs. Incorrect morphological endings were considered to be incorrect, but homophones were counted as correct.

Indexical Tasks

Regional Dialect Categorization: Participants completed a four-alternative forced-choice regional dialect categorization task. This task assessed a participant's ability to perceive and use dialect-specific information encoded in the speech signal to identify the region of origin of unfamiliar talkers. The task was based on the experimental methodology developed by Clopper and Pisoni (2004b). Unlike previous studies, however, the categorization task in the current study used only four regional dialects, based on the findings from previous studies (e.g., Clopper, 2004), which showed that native American English participants, largely from the upper Midwest region, can reliably use four broad categories of American English regional dialect categories: General American, North, South, and Mid-Atlantic. Four talkers (2F/2M) from each of those four dialect regions were selected from the Nationwide Speech Project Corpus (NSP, Clopper & Pisoni, 2006), for a total of 16 talkers. Three unique low-probability SPIN (Kalikow et al., 1977) sentences were used for each talker, for a total of 48 sentences. No sentence was ever repeated in the categorization task.

On each trial, participants heard a single sentence produced by one talker over their headphones. They were asked to identify the region of the origin of the talker using a closed set of response alternatives, which consisted of four color-coded regions (General American, North, South, and Mid-Atlantic) represented on a map of the United States. The map was displayed on a computer monitor and participants responded by clicking on a labeled box located by or within a particular region, which also shared the same color as the region. The map and four response alternatives used in the current study are shown in Figure 1. The experiment was self-paced and participants could take as long as they wanted to respond. The next trial began after they responded to the previous trial. Participants did not receive any feedback about the accuracy of

their categorization response, and they were permitted to listen to each sentence only once before responding. Sentences were pseudo-randomly presented; a talker could appear once in each third of the task, to attempt to reduce the likelihood that the same talker would be appear twice within a few trials. Thus, the task had three blocks, and the participant could take a break between each block. All sentences in the dialect categorization task were presented in quiet. Responses were coded for the dialect region selected and overall accuracy. Both incorrect and correct responses were analyzed.

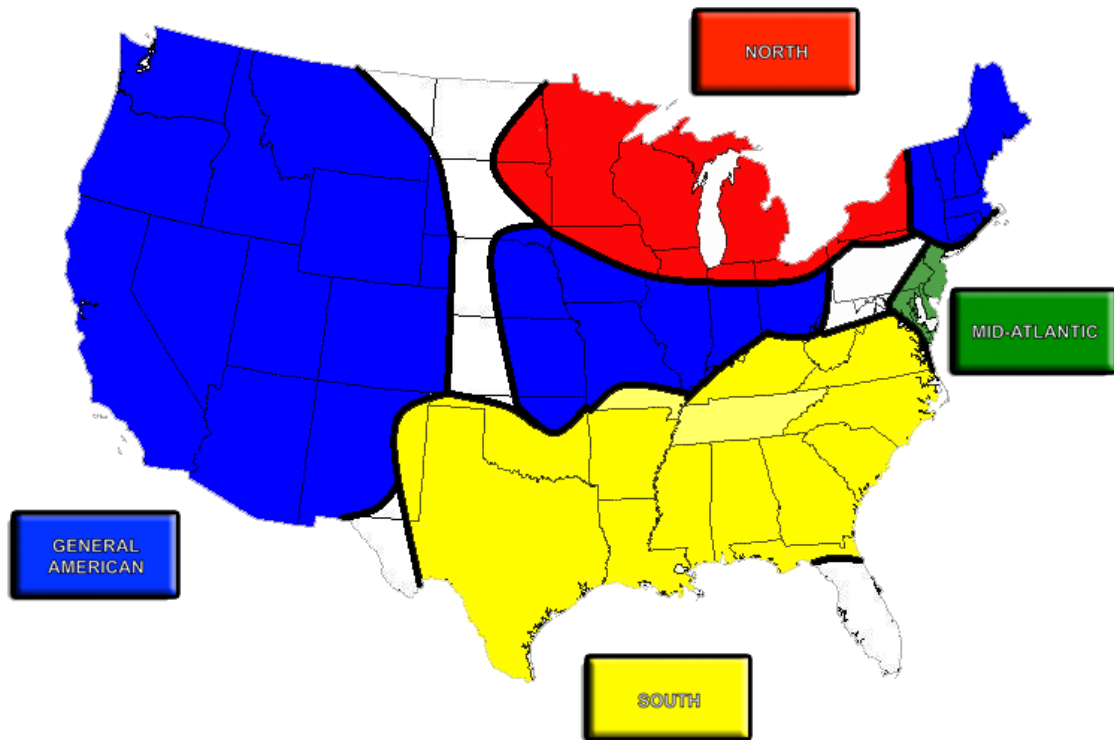


Figure 1. Regional dialect response map for the four response alternatives in the regional dialect categorization task.

Regional Dialect Classification: Participants also completed a regional dialect free classification task. This task assessed a participant’s ability to discriminate dialect-specific information in the acoustic signal. The design of the task was based on the methodology

developed by Clopper and Pisoni (2007). Four talkers (2F/2M) from each of four dialect regions (General American, North, South, and Mid-Atlantic) were selected from the Nationwide Speech Project Corpus (NSP, Clopper & Pisoni, 2006), for a total of 16 talkers. Three low-probability (LP) sentences from the SPIN test (Kalikow et al., 1977) were selected for each talker, for a total of 48 sentences. No sentence was repeated in the task and none of the talkers and sentences were used in the previous forced-choice categorization task.

In this task, digital movies were created by linking the individual sound files to still images with iMovie. Each movie had a unique image consisting of a white number, which corresponded to a number assigned to the particular sound file, displayed on a black background. The finished movies were saved as individual movie files. The movies appeared on the left side of a single PowerPoint slide in the normal edit mode. The listeners were told that each small box on the left side of the screen represented a talker and that they were to listen to each of the different talkers and drag the icons onto a grid on the right side of the screen based on the talker's region of origin, placing the images of talkers who shared a regional accent in the same group. The listeners were told that they could play the movies and rearrange the images of talkers as many times as they wanted. Additionally, they could make as many groups with as many talkers in each group as they wanted.

Each listener categorized all sixteen talkers three times, once in each of three different blocks. Sentences for each talker were randomly assigned to one of the three blocks. Each block consisted of a single powerpoint slide with all the talkers producing different phrases. Therefore, each listener received three different slides, each containing all 16 talkers producing different phrases. The computer display for the first block of the task is shown in Figure 2.

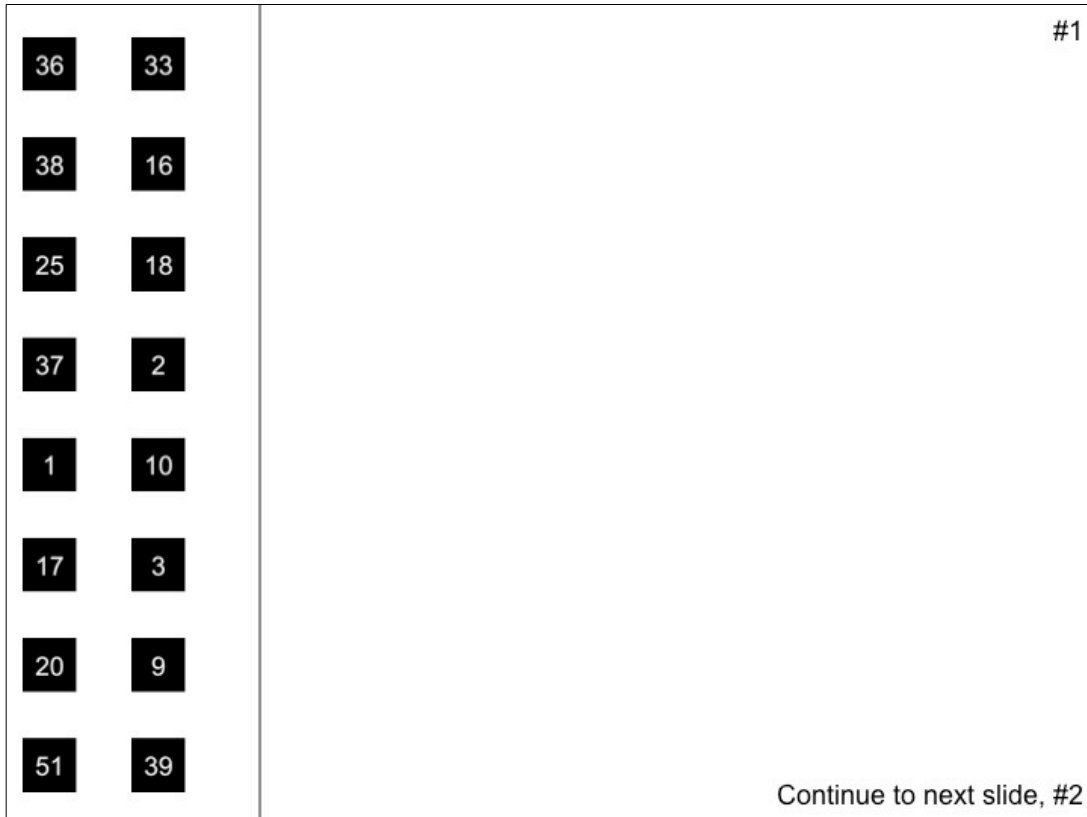


Figure 2. Computer display for the first block of the regional dialect classification task.

Data Analysis

Data analysis involved both group and individual differences analyses based on performance scores obtained for each task. Scores on the sentence recognition task included the listeners' percent correct content word accuracy scores. For the forced-choice categorization task, percent correct categorization accuracy and response patterns for each dialect were examined. Finally, for the free classification task, the mean number of groups formed, the mean number of talkers in each group, the percent of pairings that were correctly formed, and the percent of pairings that were incorrectly formed were used to analyze overall performance. Group analyses for each task included repeated-measures ANOVAs on performance scores using

the factors of: the talker's region of origin (General American, Northern, Mid-Atlantic, and Southern), listener group (AE, KLT, and KST), and SNR (+5 dB and +2 dB SNR) for the sentence recognition task. Additional analyses were also carried out to explore any significant main effects and interactions, as well as distributions in performance within and across non-native KLT and KST groups. Individual differences analyses investigated the relations among performance on all three tasks. Partial correlational analyses were carried out on performance scores across test conditions taking into consideration length of residency.

Results

Analyses of Group Differences: Native Language and Length of Residency

Sentence Recognition

Multi-Dialect Sentence Recognition: Mean content word recognition scores were calculated for all three listener groups for each of the four regional dialects at the two SNRs. Before analysis, accuracy scores were converted to rationalized arcsine units (rau) because performance was close to ceiling at +5 dB SNR for the native listeners (Studebaker, 1985). To facilitate interpretation of the findings, however, all figures and tables present the results in terms of accuracy scores. A summary of percent correct word recognition by the talker's region of origin and SNR is shown in Table 1. A series of statistical analyses revealed that word recognition accuracy differed significantly across the talker's region of origin, and by SNR. Additionally, although the overall accuracy was also significantly different for the listener groups, the effects of the talker's region of origin on word accuracy varied slightly across listener groups. Results of the statistical analyses are described below.

Table 1: Mean percent correct words for all talkers' regions of origin (General American (GA), North (NO), Mid-Atlantic (MA), South (SO)) at +5 dB and +2 dB SNR. Standard deviations are in parentheses.

Listener Group	+5 dB SNR				+2 dB SNR			
	GA	NO	MA	SO	GA	NO	MA	SO
NAE	84 (7)	93 (5)	91 (5)	90 (4)	86 (10)	87 (5)	68 (7)	87 (4)
KLT	49 (14)	55 (13)	46 (12)	46 (14)	30 (15)	51 (12)	35 (10)	32 (15)
KST	43 (20)	49 (17)	40 (22)	40 (21)	25 (15)	41 (17)	35 (12)	27 (16)

The initial statistical analysis of word recognition scores consisted of a repeated measures ANOVA with the talker's region of origin (General American, North, Mid-Atlantic, South) and SNR (+5 dB and +2 dB SNR) as within-subject factors and listener group (NAE, KLT, KST) as the between-subject factor. The analysis revealed significant main effects of the talker's region of origin [$F(3, 168) = 74.8, p < .001$], SNR [$F(1, 56) = 287.7, p < .001$], and listener group [$F(2, 56) = 115.2, p < .001$]. It also revealed significant interactions of the talker's region of origin x listener group [$F(6, 168) = 12.0, p < .001$], talker's region of origin x SNR interaction [$F(3, 168) = 25.3, p < .001$], and talker's region of origin x SNR x listener group interaction [$F(6, 168) = 25.3, p < .001$]. The SNR x listener group interaction was not significant.

The significant main effects and interactions were further explored through additional analyses. Although the main effect of listener group was significant, differences among listener groups were largely due to native language. Mean word recognition across all talkers' regions of origin and SNR conditions was 43.0% ($SD = 11.6\%$) for the KLT listener group, 37.6% ($SD = 16.8\%$) for the KST listener group, and 85.9% ($SD = 4.1\%$) for the NAE listener group. The difference between NAE listeners and both the KLT and KST listeners was significant (all $p < .001$), but no significant difference was observed between the two non-native groups. As for the

significant main effect of the talker's region of origin, a series of *t*-tests showed that Northern talkers were more intelligible than the other three dialects (General American, Mid-Atlantic, Southern) (all $p < .001$).

Overall word accuracy and accuracy for each regional dialect (across listener groups) varied by SNR. As expected, listeners were more accurate at the more favorable +5 dB SNR than at +2 dB SNR. To further explore the interaction of the talker's region of origin and SNR, paired comparison *t*-tests across the four regions at both SNRs revealed that Northern talkers were more intelligible at both SNRs (all $p < .001$). However, effects of the talker's region of origin varied by listener group. To explore both the significant two-way interaction of the talker's region of origin and listener group and the three-way interaction of the talker's region of origin with SNR and listener group, paired comparison *t*-tests were carried out across regional dialects at each SNR for each of the three listener groups. At +5 dB SNR, the NAE listeners were most accurate on Northern and Mid-Atlantic talkers, slightly less accurate on Southern talkers, and the least accurate on General American talkers. Northern talkers were significantly more intelligible than Southern talkers and General American talkers; Northern and Southern talkers were also significantly more intelligible than General American talkers (all $p < .003$). KLT listeners were significantly more accurate on Northern talkers than all other dialects (all $p < .001$), and KST listeners were significantly more accurate on Northern talkers than all other dialects ($p < .039$). At +2 dB SNR, for the NAE listeners, all talker dialects were significantly more intelligible than Mid-Atlantic talkers (all $p < .001$). None of the other dialect comparisons were significant. KLT listeners were more accurate on the Northern talkers than all other dialects (all $p < .001$), and more accurate on the Mid-Atlantic talkers than General American talkers (all $p = .009$). Similarly, KST listeners were more accurate on Northern talkers than all other dialects (all $p < .001$).

.011), and were more accurate on Mid-Atlantic talkers than both General American and Southern talkers (all $p < .001$). Each listeners group's word recognition accuracy scores for each of the four talker's region of origin at +5 dB SNR (top) and at +2 dB SNR (bottom) are displayed in Figure 3.

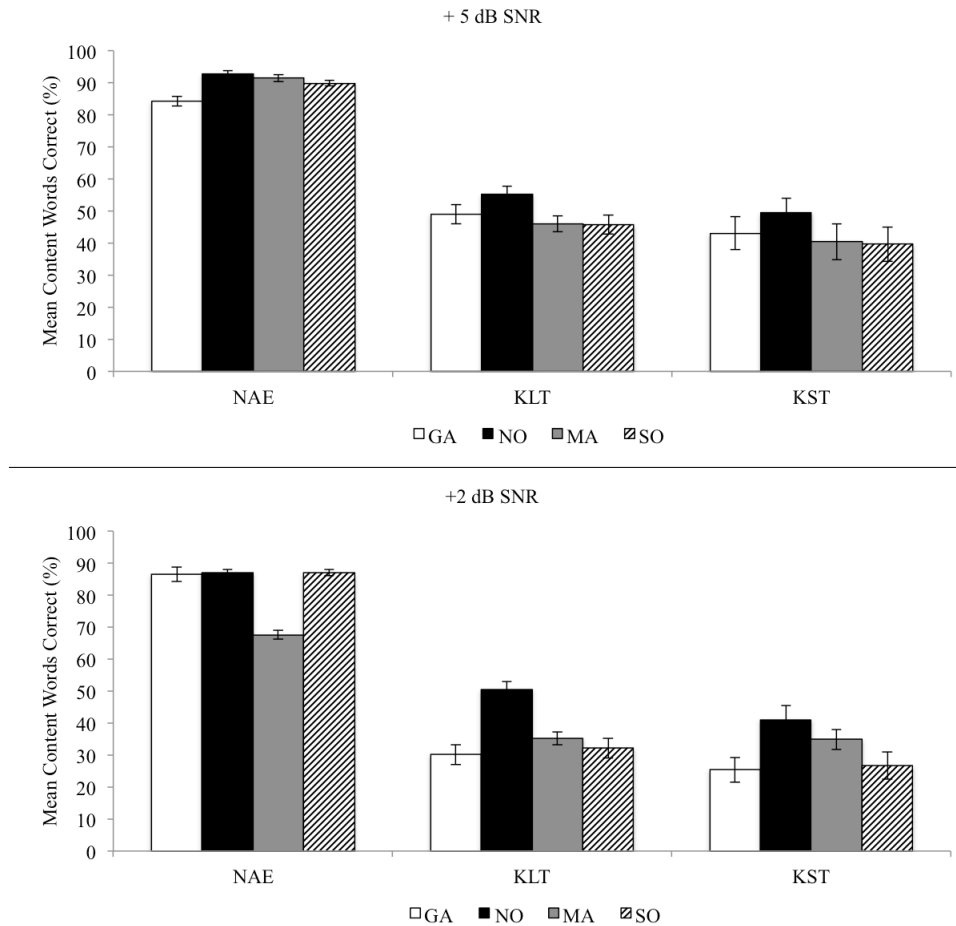


Figure 3: Mean percent correct words for talkers' regions of origin (General American (GA): empty bars, North (NO): black bars, Mid-Atlantic (MA): gray bars, South (SO): striped bars) at +5 dB (top) and +2 dB SNR (bottom).

Although the two non-native groups did not differ substantially on the sentence recognition tasks, the distribution of performance within both groups suggests that length of residency might be related to performance either through increased exposure to American

English dialects or increased proficiency. Figure 4 shows individual scores across talkers' regions of origin at both SNRs. As can be seen in Figure 4, the distribution of performance within the KLT and KST groups overlap substantially, although in a few conditions the KST group is more variable and several individual listeners performed below the range of the KLT listeners.

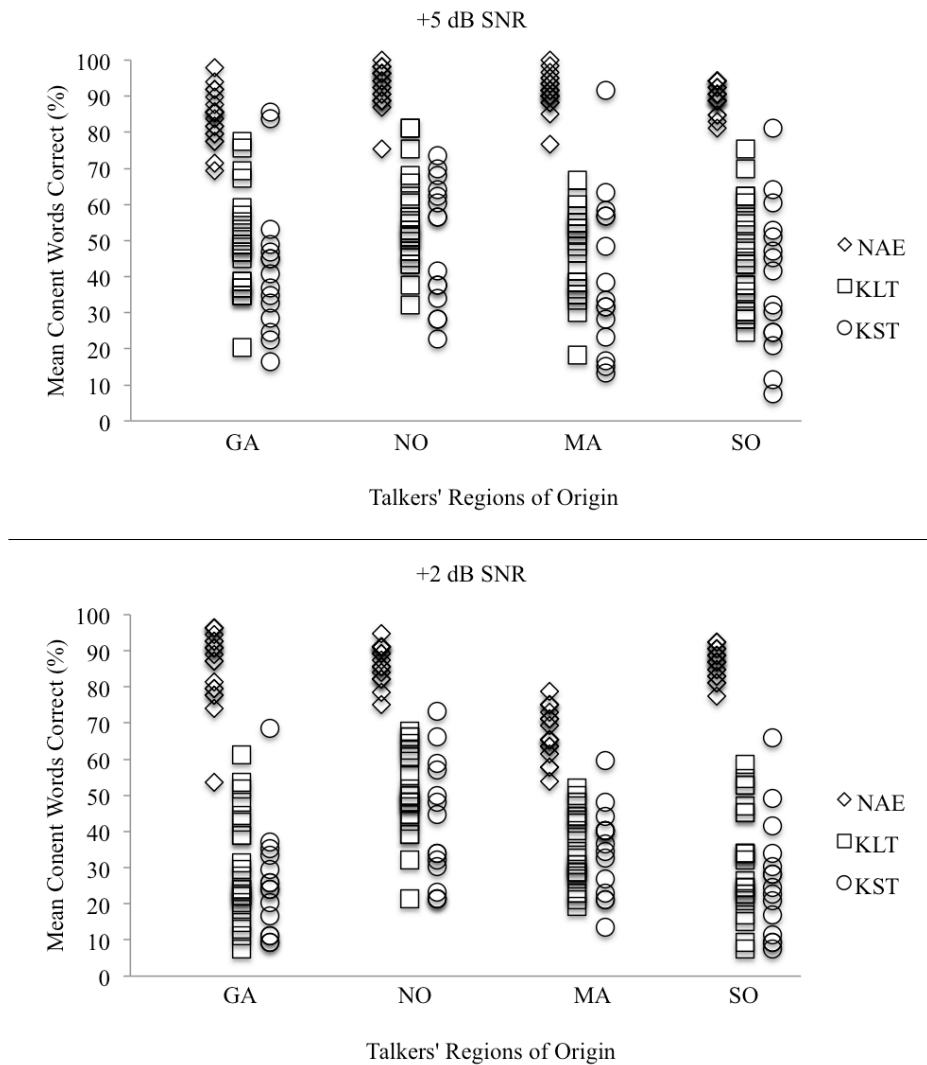


Figure 4: Mean percent correct words for each listener group (NAE: diamonds, KLT: squares, KST: circles) by the talkers' regions of origin (General American (GA), North (NO), Mid-Atlantic (MA), South (SO)) at +5 dB (top) and +2 dB SNR (bottom).

In order to investigate the effects of length of residency in further detail, correlational analyses were carried out between length of residency in the U.S. and word recognition scores on the sentence recognition task. Given that the distributions of listeners' performance seemed to vary by both the talker's region of origin and SNR, accuracy for each of the talkers' regions of origin at each SNR were examined. Length of residency in the U.S. was significantly correlated with word recognition performance across all conditions [$r = .37, p = .021$]. At +5 dB SNR, length of residency was correlated with word recognition for the Northern and Southern talkers and overall [$r = .35$ to $.38, \text{all } p \leq .029, \text{two-tailed}$], but it was not correlated with performance on General American and Mid-Atlantic talkers. At +2 dB SNR, length of residency was correlated with word recognition for the General American, Northern, and Southern talkers and overall [$r = .37$ to $.44, \text{all } p \leq .028, \text{two-tailed}$], but it was not significantly correlated with performance on the Mid-Atlantic talkers.

Indexical Tasks

Regional Dialect Categorization: Mean dialect categorization accuracy was calculated for each of the three listener groups for the four regional dialects used in the study (General American, North, Mid-Atlantic, South). Mean dialect categorization accuracy was 30.5% ($SD = 6.9\%$) for the NAE group, and 25.5% ($SD = 6.2\%$) for KLT and 29.3% ($SD = 3.6\%$) for KST non-native groups. A summary of the categorization results by regional dialect for each listener group is displayed in Table 2. Overall, performance was quite poor for all listener groups, with categorization accuracy on each dialect close to chance (25%). A repeated measures ANOVA with the talker's region of origin as the within-subject factor and listener group as the between-subjects factor revealed significant main effects of the talker's region of origin [$F(3, 168) = 58.5,$

$p < .001$] and listener group [$F(2, 56) = 4.1, p = .022$]. The interaction of the talker's region of origin and listener group was not significant. Listeners were more accurate at categorizing General American talkers than all other regional dialects (all $p < .001$), and least accurate at categorizing Southern talkers (all $p \leq .011$). NAE listeners were overall more accurate than KLT listeners [$t(42) = 2.5, p = .017$], but not significantly more accurate than KST listeners. KST listeners were also more accurate than KLT listeners [$t(36) = 2.1, p = .040$]. Because we did not expect that the KST listeners would be more accurate than KLT listeners, the relationship between length of residency in the U.S. and dialect categorization accuracy was further explored in a series of correlational analyses. Length of residency was not significantly correlated with mean regional dialect categorization performance.

Table 2: Mean percent categorization responses by regional dialect (Response Dialect) for all talkers' regions of origin (Stimulus Dialect). The four Stimulus and Response Dialects are General American (GA), North (NO), Mid-Atlantic (MA), and South (SO). Percent correct responses are in bold.

Listener Group	Stimulus Dialect	Response Dialect			
		GA	NO	MA	SO
NAE	GA	63.1	16.7	13.1	7.1
	NO	42.1	22.6	21.4	13.9
	MA	51.2	17.9	24.6	6.3
	SO	64.7	16.7	7.1	11.5
KLT	GA	47.5	21.4	16.3	14.9
	NO	26.1	21.4	15.2	37.3
	MA	48.2	19.2	16.7	15.9
	SO	52.5	15.2	15.6	16.7
KST	GA	50.0	14.4	17.8	17.8
	NO	26.7	26.7	17.2	29.4
	MA	46.7	10.0	24.4	17.8
	SO	55.0	12.8	16.1	16.1

Although performance was quite low overall, patterns in the categorization responses suggest that listeners were systematic in their responses and were clearly biased to select the

General American dialect. To explore error patterns produced by each listener group, a regional dialect confusion matrix was created for each listener group. Similarity and bias parameters were obtained using the full Similarity Choice Model (SCM; Nosofsky, 1985). The similarity parameters indicate the perceptual similarity among dialects, and the bias parameters reflect the listeners' response biases. These two parameters, taken together, can be used to model the categorization data to reveal the underlying similarity spaces of the four regional dialects (e.g., Smith, 1980; Nosofsky, 1985). In addition to the full SCM, in which both similarity and bias parameters are allowed to vary, a restricted SCM, in which the similarity parameters are held constant across listener groups and only the bias parameters are free to vary, was also carried out. Comparing the fit of the two models provides a way to examine the differences in the perceptual similarity spaces of the listener groups. If the full model did not fit the data significantly better than the restricted model, it would suggest that the perceptual similarity spaces of the regional dialects were the same for the three listener groups. In the current data, the results of the full SCM produced a better fit to the data than the restricted model, suggesting that the structure of the perceptual similarity spaces of the regional dialects was different across listener groups.

To further examine the perceptual similarity of the regional dialects, the resulting similarity parameters of the SCM for each listener group were also submitted to an ADDTREE additive clustering analysis (Sattath & Tversky, 1977; Corter, 1982). The resulting trees for each listener group are displayed in Figure 5. In this model, perceptual dissimilarity is represented by the horizontal distance between nodes, and the dissimilarity between any two dialects is the sum of the lengths of the fewest number of horizontal branches connecting them. Vertical distances are irrelevant.

The NAE listeners had two broad dialect clusters: General American and South; and North and Mid-Atlantic. Mid-Atlantic talkers were most distinct for the NAE listeners, and General American and South were perceived to be similar. The two non-native KLT and KST listener groups had similar perceptual similarity spaces. Both had two broad dialect clusters: General American and Mid-Atlantic; and North and South. Both groups perceived General American and Mid-Atlantic talkers to be similar, as well as Northern and Southern talkers. Additionally, for both groups, Northern talkers were perceptually most distinct and were farthest from the root.

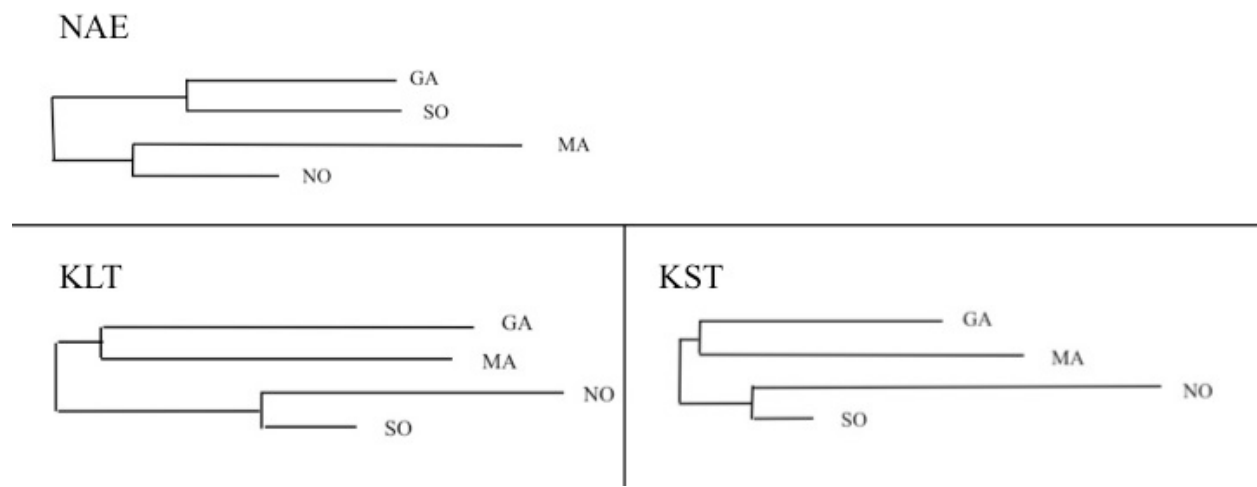


Figure 5. Clustering solutions for NAE, KLT, and KST listener groups for General American (GA), Northern (NO), Mid-Atlantic (MA), and Southern (SO) talkers.

The resulting bias parameters from the full SCM analysis are provided in Table 3. As can be seen in Table 3, the bias parameters revealed a positive bias towards responding with the General American dialect by all three listener groups. In addition to a strong positive bias to respond with the General American category, the NAE listeners also showed a negative bias towards responding with the Southern dialect. The response biases for the other dialects are close to the .25 unbiased response rate.

Table 3: Bias parameters for the four response dialects (General American (GA), North (NO), Mid-Atlantic (MA), and South (SO)) based on the full SCM for each of the three listener groups. Bias parameters close to .25 (chance) are relatively unbiased.

Listener Group	GA	NO	MA	SO
NAE	.55	.17	.18	.10
KLT	.46	.19	.16	.19
KST	.45	.18	.20	.17

Regional Dialect Classification: The mean number of groups formed, the mean number of talkers in each group, the percent of pairings that were correctly formed, and the percent of pairings that were incorrectly formed were calculated for each listener and each listener group. A summary of performance on the regional dialect free classification task is provided in Table 4. Overall, across listener groups, all listeners made on average four groups of talkers with four to five talkers per group. KLT and KST listeners made slightly fewer groups with more talkers in each group than NAE listeners.

Table 4. Classification performance for each listener group averaged across blocks.

Listener Group	Mean # Groups	Mean # Talkers/Group	% Consistent Pairings	% Inconsistent Pairings	Difference Score
NAE	4.5	4.0	30.6	28.5	2.1
KLT	3.8	4.9	30.0	32.0	-2.0
KST	3.8	4.5	28.1	29.2	-1.0

Three different measures were analyzed to assess classification strategies and performance (see, for example, Clopper and Bradlow, 2008). The percent of consistent pairings was calculated by dividing the number of times two talkers from the same dialect region were grouped together (a correct pair) by the total number of possible correct pairs. Overall, the listeners accurately grouped two talkers from the same dialect region in the same group at rate of

about 30% across all three listener groups. The percent inconsistent pairings score was calculated by dividing the number of times listeners grouped two talkers together from different dialect regions by the total number of possible pairings. Listeners also grouped two talkers from different regions at a rate of approximately 30%, as well. A difference score, calculated by subtracting the percent inconsistent score from the percent consistent pairings, was used as an indicator of overall accuracy. Given that the average percent consistent pairings and the percent inconsistent pairings score were nearly the same, the difference score was near 0 for all three groups, and no difference score was significantly different from 0, indicating poor performance on the task.

A series of one-way ANOVAs were carried out to examine the influence of native language on classification behavior. No significant effects of listener group were found for the number of groups formed, number of talkers per group, percent consistent pairings, percent inconsistent pairings, or difference scores. Additional ANOVAs on percent consistent pairings, percent inconsistent pairings, and difference scores for each of the four regions of origin also did not reveal any significant effect of listener group. Correlational analyses were also carried out between length of residency and performance scores to further examine possible influence of exposure in the U.S. No significant relations were found between length of residency and number of groups formed, number of talkers per group, consistent pairings, inconsistent pairings, or difference scores.

Clustering analyses were also carried out on the free classification data using an ADDTREE additive clustering analysis (Sattath & Tversky, 1977; Corter, 1982). The resulting trees for each of the three listener groups are provided in Figure 6. Perceptual dissimilarity is represented by the horizontal distance between nodes, and the dissimilarity between any two

dialects in the sum of the lengths of the fewest number of horizontal branches connecting them. Vertical distances are irrelevant. The NAE listeners had one dialect cluster consisting of General American and South, and two single branches for Mid-Atlantic and North. Mid-Atlantic talkers and Southern talkers were farthest from the root, and were most distinct for the NAE listeners. The non-native KLT and KST listener groups had slightly different perceptual similarity structures, along with differences in the magnitude of the dissimilarity between the dialects. The structure of the perceptual similarity space was similar for the two non-native listener groups. Both non-native groups had a cluster consisting of the General American and North. Thus, Northern and General American were perceived to be similar. Northern talkers were the also the most perceptually distinct for both groups, because the North was the farthest from the root. While the KST also had a cluster consisting of the Mid-Atlantic and South, the KLT group had two single branches for those dialects.

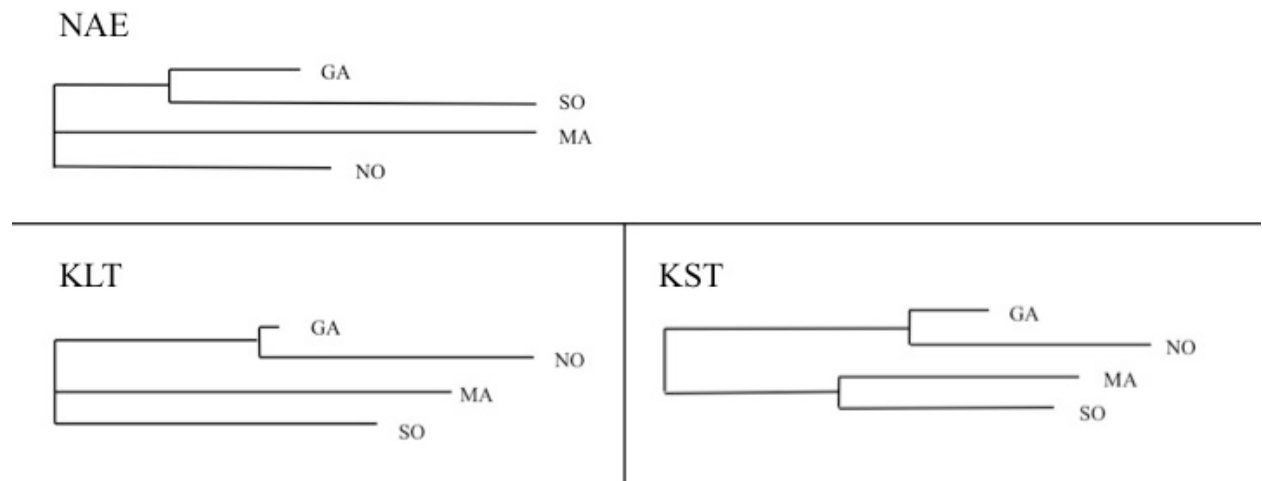


Figure 6. Clustering solutions for the NAE, KLT, and KST listener groups for General American (GA), Northern (NO), Mid-Atlantic (MA), and Southern (SO) talkers in the free classification task.

Analyses of Individual Differences: Perception of Linguistic and Indexical Information

To explore the relationship between the processing of linguistic and indexical information for individual non-native listeners, a series of correlational analyses were carried out to investigate the relations between performance on the three experimental tasks. First, the relations between an individual listener's ability to recognize speech produced by talkers from multiple regions and his/her ability to perceive and use dialect-specific information in the speech signal to identify or group talkers by region or origin was examined. Given that dialect categorization and free classification performance seemed to differ most between the standard (General American) and nonstandard (Northern, Mid-Atlantic, Southern), correlations were carried out on performance across all four regional dialects. For the NAE listeners, correlational analyses revealed no relation between overall word recognition scores from the multi-dialect sentence recognition task and regional dialect categorization accuracy.

For the analyses on non-native performance, because earlier analyses suggested that length of residency had little effect on performance in any of the three experimental tasks, the two non-native groups were combined. Instead, partial correlations taking length of residency into consideration between word recognition accuracy (words correct) on the sentence recognition task and the dialect categorization accuracy were carried out. Analyses revealed that categorization accuracy for General American talkers was positively correlated with overall word recognition accuracy [$r = .41, p = .013$], and with word recognition at both SNRs (+5 dB SNR [$r = .36, p = .028$]; +2 dB SNR [$r = .46, p = .005$]). Categorization accuracy for Southern talkers was also negatively correlated with overall word recognition accuracy [$r = -.40, p = .015$], as well as with word recognition accuracy at both SNRs (+5 dB SNR [$r = -.37, p = .026$]; +2 dB

SNR [$r = -.42, p = .009$]). Categorization accuracy for Mid-Atlantic and Northern talkers was not related to word recognition accuracy.

The relation between word recognition accuracy and free classification performance was also analyzed. For the native listeners, correlational analyses revealed no relations between word recognition and classification performance on the regional dialect free classification task. Partial correlations, taking length of residency into consideration, between overall word recognition accuracy on the sentence recognition task and the scores on free classification task were conducted for all non-native listeners. None of the free classification scores (number of groups formed, number of talkers per group, percent consistent pairings, percent inconsistent pairings, and difference scores) were found to be related to overall word recognition performance.

Analyses of Individual Differences: Perception of Indexical Information Across Tasks

Finally, correlational analyses were also carried out between scores on the dialect categorization task and the free classification task. Again, given that the categorization performance varied greatly by the region of origin of the talker, each dialect was examined separately and compared to overall trends on the free classification task. For native listeners, no significant relations emerged for categorization accuracy for the General American, Mid-Atlantic Southern and Northern dialects. Dialect categorization accuracy on the Northern dialect was only marginally correlated with the number of talkers per group [$r = .43, p = .044$]. Thus, listeners who were more likely to respond correctly with the Northern category were also more likely to make groups with more talkers in the free classification task.

For non-native listeners, partial correlations accounting for length of residency were carried out between performance on the two tasks. Dialect categorization accuracy for the

General American talkers was significantly correlated with the overall mean percent consistent pairings [$r = .48, p = .002$] and the overall mean percent inconsistent pairs [$r = .53, p = .001$]. Thus, listeners who were more accurate on categorizing General American talkers tended to make larger groups in the free classification task resulting in more consistent pairings but also more inconsistent pairings. Categorization accuracy for the Mid-Atlantic talkers was negatively correlated with the overall mean percent of inconsistent pairings [$r = -.38, p = .022$]. Thus, non-native listeners who were more accurate at categorizing Mid-Atlantic talkers tended to make fewer inconsistent pairings. Dialect categorization accuracy for the Northern and Southern dialects were not significantly correlated with free classification performance.

Discussion

The current study was designed to investigate the effects of regional dialect on non-native speech recognition, as well as the perceptual categorization and classification of regional dialects by non-native listeners. Furthermore, group (length of residency) and individual differences among non-native listeners in the processing of linguistic and indexical information in speech based on regional dialect were explored. To assess the effects of the talker's region of origin on non-native speech recognition, native speakers of Korean who had lived either a long (KLT) or short (KST) time in the U.S. completed a sentence recognition task at two SNRs (+5 dB and +2 dB SNR) using sentences selected from four U.S. regional dialects (General American, Northern, Mid-Atlantic, and Southern). Compared to the native (NAE) listeners, both non-native listener groups (KLT and KST) displayed difficulty at recognizing the words in the sentences across all talkers' regions of origin. The large differences observed in performance between the native and non-native groups suggest that the variability from regional dialect is particularly challenging for

non-native listeners. These conditions require the listener to rapidly adjust and adapt to highly variable test sentences from trial to trial, which is also a challenge to native listeners (e.g., Gilbert, Tamati, & Pisoni, 2013; Mullennix, Pisoni, & Martin, 1989; Peters, 1955) but especially to non-native listeners (e.g., Cooke, García Lecumberri, & Barker, 2008; Tamati & Pisoni, in press).

As expected, all listener groups performed more poorly at the less favorable SNR (+2 dB SNR), but no interaction between SNR and listener group was found, suggesting that differences between native and non-native performance was similar across SNRs, and non-native listeners were not disproportionately worse than native listeners at the less favorable SNR condition. Although many previous studies have found that non-native speech recognition is worse at less favorable SNRs when substantial linguistic and context information is available (e.g. Buus, Florentine, Scharf, & Canevet, 1986; Florentine, Buus, Scharf, & Canevet, 1984; Florentine, 1985; van Wijngaarden, Steeneken, & Houtgast, 2002), others studies have reported that when the use of linguistic or context information is limited or blocked, the effect of background noise is similar for native and non-native listeners (e.g., Cutler, Weber, Smits, & Cooper, 2004; Flege & Liu, 2001; Rogers, Lister, Febo, Besing, & Abrams, 2006). Although the sentences used in the current study were high-predictability SPIN sentences, the effects of sentence context and predictability may have been minor because of the short sentences (5-8 words), the way the responses were scored (e.g., content words as opposed to final word), or the moderate SNRs.

As expected, word recognition accuracy by native and non-native groups also varied by the talker's region of origin. Across all three listener groups, Northern talkers were found to be the most intelligible, while no significant differences were found among the other three dialect regions of the talkers (General American, Northern, Southern). Although General American was

assumed to be the most representative of a standard, supraregional dialect, it was not always the most intelligible. Part of the observed results may be related to use of General American multi-talker babble. Previous studies have found that speech recognition is poorer when target and competing talkers have similar voice characteristics (e.g., Brungart, 2001; Brungart, Simpson, Ericson, & Scott, 2001). However, the finding that Northern talkers were most intelligible is consistent with several previous studies using participants from the upper Midwest, which have found that Northern talkers are often most intelligible, even more than General American talkers (e.g., Clopper, Pierrehumbert, & Tamati, 2010; Clopper, Tamati, & Pierrehumbert, under review). Since the non-native listeners in this study also found the Northern talkers to be most intelligible, this suggests that Northern speech, at least when tested in laboratory settings, is inherently more intelligible. Additional studies should be carried out to examine the intelligibility of multiple regional dialects under different listening conditions and background noise, and to identify the acoustic characteristics of highly intelligible speech under those conditions.

In comparing the differences in performance between native and non-native listeners, native listeners were found to be much more consistent across the talkers' regions of origins. One possible reason is that both the SNRs were moderate, and native listeners did not have much difficulty at either SNR. Since the influence of the talker's region of origin are more deleterious at less favorable SNRs (Clopper & Bradlow, 2008), the background multitalker babble was moderate enough that the native listeners were able to learn the talker, regional dialect, and context features to recognize words in the sentences. At the less favorable condition, where their ability to do so was slightly reduced, difficulties on the Mid-Atlantic dialect, which would have been the least familiar for the NAE listeners, emerged. While non-native KLT and KST listeners

also showed similar performance for the Northern, as well as the General American and the Southern dialects, compared to native listeners, the non-native and native listeners differed markedly on the Mid-Atlantic dialect. At the less favorable SNR, non-native listeners were also more accurate on the Mid-Atlantic dialect compared to the General American and Southern dialects, unlike the native listeners who had substantial difficulty with the Mid-Atlantic dialect. Given that the Mid-Atlantic regional dialect is the least represented on the Indiana University Bloomington campus, better performance on that dialect would not be expected. One potential reason for this pattern could be the lack of experience of the non-native listeners with talkers from all regions of origin. If the non-native listeners were not receiving any benefit from experience for any of the dialects, their performance would be more likely to reflect the general intelligibility of those dialects for non-native listeners (and native speakers of Korean) and have little relationship with the distribution of talkers from various regions in their L2 environment.

The non-native KLT and KST listeners' lack of experience with American English dialects was further revealed in the effects of length of residency on word recognition performance. Although previous studies have demonstrated that length of residency in the L2 environment plays an important role in acquisition (e.g., Oyama, 1976; Mackay, Flege, Piske, & Schirru, 2001; Mackay, Meador, & Flege, 2001; Ingvalson, McClelland, & Holt, 2011), the effects of length of residency are often mediated by other language factors, like educational history or amount of language use, which have also been shown to influence speech perception abilities (e.g., Oyama, 1976; Flege, Bohn, & Jang, 1997; Flege & MacKay, 2004; Ingvalson et al., 2011). Nevertheless, a longer length of residency would be expected to improve speech recognition when dealing with multiple sources of sociolinguistic variation, like regional dialect, which would require both knowledge of the L2 phonological system as well as knowledge of

regional dialect characteristics that could be gained predominately through direct exposure. In the current study, word recognition performance was very similar between the KLT and KST groups, overall and across dialects. Correlational analyses across conditions revealed that length of residency was related to performance on several dialect and SNR conditions. However, these relations were likely driven by the KLT group. Examining performance distributions (see Figure 4), the KLT group displayed much more variability in performance, suggesting that non-native learners may only be benefiting from regional dialect familiarity in speech recognition after a long period of exposure. If this is the case, a dialect familiarity benefit should first emerge for commonly encountered dialects, which in the case of Bloomington, IN, would be the Northern (Northern Indiana), General American (Central Indiana/Bloomington/supraregional), and Southern (Southern Indiana) regional dialects.

Findings from the sentence recognition task suggest that word recognition accuracy is influenced by an interaction of native language (native or non-native), experience (especially for native listeners), dialect intelligibility, and background competition. Non-native listeners had more difficulty overall, and performance likely reflected their lack of experience with the American English dialects across most of the KLT and KST groups. In order to better understand how previous experience influences the recognition of speech produced by talkers from multiple regions of origin in an L2, future studies should be conducted to further examine the effects of length of residency and other factors, like age, education, or amount of English use, on L2 speech recognition.

To assess the perceptual categorization of regional dialects by non-native listeners, the non-native and native listener groups also completed a four-alternative forced-choice categorization task using four U.S. regional dialects (General American, Northern, Mid-Atlantic,

and Southern). Mean performance across all four dialects was poor for all listener groups. However, performance behavior varied greatly by the talker's region of origin. In particular, listeners appeared to treat the supraregional General American dialect differently than the other three nonstandard dialects. Both native and non-native groups were accurate at categorizing General American talkers as coming from a General American region, but were much poorer at categorizing talkers from the other three regions (Northern, Mid-Atlantic, and Southern). Given the response patterns (see Table 2), it was not surprising that all listeners were found to be very biased to respond with the General American response category. The present findings suggest that the talkers used in the current study may not have been as regionally marked as talkers in previous studies using different speech corpora (see also, Clopper, 2004; Clopper & Bradlow, 2008). Although categorization accuracy in the current study was lower overall than previous studies examining the perceptual categorization of regional dialects (e.g., Clopper & Pisoni, 2004a; Clopper & Pisoni, 2004b), earlier studies using the NSP materials also showed poorer dialect categorization performance compared to studies using TIMIT materials (e.g., Clopper, 2004; Clopper & Bradlow, 2008).

However, the differences in categorization accuracy for different dialects also interacted with the listener background, suggesting that the response bias may have been at least partially due to the linguistic experience of the listeners. Although the NAE listeners were slightly more accurate than the non-native KLT and KST groups overall, they were more strongly biased to respond with the General American response, even compared with the non-native groups, and were worse than chance level for the Southern dialect talkers. The native listeners in the current study have been exposed to many talkers from the northern part of the Southern dialect region, or the South Midland. The acoustic characteristics of the speech from the talkers used in the current

study are more representative of this South Midland regional dialect, which is not as marked as the Southern dialect (Clopper & Pisoni, 2006). Because native listeners, and to a lesser extent non-native listeners, would have been familiar with talkers who share similar South Midland characteristics and would have been aware that their speech is not as marked as stereotypical Southern speech, they might have been less likely to categorize a Southern talker in the current study as Southern and more likely to categorize them as General American. Follow-up questionnaires also provide converging support for this interpretation, since both native and non-native listeners often reported hearing General American that, for example, sounded like family members from Southern Indiana, but did not hear any stereotypical Southern talkers. Similarly, as reported above, word recognition scores for the General American and Southern talkers in the sentence recognition task were also very similar, demonstrating consistency across experimental tasks.

The KLT and KST groups performed similarly in the forced-choice categorization task. They showed similar response patterns and perceptual similarity spaces. While both non-native groups, like the native listeners, were most accurate on the General American talkers and poor on the Northern, Mid-Atlantic, and Southern talkers, non-native listeners were slightly less biased than native listeners to respond with the General American response option and slightly more likely to respond with the Southern response option. The KLT and KST groups' perceptual similarity spaces were also similar. Non-native listeners found General American and Mid-Atlantic talkers to be similar, and Northern and Southern talkers to be similar. Native listeners also had two broad dialect clusters, but they perceived General American and Southern talkers to be more similar to each other, and Northern and Mid-Atlantic talkers to be more similar to each other. They also perceived Mid-Atlantic talkers to be most distinct.

The response patterns and similarity spaces suggest that the native and non-native groups may have been basing their judgments on slightly different criteria in this task. For NAE listeners, the categorization responses were consistent with the word recognition data. Mid-Atlantic talkers were least intelligible in the sentence recognition task and most perceptually salient in the categorization task, and General American and Southern talkers were similarly intelligible and perceptually similar. These patterns suggest that native listeners may have been basing their judgments on perceptual saliency as well as previous knowledge, as discussed above. The non-native listeners would not have had a great amount of knowledge of American English regional dialects upon which to base their decisions. Instead, they may have been basing their judgments on features related to other talker characteristics, like intelligibility, attractiveness, etc. or they may be more heavily weighting their judgements on acoustic-phonetic features that are crucial for the perception of native-language dialectal differences or that are more easily discriminated because of their native language sound inventory. One interpretation based on the influence of the listeners' native language is that the Northern and Southern cluster involves dialects with vowel shifts and the other cluster, with General American and Mid-Atlantic dialects, involves unshifted vowels and merged low-back vowels. Features distinguishing the two dialects within each cluster involve either consonantal features that may be difficult to perceive for the Korean listeners (e.g., r-lessness or intrusive *r*'s in Southern and Mid-Atlantic dialects) and slight differences in the realization of a single vowels (e.g., /u/-fronting) rather than entire chain shifts that may leave a more general impression of being strange or different. Additional studies should be carried out to investigate the sensitivity of non-native listeners to specific dialect-specific acoustic-phonetic features given their native language and ability to discriminate or identify difficult American English speech sounds.

Listeners also completed a regional dialect free classification task in which participants heard talkers from the same four dialect regions as the other tasks in the current study and were required to make groups based on the talkers' regions of origin. This task gives more flexibility to the listener because he/she can repeatedly listen to talkers as many times as they want, form as many groups as they want, and have as many talkers in each group as they want. Furthermore, this task does not impose any a priori category structure or labels that might bias the listeners or place non-native listeners at a greater disadvantage given their lack of familiarity with U.S. geography or regional dialect category labels or names. As expected, given these task design features, even fewer differences between groups were found on the free classification task compared to the forced-choice categorization task, although it is difficult to directly compare performance on the two tasks. Overall performance was poor for all three listener groups, as indicated by the nearly equal ratio of percent consistent pairings to percent inconsistent scores. Native listeners made slightly more consistent pairings and slightly fewer erroneous pairings, but differences were not significant. Additionally, no significant effects of length of residency were found for the free classification measures.

Comparing across the indexical perception tasks, native and non-native listener groups again had slightly different perceptual similarity structures for regional dialects in the free classification task. Additionally, the similarity structures obtained from the classification data differed from the structures for the categorization data, especially for non-native listeners. In this task, non-native listeners perceived General American and Northern talkers as similar, and Northern talkers were perceived as being the most distinct. Thus, the clusters obtained in the classification task involved different regional dialects than the categorization task, suggesting that the non-native listeners perceived the regional dialects along different dimensions, relying

on different features to make their judgments in the two tasks. In the free classification task, listeners may have substantial opportunity to compare some of the more difficult features to detect or discriminate, like spectral differences indicating r-lessness or r-fullness or the height/backness of some difficult vowels, as mentioned above. If this is the case, the non-native listeners may have relied more on unique marked features rather than general impressions from more global features (i.e., across many words in a phrase or sentence) like vowel shifts. The native NAE listeners' perceptual structure was more similar to that from the forced-choice categorization task. They perceived General American and Southern talkers to be similar, but Mid-Atlantic and Northern talkers as being more distinct from one another in the classification task. Both Mid-Atlantic and Southern talkers were perceived as being generally the most distinct, as in the categorization task where Mid-Atlantic talkers were perceived as most distinct. Thus, native listeners appeared to have used similar perceptual strategies in both indexical tasks. Future studies should be conducted to further investigate the factors that influence the perceived similarity of different regional dialects by native and non-native listeners across multiple tasks and with diverse talkers and speech materials.

The current study also sought to explore the relation between the processing of regional dialect variation in both the linguistic and the indexical tasks. For the non-native KLT and KST listeners, forced-choice categorization accuracy on General American talkers was positively correlated with overall word recognition accuracy and categorization accuracy on Southern talkers was negatively correlated with overall word recognition accuracy. Given previous studies, it was not expected that poorer categorization accuracy on any dialect would be related to better recognition performance. However, comparing native and non-native performance on the categorization task revealed that native NAE listeners were both more accurate on General

American talkers and less accurate on Southern talkers than non-native listeners, likely due to a combination of previous experience and knowledge of stereotypes leading them to group the Southern and General American talkers.

Performance on the free classification task was not correlated with word recognition accuracy, suggesting that individual listener strategy on the free classification task is not related to the recognition of speech across many regional dialects. Across both indexical tasks, no relation was found between categorization accuracy and classification strategies. Given that performance on the free classification task was poor and it was not correlated with other recognition or categorization measures, the task was likely not capturing the listeners' ability to perceive the dialect-specific acoustic-phonetic differences. Instead, as mentioned above, the non-native listeners may not have been basing their judgments entirely on the acoustic similarity of the dialects, but rather at least partially on some other factor, like the intelligibility of the talkers.

Taken together, the results of the correlational analyses suggest that non-native listeners who show more native-like categorization responses (i.e., more accurate General American categorization and less accurate Southern categorization) have better word recognition abilities across multiple dialects in multi-talker babble. The present findings provide further support for earlier results suggesting that the processing of linguistic and indexical information is closely linked, and that individual differences in the ability to recognize highly variable words and sentences is related to indexical processing skills (e.g., Nygaard et al., 2004; Nygaard & Pisoni, 2008; Cleary & Pisoni, 2002; Cleary & Pisoni, 2005; Tamati & Pisoni, in press). If this interpretation is correct, non-native speech recognition will benefit from developing or obtaining broad native-like knowledge of dialect similarities and differences, rather than benefiting from improved abilities on individual dialects from task to task. Future studies on the perception of

sociolinguistic variation in an L2 should be conducted to gain a better understanding of the relation between the acquisition of L2 linguistic and sociolinguistic knowledge.

Conclusions

The goal of the current study was to investigate the perception of regional dialect variation by non-native listeners with different lengths of residency in the U.S. The results obtained using a high-variability sentence recognition task revealed that non-native KLT and KST listeners had more difficulty overall compared to native NAE listeners, and their performance likely reflected an interaction of their lack of experience with the American English dialects and dialect intelligibility. For the two indexical tasks, performance was found to be generally poor for all three listener groups, native and non-native, reflecting both the talkers' less regionally marked speech and the perceptual biases of the listeners toward classifying all of the speakers as General American. Although native NAE performance was found to be relatively stable across the two tasks, the non-native KLT and KST listeners showed different patterns of responses on the two tasks, suggesting that they may have based their judgments on a different set of acoustic-phonetic features depending on the demands of the particular task. However, longer lengths of residency in the U.S. were not found to improve performance in either task, suggesting that non-native listeners had yet to acquire native-like knowledge of L2 regional dialect characteristics or social knowledge of relations among the dialects. Finally, individual patterns across these three tasks revealed that individual listeners who had developed more native-like perception of regional dialects in the indexical processing tasks displayed better speech recognition across multiple dialects, reflecting a benefit from familiarity not just to a single dialect but a benefit from the general perceptual organization of regional dialects.

The results of the current investigation suggest potential areas for future study on the development of the L2 phonological system and its interaction with sociolinguistic categories. Additional research on the perception of regional dialect variation in an L2, particularly on the effects of length of residency, proficiency, and experience in L2 speech recognition and indexical processing abilities, should be carried out using other experimental methods and other populations of non-native listeners.

References

- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*, *109*, 1101-1109.
- Brungart, D. S., Simpson, B. D., Ericson, M. A., Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America*, *110*, 2527- 2538.
- Buus, S., Florentine, M., Scharf, B., & Canevet, G. (1986). Native French listeners' perception of American-English in noise. *Proceedings of InterNoise '86*, Cambridge, USA, 895–898.
- Cleary, M. & Pisoni, D. B. (2002). Talker discrimination by prelingually deaf children with cochlear implants: preliminary results. *Annals of Otology, Rhinology & Laryngology*, *111*, 113-118.
- Cleary, M., Pisoni, D. P., & Kirk, K. I. (2005). Influence of voice similarity on talker discrimination in children with normal hearing and children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *48*, 204-223.
- Clopper, C. G. (2004). Linguistic experience and the perceptual classification of dialect variation. Indiana University Ph.D. dissertation.
- Clopper, C. G. (2007). Auditory free classification: Methods and analysis. *Behavior Research Methods*, *40*, 575-581.
- Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech*, *51*, 175-98.
- Clopper, C. G., Pierrehumbert, J. B., & Tamati, T. N. (2010). Lexical neighborhoods and phonological confusability in cross-dialect word recognition in noise. *Laboratory Phonology*, *1*, 65-92.

- Clopper, C. G. & Pisoni, D. B. (2004a). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change*, 16, 31-48.
- Clopper, C. G. & Pisoni, D. P. (2004b). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111-140.
- Clopper, C. G. & Pisoni, D. B. (2006). The Nationwide Speech Project: A new corpus of American English dialects. *Speech Communication*, 48, 633-644.
- Clopper, C. G. & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics*, 35, 421-438.
- Clopper, C. G., Tamati, T. N., & Pierrehumbert, J. B. (under review). Typicality and linguistic relevance in the encoding of lexical information. Submitted to *Laboratory Phonology*.
- Cooke, M., García Lecumberri, M. L., Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *Journal of the Acoustical Society of America*, 123, 414-427.
- Corter, J. E. (1982). ADDTREE/P: A PASCAL program for fitting additive trees based on Sattath and Tversky's ADDTREE algorithm. *Behavior Research Methods and Instrumentation*, 14, 353-354.
- Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America*, 116, 3668-3678.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25, 437-470.

- Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23, 527–552.
- Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 24, 1–34.
- Florentine, M. (1985). Non-native listeners' perception of American-English in noise. *Proceedings of InterNoise '85*, Munich, 1021–1024.
- Florentine, M., Buus, S., Scharf, B., & Canevet, G. (1984). Speech reception thresholds in noise for native and non-native listeners. *Journal of the Acoustical Society of America*, 75, S84.
- Gilbert, J. L., Tamati, T. N., & Pisoni, D.B. (2013). Development, Reliability and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *Journal of the American Academy of Audiology*, 24, 1-11.
- Ingvalson, E. M., McClelland, J. L., & Holt, L. L. (2011). Predicting native English-like performance by native Japanese speakers. *Journal of Phonetics*, 39, 571-584.
- Kalikow, D. N., Stevens, K. N., Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61, 1337–1351.
- Levi, S. V., Winters, S. J., & Pisoni, D. B. (2011). Effects of cross-language voice training on speech perception: Whose familiar voices are more intelligible? *Journal of the Acoustical Society of American*, 130, 4053-4062.
- Mackay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America*, 110, 516–528.

- Mackay, I. R. A., Meador, D., & Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, 58, 103–125.
- Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, 85, 365-378.
- Nygaard, L. C. (2005). Perceptual Integration of Linguistic and Nonlinguistic Properties of Speech. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 390-413). Oxford, UK: Blackwell Publishing Ltd.
- Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific perceptual learning in speech perception. *Perception & Psychophysics*, 60, 355-376.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42-46.
- Oyama, S. (1976). A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research*, 5, 261-283.
- Peters, R. W. (1955). The relative intelligibility of single-voice and multiple-voice messages under various conditions of noise. In *Joint Project Report No. 56* (pp. 1-9). Pensacola, FL: US Naval School of Aviation Medicine.
- Rogers, C., Lister, J., Febo, D., Besing, J., & Abrams, H. (2006). Effects of bilingualism, noise and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics* 27, 465–485.
- Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika*, 42, 319-345.
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455-462.

Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Individual differences in speech perception in adverse listening conditions. *Journal of the American Academy of Audiology*, *24*, 616-634.

Tamati, T. N., & Pisoni, D. B. (in press). Non-native speech recognition in high-variability listening conditions using PRESTO. Submitted to the *Journal of the American Academy of Audiology*.

van Wijngaarden, S. J., Steeneken, H., Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America*, *111*, 1906–1916.

CHAPTER 3
SECOND LANGUAGE PHONOLOGICAL AND LEXICAL KNOWLEDGE
AND THE PERCEPTION OF REGIONAL DIALECT VARIATION

Introduction

Individual Differences in Speech Perception in a Second Language

Regional dialect is an important source of indexical variability in the speech signal that both native and non-native listeners frequently encounter in everyday listening environments. Listeners often communicate with talkers of diverse regional backgrounds, whose speech patterns may contain several dialect-specific features that could potentially present problems for listeners who are unfamiliar with those features. Adapting and adjusting to this source of variability in the signal may be particularly challenging to non-native listeners who have less experience with sociolinguistic variation in their second language (L2) and imperfect knowledge of the L2 phonological system compared to native listeners. Furthermore, individual differences in exposure to and knowledge of the L2 may influence listeners' ability to use regional dialect variation in communication. The findings reported in Chapter 2 showed that non-native listeners varied greatly in their perception of regional dialect variation in American English. In a multi-dialect sentence recognition task, non-native listeners were overall poor at recognizing speech produced by talkers from different regions of origin in the United States (U.S.) in multi-talker babble. Non-native listeners were worse at recognizing the content words in sentences across all regional dialects compared than native listeners, suggesting that variability from regional differences between talkers is challenging for non-native listeners. However, not all non-native listeners performed poorly and non-native listeners varied greatly in their ability to adapt to and

use dialect-specific information in the speech signal to recognize words in the high-variability sentences. Indeed, word recognition accuracy varied from 14.0% to 74.7% averaged over all conditions, demonstrating great individual variability in word recognition ability in these conditions. Additional analyses on the effects of previous exposure to American English also revealed that individual differences in word recognition ability were pervasive regardless of the listener's length of residency in the U.S.

Individual differences in the ability to adapt to and use dialect-specific information in the speech signal to identify a talker's region of origin by non-native listeners were also observed in Chapter 2. Dialect identification ability was assessed using a four-alternative forced-choice regional dialect categorization task that required listeners to perceive dialect-specific information in the speech signal and use stored knowledge of regional dialect variation to select an unfamiliar talker's region of origin from four regional dialect response alternatives: General American, Northern, Mid-Atlantic, and Southern. In the dialect categorization task, performance across all listener groups was poor, and listeners were biased to respond with the General American category. Differences among native and non-native listeners emerged in the perceptual similarity space, as native listeners perceived General American and Southern talkers as similar and Mid-Atlantic and Northern talkers as similar, while non-native listeners perceived General American and Mid-Atlantic talkers as similar and Northern and Southern talkers as similar. In addition, non-native performance was highly variable across the talker dialects. Non-native listeners' accuracy was 16.7-100% on General American talkers, 0.0-50.0% on Mid-Atlantic talkers, 0.0-50.0% on Northern talkers, and 0.0-50.0% on Southern talkers (12.5-37.5% overall), again demonstrating large individual differences in dialect identification ability. Length of residency in the U.S. also did not have a large effect on regional dialect categorization ability.

Taken together, the pattern of performance on the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task observed in Chapter 2 suggest that there were large individual differences in the perception of regional dialect variation in the participants' L2 and that while length of residency may have influenced performance, individual variability was not entirely attributable to that factor. Therefore, it is still unclear what mechanisms underlie the vast amounts of individual differences attested in those two perception tasks.

Sources of Individual Differences in L2 Speech Perception Abilities

A listener's unique developmental linguistic history influences his/her speech perception abilities across a variety of tasks. The language backgrounds of non-native listeners have consistently been shown to influence second language speech perception abilities. These factors include first language (e.g., García Lecumberri & Cooke, 2006; MacKay, Flege, Piske, & Schirru, 2001; MacKay, Meador, & Flege, 2001), age and length of exposure to L2 (e.g., Ingvalson, McClelland, & Holt, 2011; MacKay, Flege, Piske, & Schirru, 2001; MacKay, Meador, & Flege, 2001; Oyama, 1976), and frequency and amount of L2 use (e.g., Flege, Bohn, & Jang, 1997; Flege & MacKay, 2004; Ingvalson, McClelland, & Holt, 2011; Oyama, 1976). These group-related factors have also been found to affect a non-native listener's ability to recognize words in high-variability test materials, particularly foreign-accented speech (e.g., Bent & Bradlow, 2003; Imai, Walley, & Flege, 2005). Findings from these studies have generally shown that greater L2 phonological and lexical knowledge (from greater experience with or exposure to the L2 from a variety of sources) benefits the non-native listener in speech perception tasks in the L2.

However, as for Chapter 2 tasks, even when many group factors are taken into consideration, there still remains a great amount of individual variability in L2 perception abilities. Indeed, individual differences in L2 phonological or L2 vocabulary knowledge are pervasive among listeners who share a similar language background (e.g., Tamati & Pisoni, in press). Although these individual differences are widely documented, little is known about how individual listeners' perception of sociolinguistic variation in the L2 develops along with their linguistic knowledge of the L2.

Previous research on the effects of language background on the perception of indexical properties of speech may shed some light on individual differences in L2 perception of regional dialects in both linguistic and indexical perception tasks. Studies on cross-language perception of indexical information suggest that the extent to which processing depends on language-dependent information is related to the listeners' knowledge, and perhaps even familiarity with the sound system, of the language. Koster, Schiller, and Kunzel (1995) found that English listeners with knowledge of German performed better on a talker identification task in German than listeners with no knowledge of German. In addition, Koster and Schiller (1997) found that Spanish and Chinese listeners with knowledge of German performed better than Spanish and Chinese listeners with no knowledge of German, and that both groups performed worse than English listeners with knowledge of German. In a more recent study, Sullivan and Schlichting (2000) found improvement in voice recognition after initial study of English, but they did not find any additional improvement after the second semester of study. Nevertheless, taken together, these studies indicate that improvement in talker identification, and processing of nonlinguistic information occurs with increased familiarity with an L2, and with increased mastery over the sound system of the L2.

Other studies on L2 speech perception suggest important links between development of L2 phonological knowledge and the perception of indexical information in speech. Strange and Dittmann (1984) trained Japanese listeners on American English /r/ and /l/ using synthetic stimuli in a same-different discrimination task with feedback. Although the listeners improved over the course of the training and on some more difficult identification and discrimination transfer tasks with synthetic stimuli, Strange and Dittmann (1984) found that listeners did not improve in their ability to identify /r/ and /l/ in natural speech. Their findings suggest that the Japanese listeners were unable to robustly learn the non-native contrast from synthetic stimuli, which contained no variability from talker-specific voice attributes or phonetic context. Listeners who received high-variability training for /r-/l/ contrast (Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Logan, Lively, & Pisoni, 1991), on the other hand, improved at discriminating the contrast and were better at generalizing to novel utterances. In addition, Bradlow, Pisoni, Yamada, and Tohkura (1997) showed that listeners who experienced high-variability phonetic training for the /r-/l/ contrast not only improved in their perception of the contrast, but also improved in their productions. Thus, variation appears to have been retained in representations and used to classify new utterances. While the studies on talker identification in an L2 suggest that knowledge of the sound system affects the processing of the indexical properties of speech, these studies point to a role for indexical properties in the learning of new sounds and contrasts in the development and acquisition of the L2. Although the studies discussed above suggest that individual listeners with greater L2 phonological knowledge may be better at learning and using indexical information in speech, few studies have directly investigated individual differences in the perception of linguistic and indexical properties of speech in L2 acquisition.

L2 vocabulary knowledge may also underlie the ability to adapt and adjust to regional dialect variation in a speech recognition task or a dialect categorization task. Previous studies on native speech perception have found that lexicon size and lexical connectivity play important roles in speech perception and spoken word recognition abilities (e.g., Altieri, Gruenenfelder, & Pisoni, 2010; Pisoni, Nusbaum, Luce, & Slowiaczek, 1985; Samuel, 1986). Findings from these studies suggest that listeners are able to use top-down knowledge of lexical connectivity among words to recognize words from partial and degraded information in recognition tasks (e.g., Altieri et al., 2010; Pisoni et al., 1985). Tamati and colleagues (in press) also found that native listeners who had more difficulty in recognizing words in highly variable sentences differing in regional dialect, gender, and talker attributes in multitalker babble also knew fewer words than listeners who were better at recognizing words in the same sentences. Lexical knowledge has also been found to influence speech recognition in an L2. Non-native listeners who have less knowledge of English words have much more difficulty using lexical information to resolve ambiguous or degraded acoustic-phonetic information to successfully recognize spoken words (e.g., Bradlow & Pisoni, 1999; Bundgaard-Nielsen, Best, & Tyler, 2011; Ezzatian, Avivi, & Schneider, 2010). Still, while the previous studies suggest that lexical knowledge and connectivity may contribute to the ability to rapidly adapt to regional dialect variation (and variation from other sources) in a speech recognition task, these studies have not addressed the potential influence of lexical knowledge on the ability to adapt to and use indexical information in speech in a dialect categorization or classification task.

The Current Study

The current study investigated the relation between L2 phonological and lexical knowledge and the perception of L2 regional dialect information. The same three listener groups reported on in Chapter 2 (NAE = native American English group; KLT = long-term Korean group; KST = short-term Korean group) also participated in the current study. Non-native listeners were native speakers of Korean who had either lived in the U.S. for a long period of time (more than 3 years; KLT) or a relatively short period of time (6-18 months; KST). Measures of individual listeners' L2 phonological and lexical knowledge were compared to scores on the multi-dialect speech recognition task and the forced-choice regional dialect categorization task reported in Chapter 2. Although listeners had also completed a regional dialect free classification task, it was determined that performance on the task was too low to obtain meaningful variation in performance across listeners. Therefore, the scores from the free classification task obtained in Chapter 2 were not included in the current study, which focused on individual scores on the multi-dialect speech recognition task and the dialect categorization tasks.

The three listener groups from Chapter 2 completed two L2 phonological perception tasks: an ABX discrimination task and a forced-choice vowel identification task. The ABX task was designed to assess the Korean listeners' perception of American English vowel and consonant targets that are especially difficult for Korean learners of English. The American English vowel identification was designed to measure the Korean listeners' ability to identify American English monophthongal vowels. If L2 phonological knowledge underlies the ability to recognize words produced by talkers from different regions of origin, scores from an ABX and vowel identification tasks should be related to sentence recognition scores on the multi-dialect sentence recognition task. Similarly, if L2 phonological knowledge contributes to the use of

indexical information in the speech signal to discriminate or identify American English regional dialects, performance on these tasks should also be related to the Korean listeners' ability to categorize talkers by region of origin in the dialect categorization task.

Measures of English vocabulary knowledge were also collected. Listeners filled out a self-report word familiarity questionnaire that was used to measure receptive vocabulary knowledge in English. To obtain a measure of expressive vocabulary, listeners also completed a naming task in which they produced the names of objects in pictures in English. If L2 lexical knowledge underlies speech recognition ability, receptive and expressive vocabulary knowledge should be related to sentence recognition scores on the multi-dialect sentence recognition task reported in Chapter 2. If L2 lexical knowledge also contributes to the ability to categorize categorization talkers by region of origin in the L2, receptive and expressive vocabulary knowledge should also be related to dialect categorization accuracy scores on the categorization task reported in Chapter 2. Thus, the goal of the current study is to examine the contribution of L2 phonological and lexical knowledge to individual differences in perception of regional dialect variation in the multi-dialect sentence recognition task and the dialect categorization task observed in Chapter 2. Given previous findings on individual differences in speech perception, it was expected that an L2 learner's ability to identify or discriminate L2 vowels and consonants and his or her ability to recognize and produce words in the L2 (i.e., processing of linguistic information in the L2) would be related to his or her ability to perceive and use indexical information in the L2.

Methods

Listeners

Fifty-nine young adults participated in the current study. Twenty-one participants were native speakers of American English (native American English group = NAE), 23 participants were native speakers of Korean who had lived in the U.S. for approximately three years or more (long-term Korean group = KLT), and 15 participants were native speakers of Korean who had lived in the U.S. for approximately one year or less (short-term Korean group = KST). The 21 NAE participants included 15 females and 6 males, aged 18-27 years ($M = 21.0$; $SD = 1.8$ years). All NAE participants grew up in a region of the U.S. where General American English is spoken, i.e., North Midland, West, and parts of New England. The 23 KLT participants included 8 females and 15 males, aged 19-32 years ($M = 24.0$; $SD = 3.9$ years). All KLT participants were native speakers of Korean and had spent 3 years or more in the U.S. (LOR, $M = 5.1$; $SD = 1.5$ years). The 14 KST participants included 9 females and 6 males, aged 19-32 years ($M = 25.3$; $SD = 4.3$ years). All KST participants were native speakers of Korean and had spent a total of approximately 0.5 to 1.5 years in the U.S. (LOR, $M = 0.8$; $SD = 0.4$ years).

All participants reported normal hearing with no significant history of any prior hearing or speech disorders at the time of testing. They received \$20 for 120 minutes of participation (rate of \$5 per 30 minutes), which included completing several other speech perception and production tasks, neurocognitive tasks, and self-report questionnaires.

Materials and Procedures

Procedures were identical for each of the three participant groups except where explicitly mentioned. Participants were seated in front of a personal computer equipped with headphones

(Beyerdynamic DT100), a keyboard, and a mouse. Participants were tested individually in a quiet room where they completed a series of computer-based perceptual tasks, neurocognitive tasks, and self-report questionnaires. The output levels of the target sentences for all the perception tasks were calibrated to be approximately 65 dB SPL. Procedures and materials for all three tasks relevant for this study are described below.

Measures of L2 Phonological Abilities

ABX American English Vowel and Consonant Discrimination: The participants' ability to discriminate several American English vowels and consonant contrasts was examined using an ABX discrimination task. The specific ABX task used was based on an earlier study by Darcy, Park, and Yang (under review). Two female talkers from the North Midland dialect region were recorded producing English nonwords. All nonword items were not real words in either American English or Korean. Eight American English vowels and consonant contrasts that were either easy or difficult for native Korean speakers, based on the phonological systems of Korean and English, were used in this task. The vowel contrasts included the following difficult minimal pairs /i, ɪ/, /u, ʊ/, /eɪ, ε/, /æ, ε/, and /ɑ, ʌ/, and a control (easy) pair /u, o/. The consonant contrasts included the following difficult minimal pairs /l, r/, /p, f/, and /θ, s/ and the control (easy) pair /s, t/. These particular contrasts were also selected because they are involved in American English regional dialect variation and may be crucial to reliably distinguishing the four regional varieties used in the perception tasks described in Chapter 2. All target vowels were embedded in the stressed syllable of a disyllabic minimal pair frame CV(C)CVC, within three different consonantal contexts (bilabial, alveolar, and velar), e.g., [geɪ.kəʔt], [neɪ.dən], and [peɪ.bad]. All consonants were embedded in the same disyllabic minimal pair frame CV(C)CVC, in prevocalic

word-initial, intervocalic, or postvocalic word-final position, e.g., [pæs.tɪk], [pə.pɪk], and [tə.gæp].

Overall, a total of 30 pairs (5 vowel test pairs in 3 contexts and 3 consonant test pairs in 3 contexts; and 1 vowel control pair and 3 contexts and 1 consonant control pair in 3 contexts) were used in the task. All four possible ABX combinations for each pair (ABB, ABA, BAA, and BAB) were presented, for a total of 120 stimulus items (4 combinations x 30 pairs). Each combination of the 30 pairs was randomly assigned to one of four blocks so that each contrast was presented once in each context per block, for a total of 30 trials per block.

On each trial of the ABX task, one nonword pair (AB) produced by one of the female talkers was presented, followed by another nonword matching either the first or the second nonword (“X”, matching either A or B) produced by the other female talker. The stimuli in each ABX triad were separated by 250 ms of silence. The participants were asked to decide whether the third stimulus (“X”) was the same as the first (A) or the second (B) by pressing a button on a button box as quickly as possible. The left button on the button box always represented the first stimulus (A) and the right button always represented the second stimulus (B). The next ABX triad was presented after the participant responded. Within each block, the presentation of the triads was randomized. Before the task, participants received a short 8-trial practice block of ABX trials with feedback. None of the practice items were used later in the task. After the practice block of trials, participants completed the four blocks of testing. Between each block, subjects had a short break.

Responses were scored for accuracy and response time (RT). Response times were measured from the onset of the last of the three stimulus items in each ABX triad until the listener entered his or her response. Before analysis, across all listeners and trials, 16 trials for

which no response was recorded (time out and/or incorrect button press) were excluded (less than 1% of all trials).

American English Vowel Identification Task: The participants' ability to identify American English vowels was measured with an 11-alternative forced-choice vowel categorization task. Four talkers (2F/2M) from the McFAE corpus (Multi-Talker Corpus of Foreign-Accented English; Tamati, Park, & Pisoni, 2011) were selected. All talkers were native speakers of American English speakers who had grown up in the North Midland area of state of Indiana. For this task, two hVd words containing the eleven monophthongal vowels of American English /i ɪ ε eɪ æ u ʊ o ʌ ɑ ɔ/ were selected for each talker, for a total of two sets of 11 hVd words per talker and 88 hVd words overall.

On each trial, participants were randomly presented with a single hVd word, and upon being prompted by the question "Which vowel did you hear?" the subject responded by clicking on one of eleven vowel choices displayed on a computer monitor. Each vowel option contained a single hVd word, written out on the computer screen with a vowel shorthand developed by Gerstman (1968), and two high-frequency words that contained the same vowel. All response options had also been shown to the participant before the task began so that they could familiarize themselves with the shorthand and target vowels of the response alternatives. Participants in the current study had also previously completed a hVd production task (not reported on here) for which they received some training on the vowel shorthand and the target vowels of the eleven hVd words. The vowel identification task was self-paced. Participants could take as long as they wanted to respond but they only listened to one presentation of each word. The next trial began after they entered their response.

Responses were scored for accuracy. After preliminary inspection of the responses, native NAE listeners could not distinguish between the two low-back vowels /ɑ/ and /ɔ/, consistent with the low-back merger in the North Midland dialect region, which is pervasive in the English speaking community of the subjects' residence (e.g., Labov, Ash, & Boberg, 2006). Given the poor performance on these vowels, the two vowels were combined into one low-back category, which will be referred to as /ɑ/, for analysis.

Measures of L2 Vocabulary Knowledge

Boston Naming Test: Expressive vocabulary (i.e., the words one can produce in speaking or writing) was assessed using the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983). Participants saw pictures of items, ranging from low frequency to high frequency, and were asked to name the object being depicted in the pictures as quickly and as accurately as possible. The Boston Naming Test provides one long complete list of 60 items (Kaplan et al., 1983). However, for the purpose of the current study, two shorter versions of 30 items with target words roughly equivalent in lexical frequency were used (e.g., Williams, Mack, & Henderson, 1989). Non-native participants (KLT and KST) first did one short version in English (even list) and the other in Korean (odd list). Native NAE participants completed the full set only in English. On each trial, black and white images were displayed on the computer monitor. Participants were required to say the name of the object aloud. The next trial was presented after the participant gave a response (correct or incorrect) to the previous item, after the participant indicated that they did not know the object, or after approximately 5-7 seconds of silence. No cues or hints were given to the participants. The entire task was audio recorded and the vocal

responses were later scored for accuracy offline. For the purposes of the current study, only English responses were analyzed.

WordFam Word Familiarity Questionnaire: To measure receptive vocabulary (i.e., the words one can understand in listening or reading), participants also completed the WordFam test, a self-report word familiarity rating questionnaire originally developed by Lewellen, Goldinger, Pisoni, and Greene (1993). Participants were instructed to rate how familiar they were with a printed set of English words using a seven-point scale, ranging from 1 (“You have never seen or heard the word before”) to 7 (“You recognize the word and are confident that you know the meaning of the word.”). The WordFam questionnaire contains a total of 150 English words, including 50 low familiarity, 50 mid familiarity, and 50 high familiarity items based on ratings originally obtained by Nusbaum, Pisoni, and Davis (1984). The current study used an electronic version of the original WordFam questionnaire that required the participants to respond by clicking on one of seven boxes containing each of the familiarity rating options. Responses for all 150 items were collected and averaged for each of the three familiarity conditions.

Results

Group Differences: Native Language and Length of Residency

Group analyses for each task included repeated-measures ANOVAs on performance scores (accuracy or RT) with the factors 1) the sound pair (ABX) or single vowel (Vowel Identification), and 2) listener group (NAE, KLT, KST). Additional analyses were carried out to explore any significant main effects and interactions, as well as distributions in performance within and across non-native groups.

L2 Phonological Abilities

ABX American English Vowel and Consonant Discrimination: Mean accuracy scores and response time (RT) were calculated for each vowel and consonant contrast for all listener groups. Response time (RT) measures were also collected and corrected (see above). After preliminary analyses, data from one participant was excluded because he/she scored worse than chance overall on the task (31.7%), and on both the vowel and consonant control pairs (vowel control: 0%; consonant control: 33.3%). Another participant's data could not be scored because of computer failure. All groups were accurate on the control contrasts. Overall, performance was high across all three listener groups. NAE listeners scored a mean of 87.7% (SD = 6.3%) on all test sound pairs and 93.4% (SD = 7.3%) on the control sound pairs; the non-native KLT listeners scored a mean of 75.0% (SD = 8.2%) on all test sound pairs and 88.3% (SD = 10.1%) on the control sound pairs; and the non-native KST listeners scored a mean of 73.5% (SD = 8.8%) on all test sound pairs and 93.3% (SD = 9.4%) on the control sound pairs. A summary of the mean accuracy scores for the three listener groups on all vowel sound pairs are shown in Table 1, and a summary for all consonant sound pairs are shown in Table 2.

Table 1: Mean ABX accuracy responses (%) on vowel test and control sound pairs by listener group (NAE, KLT, KST). Standard deviations are in parentheses

Listener Group	Vowel Pairs					
	/i, ɪ/	/u, ʊ/	/eɪ, ε/	/æ, ε/	/ɑ, ʌ/	/u, o/ control
NAE	96.0 (5.7)	76.9 (14.1)	92.1 (8.1)	90.5 (12.2)	83.7 (12.7)	92.9 (9.6)
KLT	69.0 (13.7)	63.5 (13.8)	86.1 (14.0)	62.3 (12.8)	70.2 (15.7)	88.5 (11.9)
KST	70.0 (12.1)	65.0 (11.9)	88.9 (12.1)	58.9 (12.4)	60.6 (14.6)	92.8 (7.6)

Table 2: Mean ABX accuracy responses (%) on consonant test and control sound pairs by listener group (NAE, KLT, KST). Standard deviations are in parentheses

Listener Group	Consonant Pairs			
	/l, r/	/p, f/	/θ, s/	/s, t/ control
NAE	88.9 (10.7)	90.5 (8.9)	83.1 (11.1)	94.0 (8.0)
KLT	87.3 (12.3)	83.7 (13.8)	77.4 (12.1)	88.1 (14.6)
KST	82.2 (16.9)	85.6 (11.1)	77.2 (20.5)	93.9 (13.2)

A repeated measures ANOVA with sound pair (vowel or consonant contrast) as a within-subject factor and listener group (NAE, KLT, KST) as the between-subject factor was carried out on ABX accuracy scores. The analysis revealed significant main effects of sound pair [$F(9, 486) = 39.7, p < .001$] and listener group [$F(2, 54) = 15.2, p < .001$], and a significant sound pair x listener group interaction [$F(18, 486) = 6.8, p < .001$]. To further explore these effects, one-way ANOVAs on accuracy scores were carried out on each sound pair with listener group as the factor. As expected, no significant effect of group was found for the vowel and consonant control contrasts. In addition, no significant effects of listener group were found for any of the consonant test pairs. The main effect of listener group was significant for /i, ɪ/, /u, ʊ/, /æ, ε/, and /ɑ, ʌ/ (all $p \leq .004$), but it was not significant for the mid-front vowels /eɪ, ε/.

For the high front vowels /i, ɪ/, a series of *t*-tests revealed that NAE listeners were significantly more accurate than both KLT and KST listeners (all $p < .001$), and that the KLT and KST listeners were not significantly different. For the high back vowels /u, ʊ/, NAE listeners were more accurate than both KLT and KST listeners ($p \leq .012$), but the KLT and KST groups were not significantly different. For the low front vowels /æ, ε/, NAE listeners were significantly more accurate than both KLT and KST listeners (all $p < .001$), but the KLT and KST listeners were not different. For the low back vowels /ɑ, ʌ/, NAE listeners were more accurate than both KLT and KST listeners (all $p \leq .004$). Again, KLT and KST were not different.

A repeated measures ANOVA with sound pair (vowel or consonant contrast) as a within-subject factor and listener group (NAE, KLT, KST) as the between-subject factor was also carried out on RT measures on correct trials. The analysis revealed a significant main effect of sound pair [$F(9, 486) = 2.7, p = .005$]. The main effect of listener group and the sound pair x listener group interaction were not significant. Since RT was similar across groups for nearly all contrasts, the main effect of sound pair was not further investigated. Thus, that while mean accuracy differed for native and non-native listeners, mean response time across groups was consistent.

American English Vowel Identification: Mean accuracy scores were calculated for each American English vowel for all listener groups. Overall, participants were highly accurate, with the NAE group scoring a mean of 89.6% (SD = 10.8%) overall, the KLT group scoring a mean of 65.4% (SD = 9.9%) overall, and the KST group at 62.3% (SD = 12.9%) overall. All listener groups were above chance level across the vowels (chance on 11-alternative task = 9.1%). A repeated measures ANOVA with vowel (/i ɪ ε eɪ æ u ʊ o ʌ ɑ ɔ/) as a within subject factor and listener group (NAE, KLT, KST) as the between-subject factor was carried out on accuracy scores. The analysis revealed significant main effects of vowel [$F(9, 504) = 24.5, p < .001$] and listener group [$F(2, 56) = 35.9, p < .001$], and a significant vowel x listener group interaction [$F(18, 504) = 2.6, p < .001$]. Independent samples *t*-tests showed that NAE listeners were significantly more accurate than both KLT and KST listeners (all $p < .001$), and that the KLT and KST listeners were not significantly different.

Table 3 shows the vowel identification accuracy scores for all three listener groups (NAE, KLT, KST). As can be seen in Table 3, some vowels were clearly easier or more difficult for the non-native listener groups compared to the native listeners. To further investigate this

interaction, one-way ANOVAs were carried out on the vowel accuracy scores with listener group as the factor for each vowel. Analyses revealed significant main effects of listener group for /i/, /ɪ/, /ɛ/, /æ/, /ɑ/, /ʌ/, /oʊ/, /ʊ/, and /u/ (all $p \leq .031$), all vowels except /eɪ/. Independent samples *t*-tests on those nine vowels (/i ɪ ɛ æ u ʊ o ʌ ɑ ɔ/) showed that NAE listeners were significantly more accurate than KLT and KST groups (all $p \leq .017$), and that the KLT and KST groups were not significantly different. For front vowels, all listener groups were most accurate on the tense vowels /i/ and /eɪ/, but non-native listeners were disproportionately worse at the lax vowels /ɪ/ and /ɛ/, as would be expected given the vowel inventory of Korean. Performance on the low front vowel /æ/ was also quite low for the non-native listeners. Performance on the back vowels was less systematic. Non-native listeners performed quite poorly across all categories.

Table 3: Mean vowel identification accuracy (%) across 10 American English vowel categories by listener group (NAE, KLT, KST). Standard deviations are in parentheses.

Listener Group	Vowel				
	/i/	/ɪ/	/eɪ/	/ɛ/	/æ/
NAE	95.8 (9.9)	87.5 (13.7)	93.5 (15.6)	86.9 (17.4)	89.3 (22.5)
KLT	84.8 (18.1)	44.6 (24.1)	87.0 (18.6)	52.7 (23.8)	57.6 (25.8)
KST	82.5 (20.5)	37.5 (21.7)	81.7 (25.4)	58.3 (24.4)	54.2 (21.5)
	/ɑ/	/ʌ/	/oʊ/	/ʊ/	/u/
NAE	96.1 (7.3)	78.6 (29.4)	94.0 (12.9)	88.1 (21.8)	76.8 (24.8)
KLT	84.0 (10.5)	49.3 (29.1)	76.1 (24.4)	55.4 (25.2)	39.7 (28.4)
KST	75.0 (17.7)	45.6 (31.2)	70.0 (26.2)	55.8 (27.1)	45.8 (30.5)

L2 Vocabulary Knowledge

Boston Naming (English): The mean number of objects correctly named in the Boston Naming Test was calculated for all listener groups. To examine any effects of word frequency on

naming, the responses were divided into three blocks. The first block contained 10 trials of high frequency object names, the second block contained 10 trials of medium frequency object names, and the third block contained 10 trials of low frequency object names. A summary of the mean number of objects correctly named across these three frequency blocks (and overall) is presented in Table 4.

Table 4: Mean number of correctly named objects across three frequency blocks by listener group (NAE, KLT, KST). Standard deviations are in parentheses.

Listener Group	Frequency Block			Overall (n = 30)
	High (n = 10)	Medium (n = 10)	Low (n = 10)	
NAE	9.9 (0.4)	9.6 (0.6)	6.9 (2.1)	26.3 (2.5)
KLT	9.0 (1.0)	4.0 (1.3)	2.0 (1.0)	14.9 (2.5)
KST	8.8 (0.9)	3.2 (1.5)	1.4 (0.8)	13.5 (2.0)

A repeated measures ANOVA with frequency (high, medium, and low) as the within subject factor and listener group (NAE, KLT, KST) as the between-subject factor was carried out on the naming scores. The analysis revealed a significant main effect of frequency block [$F(2, 112) = 416.5, p < .001$], a significant main effect of listener group [$F(2, 56) = 174.3, p < .001$], and a significant frequency x listener group interaction [$F(4, 112) = 41.0, p < .001$]. Overall, more high frequency object words were correctly named than medium and low frequency words, and more medium frequency words were correctly named than low frequency words (all $p < .001$). NAE listeners were also able to name more objects overall than both KLT and KST listeners (all $p < .001$), but no difference was found between the two non-native groups. To explore the significant frequency x listener group interaction, one-way ANOVAs on word naming accuracy with listener group as the factor were carried out for each of the three frequencies. Listener group was significant for all three frequencies (all $p \leq .001$). A series of t -

tests showed that NAE listeners were significantly better at naming objects than both KLT and KST listeners on all three frequencies (all $p < .001$). No differences were found between the KLT and KST groups.

Although the NAE listeners were more accurate overall, the difference between NAE listeners and both non-native groups (KLT and KST) were greater for medium and low frequency object words than high frequency object words. Additionally, NAE listeners knew almost all high and medium frequency words and knew fewer low frequency words. In comparison, while non-native listeners also knew almost all high frequency words, they knew fewer medium and low frequency words. The non-native listeners had good knowledge of high frequency words, but their knowledge of lower frequency words was poor.

WordFam Word Familiarity Questionnaire: Mean response ratings were calculated for high, medium, and low familiarity words for all listener groups. Mean ratings are displayed in Table 5.

Table 5: Mean familiarity ratings on a scale from 1 (not familiar at all) to 7 (very familiar) for high, medium, and low familiarity English words by listener group (NAE, KLT, KST). Standard deviations are in parentheses.

Listener Group	Word Familiarity			Overall
	High	Medium	Low	
NAE	6.5 (0.4)	3.8 (1.1)	2.6 (0.8)	4.3 (0.7)
KLT	5.7 (0.5)	2.9 (0.9)	2.4 (0.8)	3.7 (0.7)
KST	5.7 (0.6)	3.0 (0.8)	2.4 (0.7)	3.7 (0.7)

A repeated measures ANOVA with word familiarity (high, medium, and low) as the within subject factor and listener group (NAE, KLT, KST) as the between-subject factor was carried out on the rating responses. The analysis revealed significant main effects of word

familiarity [$F(2, 112) = 1107.1, p < .001$] and listener group [$F(2, 56) = 5.4, p = .007$], and a significant familiarity x listener group interaction [$F(4, 112) = 5.9, p \leq .001$]. NAE listeners were overall significantly more accurate than KLT [$t(42) = 3.0, p = .005$] and KST listeners [$t(34) = 2.5, p = .016$], but KLT and KST listeners were not significantly different. Additionally, high familiarity words were rated as being significantly more familiar than medium and low familiarity words, and medium familiarity words were more familiar than low familiarity words (all $p < .001$).

To further explore the significant familiarity x listener group interaction, one-way ANOVAs on the ratings with listener group as the factor were carried on each familiarity (high, medium, and low). Analyses revealed significant main effects of listener group for high familiarity words and medium familiarity words (all $p \leq .005$), but not for low familiarity words. A series of t -tests on high and medium familiarity words revealed that NAE listeners were significantly more familiar with high familiarity words than KLT and KST listeners (all $p < .001$), but only slightly more familiar with medium familiarity words than KLT listeners ($p = .004$) and KST listeners ($p = .020$). No differences were found between the KLT and KST groups. Additionally, as shown by the standard deviations reported in Table 5, performance tended to vary little within each group for high familiarity words and more so for the medium and low familiarity words. While non-native listeners had robust knowledge of common English words, they had less knowledge of low familiarity words in American English.

Individual Differences: L2 Phonological and Lexical Knowledge and the Perception of Regional Dialect Variation

To assess the relation between the L2 phonological and lexical measures and the perception of regional dialect in an L2, a series of multiple linear regression analyses were used to predict sentence recognition accuracy scores on the multi-dialect sentence recognition task and categorization accuracy scores on the regional dialect categorization task in Chapter 2 from the L2 phonological and lexical measures described here. Taking the results of the group analyses into consideration, the following variables were used as predictors in the analyses: ABX mean accuracy across all vowel test contrasts (ABX-Vowel), ABX mean accuracy across all consonant test contrasts (ABX-CONS), ABX mean RT across all test contrasts (ABX-RT), Vowel Identification mean accuracy across all vowels (VID), Boston Naming mean overall accuracy (BN), and WordFam overall mean word familiarity ratings (WF). For the purposes of the individual differences analyses, KLT and KST groups were combined into a single non-native group. Length of residency (LOR) in the U.S. was also included as a predictor variable.

Separate analyses were also carried out with the native control group with the five predictor variables: ABX-Vowel, ABX-CONS, ABX-RT, BID, BN, and WF. These analyses were included to examine whether more general language-independent characteristics contribute to the perception of regional dialect information.

Multi-dialect Sentence Recognition

A linear regression analysis was used to assess the relations between multi-dialect sentence recognition and L2 phonological and lexical knowledge. Given that sentence recognition accuracy across all four regional dialects (General American, Northern, Mid-

Atlantic, Southern) and the two SNRs (+2 dB SNR, -2 dB SNR) was strongly correlated, the overall mean accuracy score across all conditions was first used as the dependent variable. As described in Chapter 2, accuracy scores on the sentence recognition task were converted to rationalized arcsine units (rau) because performance was close to ceiling at +5 dB SNR for the native listeners (Studebaker, 1985). To facilitate interpretation of the findings, all figures present the sentence recognition scores in terms of percent correct accuracy.

The results of the regression analysis showed that the seven predictor variables explained 53.1% of the variance [$R^2 = .53$, $F(7, 28) = 4.5$, $p = .002$]. ABX-Vowel [$\beta = .53$, $p = .006$], BN [$\beta = .43$, $p = .010$], and WF [$\beta = .47$, $p = .005$] emerged as significant predictors of multi-dialect sentence recognition scores. ABX-CONS, ABX-RT, VID, and LOR were not significant predictors. Thus, English vowel discrimination ability and English receptive and expressive vocabulary knowledge predicted multi-dialect sentence recognition in American English. Figure 1 displays the relation between sentence recognition accuracy scores and ABX vowel discrimination scores for listeners in all three listener groups (NAE, KLT, KST); Figure 2 displays the relation between sentence recognition accuracy scores and the number of objects correctly named on the Boston Naming Test for listeners in all three listener groups; and Figure 3 displays the relation between sentence recognition accuracy scores and the word familiarity ratings on the WordFam questionnaire for listeners in all three listener groups.

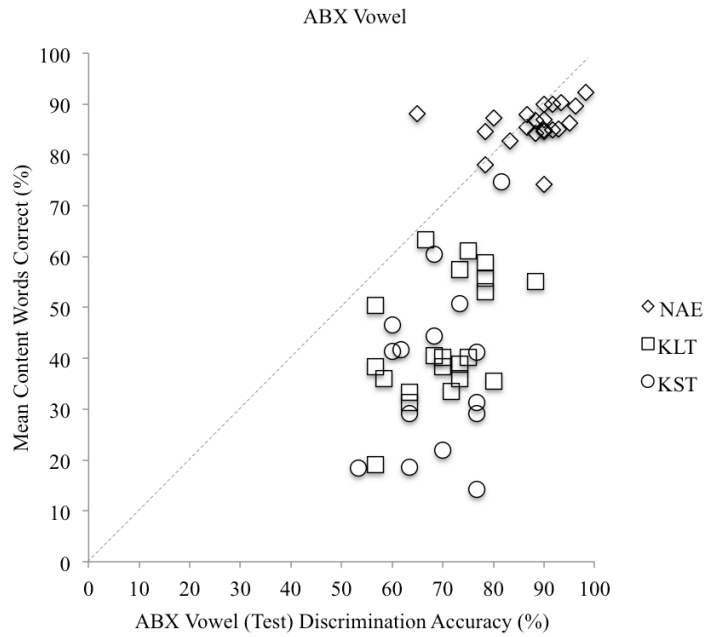


Figure 1. ABX test vowel discrimination accuracy (x-axis) and percent correct word recognition accuracy on the multi-dialect sentence recognition task (y-axis) for listeners in all three listener groups (NAE: diamonds; KLT: squares; KST: circles).

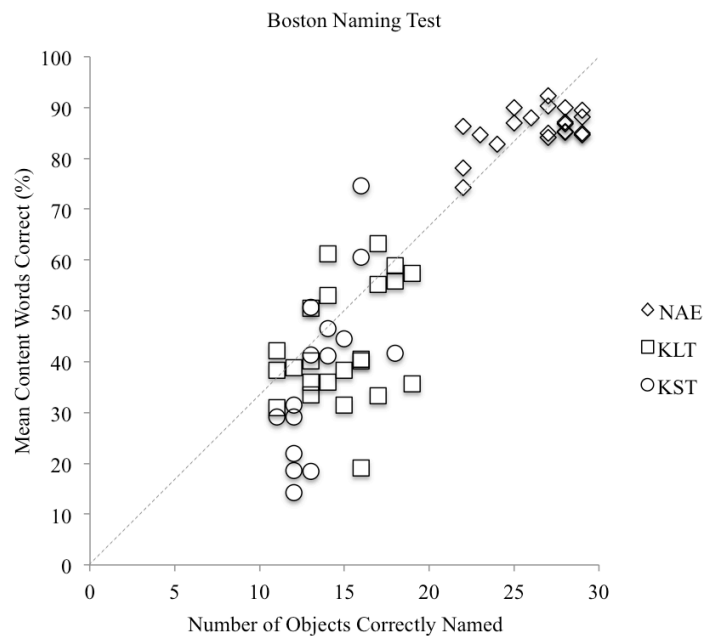


Figure 2. The number of objects correctly named on the Boston Naming Test (x-axis) and percent correct word recognition accuracy on the multi-dialect sentence recognition task (y-axis) for listeners in all three listener groups (NAE: diamonds; KLT: squares; KST: circles).

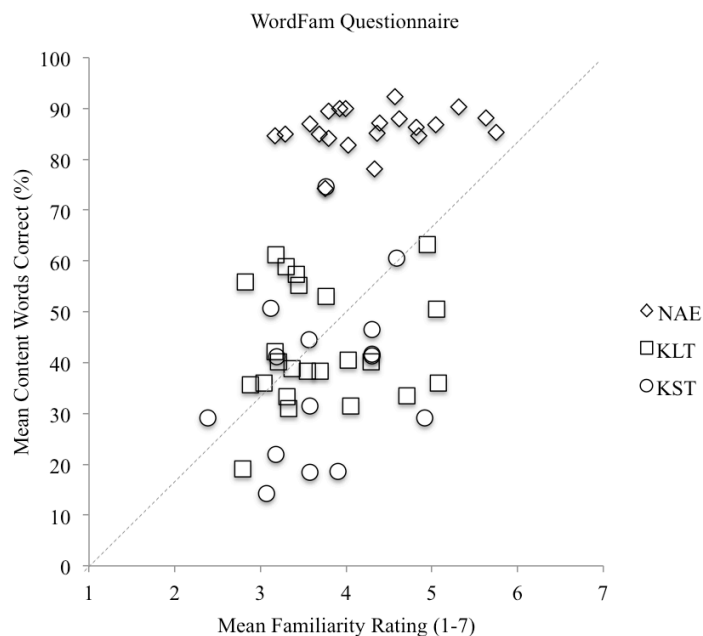


Figure 3. The mean familiarity rating on the WordFam questionnaire (x-axis) and percent correct word recognition accuracy on the multi-dialect sentence recognition task (y-axis) for listeners in all three listener groups (NAE: diamonds; KLT: squares; KST: circles).

Similar analyses were also carried out for sentence recognition accuracy for each regional dialect (General American, North, Mid-Atlantic, South) to assess whether different sets of skills may be related to recognizing words produced by talkers from more or less familiar dialect regions. For General American talkers, the results of the regression showed the seven predictor variables explained 50.2% of the variance [$R^2 = .50$, $F(7, 28) = 4.0$, $p = .004$]. ABX-Vowel [$\beta = .55$, $p = .006$], BN [$\beta = .33$, $p = .047$], and WF ratings [$\beta = .40$, $p = .021$] emerged as significant. For the less familiar regional dialects, for Northern talkers, the results of the regression showed that the seven predictor variables explained 50.3% of the variance [$R^2 = .50$, $F(7, 28) = 4.0$, $p = .004$]. ABX-Vowel [$\beta = .49$, $p = .013$], BN [$\beta = .44$, $p = .009$], and WF ratings [$\beta = .40$, $p = .018$] were significant predictors. For Mid-Atlantic talkers, the regression showed the seven predictor variables explained 46.2% of the variance [$R^2 = .46$, $F(7, 28) = 3.4$, $p = .009$]. ABX-Vowel [$\beta =$

.50, $p = .015$], BN [$\beta = .38$, $p = .027$], and WF ratings [$\beta = .55$, $p = .003$] were significant predictors. Finally, for Southern talkers, the seven predictor variables explained 57.3% of the variance [$R^2 = .57$, $F(7, 28) = 5.4$, $p = .001$]. ABX-Vowel [$\beta = .51$, $p = .006$], BN [$\beta = .49$, $p = .002$], and WF ratings [$\beta = .49$, $p = .003$] were significant predictors. ABX-CONS, ABX-RT, VID, and LOR were not significant predictors of sentence recognition scores for any of the four regional dialects.

The analysis for the native NAE listeners revealed no significant influence of the five predictor variables assessing English phonological and lexical knowledge on multi-dialect sentence recognition in a native language.

Taken together, the results of the regression analyses on sentence recognition performance demonstrate that English vowel discrimination (ABX-Vowel) and vocabulary measures contribute to L2 speech recognition across multiple, familiar and unfamiliar, American English regional dialects. Furthermore, the influence of phonological and lexical knowledge on sentence recognition was limited to L2 speech perception and word recognition performance and not native word recognition performance.

Forced-Choice Regional Dialect Categorization

A series of linear regression analyses were also used to assess the relation between regional dialect categorization accuracy and L2 phonological and lexical knowledge. Given that categorization accuracy varied by regional dialect, separate analyses were carried out on each individual regional dialect. None of the regression models reached significance, suggesting that L2 vocabulary and lexical knowledge, as measured by the seven predictor variables, were not predictors of L2 regional dialect categorization accuracy. Similarly the analyses for the native

listeners revealed no significant influence of the five predictor variables on the categorization of the four American English regional dialects.

Discussion

To investigate the relation between L2 phonological and lexical knowledge and the perception of L2 regional dialect information, the current study examined the individual differences among non-native listeners in recognizing words produced by talkers from four regions of origin and identifying regional dialects observed in Chapter 2. Non-native speakers of American English from South Korea completed two tasks assessing L2 phonological knowledge: an ABX discrimination and a vowel identification task. They also completed a self-report word familiarity questionnaire and a naming task, which were used as measures of L2 lexical knowledge. These measures were used to predict scores on the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task from Chapter 2. Before investigating the contribution of L2 phonological and lexical knowledge to regional dialect perception in the L2, the listener groups' performance was analyzed to obtain a profile of the general L2 knowledge of the non-native listeners in the current study and to motivate the use of scores from these tasks in the individual differences analyses. Overall, non-native listeners performed as expected based on earlier research. Both non-native KLT and KST groups demonstrated poorer American English vowel discrimination and identification, as well as smaller expressive and receptive vocabulary sizes in American English compared to the native NAE group. In addition, the patterns of errors in the perception tasks involving American English vowels were consistent with the influence of their native language of Korean. However, the non-native KLT and KST listeners were able to accurately discriminate American English

consonants on the ABX task, performing similarly to the native NAE group. The KLT and KST groups did not differ on any of the tasks or the questionnaire, suggesting that length of residency in the U.S. did not contribute greatly to performance.

Based on the group performances, we assessed the influence of L2 phonological and lexical knowledge on multi-dialect sentence recognition and regional dialect categorization using the following measures: vowel discrimination accuracy, consonant discrimination accuracy, RTs for difficult vowel and consonant contrasts, vowel identification accuracy, object-naming, word familiarity ratings, and length of residency. Results of the individual differences analyses revealed that vowel discrimination accuracy (i.e., accuracy on ABX task for only vowel test pairs), receptive and expressive vocabulary knowledge were significant independent predictors of multi-dialect sentence recognition accuracy. Additionally, these three variables consistently predicted sentence recognition accuracy on all four talker dialects included in the sentence recognition task (General American, North, Mid-Atlantic, South). Consonant discrimination accuracy (i.e., accuracy on ABX task for only consonant test pairs), RTs for American English ABX vowel and consonant test pairs, and length of residency in the U.S. were not significant independent predictors of multi-dialect sentence recognition accuracy, either overall or for each of the four talker dialects considered separately.

The relation between ABX vowel discrimination and sentence recognition suggests that L2 phonological knowledge improves the ability to recognize words in sentences produced by multiple male and female talkers from different regions of origin. Although the non-native listeners were able to reliably discriminate difficult consonant contrasts in American English, individual listeners' ability to discriminate difficult American English vowel contrasts may have allowed them to more easily recognize sounds and lexical items in the highly variable listening

conditions. This result supports previous findings from Meador, Flege, and MacKay (2000), who found that non-native listeners' ability to perceive L2 vowels and consonants was linked to their word recognition ability in noise. More accurate discrimination of L2 vowel and consonant perception may reflect more native-like representations for sounds and lexical items, which would support the processing of the incoming speech signal and better recognition of words in the L2. Thus, speech recognition in the L2 is improved by greater development of L2 phonological knowledge, and not just by exposure to variability in the L2 environment. However, vowel identification accuracy was not related to sentence recognition ability. The vowel identification scores may not be as reflective of L2 phonological knowledge as the ABX discrimination task. Correctly assigning a vowel label to an L2 vowel sound may involve more meta-linguistic knowledge of the English language or the listeners' ability to quickly learn the experimenter-imposed vowel category labels, rather than bottom-up perceptual abilities. The ABX task does not require the listener to have stable verbal labels for the sound categories, and instead relies more heavily on perceptual categories, and as such, may more accurately assess the listeners' knowledge of the L2 phonological system.

The results also revealed that L2 lexical knowledge was related to multi-dialect sentence recognition. This finding was expected given that previous studies have consistently found that lexicon size and connectivity play important roles in speech perception and spoken word recognition (e.g., Ganong, 1980; Pisoni et al., 1985; Samuel, 1986; Altieri et al., 2010). Some studies have suggested that lexical knowledge supports word recognition from partial or degraded information in the speech signal (e.g., Pisoni et al., 1985; Altieri et al., 2010), which would be crucial for non-native listeners who have incomplete knowledge of the L2 phonological system and limited knowledge of social-indexical variation in the L2. The finding

that L2 lexical knowledge was linked to multi-dialect sentence recognition also supports previous studies that have shown that L2 lexical knowledge is related to the ability to make use of ambiguous or compromised acoustic-phonetic information to recognize words and sentences in an L2 (e.g., Bradlow & Pisoni, 1999; Ezzatian et al., 2010; Bundgaard-Nielsen et al., 2011). The results also revealed that both receptive and expressive vocabulary knowledge contribute independently to multi-dialect sentence recognition. Both types of vocabulary knowledge may lead to more accurate responses on the sentence recognition task, which has at least two components involving listening to the spoken sentences and typing in the words they recognized. A larger L2 receptive vocabulary may support the recognition of more words in the highly variable listening conditions, and a larger L2 expressive vocabulary may support the listener's ability to recognize the words and reproduce them via typing the response. Thus, both receptive and expressive vocabulary knowledge support L2 speech recognition for sentences spoken by multiple talkers from different regions of origin in multi-talker babble.

Similar analyses were also carried out to explore individual differences in non-native listeners' performance on a regional dialect categorization task, as reported in Chapter 2. The analyses did not yield any significant predictors of regional dialect perception abilities in the listeners' L2. However, previous studies on the perception of unfamiliar regional dialects in the L1 and L2 suggest that even without substantial sociolinguistic knowledge of regional dialect characteristics, listeners are able to use reliable dialect-specific acoustic-phonetic differences to group or make perceptual similarity judgements about talkers based on region of origin (Clopper & Bradlow, 2009; Tamati, 2008). Although the forced-choice regional dialect categorization task provided labels for the listeners, the response patterns reported in Chapter 2 suggested that listeners primarily responded with the General American category, and made their judgements

based on whether they perceived the talker as standard or nonstandard. Thus, the forced-choice categorization task in the current study functioned as a discrimination task (i.e., General American vs. other), compared to other forced-choice regional dialect categorization tasks that have been used in earlier studies that draw more directly upon perceptual categories for regional dialects and previous knowledge (e.g., Clopper & Pisoni, 2004; Tamati & Pisoni, in press). As long as the non-native listeners had a least one category for standard speech (General American), they would be able to complete the task in a manner that is similar to the native NAE listeners. Thus, greater exposure to multiple American English regional dialects may not have been as beneficial on the task in the current study as it would potentially be on other regional dialect categorization tasks that draw more upon the listeners' sociolinguistic knowledge of the language. The relation between L2 phonological and lexical knowledge and regional dialect identification or categorization should be further explored in future studies using a variety of different experimental task designs to measure L2 regional dialect identification or discrimination ability.

Future studies should also explore different sources of individual differences in the perception of L2 regional dialect. In addition to language background, L2 phonological knowledge, and L2 lexical knowledge, other factors may potentially underlie differences in listeners' ability to categorize L2 regional dialects. Furthermore, although the L2 phonological and lexical measures accounted for much variability in listener performance on the multi-dialect sentence recognition task, individual differences in recognizing words produced by talkers from a variety of regions of origin may also be better explained by considering underlying neurocognitive differences. Several core neurocognitive abilities, including short-term and working memory, attention and inhibition, and executive functioning, may influence how

indexical information is perceived, encoded, and represented, and consequently how this information is used in perception tasks (Tamati, Gilbert, & Pisoni, 2013; Tamati & Pisoni, in press). Future research should examine the influence of individual differences in neurocognitive abilities on the perception of indexical variation in an L2 to be able to better account for listener variability attested on the forced-choice regional dialect categorization task and the multi-dialect sentence recognition task from Chapter 2.

Conclusions

The goal of the current study was to investigate the relations between L2 phonological and lexical knowledge and non-native perception of regional dialect information in American English. Exploring the individual listener variability among non-native KLT and KST listeners observed on the multi-dialect sentence recognition task in Chapter 2, the results of the current study revealed that the ability to discriminate difficult American English vowel contrasts, and expressive and receptive vocabulary knowledge were predictors of the non-native listeners' sentence recognition scores. The non-native listeners who were better at discriminating difficult American English vowels, and had better receptive and expressive vocabulary knowledge were also better at recognizing English words produced by talkers from different regions of origin in the U.S. The current study also examined individual differences in performance on the regional dialect categorization task by the same non-native KLT and KST listeners, as reported in Chapter 2. Unlike the results for the sentence recognition task, L2 phonological and lexical knowledge was not related to the non-native listeners' ability to categorize American English regional dialects. Taken together, these findings suggest that phonological and lexical knowledge in an L2 contribute to the ability to adapt and adjust to fine-grained acoustic-phonetic differences

characteristic of regional dialects in American English to recognize words and sentences produced by talkers from multiple regions of origin. However, these factors do not appear to be related to the ability to use those fine-grained acoustic-phonetic differences to make perceptual judgements about a talker's region of origin. Additional research should be carried out to examine the perception of social-indexical variation in a second language and the factors underlying individual differences in second language speech perception abilities.

References

- Altieri, N., Gruenenfelder, T., Pisoni, D.B. (2010). Clustering coefficients of lexical neighborhoods: Does neighborhood structure matter in spoken word recognition? *Mental Lexicon*, 5, 1-21.
- Bent, T., & Bradlow, A.R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114, 1600-1610.
- Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America*, 106, 2074-2085.
- Bradlow, A. R., Pisoni, D. B., Yamada, R. A., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310.
- Bundgaard-Nielsen, R. L., Best, C. T., Tyler, M. D., 2011. Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowels categories. *Appl. Psycholinguist.* 32, 51-67.
- Clopper, C. G. & Bradlow, A. R. (2009). Free Classification of American English dialects by native and non-native listeners. *Journal of Phonetics*, 37, 436-451.
- Clopper, C. G. & Pisoni, D. P. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111-140.
- Darcy, I., Park, H.-Y., & Yang, C.-L. (under review). Individual differences in L2 acquisition of English phonology: the relation between cognitive abilities and phonological processing. Submitted to *Studies in Second Language Acquisition*.

- Ezzatian, P., Avivi, M., & Schneider, B. A. (2010). Do nonnative listeners benefit as much as native listeners from spatial cues that release speech from masking? *Speech Communication, 52*, 919-929.
- Flege, J. E., Bohn, O. -S., Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470.
- Flege, J.E., & MacKay, I.R.A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 24*, 1–34.
- Ganong, W.F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 110-125.
- García Lecumberri, M.L., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *Journal of the Acoustical Society of America, 119*, 2445–2454.
- Gerstman, L. H. (1968). Classification of self-normalized vowels. *IEEE Transactions on Audio Electroacoustics, AU-16*, 78-80.
- Imai, S., Walley, A.C., & Flege, J.E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America, 117*, 896-907.
- Ingvalson, E.M., McClelland, J.L., & Holt, L.L. (2011). Predicting native English-like performance by native Japanese speakers. *Journal of Phonetics, 39*, 571-584.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.
- Koster, O., & Schiller, N. O. (1997). Different influences of the native language of a listener on speaker recognition. *Forensic Linguistics, 4*, 18-28.

- Koster, O., Schiller, N. O., & Kunzel, H. J. (1995). The influence of native-language background on speaker recognition. In K. Elenius, & P. Branderud (Ed.), *Proceedings of the Thirteenth International Congress of Phonetic Sciences* (vol. 4, pp. 306-309). Stockholm.
- Labov, W., Ash, S., & Boberg. (2006). *Atlas of North American English*. New York: Mouton de Gruyter.
- Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. G. (1993) Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General*, 122, 316-330.
- Lively, S. E., Logan, J. E., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America* , 94, 1242-1255.
- Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III: Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America* , 96, 2076-2087.
- Logan, J. D., Lively, S. E., & Pisoni, D. B. (1981). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* , 89, 874-886.
- Mackay, I.R.A., Meador, D., & Flege, J.E. (2001a). The identification of English consonants by native speakers of Italian. *Phonetica*, 58, 103–125.
- Mackay, I.R.A., Flege, J.E., Piske, T., & Schirru, C. (2001b). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America*, 110, 516–528.

- Meador, D., Flege, J. E., & MacKay, I. R. A. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition*, 3, 55-67.
- Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984) Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No.10 Speech Research Laboratory, Indiana University*, 357–376.
- Oyama, S. (1976). A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research*, 5, 261-283.
- Pisoni, D. B. (2007). WordFam: Rating Word Familiarity in English. Indiana University, Bloomington.
- Pisoni, D.B., Nusbaum, H.C., Luce, P.A., & Slowiaczek, L.M. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, 4, 75-95.
- Samuel, A. G. (1986). The role of the lexicon in speech perception. In E.C. Schwab & H.C. Nusbaum (Eds.), *Perception of Speech and Visual Form: Theoretical Issues, Models, and Research*. New York: Academic Press.
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455-462.
- Sullivan, K. P. H., & Schlichting, F. (2000). Speaker discrimination in a foreign language: First language environment, second language learners. *Forensic Linguistics*, 7, 95-111.
- Tamati, T. N. (2008). Perceptual similarity of unfamiliar regional dialects: Some preliminary findings. In *Research on Spoken Language Processing: Progress Report 29*. Bloomington, IN: Indiana University.
- Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Individual differences in speech perception in adverse listening conditions. *Journal of the American Academy of Audiology*, 24, 616-634.

Tamati, T. N., Park, H.-Y., & Pisoni, D. P. (2011). The development of a new corpus of foreign-accented English. Poster presented at the 2011 VLSP New Tools and Methods for Very Large-Scale Phonetics Research Workshop. Philadelphia, PA.

Tamati, T. N., & Pisoni, D. B. (in press). Non-native speech recognition in adverse listening conditions. *Journal of the American Academy of Audiology*.

Williams, B. W., Mack, W., & Henderson, V. W. (1989). Boston naming test in Alzheimer's disease. *Neuropsychologia*, 27, 1073-1079.

CHAPTER 4
COGNITIVE ABILITIES AND THE PERCEPTION OF REGIONAL DIALECT
VARIATION IN A SECOND LANGUAGE

Introduction

Individual Differences in the Perception of Variability in Speech

Like native listeners, non-native listeners must be able to adapt and use different sources of variation in speech perception to both understand the intended meaning of the utterance and to obtain relevant social information (e.g., the talker's identity, where he/she is from, his/her age). Along with indexical variability, degradation of the speech signal by noise and competition from other talkers presents an additional challenge to successful speech recognition for non-native listeners who lack robust phonological, lexical, and sociolinguistic knowledge of their second language (L2; e.g., Bradlow & Pisoni, 1999; Cooke, Lecumberri, & Barker, 2008). Previous studies have shown that listeners vary greatly in their ability to recognize words in high variability sentences or in difficult listening conditions, and have furthermore suggested that neurocognitive abilities may underlie the large individual differences attested. Recently, Gilbert, Tamati, and Pisoni (2013) found that young adult, normal-hearing native listeners varied substantially in their ability to recognize sentences produced by multiple male and female talkers from different dialect regions of the United States (U.S.) in multitalker babble. In their study, the listeners' accuracy at correctly recognizing key words ranged, from approximately 40% to 76% across four signal-to-noise ratios (SNRs; +3 dB, 0 dB, -3 dB, -5 dB SNR). Their findings demonstrate that even among listeners with hearing thresholds within normal limits, there is a large amount of variability in the ability to recognize speech in high-variability listening

conditions. Although individual differences in speech recognition tasks involving different sources of indexical variability have not been widely studied, previous research has also reported substantial individual variability in listeners' sensitivity and susceptibility to other sources of adverse listening environments and conditions (e.g., McGarr, 1983; Neff & Dethlefs, 1995; Richards & Zeng, 2001).

To examine the underlying sources of individual differences, Tamati, Gilbert, and Pisoni (2013) investigated differences in processing and memory between listeners who had performed well and listeners who had performed poorly on sentences from PRESTO, a new high-variability sentence recognition test, as described in Gilbert et al. (2013). In particular, they measured indexical processing skills related to talker, gender, and regional dialect variation, as well as several neurocognitive abilities related to short-term and working memory capacity, attentional and inhibitory control, executive functioning, and non-verbal IQ in the good and poor listeners. In their study, the good listeners who were able to recognize more words in the PRESTO sentences were better at discriminating talkers by gender and categorizing talkers by region of origin. The good listeners also had greater short-term and working memory capacity, larger vocabulary sizes, and better executive functioning skills in information processing domains related to cognitive load. Findings from the Tamati et al. (2013) study suggest that individual differences in perceptual adaptation and adjustment to indexical variability in speech may be related to the ability to perceive and encode detailed episodic and contextual information in the speech signal. Listeners with better indexical processing skills and stronger neurocognitive abilities involving immediate short-term and working memory and cognitive control may benefit from encoding more robust, highly detailed representations and memory codes in long-term memory for both linguistic and non-linguistic information in speech.

Individual Differences in L2 Speech Perception and Spoken Word Recognition

Even in groups of non-native listeners who have similar native language backgrounds and experiences, individual learners vary greatly in their speech perception abilities in their L2s. As with native listeners, the neurocognitive abilities and associated speech processing skills of these listeners may also influence individual differences in the L2. Tamati and Pisoni (in press) recently examined non-native listeners' speech recognition skills using PRESTO. In their study, non-native listeners were native speakers of Mandarin who had lived in the city of Bloomington, Indiana in the U.S. between 1 and 27 months. Tamati and Pisoni (in press) found that non-native listeners had a great deal of difficulty on the PRESTO sentence recognition task, scoring significantly lower than even the poorest listeners from the Gilbert et al. (2013) and Tamati et al. (2013) studies. Additionally, non-native listeners showed large differences in performance on the PRESTO test with scores ranging from 12.0% to 35.3% mean word recognition accuracy across for SNR conditions (+3, 0, -3, -5 dB SNR). Their performance on PRESTO was found to be related to executive function abilities, as measured by the on the BRIEF-A self-report questionnaire of executive function, suggesting that individual differences in executive function may contribute to individual listeners' ability to understand high-variability speech.

Examining individual differences in L2 speech perception, Darcy and colleagues (2010) assessed the relations between L2 phonological processing skills and neurocognitive abilities using several conventional L2 phonological processing tasks. They found that larger working memory capacity, faster processing and lexical retrieval speed, and executive function were related to better performance on the conventional L2 speech perception tasks, although the type and strength of the relation between neurocognitive abilities and speech perception abilities varied from task to task. Many other studies have also provided evidence that core, underlying

neurocognitive abilities such as attention, inhibition, working memory, and cognitive control may influence speech perception and spoken word recognition in an L2, along with the acquisition of other types of L2 grammatical knowledge (e.g., Salthouse, 1996; Segalowitz, 1997; Miyake & Friedman, 1998). Taken together, studies on both native and non-native speech perception suggest that several underlying core neurocognitive processes play a key role in L2 acquisition and explain individual differences in L2 speech perception abilities.

Individual Differences in L2 Perception of Regional Dialects

In the study described in Chapter 2, non-native listeners completed three perceptual tasks involving regional dialect variability in speech. The non-native listeners displayed great variability in their ability to recognize English words produced by talkers from different dialect regions in the U.S. and in their ability to identify regional dialects. Chapter 3 explored the contributions of L2 phonological and lexical knowledge to individual differences in the perception of L2 regional dialects. In that study, non-native listeners completed two conventional L2 phonological tasks, an ABX vowel and consonant discrimination task, and a vowel identification task. All listeners also filled out a word familiarity questionnaire designed to obtain a measure of receptive vocabulary and an object naming task designed to obtain a measure of expressive vocabulary. Comparing the L2 phonological and lexical knowledge scores to regional dialect perception measures from Chapter 2, the findings suggested that L2 phonological and lexical knowledge contributes to the recognition of words in sentences produced by talkers from multiple U.S. dialect regions in the multi-dialect sentence recognition task. However, there was little relation between L2 phonological and lexical knowledge and the identification of regional dialects in the regional dialect categorization task. The contribution of L2 phonological and

lexical knowledge may be more global, benefiting speech recognition in adverse conditions more generally but not necessarily relating to a specific indexical processing or encoding skill. Thus, although L2 listeners may be able to use knowledge of phonology and vocabulary in the L2 to uncover the linguistic information encoded in an utterance produced by talkers from different L2 dialect regions, other processing factors may underlie the L2 listener's ability to adapt and use L2 regional dialect information in sentence recognition and indexical discrimination or identification tasks.

The Current Study

The current study investigated the contribution of elementary neurocognitive abilities to individual differences in the perception of L2 regional dialect information. The same three listener groups reported on in Chapters 2-3 (NAE = native American English group; KLT = long-term Korean group; KST = short-term Korean group) also participated in the current study. The non-native KLT and KST listeners were native speakers of Korean who had either lived in the U.S. for a long period of time (more than 3 years) or a short period of time (6-18 months). All participants completed a battery of behavioral tasks and self-report questionnaires that assessed short-term and working memory capacity, attention/inhibition, and executive functioning. The neurocognitive measures were compared to scores on the multi-dialect speech recognition task and the forced-choice regional dialect categorization task reported in Chapter 2. Again, it was determined that performance on the regional dialect free classification task described in Chapter 2 was too low to obtain meaningful variation in performance across listeners. Therefore, the scores from the free classification task were not included in the current

study, which focused on individual scores on the multi-dialect speech recognition task and the dialect categorization tasks.

Short-term and working memory capacity may underlie the observed individual differences in the perception of regional dialect variation in the L2. To recognize words produced by male and female talkers from different regions of origin, a listener must devote additional processing resources to be able to encode detailed indexical information encoded in the speech signal (e.g., Martin, Mullennix, Pisoni, & Summers, 1989; Mullennix, Pisoni, & Martin, 1989; Mullennix & Pisoni, 1990). Listeners with larger short-term and working memory capacity may be able to allocate more processing resources to these challenging tasks, resulting in a better ability to use both the linguistic and indexical information in the speech signal. In both the linguistic and indexical perception tasks, listeners must be able to resolve acoustic-phonetic ambiguity to use the indexical and linguistic information in the speech signal; a process that requires holding and manipulating phonetic and phonological memory codes in verbal short-term memory (e.g., Cowan, 1988, 1997; Gathercole & Baddeley, 1993; Nusbaum & Magnuson, 1997; Baddeley et al., 1998; Baddeley, 2003). Devoting more processing resources to encoding detailed indexical information in speech may be especially helpful when a listener is trying to understand unfamiliar regional dialects, because he/she must rapidly learn to perceive and encode reliable dialect-specific patterns in the speech signal. As for the indexical tasks, listeners exposed to regional dialect variation must be able to hold that channel of information in memory, and analyze it to extract cues to dialect membership. Listeners who are able to dedicate additional processing resources to this task may be able to more quickly acquire sociolinguistic knowledge of the L2 regional dialects because they can form more robust phonological and lexical representations of dialects that they can use to facilitate the categorization of regional

varieties. Given these initial considerations, non-native listeners with greater short-term and working memory capacity were expected to better recognize speech produced by talkers from different regions of origin and identify regional dialects in the L2, especially unfamiliar regional dialects, better than non-native listeners with lower short-term and working memory capacity.

To encode and process both the indexical and linguistic properties in the signal, listeners need to be able to simultaneously focus and attend to the vocal source and the linguistic content of the message. More flexible attentional and inhibitory control abilities may help listeners to rapidly adjust to new talkers and regional dialect characteristics. Better attentional and inhibitory skills would enable listeners to reliably attend selectively to the linguistic and indexical information in the target utterance and quickly adjust to new constantly changing talker and linguistic information to both recognize the linguistic content of the utterance and use acoustic-phonetic information in the signal to make reliable judgements about talkers. Additionally, those skills might be especially beneficial to non-native listeners with little exposure to the L2 in naturalistic environments, since they would have the difficult task of dealing with novel linguistic and indexical information encountered on a daily basis. Non-native listeners with greater attentional and inhibitory control abilities were expected to be more accurate at recognizing speech produced by talkers from different regions of origin and identify regional dialects in the L2, especially unfamiliar regional dialects, compared to non-native listeners with poor attentional and inhibitory control processes.

Executive functioning and cognitive control may also contribute to the perception of regional dialect variation in an L2. As mentioned above, individual differences in executive functioning in native and non-native listeners have been found to be associated with speech perception skills, including speech recognition in adverse listening conditions (Beer,

Kronenberger, & Pisoni, 2011; Tamati et al., 2013; Tamati & Pisoni, in press). Executive function is heavily involved in a wide range of daily activities because it is thought of as the cognitive control system that manages other cognitive processes, including working memory and attentional and inhibitory process (Barkley, 1997, 2012). Executive functioning skills may be crucial for learning new patterns of sociolinguistic information in an L2. These skills are involved in the rapid adaptation and adjustment to new or changing information, focusing or changing attention to different sources of information, and allocating processing resources to tasks or behaviors. Successful and robust adaptation to the indexical and linguistic information in the speech signal may draw heavily on these processes.

For non-native listeners, recognizing words in sentences produced by multiple talkers who come from several different dialect regions would be especially difficult because they lack exposure and knowledge of the L2 phonological system as well as L2 sociolinguistic information. To complete both the linguistic and the indexical processing tasks, listeners must be able to successfully control their attention to indexical information in the signal to acquire knowledge of the dialect-specific characteristics of regional dialects of American English. Furthermore, listeners must be able to control their attention to the relevant linguistic-symbolic information in the speech signal, while simultaneously adapting to the changing talkers, genders, and regional dialects. This may be particularly important with unfamiliar regional dialects that present the listeners with novel variability to which they must attend and learn. Thus, non-native listeners with stronger executive functioning abilities were expected to recognize words produced by talkers from different regions of origin and identify regional dialects in the L2, especially unfamiliar regional dialects, better than non-native listeners with weak or compromised executive functioning skills.

To summarize, the goal of the current study was to investigate the contributions of several core neurocognitive abilities to individual differences in L2 perception of regional dialect variation in the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task observed in Chapter 2. Given previous findings on individual differences in speech perception, L2 listeners who were more accurate at recognizing words in the multi-dialect sentence recognition task and more accurate at identifying American English regional dialects in the dialect categorization task were expected to show greater short-term and working memory capacity, better control of attention and inhibition, and stronger executive functioning skills. Thus, basic differences in neurocognitive abilities were predicted to be related to the processing of regional dialect information in the L2.

Methods

Listeners

Fifty-nine young adults participated in the current study. All of the listeners had participated in the studies described in Chapter 2 and Chapter 3. Twenty-one participants were native speakers of American English (native American English group = NAE), 23 participants were native speakers of Korean who had lived in the U.S. for approximately three years or more (long-term Korean group = KLT), and 15 participants were native speakers of Korean who had lived in the U.S. for approximately one year or less (short-term Korean group = KST). The 21 NAE participants included 15 females and 6 males, aged 18-27 years ($M = 21.0$; $SD = 1.8$ years). All NAE participants grew up in a region of the U.S. where General American English is spoken, i.e., North Midland, West, and parts of New England. The 23 KLT participants included 8 females and 15 males, aged 19-32 years ($M = 24.0$; $SD = 3.9$ years). All KLT participants were

native speakers of Korean and had spent 3 years or more in the U.S. (LOR, $M = 5.1$; $SD = 1.5$ years). The 14 KST participants included 9 females and 6 males, aged 19-32 years ($M = 25.3$; $SD = 4.3$ years). All KST participants were native speakers of Korean and had spent a total of approximately 0.5 to 1.5 years in the U.S. (LOR, $M = 0.8$; $SD = 0.4$ years).

All participants reported normal hearing with no significant history of any prior hearing or speech disorders at the time of testing. They received \$20 for 120 minutes of participation (rate of \$5 per 30 minutes), which included completing several other speech perception and production tasks, neurocognitive tasks, and self-report questionnaires.

Materials and Procedures

Procedures were identical for each of the three participant groups except where explicitly mentioned. Participants were seated in front of a personal computer equipped with headphones (Beyerdynamic DT100), a keyboard, and a mouse. Participants were tested individually in a quiet testing room, where they completed a series of computer-based perceptual tasks, neurocognitive tasks, and self-report questionnaires. The output level of the target sentences for all the perception tasks were calibrated to be approximately 65 dB SPL. Procedures and materials for all three tasks are described below.

Neurocognitive Measures

Visual Forward and Backward Digit Span Task: Participants completed two visual digit span tasks to measure short-term and working memory capacity. The visual span tasks were used instead of conventional auditory tasks in order to reduce the possible effects of L2 linguistic experience on performance. Native speakers of Korean use Arabic numerals in their native

context. In the visual digit span task, digits (e.g., “1”, “2”) were presented one-at-a-time in the middle of a computer screen. Each digit was presented for 1 s. After each digit, the computer screen was blank for 500 ms before the next digit appeared. The list length of digit sequences ranged from two to nine items in length. The maximum list length of the forward span task was 8 digits; the maximum list length of the backward span task was 7 digits. Lists were constructed from the digit span tests used in the Wechsler Memory Scale 3rd Edition (Wechsler, 1997). After seeing a list of digits, participants were required to type in the sequence of digits in the correct order on a keyboard as they were instructed (either in forward order for the Digit Span Forward task or in backward order for the Digit Span Backward task). Both digit span tasks started with a list of two numbers. Each list length is presented twice. The task continued until the lists were exhausted (Forward: length of 8 digit sequences, total 16 possible trials; Backward: length of 7 digit sequences, total 14 possible trials). The participants were required to reproduce all of the digits in the correct order for the trial to be counted as correct. Responses were scored offline for correctness. The longest list length correct, i.e., the longest list before two incorrect responses on the same list length was tallied and used for analysis.

STROOP Paper and Pencil Task: A paper and pencil version of the Stroop Color and Word Test (Golden, 1978; Golden & Freshwater, 2002) was used to obtain measures of attention and inhibition. The standard STROOP test is divided into three parts. In the first part (the Word reading condition), participants are instructed to read a list of color words (‘blue’, ‘green’, ‘red’) typed with black ink aloud as quickly as possible. Any mistakes could be corrected by returning to the word and reading it over again. Participants have 45 seconds to read as many words aloud as possible. In the second part (the Color naming condition), participants are instructed to name the color of three *XXXs* written with blue, green, or red ink in a list aloud, going through as many

as possible in 45 seconds. In the third part (the Color-Word naming condition), participants are required to name the color of the printed ink for the color words ('blue' 'green', 'red') in a list. In this condition, the color of the ink does not match the color word (e.g., the color word 'blue' is printed in red ink). Again, participants have 45 seconds to go through as many color-word trials in the list as possible, while correcting any errors made by repeating the item.

Each part of the test was scored by counting the number of items that the participant was able to produce in the 45 seconds. Raw Word reading, Color naming, and Color-Word naming scores were calculated from the number of items correctly produced during each part of the test. Raw STROOP Interference scores were then calculated by subtracting a predicted Color-Word score ($\text{Word} \times \text{Color} / \text{Word} + \text{Color}$) from the raw Color-Word score (Golden, 1978).

Trail Making Task: To measure of executive functioning and processing speed, participants completed a version of the paper and pencil Trail Making Test (Army Individual Test Battery, 1944). The Trail Making Test, which is comprised of two parts, Part A and Part B, is used as a measure of visual search and processing speed and visual search abilities (Part A), working memory and task-switching ability (Part B), and executive control abilities (Parts A and B) (e.g., Sánchez-Cubillo et al., 2009). In Part A of the task, after a short practice set of 8 items, the participant is required to connect 25 dots containing numbers 1-25 in sequential order as quickly as possible without compromising accuracy. In Part B, after a short practice set of 8 items, the participant is asked to connect 25 dots containing numbers 1-13 and letters A-L, alternating between numbers and letters (i.e., 1-A-2-B-3-C etc.). In both Part A and B, the participant was instructed that if he/she made a mistake, he/she should correct the mistake before moving on. The experimenter recorded the time it took the participant to complete both parts of

the test using a hand-held stopwatch. The time to complete Part A and Part B, along with a difference score (Part B – Part A) were calculated and analyzed.

BRIEF-A: To obtain a measure of executive functioning in everyday real-world situations outside the clinic or laboratory, participants also completed the Behavioral Rating Inventory of Executive Function – Adult Version (BRIEF-A; Roth, Isquith, & Gioia, 2005; Psychological Assessment Resources). The BRIEF-A is a self-report questionnaire that consists of 75 statements about executive functions in everyday life to which participants respond if the issue or behavior has been problematic for them: never (1 point), sometimes (2 points), or often (3 points) in the past six months. Nine clinical domains of executive functions (Shift, Inhibit, Emotional Control, Self-Monitor, Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials), two aggregate indexes (Behavioral Regulation Index (BRI) and a Metacognitive Index (MI)), and an overall global General Executive Composite (GEC) score are created from the individual test items on the BRIEF-A questionnaire. A summary of the nine clinical domains and composite scores is given in Table 1.

For data analysis, raw scores collected from the responses were compared across groups and individuals for the GEC, BRI, and MI composite scores and each of the nine clinical scales. Note that higher scores on the BRIEF-A reflect greater disturbances in a particular executive function domain.

Table 1: Description of measures from the BRIEF-A self-report questionnaire on executive functioning.

Score	General Description of Behaviors Assessed
Global Executive Composite (GEC)	<i>Overall measure of executive functioning in everyday life based on the aggregate of scores on all nine individual scales</i>
Behavior Regulation Index (BRI)	<i>Measure of the ability to appropriately change or adapt their emotional behavior, demonstrating good inhibitory control, based on the aggregate score from the Shift, Inhibit, and Emotional Control scales</i>
Shift	<i>Moving or changing from one situation, topic, or task to another</i>
Inhibit	<i>Inhibiting or stopping oneself from acting on impulse in different situations</i>
Emotional Control	<i>Controlling or modulating emotional behavior</i>
Metacognitive Index (MI)	<i>Measure of the ability to appropriately manage and complete different tasks and actions based on the aggregate score from the Self-Monitor, Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials scales</i>
Self-Monitor	<i>Evaluation of how one's own actions or behaviors affect others</i>
Initiate	<i>Starting a new task or generate new, independent thoughts or ideas</i>
Working Memory	<i>Holding information in memory for completing, or taking the necessary steps to complete, a task or goal</i>
Plan/Organize	<i>Management of information to complete current tasks or in anticipation of future tasks</i>
Task Monitor	<i>Evaluation of one's own actions or behaviors during or after a task or activity</i>
Organization of Materials	<i>The orderliness or organization of one's belongings and space or actions</i>

Results

Individual Differences: Neurocognitive Abilities and L2 Perception of Regional Dialect Variation

To assess the relations between the neurocognitive measures and the perception of regional dialect in an L2, a series of stepwise multiple linear regression analyses were used to identify factors that predict performance on the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task summarized in Chapter 2. Scores calculated for each of the measures described above were used as predictor variables in the regression analyses.

These measures included:

- Forward Digit Span: Longest consecutive list length correctly reproduced, i.e., the longest list before making two incorrect responses on the same list length
- Backward Digit Span: Longest consecutive list length correctly reproduced in reverse order, i.e., the longest list before making two incorrect responses on the same list length
- STROOP Paper and Pencil Task: Color, Word, Color-Word, and the calculated Interference scores
- Trail Making Task: Part A duration, Part B duration, Difference Score (B-A)
- BRIEF-A: General Executive Composite (GEC) score

For the purposes of the individual differences analyses, KLT and KST groups were combined into a single non-native group. Separate analyses were also carried out with the native NAE control group. These analyses were carried out in order to determine if more general language-independent measures contribute to the perception of regional dialect information.

Multi-Dialect Sentence Recognition

A stepwise linear regression analysis was carried out to assess the relation between multi-dialect sentence recognition and any core, underlying neurocognitive abilities assessed in this study. Given that sentence recognition scores for each of the four regional dialects (General American, Northern, Mid-Atlantic, Southern) at the two SNRs (+2 dB SNR, -2 dB SNR) were highly correlated, the overall mean accuracy score across all conditions was first used as the dependent variable. The analysis of the scores obtained for the native NAE listeners revealed no significant influence of the neurocognitive factors on multi-dialect sentence recognition in their native language. For the non-native KLT and KST listeners, only the STROOP Color-Word naming score emerged as a significant predictor of sentence recognition performance, explaining 25.9% of the variance [$R^2 = .26$, $\beta = .51$, $F(1, 32) = 12.23$, $p = .001$]. None of the other measures emerged as significant.

Similar analyses were also carried out on sentence recognition accuracy for each regional dialect (General American, North, Mid-Atlantic, South) to assess whether different sets of skills may be related to recognizing speech produced by talkers from different unfamiliar or familiar regional dialects. For the General American talkers, only the STROOP Color-Words naming score emerged as significant, explaining 19.4% of the variance [$R^2 = .19$, $\beta = .44$, $F(1, 35) = 8.42$, $p = .006$]. For the Northern talkers, the STROOP Color-Word naming score was a significant predictor, explaining 31.4% of the variance [$R^2 = .31$, $\beta = .56$, $F(1, 35) = 16.0$, $p < .001$]. For the Mid-Atlantic talkers, the STROOP Color-Word naming score was also a significant predictor, explaining 24.6% of the variance [$R^2 = .25$, $\beta = .50$, $F(1, 35) = 11.41$, $p = .002$]. Finally, for the Southern talkers, again the STROOP Color-Word variable was a significant predictor of sentence recognition performance, explaining 22.6% of the variance [$R^2 = .23$, $\beta = .48$, $F(1, 35) = 10.24$, p

= .003]. Figure 1 displays the relation between mean overall sentence recognition accuracy scores and the STROOP Color-Word naming score for listeners in all three listener groups (NAE, KLT, KST).

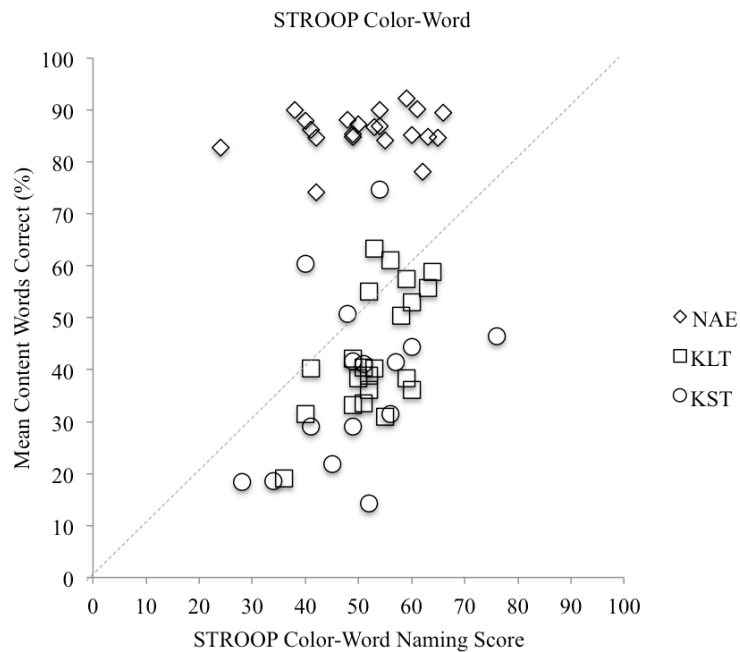


Figure 1. STROOP Color-Word naming scores (x-axis) and percent correct word recognition accuracy on the multi-dialect sentence recognition task (y-axis).

Forced-Choice Regional Dialect Categorization

A series of linear regression analyses were also used to assess the relations between regional dialect categorization accuracy and neurocognitive abilities. Given that categorization accuracy varied by regional dialect (General American, Northern, Mid-Atlantic, Southern), separate analyses were carried out to assess performance on each of the regional dialects by the non-native KLT and KST listeners. For the native NAE listeners, the analyses revealed no significant influence of the neurocognitive factors on the categorization of the General American

and the Southern talkers. For the Northern talkers, however, two measures were significant predictors, accounting for 44.6% of the variance [$R^2 = .45$, $F(2,18) = 7.24$, $p = .005$].

Categorization accuracy for the Northern talkers was related to the STROOP Color naming score [$\beta = .91$, $p = .001$] and Forward Digit Span [$\beta = -.69$, $p = .010$]. For Mid-Atlantic talkers, only the STROOP Color-Word naming score emerged as significant, explaining 21.1% of the variance [$R^2 = .21$, $\beta = -.46$, $F(1,19) = 5.12$, $p = .036$].

For the non-native KLT and KST listeners, none of the variables emerged as significant predictors of regional dialect categorization performance for both the General American and the Northern talkers. For Mid-Atlantic talkers, only the BRIEF-A GEC score emerged as a significant predictor in the model, explaining 12.5% of the variance [$R^2 = .13$, $\beta = -.35$, $F(1, 35) = 4.99$, $p = .032$]. To further explore the relation between executive functioning and the categorization of the Mid-Atlantic talkers, correlations were carried out between categorization accuracy and the two BRIEF-A aggregate scores, the Metacognitive Index (MI) and Behavioral Regulation Index (BRI). Correlational analyses revealed that categorization accuracy for the Mid-Atlantic talkers was marginally correlated with both the MI aggregate score $r = -.32$, $p = .024$, and with the BRI aggregate score [$r = -.32$, $p = .026$]. All reported correlations are one-tailed, given the predication that participants with better executive functioning abilities would be better able to rapidly adapt and adjust to variability in the PRESTO sentences. For Southern talkers, the STROOP Interference score was a significant predictor, explaining 13.9% of the variance [$R^2 = .14$, $\beta = -.37$, $F(1, 35) = 5.63$, $p = .023$]. Figure 2 displays the relation between categorization accuracy for the Mid-Atlantic talkers and the BRIEF-A GEC score (top) and the relation between categorization accuracy for the Southern talkers and the STROOP interference score (bottom) for listeners in all three listener groups (NAE, KLT, KST).

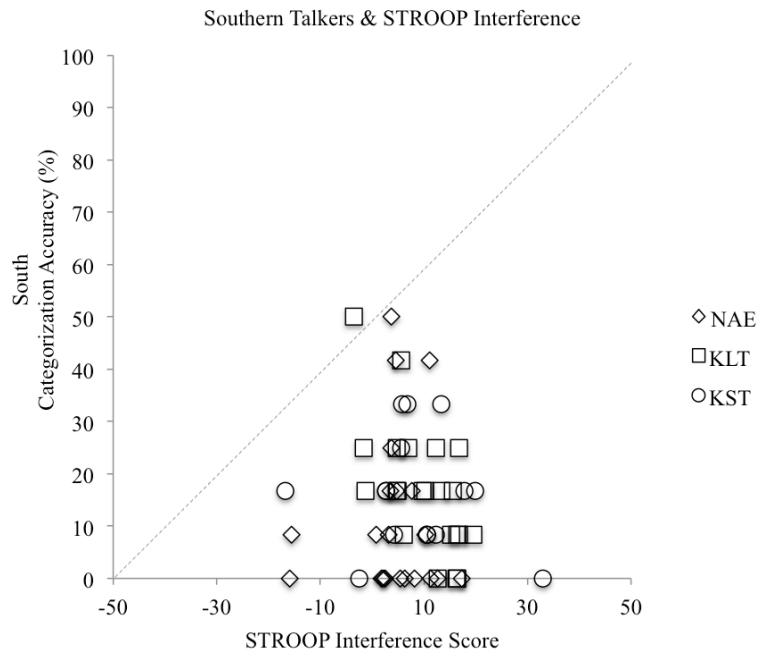
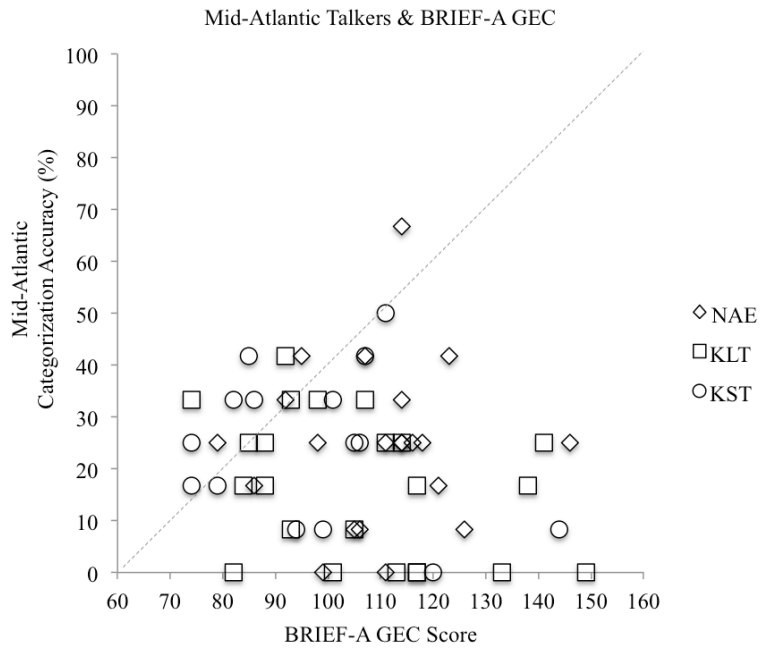


Figure 2. The BRIEF-A GEC score (x-axis) and percent correct categorization accuracy on the Mid-Atlantic talkers (y-axis) (top); and the STROOP Interference score (x-axis) and percent correct categorization accuracy on the Southern talkers (y-axis) (bottom).

Discussion

To investigate the relation between elementary, underlying neurocognitive abilities and the perception of L2 regional dialect information, the current study examined individual differences in recognizing words in sentences produced by talkers from different U.S. dialect regions and identifying American English regional dialects. Non-native speakers of American English from South Korea completed several behavioral tasks and self-report questionnaires assessing short-term and working memory capacity, attention/inhibition, and executive functioning. These measures were compared to scores on the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task from Chapter 2.

Results of the individual differences analyses revealed that a single measure of attention and inhibition was a significant predictor of multi-dialect sentence recognition accuracy. The STROOP Color-Word naming score was related to the overall mean word recognition accuracy. Examining performance on each talker dialect for consistent contributors to accuracy on the task, the Color-Word score also emerged as a significant predictor of word recognition accuracy for the General American, Northern, Mid-Atlantic, and Southern talkers. The consistency of the Color-Word score across the four regional dialects suggests that the skills required to correctly say the color in the presence of conflicting text is related to the recognition of words in sentences produced by multiple female and male talkers from different U.S. dialect regions in multitalker babble. Given that the subject must inhibit responding to the color word, while producing the name of the color of the ink, the Color-Word portion of the STROOP task may have both attentional and inhibitory control components as well as a response organization and speech production component. Similarly, the sentence recognition task requires the listener to focus attentional and control resources on the target sentence and talker while simultaneously ignoring

the competing background talkers to reproduce the target sentence. Furthermore, variability from novel unfamiliar regional dialects may place greater processing demands on the listener (Johnsrude et al., 2013). Based on findings from the current study, these demands may be better met by listeners who have developed better attentional and inhibitory abilities. Findings from previous studies have also suggested that attentional and inhibitory abilities play a role on various aspects of L2 acquisition (e.g., Lev-Ari & Peperkamp, 2013; Segalowitz & Frenkiel-Fishman, 2005) that involve the perception of individual L2 sounds or contrasts. The findings from the current study, taken together with the findings from the previous studies, suggest that control of attention and inhibition contributes to the encoding of detailed acoustic-phonetic information in the signal, crucial for the successful discrimination and/or identification of novel sounds and L2 lexical access.

Similar analyses were also carried out for the non-native listeners' scores on the forced-choice regional dialect categorization observed in Chapter 2. The results of the regression analyses on the categorization accuracy scores were less conclusive, and the relations between categorization accuracy and neurocognitive measures varied by dialect. No score emerged as a significant predictor of categorization accuracy for the General American and the Northern talkers. Executive functioning skills were related to the categorization of the Mid-Atlantic talkers in the forced-choice regional dialect categorization task. Lower scores on the BRIEF-A reflect greater disturbances in a particular executive function domain, so the negative relation between the BRIEF-A GEC and categorization accuracy for the Mid-Atlantic talkers shows that non-native listeners with stronger executive functioning skills were more accurate at categorizing the Mid-Atlantic talkers. Better attentional and inhibitory control skills were related to less accurate categorization of Southern talkers. The relation between attentional and inhibitory control skills

and poorer categorization of Southern talkers may seem unexpected because stronger neurocognitive skills were predicted to contribute to more accurate categorization performance on any of the four regional dialects. However, as discussed in Chapter 2, non-native categorization of the Southern talkers was negatively correlated with multi-dialect sentence recognition accuracy. Furthermore, native NAE listeners were also less accurate on Southern talkers than non-native listeners and more accurate on General American talkers. Thus, the non-native listeners who were less accurate on the Southern talkers were showing more native-like response patterns for that dialect, as well as more native-like sentence recognition. Having better attentional and inhibitory control skills may help non-native listeners to selectively attend to and use the regional dialect variation in the target utterance to be able to acquire sociolinguistic knowledge of regional dialects and regional stereotypes. Interestingly, some of the neurocognitive measures were also related to native NAE regional dialect categorization performance. NAE listeners with greater STROOP Color naming scores and larger Forward Digit Span scores were also more accurate at categorizing the Northern talkers. Additionally, NAE listeners with greater STROOP Color-Word naming scores were more accurate at categorizing the Mid-Atlantic talkers. Although the patterns of results for the categorization task were not as consistent as those for the sentence recognition task, core, elementary cognitive processes also play a role in non-native and native categorization of unfamiliar American English regional dialects.

Neurocognitive abilities may be involved in regional dialect categorization in a variety of ways. First, a listener with better attentional and inhibitory control skills and executive functioning may be better at shifting their attention to the indexical information encoded in the speech signal when this channel of information is relevant to the communicative environment.

Additionally, listeners with strong attentional and inhibitory control and executive function skills may also be able to ignore other sources of competition and variability in the speech signal, such as variability in the linguistic content of the utterance, to form more robust representations of L2 regional dialect variation. Similarly, within the task itself, those listeners may also more easily adapt to the changing linguistic content and/or additional talker variability from trial-to-trial to be able to focus their attention on the dialect-specific cues present in a particular trial to make use of this source of context to facilitate regional dialect categorization.

Other neurocognitive abilities tested in the current study were not related to individual differences in performance on the sentence recognition task or the indexical tasks. Results of the sentence recognition task were surprising given that previous studies have found that underlying, elementary neurocognitive abilities, such as working memory capacity and executive functioning, influence individual listeners' ability to understand high-variability speech (Tamati & Pisoni, in press; Tamati et al., 2013). However, as reported in Chapter 3, performance on the multi-dialect sentence recognition task was strongly related to L2 phonological and lexical knowledge. While there may also be a relation between L2 phonological and lexical knowledge and elementary neurocognitive abilities, it may be that for this particular group of listeners that the differences in knowledge of the L2, regardless of how long they were in the U.S., were the main contributors to high-variability sentence recognition. In that case, large differences in knowledge of the L2 within the group of listeners would account for much of the variability in sentence recognition performance, compared to the rather minor differences in the range of neurocognitive abilities in highly educated, typically developed young adults.

For the regional dialect categorization task, it is still unclear what factors contribute to individual differences in performance. Although the findings from the present study suggest that

attentional and inhibitory control and executive functioning skills may be related to the perceptual categorization of regional dialect variation, these measures account for very little additional unique variance observed in performance. Some of the patterns observed for both the linguistic and indexical tasks may be accounted for by the tasks selected for the current study. Several tasks did not yield much variability in performance among the non-native listeners. Furthermore, the test battery used in the current study was limited and only covered a narrow range of information processing skills. Future studies on individual differences in L2 speech perception should include other performance-based tasks assessing neurocognitive performance measures, such as cognitive control processes used in attention/inhibition, short-term and working memory, and processing speed. Additionally, other sources of indexical information, such as foreign accent, talkers' voices, or emotions in speech, could also be investigated to provide converging evidence of the contribution of cognitive control processes and executive function in the perception of indexical information in speech.

Conclusions

The goal of the current study was to investigate the contributions of neurocognitive abilities to individual differences observed in the processing of regional dialect variation in a multi-dialect sentence recognition task and a regional dialect categorization task. The results revealed that attentional and inhibitory skills were related to individual differences in non-native listeners in sentence recognition performance across four American English talker dialects. Given findings from Chapter 3, L2 phonological and lexical knowledge likely plays a greater role in sentence recognition, at least among this group of highly educated, typically developed young adults. These findings suggest that difficulties in processing high-variability speech are

largely related to lack of knowledge and exposure to sociolinguistic variation in the L2.

Attentional and inhibitory control, along with executive functioning, were also found to be related to individual differences in the ability to identify the region of origin of an unfamiliar talker in a forced-choice regional dialect categorization task. Thus, attentional and inhibitory control and executive function skills may contribute to the perceptual analysis of the fine-grained acoustic-phonetic differences. However, the amount of variability accounted for in performance on the regional dialect categorization task by the neurocognitive measures was relatively minor. Nevertheless, taken together, findings from the linguistic and indexical perception tasks suggest that neurocognitive abilities, in particular attentional and inhibitory abilities and cognitive control strategies, play a role in adapting and adjusting to fine-grained acoustic-phonetic regional dialect differences. Future research should be carried out using additional experimental methods and other populations of non-native listeners to further explore other neurocognitive and perceptual factors underlying individual differences in the acquisition, perception, and use of indexical information in speech.

References

- Army Individual Test Battery (1944). *Manual of directions and scoring*. Washington, D.C.: War Department, Adjutant General's Office.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158-173.
- Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*, 189-208.
- Barkley, R.A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65-94.
- Barkley, R.A. (2012). *Executive Functions: What They Are, How They Work, and Why They Evolved*. New York: The Guilford Press.
- Beer, J., Kronenberger, W.G., & Pisoni, D.B. (2011). Executive function in everyday life: Implications for young cochlear implant users. *Cochlear Implants International*, *12(Suppl. 1)*, S89-S91.
- Bradlow, A.R., & Pisoni, D.B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America*, *123*, 414-427.
- Cooke, M. Garcia Lecumberri, M.L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *Journal of the Acoustical Society of America*, *106*, 2074-2085.
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163-191.

- Cowan, N. (1997). *Attention and memory: An integrated framework*. New York: Oxford University Press.
- Darcy, I., Park, H.-Y., Yang, C.-L., & Gleiser, A. (2012, March). Individual differences in the development of L2 phonological processing: The contribution of cognitive abilities and executive function. Poster presented at the Conference on Sources of Individual Linguistic Differences, Ottawa, Canada.
- Gathercole, S.E., & Baddeley, A.D. (1993). *Working memory and language*. Mahway, NJ: Lawrence Erlbaum Associates.
- Gilbert, J.L., Tamati, T.N., & Pisoni, D.P. (2013). Development, Reliability and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *Journal of the American Academy of Audiology, 24*, 26-36.
- Golden, C.J. (1978). *Stroop Color and Word Test: A manual for clinical and experimental uses*. Wood Dale, IL: Stoelting Co.
- Golden, C.J., & Freshwater, S.M. (2002). *The Stroop Color and Word Test: Revised examiner's manual*. Wood Dale, IL: Stoelting Co.
- Johnsrude, I.S., Mackey, A., Hakyemez, H., Alexander, E., Trang, H.P., & Carlyon, R.P. (2013). Swinging at a cocktail party: Voice familiarity aids speech perception in the presence of a competing voice. *Psychological Science, 24*, 1995-2004.
- Martin, C.S., Mullennix, J.W., Pisoni, D.B., & Summers, W.V. (1989) Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 676-681.
- McGarr, N.S. (1983). The intelligibility of deaf speech to experience and inexperienced listeners. *Journal of Speech and Hearing Research, 26*, 451-458.

- Miyake, A., & Friedman, N.P. (1998). Individual differences in second language proficiency: working memory as language aptitude. In Healy, A.F., & Bourne, Jr., L.E. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339-364). Mahawah, N.J.: Lawrence Erlbaum.
- Mullennix, J.W., Pisoni, D.B., & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, 85, 365-378.
- Mullennix, J.W., & Pisoni, D.B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics*, 47, 379-390.
- Neff, D.L., & Dethlefs, T.M. (1995). Individual differences in simultaneous masking with random-frequency, multicomponent maskers. *Journal of the Acoustical Society of America*, 98, 125–134.
- Nusbaum, H.C., & Magnuson, J.S. (1997). Phonetic constancy as a cognitive process. In Johnson, K., & Mullennix, J.W. (Eds.), *Talker Variability in Speech Processing* (pp. 109-132). San Diego: Academic Press.
- Richards, V.M., & Zeng, T. (2001). Informational masking in profile analysis: Comparing ideal and human observers. *Journal of the Association for Research in Otolaryngology*, 02, 189-198.
- Roth, R., Isquith, P., & Gioia, G. (2005) BRIEF-A: Behavioral Rating Inventory of Executive Function – Adult Version: Professional Manual. Lutz, FL, Psychological Assessment Resources.
- Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403-428.

- Segalowitz, N. (1997). Individual differences in second language acquisition. In de Groot, A., & Kroll, J.F. (Eds.), *Tutorials in bilingualism* (pp. 85-112). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sánchez-Cubillo, I., Periáñez, J.A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., & Barceló, R. (2009). Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society, 15*, 438-450.
- Tamati, T.N., & Pisoni, D.B. (in press). Non-native Listeners' Recognition of High-Variability Speech Using PRESTO. *Journal of the American Academy of Audiology*.
- Tamati, T.N., Gilbert, J.L., & Pisoni, D.B. (2013). Individual differences in speech perception in adverse listening conditions. *Journal of the American Academy of Audiology, 24*, 616-634.
- Wechsler, D. (1997) *Wechsler Memory Scale—Third Edition (WMS-III) administration and scoring manual*. San Antonio, TX: Harcourt Assessment.

CHAPTER 5

REGIONAL DIALECT CATEGORIZATION TRAINING: GROUP AND INDIVIDUAL PATTERNS OF LEARNING

Introduction

In everyday situations, listeners will often encounter difficult environments that impede successful speech communication. While real-world listening conditions can involve background noise or competition from other talkers, natural variability in speech related to talker and group differences encoded in the speech signal also plays an important role in speech perception. Listeners can use this information to make judgments about the talker, such as his/her identity (e.g., Van Lancker, Kreiman, & Emmorey, 1985; Van Lancker, Kreiman, & Wickens, 1985), gender (e.g., Lass, Hughes, Bowyer, Waters, & Bourne, 1976), age (e.g., Ptacek & Sander, 1966), and region of origin (e.g., Clopper & Pisoni, 2004c; Van Bezooijen & Gooskens, 1999). However, indexical variability may also present a challenge to successful speech recognition. A real-world environment involving recognition of speech from multiple talkers with different language backgrounds may contribute to a more complex and challenging environment because it requires listeners to repeatedly adjust to and use the indexical information in the speech signal. Individual talkers vary substantially in their inherent intelligibility (Bradlow, Toretta, & Pisoni, 1996). Beyond individual talker characteristics, speech produced by talkers from unfamiliar or marked dialect regions is often more difficult to recognize and comprehend (e.g., Mason, 1946; Labov & Ash, 1997; Clopper & Bradlow, 2008), and foreign-accented speech is also less intelligible than speech from native speakers (e.g., Bradlow & Pisoni, 1999). Thus, a speaker of a language, including second language (L2) learners, must be able to use previous knowledge and

experience to rapidly adapt to and use indexical variability in the speech signal for successful communication in real-world environments.

Perception of Regional Dialect Variation by Native Listeners

Perceiving and using detailed indexical information about talkers' region of origin may be particularly challenging for non-native listeners, who have imperfect knowledge of the L2 and less experience with L2 sociolinguistic variation. However, regional dialect variation would likely be important for non-native listeners living in an L2 environment. Both native and non-native listeners often communicate with talkers of diverse regional backgrounds, whose speech patterns may contain several dialect-specific acoustic-phonetic features characteristic of their region of origin. Studies on the perception of regional dialect variation by native listeners have shown that they have knowledge of regional dialect variation and can use their knowledge to successfully perceive this source of information in speech communication tasks. Indeed, native listeners are able to explicitly categorize talkers by geographical region of origin (e.g., Clopper & Pisoni, 2004c; Van Bezooijen & Gooskens, 1999), group talkers by region of origin (e.g., Clopper & Pisoni, 2007), and make explicit judgments about the similarity of talkers' voices based on regional dialect (e.g., Clopper, Levi, & Pisoni, 2006). Findings from these studies demonstrate that listeners have knowledge of dialect-specific acoustic-phonetic differences in their native language, and they are able to use this knowledge to make reliable perceptual judgments about talkers based on broad regional dialect categories in American English.

However, listeners vary greatly in their ability to recognize dialect-specific features and use them in speech perception tasks. This ability appears to be partially mediated by the prior linguistic experience and developmental history of the listener. Greater familiarity with regional

varieties (for example, by having lived in many regions or having lived near more than one region) leads to easier identification and greater perceived distinctiveness of the familiar varieties and other varieties (e.g., Clopper & Pisoni, 2004b; Clopper & Pisoni, 2004c; Clopper & Pisoni, 2006). Similarly, previous studies have shown that familiar or standard dialects are more intelligible, and even native listeners may have some difficulty in recognizing speech produced by a talker from an unfamiliar or nonstandard dialect region. For example, Labov and Ash (1997) found that listeners benefited from being from the same region as the talkers, since listeners who shared a dialect with the talkers were generally more accurate in a transcription task than listeners who did not share a dialect with the talkers (see also Mason, 1946). Listeners are not only more accurate when presented with talkers from their own geographical regions in speech recognition and comprehension tasks, but they are also more accurate when presented with talkers of standard or supraregional varieties. For example, Clopper and Bradlow (2008) found that talkers from General American dialect regions were the most intelligible for native speakers of American English of all dialect backgrounds in a sentence recognition task. Floccia, Goslin, Girard, and Konopczynski (2006) also found that listeners were faster to respond in a lexical decision task to talkers from the same region and to talkers who spoke the standard Parisian dialect, compared to unfamiliar regional French dialects. These studies suggest that listeners display greater facilitation in recognizing speech produced by talkers from familiar local and standard dialect regions in their native language, but they may also have difficulty in recognizing speech produced by talkers from unfamiliar or nonstandard dialect regions.

Perception of Regional Dialect Variation by Non-native Listeners

Being able to effectively perceive and use dialect-specific acoustic-phonetic information in speech communication would be beneficial in developing native-like communicative skills in an L2. However, because early linguistic experience plays a crucial role in being able to deal with regional dialect variation, non-native listeners would likely be at a disadvantage compared to native listeners. Beyond basic issues of language proficiency and knowledge, L2 learners also have not been exposed to regional dialects in the L2 to the same extent as a native speaker. L2 learners do not have detailed knowledge of dialect-specific differences, and their limited knowledge would make regional dialect recognition or discrimination tasks especially difficult (Clopper & Bradlow, 2009). Although most studies on L2 acquisition have focused on the perception of individual speech sounds or sound categories, only a few studies have been carried out on the perception of indexical information in the speech signal. Findings from these studies have suggested that non-native listeners are sensitive to dialect-specific information and are capable of accommodating to and learning new patterns of regional dialect variation in a second language (e.g., Clopper & Bradlow, 2009; Ikeno & Hansen, 2007). However, non-native listeners are not as accurate as native listeners and likely rely on different perceptual features in making their decisions.

Clopper and Bradlow (2009) found that, although non-native listeners were fairly accurate in grouping similar talkers together by regional dialect using a free classification task in their L2 (English), native speakers were overall more accurate. Examining native and non-native performance, Clopper and Bradlow (2009) also pointed out that non-native listeners relied more heavily on talker-specific acoustic-phonetic differences rather than knowledge of how different features group together to index dialects especially when there was considerable within-dialect

variability. The non-native listeners likely did not have robust knowledge of the distributions of dialect-specific features that could be used to make reliable perceptual similarity judgments about a talker's region of origin.

Other studies have suggested that non-native listeners can improve their ability to perceive and use dialect-specific acoustic-phonetic information to make judgements about a talker's region of origin. Eisenstein (1982) found that beginner and intermediate learners of English in New York City were less accurate than native listeners in a dialect discrimination task with several types of American English dialects, but the most proficient learners performed similarly to native listeners. Results from Stephan (1997) also suggested that L2 learners might be capable of gaining greater sensitivity to multiple L2 regional dialects with exposure and experience. In their study, native German learners of English were much better at identifying more familiar or salient varieties of English, such as Southern American English, although they were very poor at identifying other less familiar varieties of English, such as South African English. Taken together, the Eisenstein (1982) and Stephen (1997) studies suggest that, although non-native listeners may have more initial difficulty in discriminating and identifying regional dialects, greater experience and exposure to these varieties may lead to better performance.

However, the findings reported in Chapter 2 indicate that acquiring explicit knowledge of dialect-specific acoustic-phonetic differences in an L2 may be quite difficult, even after many years in the L2 environment. In a regional dialect categorization task, non-native listeners who had lived in the United States (U.S.) for three years or more (long-term Korean group = KLT) and non-native listeners who had lived in the U.S. between 6 and 18 months (short-term Korean group = KST) performed more poorly than native listeners at categorizing talkers from four different U.S. dialect regions (General American, Northern, Mid-Atlantic, and Southern) by

region of origin. The non-native listeners' performance on the task was very poor and close to chance (25%), with overall accuracy at 25.5% ($SD = 6.2\%$) for the KLT group and at 29.3% ($SD = 3.6\%$) for the KST groups. Additionally, categorization accuracy and response patterns were very similar for both non-native KLT and KST groups. These results suggest that the long-term KLT listeners had not acquired very much detailed explicit knowledge of American English regional dialect differences even after many years of living in the U.S. Thus, although non-native listeners may be able to acquire some knowledge of regional dialect variation in their L2 in some learning situations and environments, acquiring sociolinguistic knowledge of the L2 may be extremely challenging. Furthermore, it is not clear what factors might facilitate the perception and use of L2 regional dialect variation.

Other studies investigating the ability of non-native listeners to recognize speech produced by talkers from different L2 dialect regions have also suggested that increased exposure to a regional variety may not always be sufficient for documenting improvement in recognition ability. Fox and McGory (2007) found that native Japanese learners of English (as well as native English speakers), regardless of location of residence (Ohio or Alabama), were more accurate in identifying vowels produced by a General American English talker than a Southern American English talker. However, unlike native listeners from Alabama, who were also quite accurate for the Southern vowels, Japanese listeners living in Alabama did not exhibit a benefit for the Southern variety. Similarly, Eisenstein and Verdi (1985) reported that English learners in New York City had much more difficulty in understanding African American English than both New York English and General American English, even though the listeners were familiar with all three varieties. The findings from these studies suggest that while non-native

listeners may show a benefit for a standard variety, they also may have difficulty with nonstandard varieties, despite having had substantial contact with them.

Results from the multi-dialect sentence recognition task in Chapter 2 also suggest that non-native listeners may not show improvements in their ability to recognize speech produced by talkers from specific dialect regions with increased exposure to those regions in the L2. In that study, non-native listeners completed a multi-dialect sentence recognition task with talkers from four different U.S. dialect regions in multitalker babble at two different signal-to-noise ratios (SNRs). Non-native word recognition accuracy varied from 14.0% to 74.7% averaged over all four regional dialects and both SNR conditions. Regardless of the amount of time the L2 listeners had lived in the U.S., Northern talkers were the most intelligible even though listeners were residing in an area where the speech characteristics of the General American and Southern varieties were predominately heard. Furthermore, while word recognition accuracy within the non-native group was greater for listeners with longer lengths of residency consistently across all talkers' regions of origin, the benefit was not specific to the dialect(s) with which the listener had the most experience. Thus, the non-native listeners in that study only showed a general overall improvement in their ability to adapt and adjust to L2 regional dialect variation.

Explicit Regional Dialect Categorization Training

Taken together, studies examining the perception of regional dialect variation by native and non-native listeners demonstrate that native and non-native listeners can reliably use dialect-specific information in the speech signal to identify or categorize talkers by region of origin. However, even native listeners have some difficulty in explicitly identifying unfamiliar or nonstandard dialects and recognizing speech produced by talkers from unfamiliar or nonstandard

dialect regions. For non-native listeners, successfully adapting to regional dialect variation in their second language appears to be even more of a challenge, possibly due to having less knowledge and experience with their second language, both in terms of general linguistic knowledge and more specific sociolinguistic knowledge of the various regional dialects of their L2. Furthermore, improving L2 listeners' ability to identify the region of origin of talkers and to recognize speech produced by talkers from various regions of origin may be extremely difficult to do in a naturalistic L2 language learning environment.

In the absence of acquiring sociolinguistic knowledge of L2 regional dialect in everyday communication environments, several laboratory training methods may provide a way of improving non-native listeners' sensitivity to and awareness of regional dialect variation in their L2. Second language learners must develop certain skills that are necessary for successfully adapting to the variability in the speech signal. These crucial skills include sensitivity to sublexical and subphonemic variability to reliably distinguish between varieties and knowledge of how specific acoustic-phonetic features group together to index the different varieties of an L2 (Clopper & Bradlow, 2009). Thus, any training regimen that is designed to improve dialect categorization and high-variability speech recognition must be able to improve a listener's sensitivity to these features and help them learn how features group together for each dialect. Learning to identify and categorize different sources of systematic variation can lead to increased sensitivity to that variation. Processes of differentiation, whereby one's sensitivity to previously unnoticed differences among items is increased through experience with those items, and dimensionalization, whereby experience with systematic variation can cause one to create new perceptual dimensions by which the variation can be ordered, interact to both create and expand/contract the perceptual dimensions by which we perceive items in the world (see, e.g.,

Lawrence, 1949; Goldstone, 1994; Goldstone, 1998; Goldstone, 2003). In other words, with substantial experience these perceptual processes and strategies may help order and break up stimulus items that were previously perceived as consisting of random variation and noise along relevant perceptual dimensions, as well as make perceptual differences more apparent across categories (and similarities to become more apparent within categories).

In the perceptual analysis of speech, we would also expect that with experience with this type of variation, and in particular learning upon which dimensions to organize the dialect-specific variation through dialect categorization training, listeners may gain increased sensitivity to dialect-specific differences and improved knowledge of dialect-specific features. Gaining this experience would facilitate not only the explicit identification of the regional dialects, but also the perceptual analysis of the linguistic content encoded in a speech signal where regional dialect variation is present. Previous studies have been successful in using indexical training to improve listeners' ability to identify different sources of indexical properties of speech, such as the attributes of a talker's voice (e.g., Nygaard, Sommers, & Pisoni, 1994; Nygaard & Pisoni, 1998) or the regional dialect of a talker (Clopper & Pisoni, 2004a). In addition, high-variability phonetic training has also been successful in promoting learning of specific speech sounds or sound contrasts in a second language (see, for example, the seminal r/l studies of Bradlow, Pisoni, Yamada, & Tohkura, 1997; Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994). Furthermore, other studies have suggested that increased sensitivity to different sources of variability in speech may improve the recognition of spoken words. Nygaard et al. (1994) and Nygaard and Pisoni (1998) used novel talker identification training techniques to improve the listeners' sensitivity to talker-specific information, resulting in better discrimination of familiar and unfamiliar talkers' voices as well as a benefit in recognizing

speech produced by those familiar talkers.

The Current Study

Taking these previous studies into consideration, the current study investigated the use of a regional dialect categorization training task to improve listeners' sensitivity to dialect-specific information in the speech signal and knowledge of regional dialect variation, resulting in a better ability to identify the region of origin of unfamiliar talkers. Furthermore, the present study also assessed the effects of regional dialect categorization training on the recognition of words produced by talkers from different U.S. dialect regions as well as the discrimination and identification of American English vowels and consonants. To explore these issues, native and non-native listeners who had participated in a previous study, reported on in Chapters 1-4, returned to the laboratory and completed explicit regional dialect categorization training. The non-native listeners were all native speakers of Korean who had either lived in the U.S. for a long period of time (more than 3 years) or a relatively short period of time (6-18 months). These two groups of listeners completed explicit regional dialect training and several pre-training and post-training perceptual tasks.

The design of the categorization training protocol used here incorporated different aspects of the methodologies used in previous work that relied on high-variability stimulus materials in training on linguistic and indexical aspects of speech. In particular, the current protocol drew heavily from the only previous dialect categorization training study (Clopper & Pisoni, 2004a). This explicit regional dialect categorization training procedure was used to assess generalization and transfer effects from categorization training, as well as to achieve the practical goal of improving L2 speech perception and spoken word recognition. Thus, the current study

investigated not only L2 learners' ability to recognize speech produced by talkers from different regions and make explicit dialect identification judgments about talkers based on their region of origin in their L2, but also how these abilities might be improved. In particular, the current study addressed the efficacy of regional dialect categorization training in building robust talker-independent perceptual dialect categories that may improve regional dialect identification as well as facilitate the perception of linguistic properties of speech.

Methods

Listeners

A total of 38 young adults participated as test (Exp) and active control (Control-1) listeners in the training study. All of the Exp and Control-1 listeners had previously participated in an experiment examining regional dialect perception in American English, which involved completing several speech perception tasks, neurocognitive tasks, and self-report questionnaires. For information on the tasks and questionnaires, see Chapters 2-4. Fifteen participants were native speakers of American English (NAE) and 23 participants were native speakers of Korean (NNK). The NAE participants included 10 females and 5 males, aged 18-27 years ($M = 20.9$; $SD = 2.1$ years). The NNK participants included 9 females and 14 males, aged 19-32 years ($M = 25.3$; $SD = 4.1$ years). All NNK participants were native speakers of Korean and had an average length of residency in the U.S. (LOR) of 2.9 years ($SD = 2.6$ years). In the previous study, 11 of the NNK listeners were classified as long-term Korean listeners (KLT), who had lived three or more years in the U.S. and 12 of the NNK were classified as short-term Korean listeners (KST), who had lived between 6 and 18 months in the U.S. Before testing, participants were divided into two experimental groups: Exp and Control-1. The Exp group consisted of 9 NAE (7 female; ages

18-23) and 14 NNK listeners (4 female; ages 19-32; 7 KLT; LOR $M = 2.9$ years). The Control-1 group consisted of 6 NAE (3 female; ages 19-27) and 9 NNK listeners (5 female; ages 19-32; 4 KLT; LOR $M = 2.9$ years).

In addition to the Exp and Control-1 listener groups, a second control group (Control-2) that did not complete the categorization training was included. The Control-2 group consisted of 13 native speakers of American English. The Control-2 participants included 8 females and 5 males, aged 19-29 years ($M = 21.1$; $SD = 2.8$ years). None of the Control-2 participants had participated in the earlier studies reported in Chapters 2-4.

All participants had normal hearing with no significant history of hearing or speech disorders at the time of testing based on self-reports. They received \$20 for 120 minutes of participation (rate of \$5 per 30 minutes), which included completing several speech perception tasks, neurocognitive tasks, and self-report questionnaires. The three listener groups completed different sets of tasks, which are described below.

Materials and Procedures

Procedures were identical for each of the three listener groups except where explicitly mentioned. All participants were tested individually, and were seated in front of a personal computer equipped with headphones (Beyerdynamic DT100), a keyboard, and a mouse. The output levels of the target sentences for all the perception tasks were calibrated to be approximately 65 dB SPL. Procedures and materials for all tasks are described below.

Explicit Regional Dialect Categorization Training

The regional dialect categorization training protocol consisted of a pre-training, training, learning, and a generalization phase. Table 1 describes the core procedure for the explicit regional dialect categorization training. All three listener groups completed the pre-training and post-training regional dialect categorization tasks involving explicit regional dialect categorization training. However, only the Exp and Control-1 listener groups completed the training blocks of the categorization training, while the passive Control-2 listener group did not.

Table 1. Core experimental procedures for the explicit regional dialect categorization training.

Block:	Pre-Training	Train 1	Train 2	Learning	Generalization
Listener Groups	Exp Control-1 Control-2	Exp Control-1	Exp Control-1	Exp Control-1 Control-2	Exp Control-1 Control-2
Feedback for Exp Group	No	Yes	Yes	No	No
Materials	4 talkers/ dialect 2 unique sentences	4 talkers/ dialect 2 reps of 1 sentence (1)	4 talkers/ dialect 2 reps of 1 sentence (2)	4 talkers/ dialect 2 old and 2 unique sentences	4 new talkers/ dialect 3 unique sentences

Pre-Training Regional Dialect Categorization: All listener groups completed a pre-training regional dialect categorization task, which was used to assess the listeners' baseline categorization ability before any training. The six-alternative forced-choice dialect categorization task used in this study was similar to the four-alternative forced-choice regional dialect categorization described in Chapter 2. Twenty-four talkers (12F/12M) were selected from the

TIMIT acoustic-phonetic speech corpus (Garofolo et al., 1993). Four talkers (2F/2M) were from each of the following six dialect regions: New England, North Midland, South Midland, North, South, and West. Two unique sentences were selected for each talker. On each trial, listeners heard a single talker and were asked to choose the geographical region where they thought the talker was from using a closed set of six response alternatives. Participants were required to choose from the six dialect regions (New England, North Midland, South Midland, North, South, and West) represented on a colored graphical map of the U.S. displayed on a computer monitor. Subjects entered their responses by clicking a cursor on a labeled box located within each dialect region. The map and six response alternatives used in the current study are shown in Figure 1. Participants could take as long as they wanted to respond. Once they responded, the next trial began. Responses were scored for accuracy.

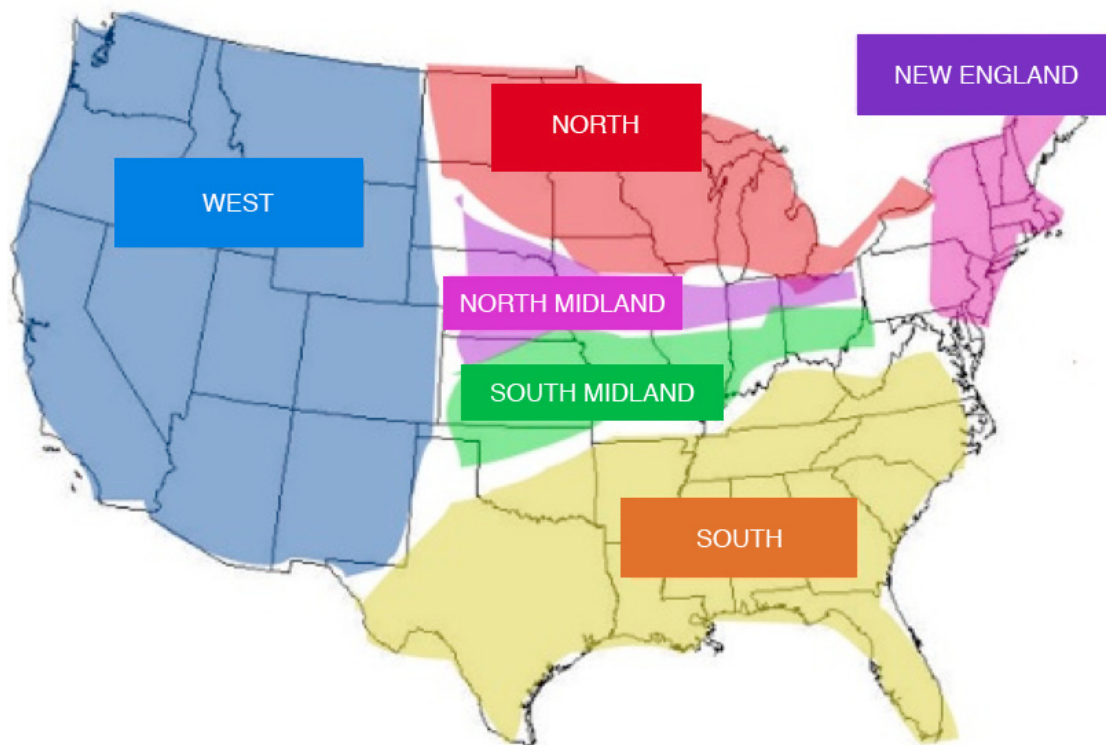


Figure 1. Regional dialect response map with six response alternatives used in all categorization tasks in the current study.

Regional Dialect Categorization Training: After the baseline pre-training categorization task, the Exp and Control-1 listeners completed two training blocks ('Train 1' and 'Train 2' blocks, as shown in Table 1). For the Exp listeners, training blocks consisted of a regional dialect categorization task with feedback. The same 24 talkers from the pre-test were used in the training phase. Again, these 24 talkers (12F/12M) included 4 talkers (2F/2M) from each of the following six TIMIT dialect regions: New England, North Midland, South Midland, North, South, and West. Two sentences were used for each talker for the training blocks: (1) "She had your suit in greasy wash water all year" and (2) "Don't ask me to carry an oily rag like that." Both sentences are "baseline" calibration sentences collected from all talkers in the TIMIT database, which were designed to obtain dialectal differences in American English (Garofolo et al., 1993; Clopper & Pisoni, 2004c).

On each trial, participants listened to one talker producing a single sentence over the headphones. Again, they were asked to indicate the dialect region they thought the talker was from by selecting one of six dialect regions (New England, North Midland, South Midland, North, South, and West) represented on a map of the continental U.S. displayed on a computer monitor. The map was the same as the one used in the pre-training task. Participants again responded by clicking on a labeled box located within each dialect region. They could take as long as they desired to respond, and once they responded they received the explicit training feedback. Feedback consisted of 1) an indication of whether the listener's response was correct or incorrect and 2) the correct response with a repetition of the sentences over the headphones. The listeners received both types of feedback whether or not they had correctly responded to the trial or not. After the feedback, participants clicked on the screen to continue to the next trial. This procedure was the same across both of the two training blocks.

In the two training blocks, listeners in the Exp group were presented with three repetitions of the same sentence for each talker, for a total of 72 trials per training block. Unlike the Exp Group listeners, the Control-1 listeners did not receive any feedback. Control-1 listeners were presented with the same stimuli and performed the same regional dialect categorization task with the same map. After entering their response, they did not get to see the correct answer. However, the Control-1 listeners did hear the second repetition of the sentence, just like the Exp listeners. All listeners were able to take breaks between experimental blocks. Additionally, after the training blocks, they were asked to describe some of the characteristics of each American English regional dialect to the best of their ability.

The passive Control-2 listener group did not complete the training blocks. Instead, they filled out a language background questionnaire and completed a short interview about their knowledge of the characteristics of the American English regional dialects included in the current study.

Post-training Regional Dialect Categorization: All three listener groups then completed a post-training regional dialect categorization task. A learning block was included to make sure that listeners had learned the regions of origins of the talkers included in the pre-training and training blocks ('Learning' block, as shown in Table 1). The learning block followed the same procedure as the pre-training block using the same six response alternatives and map, although no feedback was provided. Pre-test sentences were presented again during this phase, along with two new unique sentences for each talker. Exp, Control-1, and Control-2 listener groups performed this task with identical instructions. After the learning block, a generalization block ('Generalization' block, as shown in Table 1) was included to examine talker-independent learning of regional dialect categories by the listeners. The generalization block consisted of a

new set of 24 talkers (4 talkers (2M/2F) per dialect region) using the same six response alternatives and map. Two unique sentences per talker were used in the generalization block. All listener groups also performed this task with identical instructions, as well. After the two post-training regional dialect categorization blocks, all listeners took a short break.

Pre-Training and Post-Training Perception Tasks

Listeners completed several perception tasks before training (*pre-training* tasks) and after training (*post-training* tasks). A multi-dialect sentence recognition task was used to measure the listeners' ability to recognize spoken words produced by talkers from multiple American English dialect regions. An ABX vowel and consonant discrimination task was used to measure their ability to discriminate American English vowels and consonants. A vowel identification task was used to measure their ability to identify American English vowels. Measures were obtained both before and after the regional dialect categorization training. These tasks were included to examine any transfer effects from the dialect categorization training to novel perceptual tasks.

The three listener groups did not complete the same set of pre-training and post-training perception tasks. The Exp and Control-1 listener groups completed the pre-training multi-dialect sentence recognition task and the vowel identification task in a previous testing session for another study, which is described in Chapter 3. They completed the post-training multi-dialect sentence recognition task, the ABX discrimination task, and the vowel identification task within the same testing session as the categorization training. The Control-2 listener group only completed the pre-training and post-training sentence recognition tasks for the purpose of documenting the intelligibility of those lists, since they had not been previously used in the earlier study. The pre- and post-training versions of the ABX discrimination task and the vowel

identification tasks were identical, so it was determined that it was unnecessary to obtain two scores on the task within the same testing session.

Multi-Dialect Sentence Recognition: The listeners' ability to recognize spoken words in sentences produced by male and female talkers from different U.S. dialect regions was assessed using a multi-dialect sentence recognition task. Four talkers (2F/2M) from each of the New England, North, North Midland, South Midland, South, and West dialect regions were selected from the TIMIT acoustic-phonetic speech corpus (Garofolo et al., 1993). Four unique utterances were selected for each talker, for a total of 96 sentences. Two randomly selected sentences were mixed with speech-shaped noise at +2 dB SNR, and two randomly selected sentences were mixed with speech-shaped noise at -2 dB SNR. On each trial, listeners were presented with one sentence over headphones and were asked to type what they heard into a dialog box on the computer screen. All listeners were encouraged to give partial responses or to guess if they were unsure what they heard.

Sentences were blocked by SNR. Sentences at the more moderate SNR (+2 dB) were presented first, and the sentences at the less favorable SNR (-2 dB) were presented in a second block. Within each SNR block, the two sentences from each talker were randomly placed into two sub-blocks. Sentences were pseudo-randomly presented within each of the sub-blocks. Overall, a talker only appeared once in each of four sub-blocks of the task. This was designed to reduce the likelihood that the same talker would appear twice within a few trials. Listeners did not receive any feedback, and were able to proceed to the next trial by either pressing the return button or clicking on an on-screen button. Listeners had the option of taking a break halfway through the task. Responses were scored by content words correct, which included all nouns,

verbs, adjectives, and adverbs. Incorrect morphological endings were considered incorrect, but homophones were counted as correct responses.

The procedures used in the pre-training and post-training multi-dialect sentence recognition tasks were identical. However, the post-training task included a set of 4 new talkers (2M/2F) for each of the four dialects producing 4 unique sentences each (2 at +2 dB SNR and 2 at -2 dB SNR), for a total of 96 sentences. Thus, all test materials in the post-training task were novel.

ABX American English Vowel and Consonant Discrimination: The participants' ability to discriminate several American English vowels and consonant contrasts was examined using an ABX discrimination task. The specific ABX task used was based on an earlier study by Darcy, Park, and Yang (under review). Two female talkers from the North Midland dialect region were recorded producing English nonwords. All nonword items were not real words in either American English or Korean. Eight American English vowels and consonant contrasts that were either easy or difficult for native Korean speakers, based on the phonological systems of Korean and English, were used in this task. The vowel contrasts included the following difficult minimal pairs /i, ɪ/, /u, ʊ/, /eɪ, ε/, /æ, ε/, and /ɑ, ʌ/, and a control (easy) pair /u, o/. The consonant contrasts included the following difficult minimal pairs /l, r/, /p, f/, and /θ, s/ and the control (easy) pair /s, t/. These particular segmental contrasts were also selected because they are involved in American English regional dialect variation and may be crucial to reliably distinguishing the four regional varieties used in the perception tasks described in Chapter 2. All target vowels were embedded in the stressed syllable of a disyllabic minimal pair frame CV(C)CVC, within three different consonantal contexts (bilabial, alveolar, and velar), e.g., [geɪ.kəʔt], [neɪ.dən], and [peɪ.bad]. All consonants were embedded in the same disyllabic minimal pair frame CV(C)CVC, in prevocalic

word-initial, intervocalic, or postvocalic word-final position, e.g., [pæs.tɪk], [pə.pɪk], and [tə.gæp].

Overall, a total of 30 pairs (5 vowel test pairs in 3 contexts and 3 consonant test pairs in 3 contexts; and 1 vowel control pair and 3 contexts and 1 consonant control pair in 3 contexts) were used in the task. All four possible ABX combinations for each pair (ABB, ABA, BAA, and BAB) were presented, for a total of 120 stimulus items (4 combinations x 30 pairs). Each combination of the 30 pairs was randomly assigned to one of four blocks so that each contrast was presented once in each context per block, for a total of 30 trials per block.

On each trial of the ABX task, one nonword pair (AB) produced by one of the female talkers was presented, followed by another nonword matching either the first or the second nonword (“X”, matching either A or B) produced by the other female talker. The stimuli in each ABX triad were separated by 250 ms of silence. The participants were asked to decide whether the third stimulus (“X”) was the same as the first (A) or the second (B) by pressing a button on a button box as quickly as possible. The left button on the button box always represented the first stimulus (A) and the right button always represented the second stimulus (B). The next ABX triad was presented after the participant responded.

Within each block, the presentation of the triads was randomized. Before the task, participants received a short 8-trial practice block of ABX trials with feedback. None of the practice items were used later in the task. After the practice block of trials, participants completed the four blocks of testing. Between each block, subjects had a short break. The pre-training and post-training ABX tasks were identical, except for the presentation order within each block, which was randomized. Responses were scored for accuracy and response time (RT).

Response times were measured from the onset of the last of the three stimulus items in each ABX triad until the listener entered his/her response.

American English Vowel Identification Task: The participants' ability to identify American English vowels was measured with an 11-alternative forced-choice vowel categorization task, which was described in Chapter 3. Four talkers (2F/2M) from the McFAE corpus (Multi-Talker Corpus of Foreign-Accented English; Tamati, Park, & Pisoni, 2011) were selected. All talkers were native speakers of American English speakers who had grown up in the North Midland area of state of Indiana. For this task, two hVd words containing the eleven monophthongal vowels of American English /i ɪ ε eɪ æ u ʊ o ʌ ɑ ɔ/ were selected for each talker, for a total of two sets of 11 hVd words per talker and 88 hVd words overall.

On each trial, participants were randomly presented with a single hVd word, and upon being prompted by the question "Which vowel did you hear?" the subject responded by clicking on one of eleven vowel choices displayed on a computer monitor. Each vowel option contained a single hVd word, written out on the computer screen with a vowel shorthand notation developed by Gerstman (1968), and two high-frequency words that contain the same vowel. All response options had also been shown to the participant before the task began so that they could familiarize themselves with the shorthand and target vowels of the response alternatives. Participants in the currently study had also previously completed a hVd production task (not reported on here) for which they received some training on the vowel shorthand and the target vowels of the eleven hVd words. The vowel identification task was self-paced. Participants could take as long as they wanted to respond but they only listened to one presentation of each word. The next trial began after they entered their response.

All responses were scored for accuracy. After preliminary inspection of the response patterns, the native listeners could not distinguish between the two low-back vowels /ɑ/ and /ɔ/, consistent with the low-back merger in the North Midland dialect region, which is pervasive in the English speaking community of the subjects' residence (e.g., Labov, Ash, & Boberg, 2006). Given the poor performance on these vowels by the native listeners, the two vowels were combined into one low-back category, which will be referred to as /ɑ/, for analysis. The pre-training vowel identification task consisted of the full set of 88 hVd words, whereas the post-training vowel identification consisted of only one production of each of the 11 hVd words per talker, for a total of 44 hVd words.

Results

Explicit Regional Dialect Categorization Training

Mean dialect categorization accuracy was calculated for the pre-training regional dialect categorization and the post-training learning and generalization categorization blocks for each of the three listener groups (Exp, Control-1, Control-2). Figure 2 displays the mean categorization accuracy on the pre-training categorization block, the post-training learning block with sentences produced by training talkers, and the post-training generalization block with new sentences and new talkers for all three listener groups (Exp, Control-1, Control-2). The Exp and Control-1 groups are also further broken down by native language.

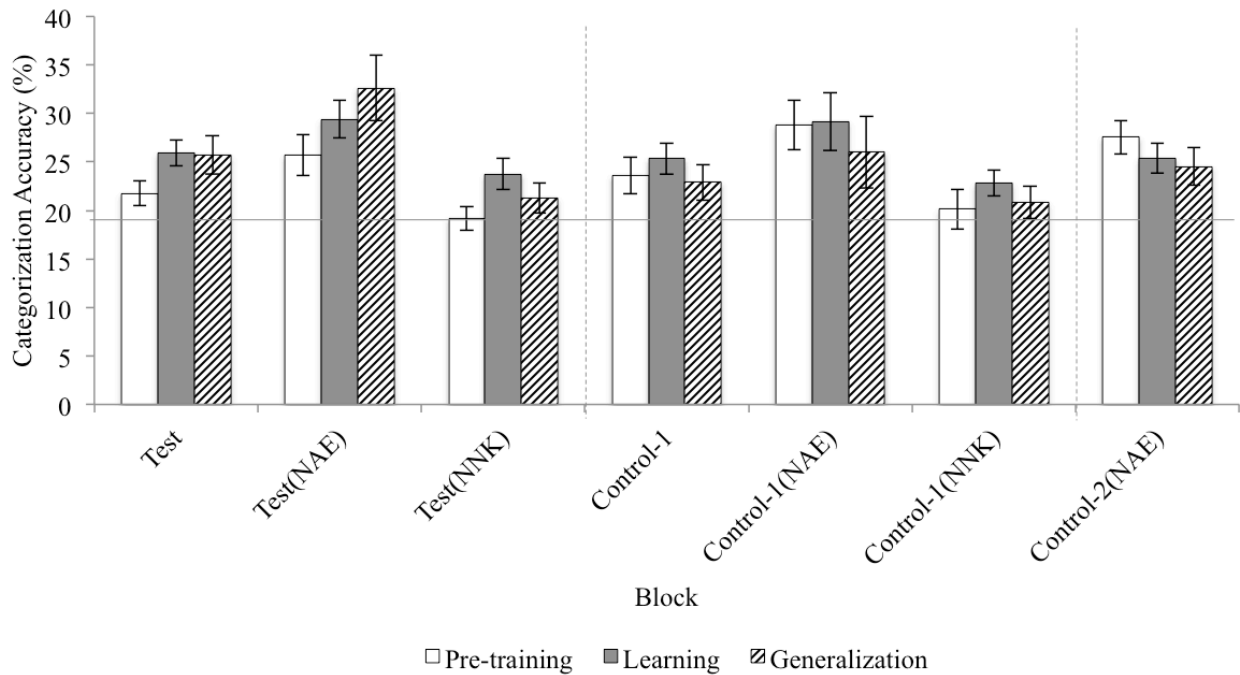


Figure 2. Mean categorization accuracy on all categorization blocks in the training procedure (Pre-training block – empty bars; Learning block – gray filled bars; Generalization – striped bars) for the three listener groups (Exp; Control-1; Control-2) broken down by native language (NAE – native listeners; NNK – non-native listeners). Chance performance (16.7%) is indicated by the solid gray horizontal line. Errors bars are +/- 1 SE.

To assess whether the listeners had learned the region of origins of the talkers in the pre-training, training, and learning blocks of the categorization training, difference scores were calculated by subtracting the mean pre-training categorization accuracy from the categorization of both old and new sentences in the learning block. Figure 3 displays the average difference scores for performance between the pre-training task and the post-training learning block for the three listener groups (Exp, Control-1, Control-2). The Exp and Control-1 groups are also further broken down by native language. As can be seen in Figure 3, although performance was highly variable within each group, the Exp listeners clearly showed a pattern of improvement on the training talkers. A one-sample t -test confirmed that the differences were greater than 0 [$t(22) = 2.57, p = .017$]. The Control-1 listener group showed only modest improvement on categorizing

training talkers after exposure during the training blocks and no feedback. The difference scores for the Control-1 listeners were slightly greater than 0 [$t(14) = 2.14, p = .050$]. Finally, the Control-2 group showed no improvement at categorizing the training talkers on the old sentences. Again, the difference scores for the Control-2 listeners were not greater than 0.

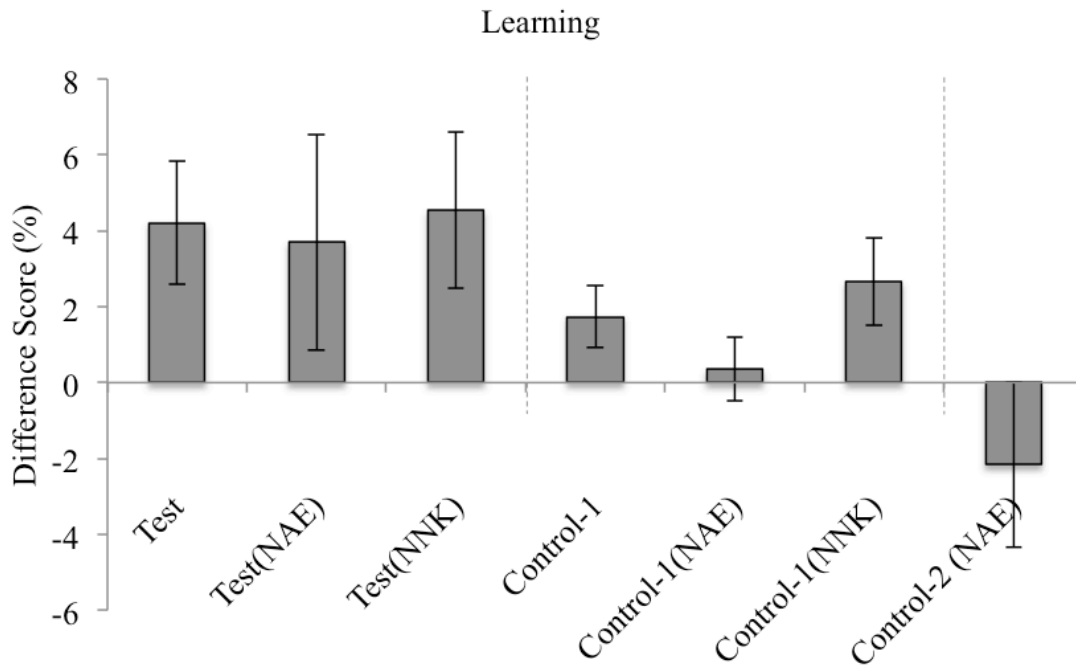


Figure 3. Learning Block: Mean differences (%) between pre-training categorization accuracy and learning block accuracy for the three listener groups (Exp, Control-1, Control-2), broken down by native language (NAE – native listeners, NNK – non-native listeners). Errors bars are +/- 1 SE.

Categorization accuracy on new sentences produced by new talkers in the generalization block was also examined. Given the overall difference in accuracy among the groups across all blocks, difference scores were again calculated by subtracting the mean pre-training categorization accuracy from the categorization of the new talkers in the generalization block. Figure 4 displays the average difference scores for performance between the pre-training task and the post-training generalization block for the three listener groups (Exp, Control-1, Control-

2). Again, the Exp and Control-1 groups are also further broken down by native language. The difference scores for the Exp listeners were greater than 0 by a one-sample *t*-test [$t(22) = 2.66, p = .014$]. Difference scores for the Control-1 and Control-2 groups were not significantly different 0. Thus, the Exp listeners also showed moderate improvement at categorizing novel talkers by their region of origin.

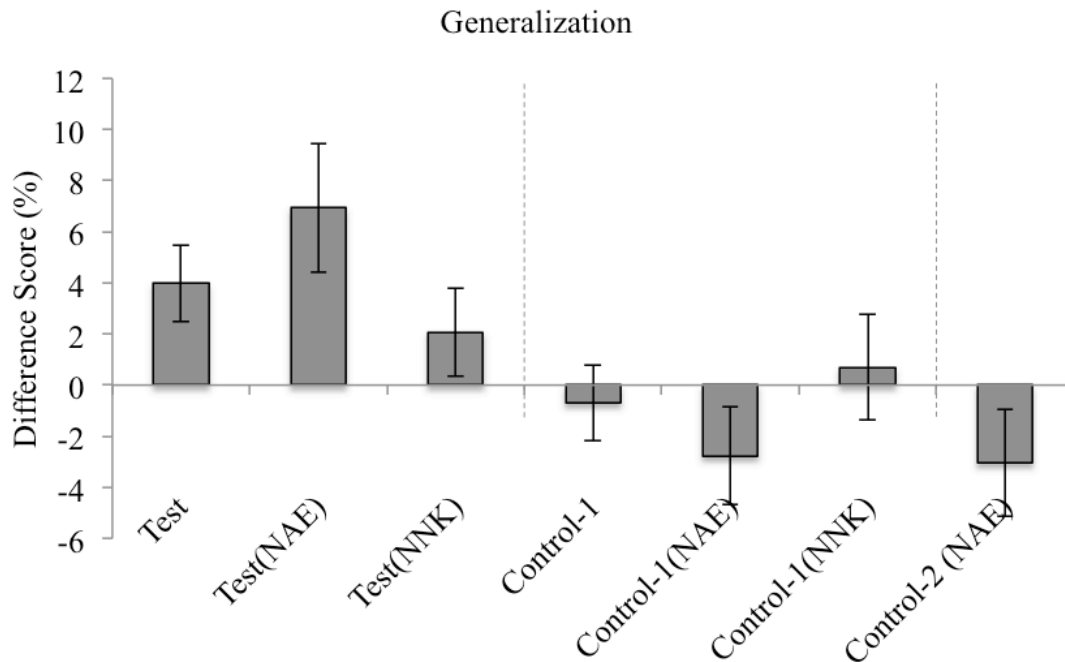


Figure 4. Generalization Block: Mean differences (%) between pre-training categorization accuracy and generalization block categorization accuracy for the three listener groups (Exp, Control-1, Control-2), broken down by native language (NAE – native listeners, NNK – non-native listeners). Errors bars are +/- 1 SE.

Pre-Training and Post-Training Perception Tasks

To assess whether the Exp listeners also showed improvement on the transfer tasks, mean accuracy scores were calculated for the three pre-training and post-training perception tasks: the multi-dialect sentence recognition task, the ABX American English vowel and consonant

discrimination task, and the American English vowel identification task. A series of *t*-tests on difference scores between pre- and post-training content word recognition accuracy on the multi-dialect sentence recognition task revealed no significant gains in performance after completing the explicit categorization training for any of the three listener groups (Exp, Control-1, Control-2). Similarly, *t*-tests on difference scores between pre- and post-training on the ABX Discrimination Task and the Vowel Identification Task for the two active listener groups (Exp, Control-1) provided no evidence of any significant gain in discrimination or identification accuracy after completing the regional dialect categorization training. Thus, the Exp listeners' performance on the pre-training and post-training perception tasks showed no indication of any transfer effects from the explicit regional dialect categorization training.

Individual and Group Differences in Training

To further examine patterns of learning and generalization, the Exp listeners were divided into two groups based on their difference scores for the learning and pre-training regional dialect categorization training blocks. High-learners were the 12 Exp listeners who showed the most improvement from the pre-training block to the learning block during the categorization training, and Low-learners were the 11 Exp listeners who showed the least improvement (or no improvement) in the learning block compared to the pre-training block. Table 2 displays the language backgrounds of the High- and Low-learners from the Exp listener group.

Table 2. High- and Low-learners from the Exp group, with participants' language backgrounds.

	High-learners	Low-learners
# Total Listeners	12	11
# Native Listeners	5	4
# Non-Native Listeners	7	7
- Average LOR in U.S.	42 mo.	27 mo.
- LOR group	4 KLT; 3 KST	3 KLT; 4 KST

As shown in Table 2, the High- and Low-learner groups were roughly equally represented by native, long-term non-native (KLT), and short-term non-native listeners (KST). Furthermore, although learning varied by regional dialect, native and non-native listeners within the High- and Low-learner groups consistently showed similar patterns of learning across the regional dialects (see Appendix A). Thus, learning during the categorization training was not substantially influenced by language background of the participant.

The regional dialect categorization and multi-dialect sentence recognition scores for the High and Low groups were further compared to uncover any differences between the listeners who were able to benefit from the categorization training (High-learners) and those listeners who were not (Low-learners). Before training, the High learners were already better at categorizing talkers by regional dialect [$t(21) = 3.01, p = .007$]. After training, the High-learners showed significantly greater improvement from pre-training to both the learning [$t(21) = 9.02, p < .001$] and the generalization blocks [$t(21) = 2.77, p = .012$] compared to the Low-learners. Figure 5 displays the relation between pre-training dialect categorization accuracy and post-training dialect categorization accuracy scores; the learning block scores are presented in the plot on the top and the generalization block scores are presented in the plot on the bottom.

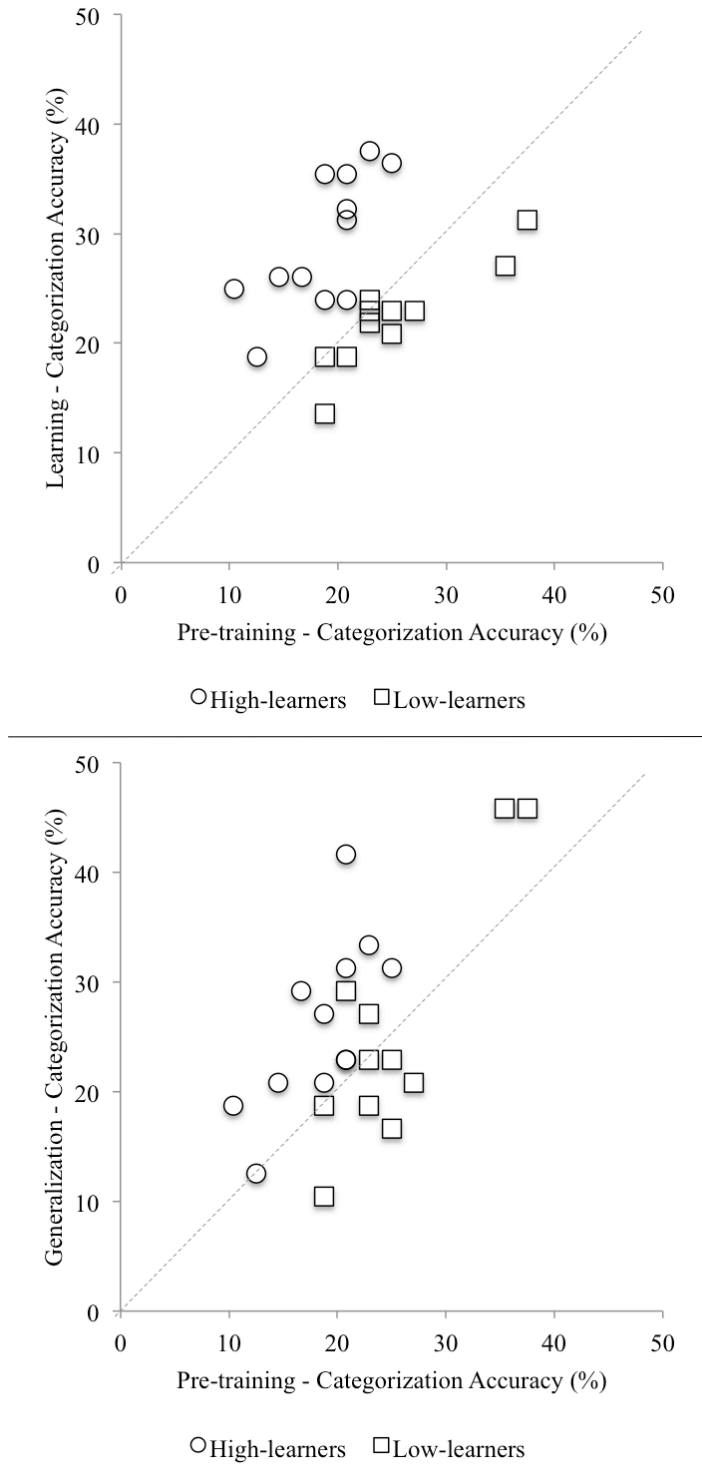


Figure 5. Percent correct dialect categorization accuracy (%) in the pre-training block (x-axis) and in the learning block (y-axis) (top); and in the generalization block (y-axis) (bottom). Empty circles represent High-learners from the Exp listener group and the empty squares represent the Low-learners from the Exp listener group.

The High-learners were also initially more accurate at recognizing words produced by talkers from different U.S. dialect regions on the pre-training multi-dialect sentence recognition task [mean = 61.7%] than the Low-learners [mean = 51.9%], but the difference did not reach significance [$t(19) = .933, p > .05$] as performance was highly variable. Neither group improved in sentence recognition accuracy from the pre-training multi-dialect sentence recognition task to the post-training sentence recognition task. As such, no difference was found between groups in the amount of improvement from the pre-training multi-dialect sentence recognition task to the post-training sentence recognition task.

Discussion

The goal of the current study was to investigate the effects of regional dialect categorization training on listeners' ability to identify the region of origin of an unfamiliar talker and on their ability to recognize spoken words and discriminate speech sounds in American English. To explore these issues, native speakers of American English and native speakers of Korean completed several blocks of high-variability explicit regional dialect categorization training, as well as several pre- and post-training perceptual tasks. Comparing regional dialect categorization accuracy before and after the training blocks, Exp listeners' positive difference scores suggested that they had benefited from the explicit feedback in the training blocks to learn new features of the American English dialects that they were then able to use to categorize both familiar and unfamiliar talkers by region of origin. Furthermore, both native and non-native Exp listeners performed similarly, suggesting that the language background of the listener did not limit the effectiveness of the high-variability training protocol. Unlike the Exp listeners, however, neither the passive Control-1 listeners, who had the same exposure to the training

talkers during the training blocks but received no feedback about the talkers' regions of origin, nor the Control-2 listeners, who had no exposure to the training talkers during the training blocks and received no feedback, showed much improvement in regional dialect categorization accuracy from pre-training to post-training categorization blocks. Although the Exp listeners' gains were very small, with the average difference score around 4%, these findings suggest that the explicit regional dialect categorization training was modestly effective at increasing the Exp listeners' sensitivity to dialect-specific differences and knowledge of dialect-specific features. These findings are similar to previous studies that successfully used high-variability training on talkers' voices (e.g., Nygaard et al., 1994; Nygaard & Pisoni, 1998) and regional dialects (Clopper & Pisoni, 2004a).

To assess whether the categorization training transferred to other perceptual tasks, Exp and Control-1 listeners also completed three sets of pre- and post-training speech perception tasks: a multi-dialect sentence recognition task, an ABX vowel and consonant discrimination task, and a vowel identification task. Examining the difference in accuracy between pre- and post-training on each task, the Exp listeners' showed no evidence of any improvement in word recognition in a linguistic task that also involved variability from regional dialects (multi-dialect sentence recognition task) or in vowel and consonant discrimination and identification in linguistic tasks that did not contain substantial variability from regional dialects (ABX vowel and consonant discrimination task and the vowel identification task). Unlike previous studies that trained listeners to identify talkers' voices (Nygaard et al., 1994; Nygaard & Pisoni, 1998), increased sensitivity to dialect-specific information from the categorization training protocol did not affect the perception of American English speech sounds and words in sentences.

Nygaard et al. (1994) and Nygaard and Pisoni (1998) both found evidence that learning

talkers' voices facilitated the recognition of words and sentences produced by familiar talkers. The lack of any transfer effects in the current study may not definitively indicate that the perceptual learning of regional dialect features in a training task does not transfer to other perceptual tasks. Instead, the design of the training task may have provided limited opportunity for learning and transfer of learning. Compared to the previous studies, which trained listeners on talkers' voices over several days, the training blocks in the current study were relatively short, around 45 minutes in total on one day. The short training period may not have given the Exp listeners enough exposure and feedback for them to form more robust representations of regional dialect variation in American English, or enough time for consolidation of learning.

The type of stimulus materials used also may not have presented the listeners with enough variability. Regional dialects of American English are characterized by a great deal within-dialect variability, and listeners must rely on signal-independent knowledge of the distribution of reliable dialect-specific features that are crucial for categorizing talkers from the same dialect region but whose speech does not display the identical set of features (e.g., Clopper & Bradlow, 2009). To be successful, the regional dialect categorization training must not only increase the sensitivity of unfamiliar listeners to dialect-specific differences, but it must also improve the knowledge of the distribution of those features across multiple talkers from the same region in order to create new or expand/contract old perceptual dimensions for regional dialect variation in American English (e.g., Lawrence, 1949; Goldstone, 1994; Goldstone, 1998; Goldstone, 2003). Thus, the training in the current study, which used only two sentences in the training blocks, may not have provided enough exposure and variability to result in robust learning. Therefore, although the listeners may have learned some features of the less familiar dialects, resulting in an overall improvement on the post-training categorization blocks, they may

not have gained enough knowledge of the dialect-specific differences to facilitate the recognition of words in sentences produced by talkers from the same regions as the talkers in the training task. In future studies, the length of training and the type of training materials should be explored in greater depth in order to determine the optimal design that will result in robust learning of dialect-specific acoustic-phonetic features and the transfer of perceptual learning of those features to the linguistic processing tasks.

Patterns of learning and transfer of the good (High-learners) and poor (Low-learners) learners from the Exp listener group were also compared. As in Nygaard et al. (1994) and Nygaard and Pisoni (1998), it was expected that the good learners would show considerable transfer of learning to the novel perception tasks. The High-learners also did not show any improvement in the transfer perception tasks. However, further comparison of the High- and Low-learners showed that the High-learners were already more accurate than the Low-learners on the pre-training categorization task. The difference in pre-training categorization accuracy between the two groups suggests that High-learners were able to use their previous knowledge of features of American English regional varieties to aid further learning in the training blocks. The listeners who were able to learn the most during the training were also the ones who were initially better at categorizing talkers by region of origin and, to a lesser extent, more accurate at recognizing words in sentences produced by talkers from different U.S. regions. Interestingly, the High- and Low-learners were nearly equally represented by native and non-native listeners, and further by non-native listeners who had lived in the U.S. a relatively long period of time (KLT) and by non-native listeners who had lived in the U.S. a relatively short period of time (KST). Thus, individual differences in the amount of perceptual learning accomplished during the categorization training were not related to language background. Instead of background and

exposure, individual differences in learning on the regional dialect categorization training may be related to other, more basic perceptual skills (Chapter 3) or foundational neurocognitive abilities (Chapter 4). Other additional social, perceptual, and cognitive factors may influence individual differences in the perceptual learning of regional dialect variation and these should be further explored in future studies.

Conclusions

The current study investigated the effects of regional dialect categorization training on the learning of dialect-specific features in a native language and L2 to improve the listeners' ability to identify unfamiliar talkers' regions of origin. A comparison of good and poor learners in the training task demonstrated that the amount of benefit a listener received from the training was not related to his or her native language (American English or Korean) or for non-native listeners, how long they had lived in the U.S. However, the good learners had a greater overall ability to perceive and use reliable dialect-specific features to categorize talkers by region of origin. These results suggest that language background and experience was not likely related to the perceptual learning of regional dialect variation. Instead, perceptual learning of regional dialects may be related to individual perceptual and/or neurocognitive differences. Future studies should not only explore more effective training methods for learning regional dialect variation or other sources of indexical variability in speech, but should also further investigate sources of individual differences in learning ability among listeners. Better understanding sources of individual and group differences would help to identify foundational processes that underlie perceptual learning and adaptation in speech perception and spoken word recognition.

The study also examined whether learning in the regional dialect categorization training

task would transfer to the recognition of novel words and speech sounds in American English using several transfer perceptual tasks. In the high-variability explicit regional dialect categorization training task, both native and non-native listeners who had received the training demonstrated improvement in their ability to categorize familiar and unfamiliar talkers after the categorization training. However, the listeners showed no evidence of improvement on a multi-dialect sentence recognition task, an ABX American English vowel and consonant discrimination task, or a forced-choice American English vowel identification task. Further analyses on good and poor learners also showed that learning in the training task did not facilitate the perception of words and speech sounds in the transfer tasks for even the good learners. The lack of far transfer effects may be related to the design of the current training study, which may not have resulted in substantial learning of American English regional dialect characteristics. Other training protocols involving different amounts and types of exposure and feedback should be considered to determine whether learning of dialect-specific acoustic-phonetic variability affects the processing of other linguistic properties of speech in tasks that differ from those used in training.

Appendix A

Difference scores (%) calculated from subtracting the pre-training regional dialect categorization accuracy from the post-training regional dialect categorization accuracy in the learning block, overall and for each individual regional dialect, are presented in the Appendix A table for the High-learners and the Low-learners from the Exp group.

Difference Score (%) Post-training Learning – Pre-training	High-learners	Low-learners
Overall Difference Score	10.8	-2.9
Native Listeners	10.2	-4.4
Non-native Listeners	11.2	-2.1
New England	8.0	2.7
Native Listeners	15.8	2.1
Non-native Listeners	2.4	3.0
North	9.0	-1.5
Native Listeners	8.3	3.1
Non-native Listeners	9.5	-4.2
North Midland	13.5	1.9
Native Listeners	14.2	2.1
Non-native Listeners	13.1	1.8
South Midland	11.1	6.1
Native Listeners	11.7	15.6
Non-native Listeners	10.7	0.6
South	3.8	-14.8
Native Listeners	-9.2	-24.0
Non-native Listeners	13.1	-9.5
West	12.5	-5.7
Native Listeners	20	-10.4
Non-native Listeners	7.1	-3.0

References

- Bradlow, A. R., Pisoni, D. B., Yamada, R. A., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310.
- Bradlow, A. R., Toretta, G. M., Pisoni, D. B. (1996) Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 20, 255-272.
- Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech*, 51, 175-98.
- Clopper, C. G. & Bradlow, A. R. (2009). Free Classification of American English dialects by native and non-native listeners. *Journal of Phonetics*, 37, 436-451.
- Clopper, C. G. Levi, S. V., & Pisoni, D. B. (2006). Perceptual similarity of regional dialects of American English. *Journal of the Acoustical Society of America*, 119, 566-574.
- Clopper, C. G., & Pisoni, D. B. (2004a). Effects of talker variability on perceptual learning of dialects. *Language and Speech*, 47, 207-239.
- Clopper, C. G. & Pisoni, D. B. (2004b). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change*, 16, 31-48.
- Clopper, C. G. & Pisoni, D. P. (2004c). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111-140.
- Clopper, C. G., & Pisoni, D. B. (2006). Effects of region of origin and geographic mobility on perceptual dialect categorization. *Language Variation and Change*, 18, 193-221.

- Clopper, C. G. & Pisoni, D. B. (2007). Free classification of regional dialects of 14 American English. *Journal of Phonetics*, 35, 421-438.
- Darcy, I., Park, H.-Y., & Yang, C.-L. (under review). Individual differences in L2 acquisition of English phonology: the relation between cognitive abilities and phonological processing. Submitted to *Studies in Second Language Acquisition*.
- Eisenstein, M. (1982). A study of social variation in adult second language acquisition. *Language Learning*, 32, 367-391.
- Eisenstein, M. & Verdi, G. (1985). The intelligibility of social dialects for working-class adult learners of English. *Language Learning*, 35, 287-298.
- Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? Human Perception and Performance. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1276-1293.
- Fox, R. A. & McGory, J. T. (2007). Second language acquisition of a regional dialect of American English by native Japanese speakers. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 117-134). Amsterdam: John Benjamins.
- Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., & Dahlgren, N. L. (1993). The DARPA TIMIT acoustic-phonetic continuous speech corpus. Linguistic Data Consortium, Philadelphia.
- Gerstman, L. H. (1968). Classification of self-normalized vowels. *IEEE Transactions on Audio Electroacoustics*, AU-16, 78-80.
- Goldstone, R.L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, 123, 178-200.

- Goldstone, R.L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585-612.
- Goldstone, R.L. (2003). Learning to perceive while perceiving to learn. In R. Kimchi, M. Behrmann, & C. Olson (Eds.) *Perceptual organization in vision: Behavioral and neural perspectives* (pp. 233-278). New Jersey: Lawrence Erlbaum Associates.
- Ikeno, A. & Hansen, J. H. L. (2007). The Effect of Listener Accent Background on Accent Perception and Comprehension. *EURASIP Journal on Audio, Speech, and Music Processing: Special Issue on Perception*, 1-8.
- Lawrence, D. H. (1949). Acquired distinctiveness of cues: 1. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology*, 39, 770-784.
- Labov, W., & Ash, S. (1997). Understanding Birmingham. In T. Bernstein, T. Nunnally, & R. Sabino (Eds.), *Language Variety in the South Revisited* (pp. 508-573). Tuscaloosa, AL: Alabama University Press.
- Lass, N. J., Hughes, K. R., Bowyer, M. D., Waters, L. T., & Bourne, V. T. (1976). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *Journal of the Acoustical Society of America*, 59, 675-678.
- Lively, S.E., Logan, J.S., & Pisoni, D.B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242-1255.
- Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y. I., Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, 96, 2076-2087.

- Mason, H. M. (1946). Understandability of speech in noise as affected by region of origin of speaker and listener. *Speech Monographs*, 13, 54-68.
- Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific perceptual learning in speech perception. *Perception & Psychophysics*, 60, 355-376.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5 (42-46).
- Ptacek, P. H. & Sander, E. K. (1966). Age recognition from voice. *Journal of Speech and Hearing Research*, 9, 273-277.
- Stephan, C. (1997). The unknown Englishes? Testing German students' ability to identify varieties of English. In E. W. Schneider (Ed.), *Englishes around the world* (pp. 93-108). Amsterdam: John Benjamins.
- Tamati, T. N., Park, H.-Y., & Pisoni, D. P. (2011). The development of a new corpus of foreign-accented English. Poster presented at the 2011 VLSP New Tools and Methods for Very Large-Scale Phonetics Research Workshop. Philadelphia, PA.
- Van Bezooijen, R., & Gooskens, C. (1999). Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology*, 18, 31-48.
- Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: Patterns and parameters: Part I. Recognition of backward voices. *Journal of Phonetics*, 13, 19-38.
- Van Lancker, D., Kreiman, J., & Wickens, T. (1985). Familiar voice recognitions: Patterns and parameters: Part II. Recognition of rate-altered voices. *Journal of Phonetics*, 13, 39-52.

CHAPTER 6
SUMMARY OF FINDINGS ON THE PERCEPTION OF
REGIONAL DIALECT VARIATION IN A SECOND LANGUAGE

Introduction

In everyday situations, listeners will often encounter adverse listening conditions that hinder successful speech communication. These real-world listening conditions not only involve noisy environments involving background noise or competition from other talkers, but also reflect substantial natural variability from multiple talkers with diverse language and developmental histories (e.g., Gilbert, Tamati, & Pisoni, 2013). Indeed, indexical variability encoded in the speech signal plays an important role in speech perception and spoken word recognition (e.g., Pisoni, 1997; Mattys, Davis, Bradlow, & Scott, 2012). Listeners make use of previous linguistic knowledge and experience to rapidly adapt to and use talker- and group-related variability encoded in the speech signal. However, high-variability listening conditions with multiple sources of variability may also make successful speech recognition and comprehension more difficult or effortful, and even young, normal-hearing adult native listeners differ substantially in their ability to recognize speech under such conditions (Gilbert et al., 2013).

Real-world listening conditions involving multiple talkers with different regions of origin may be particularly challenging to non-native listeners who have imperfect knowledge of their second language (L2) and less experience with sociolinguistic variation in the L2. While a great deal of research on non-native speech recognition in adverse listening conditions has focused on recognizing speech in different types of background noise and/or competition (see García

Lecumberri, Cooke, & Cutler, 2010), little is currently known about how non-native listeners perceive different sources of indexical variability and acquire knowledge about sociolinguistic variation. Furthermore, previous studies on non-native speech perception in high-variability listening conditions have not addressed the core foundational processes underlying the perception of indexical variability in speech, as well as individual listener differences in basic speech perception and spoken word recognition abilities in an L2.

Non-native listeners are well known for having difficulty in communicating in high-variability, adverse listening conditions (e.g., Tamati & Pisoni, in press). However, few studies have explored the factors that underlie non-native speech perception and spoken word recognition in different conditions characteristic of real-world speech communication. Similar to observations on the study of human memory (Jenkins, 1979; Roediger, 2008), speech perception and spoken word recognition performance is multi-dimensional and can be influenced by a variety of environmental and task design elements. Figure 1 shows one way to characterize the relations among these variables. Among the potential factors, speech perception performance across multiple environments and tasks depends on the task's objective (e.g., keyword recognition, true/false judgments, indexical judgements), the contents of the target signal (e.g., talker characteristics, linguistic material), the characteristics of the signal and background (e.g., type of noise and/or background competition, signal-to-noise ratio (SNR), degraded speech signals), the individual perceiver (e.g., native language, linguistic background, cognitive resources), and the interactions among these factors (e.g., Gilbert et al., 2013; Theunissen et al., 2009). Furthermore, these factors can potentially contribute to the difficulty of communicating in a particular listening environment, making it less-than-ideal. Understanding how these factors influence speech perception and spoken word recognition, both independently and in

combination with one another, can help to describe the processes underlying real-world speech communication. In order to learn about how non-native listeners perceive, encode, and represent L2 speech and deal with different sources of variability in speech, it is important to explore the type of noise and/or competition, the characteristics of the target materials, the design and goals of the task, and the individuals carrying out the task.

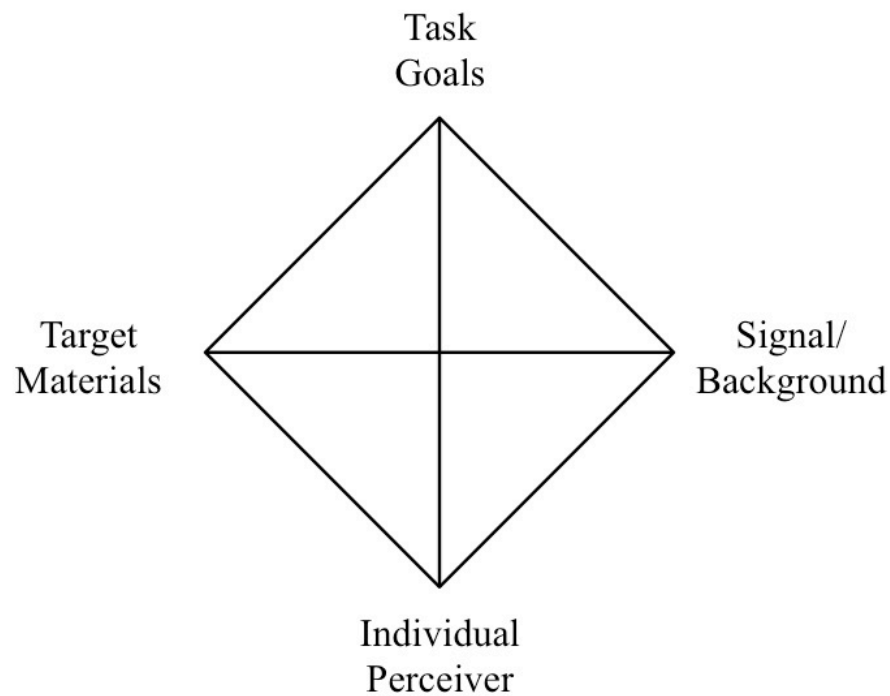


Figure 1. Characterization of the relations among different factors affecting speech perception performance on any given task.

The current study investigated several of these factors and their interactions to better understand non-native speech perception in real-world, everyday communicative environments. In particular, the study addressed the perception of regional dialect variation in an L2, as well as several factors that contribute to non-native listeners' ability to perceive and use robust dialect-specific information in L2 speech. While earlier research suggests that the listener's native

language and previous experience with native or L2 regional dialects affects the perception of regional dialect variation, it is not clear how these factors contribute to and interact with the developing second language phonological system and underlying cognitive functions. Furthermore, the contributions of these factors to the perception of regional dialect variation may vary depending on the type of perception task and its goals. To address these issues, the research reported in this dissertation expands on the previous work dealing with perception and acquisition of regional dialect variation by investigating: the ability of non-native listeners to recognize speech produced by talkers from different regions of the United States (U.S.) and their ability to identify or discriminate regional dialects of American English, sources of individual differences in these perceptual abilities, and the benefits of perceptual regional dialect categorization training in an L2.

The Perception of Regional Dialect Variation in L2 Acquisition

One of the basic objectives of the current study was to investigate the perception of regional dialects of American English by L2 listeners with different amounts of exposure to American English using several different perception tasks. Results from a multi-dialect sentence recognition task, a forced-choice regional dialect categorization task, and an auditory regional dialect free classification task reported in Chapter 2 showed that non-native listeners' perception of L2 regional dialects was greatly influenced by the task and the task demands, the regional dialects used in the tasks, and the listeners' prior language background and experience.

In a multi-dialect sentence recognition task, non-native speakers of American English from South Korea who had resided in the U.S. for a relatively long period of time (KLT: more than three years in the U.S.) and Korean non-native speakers of American English who had

resided in the United States a relatively short period of time (KST: between 6 and 18 months in the U.S.) had more difficulty on the multi-dialect sentence recognition task compared to native speakers of American English (NAE). Although non-native listeners were less accurate than the native listeners on recognizing speech produced by talkers from all four regional dialects used in the task, the non-native Korean listeners' response accuracy suggested that they were influenced by the talker's region of origin, the signal-to-noise ratio (SNR), and by the type of background noise used in the task (i.e., multi-talker babble). Additionally, although performance was strongly influenced by the native language of listener (native American English vs. non-native Korean), the non-native listeners' test scores were only marginally related to length of residency. Taking these findings into consideration, the non-native listeners' performance reflected an interaction of their lack of experience with the American English dialects and each dialect's inherent intelligibility in General American multitalter babble.

For the two indexical processing tasks that were used to assess the listeners' ability to categorize talkers by region of origin and group talkers by region of origin, performance was found to be generally poor for the native NAE listener group and the two non-native KLT and KST listener groups. Categorization accuracy from the categorization task and responses distributions from both tasks reflected perceptual biases of the listeners toward classifying all of the speakers as General American. Although response patterns in both tasks reflected the general bias toward responding with the General American category regardless of the talker's region of origin, this response bias was particularly strong for responses to the Southern talkers who produced variants that were representative of the type of the South Midland or Southern variety spoken in the region where the listeners were residing at the time of testing. In other words, the listeners were more likely to perceive the more frequently encountered Southern or South

Midland speech as General American, compared to Northern and Mid-Atlantic speech samples.

Comparing performance on the forced-choice categorization task and the free classification task, NAE listeners were quite consistent in their responses and they appeared to have used similar perceptual strategies in both tasks. However, both groups of non-native listeners showed different patterns of responses on the two tasks, suggesting that they may have based their judgments on different sets of acoustic-phonetic features depending on the demands of the particular task. The free classification task may have given listeners more opportunity to compare some of the more difficult features to detect or discriminate, allowing them to use more unique idiosyncratic acoustic-phonetic features. In contrast, the forced-choice categorization task may have constrained them to rely more upon general impressions of the accent varieties across large sets of features. Regardless of the small differences in performance on the two indexical processing tasks, a longer length of residency in the U.S. was not found to improve performance in either task, suggesting that these non-native listeners were still in the process of acquiring knowledge of L2 regional dialect characteristics and social knowledge of relations among the dialects.

The findings from the three experiments reported in Chapter 2 indicate that language experience and exposure play important roles in the perception of regional dialects in an L2. The lack of experience and knowledge of the different regional dialects of American English and the absence of robust stable dialect categories in long-term memory made encoding and processing of indexical information in English more difficult for the two groups of non-native listeners compared to the native NAE listeners. However, it is still unclear what type(s) of experiences contribute to better perception of indexical information in the L2. Length of residency did not greatly influence the perception of familiar L2 regional varieties, especially in the two indexical

processing tasks in which the listeners were required to make judgements about the talkers' regions of origin. Furthermore, the non-native listeners did not show a benefit in word recognition for familiar L2 varieties compared to unfamiliar varieties. Since length of residency did not result in greatly improved performance on any of the three perceptual tasks, the findings suggest that the listeners were not exposed to a great amount of regional dialect variation in their L2 language-learning environment that consisted of a college campus with a large Korean student population. Although this study suggests that length of residency did not greatly benefit this group of Korean learners of English in the perception tasks, this may be a special case limited to students on college campuses or living in environments where their native language is frequently spoken. A longitudinal study tracking the perception of L2 regional dialects by a diverse group of L2 learners living in different types of language learning environments would help to further elucidate how L2 learners develop knowledge of L2 sociolinguistic information over time.

Additionally, if the non-native listeners' language experience in the L2 environment had not provided them with sufficient exposure to different regional varieties, their performance on these tasks may reflect other sources of knowledge or language experience. The perception of L2 regional dialects may not only reflect the phonological system of the listeners' native language, but it may also reflect biases from regional dialect characteristics in their native language. Biases may involve expectations for types of acoustic-phonetic differences characteristic of different dialects (e.g., intonational, phonemic, sub-phonemic differences), as well as a priori expectations about the degree to which varieties differ. These expectations would be different for other sources of variability, such as foreign accent, for which non-native listeners may have similar expectations in both their native language and L2. For example, Korean listeners in the current

study may have paid more attention to the features that are characteristic of dialects in their own native language, as reported in a previous study that examined the perception of regional dialects by Mandarin learners of American English (Clopper & Bradlow, 2009). Furthermore, the Korean listeners in the present study may also have expected that the regional dialects of American English would have had more unique characteristics and would have been more perceptually dissimilar or socially marked, since regional varieties of their native Korean may also be perceived as being highly perceptually dissimilar (Long & Yim, 2002). Thus, for the Korean listeners, the dialects in the current study, and potentially in everyday listening environments, may have sounded too similar to be considered different varieties. A larger study with listeners with native languages that differ not only in the characteristics of their sound systems but also in the characteristics of regional dialects would allow us to better understand the native language effects in the perception of L2 sociolinguistic variation. Furthermore, comparing listeners' perception of regional dialects with their perception of other sources of variability in speech, such as foreign accents or international varieties of English, may also shed light on the influence of the listeners' native language on L2 perception.

Sources of Individual Differences on L2 Speech Perception Tasks

Although the amount of exposure to American English regional varieties in the L2 language environment, as measured by length of residency, was not related to performance on the three perception tasks discussed in Chapter 2, other experiences or skills may influence an individual listener's perception of L2 regional dialect variation. Indeed, similar to previous studies on non-native speech recognition (e.g., Tamati & Pisoni, in press), individual non-native KLT and KST listeners varied greatly in their perception of regional dialect variation in English.

As such, analyses of individual listener performance may shed light on the underlying factors that play a role in L2 regional dialect perception, beyond traditional demographic measures like length of residency or age of arrival. Although performance was poor overall for the KST and KLT listener groups, not all non-native listeners performed poorly, and there was substantial listener variability on both the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task. Indeed, taken together, patterns in performance on the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task observed in Chapter 2 suggest that there were large individual differences in the perception of regional dialect variation in the participants' L2 and that while length of residency may have influenced performance, individual variability was not entirely attributable to that factor alone.

To examine the factor(s) that may contribute to the individual differences attested in the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task, the study reported on in Chapter 3 investigated the relation between L2 phonological and lexical knowledge and performance on the multi-dialect sentence recognition task and the dialect categorization task and Chapter 4 examined the relations between several core, underlying neurocognitive abilities and performance on the two perception tasks. Findings from the individual differences analyses showed that that L2 speech perception and spoken word recognition abilities were related to both L2 phonological and lexical knowledge and core, underlying neurocognitive abilities. However, again, sources of individual differences in non-native perception of L2 regional dialects were greatly influenced by the task and the task demands, because the contributions of these sources of individual listener variability to speech perception processes depended greatly on the specific perception task listeners completed. In particular, L2 phonological and lexical knowledge was found to be related to scores on the multi-

dialect sentence recognition task, while several core, foundational neurocognitive abilities were found to be related to scores on both the multi-dialect sentence recognition task and the forced-choice regional dialect categorization task.

Multi-dialect sentence recognition accuracy was strongly related to L2 phonological and lexical knowledge, but was only weakly related to general neurocognitive abilities. The listeners' ability to discriminate difficult American English vowel contrasts may have allowed them to more easily recognize sounds and lexical items in the highly variable listening conditions, because more accurate discrimination of L2 vowel and consonant perception may reflect more native-like representations for sounds and lexical items. Thus, speech recognition in an L2 may be improved by greater development of L2 phonological knowledge, and not just by exposure to the variability in the L2 environment. The results of the present studies also revealed that L2 lexical knowledge was related to multi-dialect sentence recognition, suggesting that lexical knowledge supports non-native word recognition from partial or degraded information in the speech signal. Greater lexical knowledge may allow the listener to make better use of ambiguous or compromised acoustic-phonetic information to recognize words in high-variability sentences in multi-talker babble on the multi-dialect sentence recognition task. These findings suggest that listeners who are better able to acquire knowledge of the vocabulary and phonological system of their L2 may be able to use that information in rapidly adapting and adjusting to multiple sources of variability in speech, and not necessarily variability specific to any particular regional dialect. Performance on the multi-dialect sentence recognition was also related to a few core neurocognitive abilities of the listeners. The sentence recognition task requires the listener to focus attention and control resources on the target sentence and vocal source (talker) while simultaneously ignoring the competing background talkers to reproduce the target sentence.

Thus, core attentional control and inhibitory processes may contribute to the encoding of detailed acoustic-phonetic information in the signal that is crucial for the successful discrimination and/or identification of novel sounds and L2 lexical access.

Analyses of individual differences on the forced-choice regional dialect categorization task showed somewhat contrasting trends compared to the results obtained for performance on the sentence recognition task. Based on findings reported on in Chapter 3, greater phonological and lexical knowledge of the L2 may not help the listener in perceiving and using regional dialect information to make judgements about a talker based on his/her region of origin, suggesting that listeners, regardless of how accurately they were able to discriminate or identify broad sound categories of the L2 phonological system, were sensitive to some of the subphonemic vowel or consonant variability characteristic of different regional varieties. Still, some individual listeners may have benefited from using neurocognitive strategies in the categorization task. Neurocognitive abilities are involved in regional dialect categorization in a variety of ways. First, a listener with better attentional and inhibitory control skills and executive functioning may be better at shifting their attention to the indexical information encoded in the speech signal when this channel of information is relevant to the communicative environment and task demands. Additionally, listeners with strong attentional and inhibitory control and executive function skills may also be able to ignore other sources of competition and variability in the speech signal, such as variability in the linguistic content of the utterance, to form more robust representations of the L2 regional dialect variation. Thus, having better foundational attentional and inhibitory control skills may help non-native listeners to selectively attend to and use the regional dialect information in the target utterance to be able to acquire sociolinguistic knowledge of regional dialects and regional stereotypes. Similarly, within the task itself, those

listeners may also more easily adapt to the changing linguistic content and/or additional talker variability from trial-to-trial to focus their attention on the dialect-specific cues present on a particular trial to make use of this source of context to facilitate regional dialect categorization.

Although the findings from the present research suggest that attentional and inhibitory control and executive functioning skills may be related to the perceptual categorization of regional dialect variation, these measures accounted for very little variance observed in performance. As such, the precise sources of individual differences in regional dialect identification or categorization still remain unclear. In addition to using other behavioral tasks to measure L2 regional dialect identification or discrimination ability, future studies on individual differences in L2 speech perception should include other performance-based tasks assessing neurocognitive performance measures, such as cognitive control processes used in attention/inhibition, short-term and working memory, and processing speed. The studies reported in Chapters 3 and 4 used a limited test battery covering a narrow range of L2 perceptual abilities and information processing skills to begin to elucidate many potential sources of individual differences. Individual listener variability in non-native speech perception and spoken word recognition skills likely reflects the use of multiple neurocognitive processes and associated perceptual and cognitive abilities working together as an integrated system. Indexical processing skills of non-native listeners using other sources of indexical information, such as foreign accent or emotions in speech, as well as tests of L2 phonological awareness should be explored further using a broad range of information processing and psycholinguistic tasks under various settings or conditions. Similarly, other neurocognitive performance measures, such as cognitive control processes used in attention/inhibition, short-term and working memory, and processing speed,

should be assessed in future studies of individual differences in non-native speech recognition in real-world listening environments and the acquisition of L2 sociolinguistic knowledge.

Processing of Linguistic and Indexical Information in Speech

Another goal of the studies reported here was to explore the relation between the perception of linguistic information in speech and the perception of nonlinguistic (indexical) information in speech. Given the findings from previous studies on individual differences in speech perception, the individual listener's ability to recognize words produced by talkers from different regions of origins and his/her ability to discriminate and identify regional dialects were expected to be closely related. Individual listeners may benefit from more robust, highly detailed phonological and lexical representations and memory codes in long-term memory for both signal and context information in processing linguistic and indexical information, as used in sentence recognition and indexical categorization or discrimination tasks.

Examination of patterns of performance on the multi-dialect sentence recognition task and the regional dialect categorization task described in Chapter 2 revealed that individual listeners who had developed more native-like perception of regional dialects in the indexical processing tasks displayed better speech recognition scores across multiple dialects, reflecting a benefit from familiarity not just to a single dialect but from the general perceptual organization of several regional dialects of the L2. The analyses on individual differences suggest that the ability to recognize highly variable words and sentences may be related to indexical processing skills because individual listeners may benefit from having more robust, highly detailed representations and memory codes for American English regional dialect variation in long-term memory on both linguistic and indexical processing tasks.

Although performance on the multi-dialect sentence recognition task and the regional dialect categorization task was found to be related, as shown in the study reported in Chapter 2, sources of individual differences in overall patterns of performance on the linguistic and indexical processing tasks were quite diverse in the studies reported in Chapters 3 and 4. As mentioned above, differences among listeners on those two tasks were found to be related to different sources of individual variability. The results from Chapter 3 showed that linguistic knowledge was related to multi-dialect sentence recognition accuracy, although it was not related to regional dialect categorization accuracy. Additionally, results from Chapter 4 showed that a few neurocognitive abilities were moderately related to individual listener performance on both the linguistic and indexical processing tasks. Thus, recognizing words produced by talkers from several U.S. dialect regions was related to both linguistic knowledge and neurocognitive abilities, but the ability to explicitly categorize talkers by region of origin was only related to neurocognitive abilities.

The dissociation between linguistic processing tasks and indexical processing tasks was replicated in the study reported in Chapter 5. In that study, native and non-native listeners completed a regional dialect categorization training task. The training study investigated the benefit of regional dialect categorization training on the learning of dialect-specific features in the native language and L2 to improve the listeners' ability to identify unfamiliar talkers' regions of origin. A comparison of good and poor learners in the training task demonstrated that the amount of benefit a listener received from the training experience was not related to his or her native language (American English or Korean) or, for non-native listeners, how long they had lived in the U.S. However, the good regional dialect learners had a greater overall ability to perceive and use reliable dialect-specific features to categorize talkers by region of origin. These

results suggest that language background and experience was not likely related to the perceptual learning of regional dialect variation, and instead regional dialect perceptual learning may be related to individual perceptual and/or neurocognitive skills.

The study reported in Chapter 5 also examined whether the perceptual learning in the training task would transfer to the recognition of words and speech sounds in American English in several transfer perceptual tasks. In the high-variability explicit regional dialect categorization training task, both native and non-native listeners who had received the training demonstrated modest improvement in their ability to explicitly categorize familiar and unfamiliar talkers after the dialect categorization training. However, the listeners showed no evidence improvement on a multi-dialect sentence recognition task, an ABX American English vowel and consonant discrimination task, or an American English vowel identification task. Further analyses of the good and poor learners also showed that the perceptual learning in the training task did not facilitate the perception of words and speech sounds in the transfer tasks even for the good learners. Although the lack of transfer of training effects may be related to the specific design of the current training study, which may not have resulted in substantial learning of regional dialect variation or ample opportunity for consolidation of learning, these findings suggest that the amount of transfer from learning social-indexical categories to the processing of linguistic information may be limited, at least under these specific conditions.

Taken together, the studies reported in Chapters 2-5 suggest that although there may be some overlap because performance on the two tasks both benefit from robust representations for linguistic and indexical variation in speech, different core, underlying processes may support the perception of indexical information in indexical processing tasks and linguistic information in linguistic processing tasks even when the task materials include a large amount of indexical

variability. Although the current study does not directly address this issue, L2 social-indexical knowledge and L2 phonological and lexical knowledge may be acquired at different stages in acquisition. In other words, learners who have gained more knowledge of the phonological and phonotactic system of the L2 and the vocabulary of the L2 may not display improvements in social-indexical knowledge, and vice versa. Although greater exposure and experience with the language may lead to more robust representations for L2 words and sound segments, as well as social-indexical categories, it is not case that greater exposure and experience necessarily result in greater knowledge in both areas. This may be what occurred for the non-native listeners in the current study, who received a great amount of instruction about the English language but likely have not been exposed sufficiently to many American English regional varieties due to the university and community environment, where there are many foreign students, including Korean students, and where a more standard variety of English may be most frequency encountered. Again, a longitudinal study with L2 learners with different types and amounts of experience and exposure to the target language would be helpful in exploring patterns of language acquisition and the relation between the development of the perception of linguistic and indexical information in speech in an L2.

Final Thoughts and Conclusions

In order to gain a better understanding of how regional dialect variation in an L2 is perceived and processed, the studies reported in Chapter 2-5 investigated the perception of American English regional dialects by Korean learners of English residing in the U.S. In particular, the research reported here examined non-native listeners' ability to recognize speech produced by talkers from different U.S. dialect regions as well as their ability to identify and

discriminate regional dialects of American English. Furthermore, the present research expands on the previous studies on the perception of regional dialect variation by exploring many potential sources of individual differences in these perceptual abilities, specifically those related to L2 phonological and lexical knowledge and several core, underlying neurocognitive abilities individual differences. Finally, the benefits of perceptual regional dialect categorization training on the learning of acoustic-phonetic features of L2 regional dialects were also assessed. Findings from the studies reported here provide new knowledge about the perception of regional dialect variation in L2 acquisition, sources of individual differences in L2 speech perception, and the relations between the processing of linguistic and indexical information in speech in an L2.

In addition to the suggestions for future research directions sketched above regarding each of these areas, additional research should be carried out to explore the factors that contribute to non-native listeners' ability to perceive, encode, and store robust dialect-specific information in real-world, day-to-day life. The current study used regional dialect variation to explore the perception of L2 social-indexical variation using both indexical and linguistic processing tasks. However, it is not clear to what extent regional dialect variation impacts everyday speech communication in an L2 language learning environment. Identifying potential areas of difficulty in real-world speech perception and spoken word recognition may help create individualized targeted instruction methods and can serve as a basis to motivate the development of novel training programs to improve speech recognition and understanding for L2 listeners.

References

- Clopper, C. G. & Bradlow, A. R. (2009). Free Classification of American English dialects by native and non-native listeners. *Journal of Phonetics*, 37, 436-451.
- García Lecumberri, M. L, Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, 52, 864– 886.
- Gilbert, J. L., Tamati, T. N., & Pisoni, D.B. (2013). Development, Reliability and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *Journal of the American Academy of Audiology*, 24, 1-11.
- Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of Processing in Human Memory* (pp. 429-446). Hillsdale, NJ: Erlbaum.
- Long, D., & Yim, Y.-C. (2002). Regional differences in the perception of Korean dialects. In D. Long & D. R. Preston (Eds.), *Handbook of Perceptually Dialectology, Volume 2* (pp. 249-275). Amsterdam: John Benjamins.
- Mattys, S., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse listening conditions: A review. *Language and Cognitive Processes*, 2, 953-978.
- Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 9-32). San Diego, CA: Academic Press.
- Roediger, H. L. (2008). Relativity of remembering: why the laws of memory vanished. *Annual Review of Psychology*, 2, 189-198.

Tamati, T. N., & Pisoni, D. B. (in press). Non-native speech recognition in high-variability listening conditions using PRESTO. Submitted to the *Journal of the American Academy of Audiology*.

Theunissen, M., Swanepoel, D. W., & Hanekom, J. (2009). Sentence recognition in noise: Variables in compilation and interpretation of tests. *International Journal of Audiology*, 48, 743-757.